

1 Polygenic Risk Modelling for Prediction of Epithelial Ovarian Cancer Risk

2 Eileen O. Dareng^{1*}, Jonathan P. Tyrer^{2*}, Daniel R. Barnes¹, Michelle R. Jones³, Xin Yang¹,
3 Katja K.H. Aben^{4, 5}, Muriel A. Adank⁶, Simona Agata⁷, Irene L. Andrulis^{8, 9}, Hoda Anton-
4 Culver¹⁰, Natalia N. Antonenkova¹¹, Gerasimos Aravantinos¹², Banu K. Arun¹³, Annelie
5 Augustinsson¹⁴, Judith Balmaña^{15, 16}, Elisa V. Bandera¹⁷, Rosa B. Barkardottir^{18, 19}, Daniel
6 Barrowdale¹, Matthias W. Beckmann²⁰, Alicia Beeghly-Fadiel²¹, Javier Benitez^{22, 23}, Marina
7 Bermisheva²⁴, Marcus Q. Bernardini²⁵, Line Bjorge^{26, 27}, Amanda Black²⁸, Natalia V.
8 Bogdanova^{11, 29, 30}, Bernardo Bonanni³¹, Ake Borg³², James D. Brenton³³, Agnieszka
9 Budzilowska³⁴, Ralf Butzow³⁵, Saundra S. Buys³⁶, Hui Cai²¹, Maria A. Caligo³⁷, Ian Campbell³⁸,
10 ³⁹, Rikki Cannioto⁴⁰, Hayley Cassingham⁴¹, Jenny Chang-Claude^{42, 43}, Stephen J. Chanock⁴⁴,
11 Kexin Chen⁴⁵, Yoke-Eng Chiew^{46, 47}, Wendy K. Chung⁴⁸, Kathleen B.M. Claes⁴⁹, Sarah
12 Colanna³⁶, GEMO Study Collaborators⁵⁰⁻⁵², GC-HBOC study Collaborators⁵³, EMBRACE
13 Collaborators¹, Linda S. Cook^{54, 55}, Fergus J. Couch⁵⁶, Mary B. Daly⁵⁷, Fanny Dao⁵⁸, Eleanor
14 Davies⁵⁹, Miguel de la Hoya⁶⁰, Robin de Putter⁴⁹, Joe Dennis¹, Allison DePersia^{61, 62}, Peter
15 Devilee^{63, 64}, Orland Diez^{65, 66}, Yuan Chun Ding⁶⁷, Jennifer A. Doherty⁶⁸, Susan M. Domchek⁶⁹,
16 Thilo Dörk³⁰, Andreas du Bois^{70, 71}, Matthias Dürst⁷², Diana M. Eccles⁷³, Heather A. Eliassen⁷⁴,
17 ⁷⁵, Christoph Engel^{76, 77}, D. Gareth Evans^{78, 79}, Peter A. Fasching^{20, 80}, James M. Flanagan⁸¹,
18 Lenka Foretova⁸², Renée T. Fortner⁴², Eitan Friedman^{83, 84}, Patricia A. Ganz⁸⁵, Judy Garber⁸⁶,
19 Francesca Gensini⁸⁷, Graham G. Giles⁸⁸⁻⁹⁰, Gord Glendon⁸, Andrew K. Godwin⁹¹, Marc T.
20 Goodman⁹², Mark H. Greene⁹³, Jacek Gronwald⁹⁴, OPAL Study Group⁹⁵, AOCS Group^{38, 46},
21 Eric Hahnen^{53, 96}, Christopher A. Haiman⁹⁷, Niclas Håkansson⁹⁸, Ute Hamann⁹⁹, Thomas V.O.
22 Hansen¹⁰⁰, Holly R. Harris^{101, 102}, Mikael Hartman^{103, 104}, Florian Heitz^{70, 71, 105}, Michelle A.T.
23 Hildebrandt¹⁰⁶, Estrid Høgdall^{107, 108}, Claus K. Høgdall¹⁰⁹, John L. Hopper⁸⁹, Ruea-Yea
24 Huang¹¹⁰, Chad Huff¹⁰⁶, Peter J. Hulick^{61, 62}, David G. Huntsman¹¹¹⁻¹¹⁴, Evgeny N. Imyanitov¹¹⁵,
25 KConFab Investigators³⁸, HEBON Investigators¹¹⁶, Claudine Isaacs¹¹⁷, Anna Jakubowska^{94, 118},
26 Paul A. James^{39, 119}, Ramunas Janavicius^{120, 121}, Allan Jensen¹⁰⁷, Oskar Th. Johannsson¹²², Esther
27 M. John^{123, 124}, Michael E. Jones¹²⁵, Daehee Kang¹²⁶⁻¹²⁸, Beth Y. Karlan¹²⁹, Anthony
28 Karnezis¹³⁰, Linda E. Kelemen¹³¹, Elza Khusnutdinova^{24, 132}, Lambertus A. Kiemeny⁴, Byoung-
29 Gie Kim¹³³, Susanne K. Kjaer^{107, 109}, Ian Komenaka¹³⁴, Jolanta Kupryjanczyk³⁴, Allison W.
30 Kurian^{123, 124}, Ava Kwong¹³⁵⁻¹³⁷, Diether Lambrechts^{138, 139}, Melissa C. Larson¹⁴⁰, Conxi
31 Lazaro¹⁴¹, Nhu D. Le¹⁴², Goska Leslie¹, Jenny Lester¹²⁹, Fabienne Lesueur^{51, 52, 143}, Douglas A.
32 Levine^{58, 144}, Lian Li⁴⁵, Jingmei Li¹⁴⁵, Jennifer T. Loud⁹³, Karen H. Lu¹⁴⁶, Jan Lubiński⁹⁴, Eva
33 Machackova⁸², Phuong L. Mai¹⁴⁷, Siranoush Manoukian¹⁴⁸, Jeffrey R. Marks¹⁴⁹, Rayna Kim
34 Matsuno¹⁵⁰, Keitaro Matsuo^{151, 152}, Taymaa May²⁵, Lesley McGuffog¹, John R. McLaughlin¹⁵³,

35 Iain A. McNeish^{154, 155}, Noura Mebirouk^{51, 52, 143}, Usha Menon¹⁵⁶, Austin Miller¹⁵⁷, Roger L.
36 Milne⁸⁸⁻⁹⁰, Albina Minlikeeva¹⁵⁸, Francesmary Modugno^{159, 160}, Marco Montagna⁷, Kirsten B.
37 Moysich¹⁵⁸, Elizabeth Munro^{161, 162}, , Katherine L. Nathanson⁶⁹, Susan L. Neuhausen⁶⁷, Heli
38 Nevanlinna¹⁶³, Joanne Ngeow Yuen Yie^{164, 165}, Henriette Roed Nielsen¹⁶⁶, Finn C. Nielsen¹⁰⁰,
39 Liene Nikitina-Zake¹⁶⁷, Kunle Odunsi¹⁶⁸, Kenneth Offit^{169, 170}, Edith Olah¹⁷¹, Siel Olbrecht¹⁷²,
40 Olufunmilayo I. Olopade¹⁷³, Sara H. Olson¹⁷⁴, Håkan Olsson¹⁴, Ana Osorio^{23, 175}, Laura Papi⁸⁷,
41 Sue K. Park¹²⁶⁻¹²⁸, Michael T. Parsons¹⁷⁶, Harsha Pathak⁹¹, Inge Sokilde Pedersen¹⁷⁷⁻¹⁷⁹, Ana
42 Peixoto¹⁸⁰, Tanja Pejovic^{161, 162}, Pedro Perez-Segura⁶⁰, Jennifer B. Permut¹⁸¹, Beth Peshkin¹¹⁷,
43 Paolo Peterlongo¹⁸², Anna Piskorz³³, Darya Prokofyeva¹⁸³, Paolo Radice¹⁸⁴, Johanna Rantala¹⁸⁵,
44 Marjorie J. Riggan¹⁸⁶, Harvey A. Risch¹⁸⁷, Cristina Rodriguez-Antona^{22, 23}, Eric Ross¹⁸⁸, Mary
45 Anne Rossing^{101, 102}, Ingo Runnebaum⁷², Dale P. Sandler¹⁸⁹, Marta Santamariña^{175, 190, 191}, Penny
46 Soucy¹⁹², Rita K. Schmutzler^{53, 96, 193}, V. Wendy Setiawan⁹⁷, Kang Shan¹⁹⁴, Weiva Sieh^{195, 196},
47 Jacques Simard¹⁹⁷, Christian F. Singer¹⁹⁸, Anna P Sokolenko¹¹⁵, Honglin Song¹⁹⁹, Melissa C.
48 Southey^{88, 90, 200}, Helen Steed²⁰¹, Dominique Stoppa-Lyonnet^{50, 202, 203}, Rebecca Sutphen²⁰⁴,
49 Anthony J. Swerdlow^{125, 205}, Yen Yen Tan²⁰⁶, Manuel R. Teixeira^{180, 207}, Soo Hwang Teo^{208, 209},
50 Kathryn L. Terry^{74, 210}, Mary Beth Terry²¹¹, Mads Thomassen¹⁶⁶, Pamela J. Thompson⁹², Liv
51 Cecilie Vestrheim Thomsen^{26, 27}, Darcy L. Thull²¹², Marc Tischkowitz^{213, 214}, Linda Titus²¹⁵,
52 Amanda E. Toland²¹⁶, Diana Torres^{99, 217}, Britton Trabert²⁸, Ruth Travis²¹⁸, Nadine Tung²¹⁹,
53 Shelley S. Tworoger^{181, 220}, Ellen Valen^{26, 27}, Anne M. van Altena⁴, Annemieke H. van der
54 Hout²²¹, Els Van Nieuwenhuysen¹⁷², Elizabeth J. van Rensburg²²², Ana Vega²²³⁻²²⁵, Digna Velez
55 Edwards²²⁶, Robert A. Vierkant¹⁴⁰, Frances Wang^{227, 228}, Barbara Wappenschmidt^{53, 96},
56 Penelope M. Webb⁹⁵, Clarice R. Weinberg²²⁹, Jeffrey N. Weitzel²³⁰, Nicolas Wentzensen²⁸,
57 Emily White^{102, 231}, Alice S. Whittemore^{123, 232}, Stacey J. Winham¹⁴⁰, Alicja Wolk^{98, 233}, Yin-Ling
58 Woo²³⁴, Anna H. Wu⁹⁷, Li Yan²³⁵, Drakoulis Yannoukacos²³⁶, Katia M. Zavaglia³⁷, Wei
59 Zheng²¹, Argyrios Ziogas¹⁰, Kristin K. Zorn¹⁴⁷, Douglas Easton^{1,2}, Kate Lawrenson^{3, 237}, Anna
60 DeFazio^{46, 47}, Thomas A. Sellers²³⁸, Susan J. Ramus^{239, 240}, Celeste L. Pearce^{241, 242}, Alvaro N.
61 Monteiro¹⁸¹, Julie Cunningham²⁴³, Ellen L. Goode²⁴³, Joellen M. Schildkraut²⁴⁴, Andrew
62 Berchuck¹⁸⁶, Georgia Chenevix-Trench¹⁷⁶, Simon A. Gayther³, Antonis C. Antoniou¹, Paul D.P.
63 Pharoah^{1,2}

64

65 ¹ University of Cambridge, Centre for Cancer Genetic Epidemiology, Department of Public
66 Health and Primary Care, Cambridge, UK.

67 ² University of Cambridge, Centre for Cancer Genetic Epidemiology, Department of Oncology,
68 Cambridge, UK.

- 69 ³ Center for Bioinformatics and Functional Genomics, Cedars-Sinai Medical Center, Los
70 Angeles, CA, USA.
- 71 ⁴ Radboud University Medical Center, Radboud Institute for Health Sciences, Nijmegen, The
72 Netherlands.
- 73 ⁵ Netherlands Comprehensive Cancer Organisation, Utrecht, The Netherlands.
- 74 ⁶ The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Family Cancer Clinic,
75 Amsterdam, The Netherlands.
- 76 ⁷ Veneto Institute of Oncology IOV - IRCCS, Immunology and Molecular Oncology Unit,
77 Padua, Italy.
- 78 ⁸ Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Fred A. Litwin Center for
79 Cancer Genetics, Toronto, ON, Canada.
- 80 ⁹ University of Toronto, Department of Molecular Genetics, Toronto, ON, Canada.
- 81 ¹⁰ University of California Irvine, Department of Epidemiology, Genetic Epidemiology
82 Research Institute, Irvine, CA, USA.
- 83 ¹¹ N.N. Alexandrov Research Institute of Oncology and Medical Radiology, Minsk, Belarus.
- 84 ¹² 'Agii Anargiri' Cancer Hospital, Athens, Greece.
- 85 ¹³ University of Texas MD Anderson Cancer Center, Department of Breast Medical Oncology,
86 Houston, TX, USA.
- 87 ¹⁴ Lund University, Department of Cancer Epidemiology, Clinical Sciences, Lund, Sweden.
- 88 ¹⁵ Vall d'Hebron Institute of Oncology, Hereditary cancer Genetics Group, Barcelona, Spain.
- 89 ¹⁶ University Hospital of Vall d'Hebron, Department of Medical Oncology, Barcelona, Spain.
- 90 ¹⁷ Rutgers Cancer Institute of New Jersey, Cancer Prevention and Control Program, New
91 Brunswick, NJ, USA.
- 92 ¹⁸ Landspítali University Hospital, Department of Pathology, Reykjavik, Iceland.
- 93 ¹⁹ University of Iceland, BMC (Biomedical Centre), Faculty of Medicine, Reykjavik, Iceland.
- 94 ²⁰ University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nuremberg,
95 Department of Gynecology and Obstetrics, Comprehensive Cancer Center ER-EMN, Erlangen,
96 Germany.
- 97 ²¹ Vanderbilt University School of Medicine, Division of Epidemiology, Department of
98 Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Nashville, TN,
99 USA.
- 100 ²² Biomedical Network on Rare Diseases (CIBERER), Madrid, Spain.
- 101 ²³ Spanish National Cancer Research Centre (CNIO), Human Cancer Genetics Programme,
102 Madrid, Spain.

- 103 ²⁴ Ufa Federal Research Centre of the Russian Academy of Sciences, Institute of Biochemistry
104 and Genetics, Ufa, Russia.
- 105 ²⁵ Princess Margaret Hospital, Division of Gynecologic Oncology, University Health Network,
106 Toronto, Ontario, Canada.
- 107 ²⁶ Haukeland University Hospital, Department of Obstetrics and Gynecology, Bergen, Norway.
- 108 ²⁷ University of Bergen, Centre for Cancer Biomarkers CCBIO, Department of Clinical Science,
109 Bergen, Norway.
- 110 ²⁸ National Cancer Institute, Division of Cancer Epidemiology and Genetics, Bethesda, MD,
111 USA.
- 112 ²⁹ Hannover Medical School, Department of Radiation Oncology, Hannover, Germany.
- 113 ³⁰ Hannover Medical School, Gynaecology Research Unit, Hannover, Germany.
- 114 ³¹ IEO, European Institute of Oncology IRCCS, Division of Cancer Prevention and Genetics,
115 Milan, Italy.
- 116 ³² Lund University and Skåne University Hospital, Department of Oncology, Lund, Sweden.
- 117 ³³ Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK.
- 118 ³⁴ Maria Skłodowska-Curie National Research Institute of Oncology, Department of Pathology
119 and Laboratory Diagnostics, Warsaw, Poland.
- 120 ³⁵ University of Helsinki, Department of Pathology, Helsinki University Hospital, Helsinki,
121 Finland.
- 122 ³⁶ Huntsman Cancer Institute, Department of Medicine, Salt Lake City, UT, USA.
- 123 ³⁷ University Hospital, SOD Genetica Molecolare, Pisa, Italy.
- 124 ³⁸ Peter MacCallum Cancer Center, Melbourne, Victoria, Australia.
- 125 ³⁹ The University of Melbourne, Sir Peter MacCallum Department of Oncology, Melbourne,
126 Victoria, Australia.
- 127 ⁴⁰ Roswell Park Cancer Institute, Cancer Pathology & Prevention, Division of Cancer
128 Prevention and Population Sciences, Buffalo, NY, USA.
- 129 ⁴¹ Division of Human Genetics, The Ohio State University, Department of Internal Medicine,
130 Columbus, OH, USA.
- 131 ⁴² German Cancer Research Center (DKFZ), Division of Cancer Epidemiology, Heidelberg,
132 Germany.
- 133 ⁴³ University Medical Center Hamburg-Eppendorf, Cancer Epidemiology Group, University
134 Cancer Center Hamburg (UCCH), Hamburg, Germany.
- 135 ⁴⁴ National Cancer Institute, National Institutes of Health, Department of Health and Human
136 Services, Division of Cancer Epidemiology and Genetics, Bethesda, MD, USA.

- 137 ⁴⁵ Tianjin Medical University Cancer Institute and Hospital, Department of Epidemiology,
138 Tianjin, China.
- 139 ⁴⁶ The University of Sydney, Centre for Cancer Research, The Westmead Institute for Medical
140 Research, Sydney, New South Wales, Australia.
- 141 ⁴⁷ Westmead Hospital, Department of Gynaecological Oncology, Sydney, New South Wales,
142 Australia.
- 143 ⁴⁸ Columbia University, Departments of Pediatrics and Medicine, New York, NY, USA.
- 144 ⁴⁹ Ghent University, Centre for Medical Genetics, Gent, Belgium.
- 145 ⁵⁰ INSERM U830, Department of Tumour Biology, Paris, France.
- 146 ⁵¹ Institut Curie, Paris, France.
- 147 ⁵² Mines ParisTech, Fontainebleau, France.
- 148 ⁵³ Faculty of Medicine and University Hospital Cologne, University of Cologne, Center for
149 Familial Breast and Ovarian Cancer, Cologne, Germany.
- 150 ⁵⁴ University of New Mexico, University of New Mexico Health Sciences Center, Albuquerque,
151 NM, USA.
- 152 ⁵⁵ Alberta Health Services, Department of Cancer Epidemiology and Prevention Research,
153 Calgary, AB, Canada.
- 154 ⁵⁶ Mayo Clinic, Department of Laboratory Medicine and Pathology, Rochester, MN, USA.
- 155 ⁵⁷ Fox Chase Cancer Center, Department of Clinical Genetics, Philadelphia, PA, USA.
- 156 ⁵⁸ Memorial Sloan Kettering Cancer Center, Gynecology Service, Department of Surgery, New
157 York, NY, USA.
- 158 ⁵⁹ Cambridge, Cambridge, UK.
- 159 ⁶⁰ CIBERONC, Hospital Clinico San Carlos, IdISSC (Instituto de Investigación Sanitaria del
160 Hospital Clínico San Carlos), Molecular Oncology Laboratory, Madrid, Spain.
- 161 ⁶¹ NorthShore University Health System, Center for Medical Genetics, Evanston, IL, USA.
- 162 ⁶² The University of Chicago Pritzker School of Medicine, Chicago, IL, USA.
- 163 ⁶³ Leiden University Medical Center, Department of Pathology, Leiden, The Netherlands.
- 164 ⁶⁴ Leiden University Medical Center, Department of Human Genetics, Leiden, The Netherlands.
- 165 ⁶⁵ Vall dHebron Institute of Oncology (VHIO), Oncogenetics Group, Barcelona, Spain.
- 166 ⁶⁶ University Hospital Vall dHebron, Clinical and Molecular Genetics Area, Barcelona, Spain.
- 167 ⁶⁷ Beckman Research Institute of City of Hope, Department of Population Sciences, Duarte, CA,
168 USA.

169 ⁶⁸ University of Utah, Huntsman Cancer Institute, Department of Population Health Sciences,
170 Salt Lake City, UT, USA.

171 ⁶⁹ University of Pennsylvania, Basser Center for BRCA, Abramson Cancer Center,
172 Philadelphia, PA, USA.

173 ⁷⁰ Ev. Kliniken Essen-Mitte (KEM), Department of Gynecology and Gynecologic Oncology,
174 Essen, Germany.

175 ⁷¹ Dr. Horst Schmidt Kliniken Wiesbaden, Department of Gynecology and Gynecologic
176 Oncology, Wiesbaden, Germany.

177 ⁷² Jena University Hospital - Friedrich Schiller University, Department of Gynaecology, Jena,
178 Germany.

179 ⁷³ University of Southampton, Faculty of Medicine, Southampton, UK.

180 ⁷⁴ Harvard T.H. Chan School of Public Health, Department of Epidemiology, Boston, MA,
181 USA.

182 ⁷⁵ Brigham and Women's Hospital and Harvard Medical School, Channing Division of Network
183 Medicine, Boston, MA, USA.

184 ⁷⁶ University of Leipzig, Institute for Medical Informatics, Statistics and Epidemiology, Leipzig,
185 Germany.

186 ⁷⁷ University of Leipzig, LIFE - Leipzig Research Centre for Civilization Diseases, Leipzig,
187 Germany.

188 ⁷⁸ University of Manchester, Manchester Academic Health Science Centre, Division of
189 Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine
190 and Health, Manchester, UK.

191 ⁷⁹ St Mary's Hospital, Manchester University NHS Foundation Trust, Manchester Academic
192 Health Science Centre, North West Genomics Laboratory Hub, Manchester Centre for Genomic
193 Medicine, Manchester, UK.

194 ⁸⁰ University of California at Los Angeles, David Geffen School of Medicine, Department of
195 Medicine Division of Hematology and Oncology, Los Angeles, CA, USA.

196 ⁸¹ Imperial College London, Division of Cancer and Ovarian Cancer Action Research Centre,
197 Department of Surgery and Cancer, London, UK.

198 ⁸² Masaryk Memorial Cancer Institute, Department of Cancer Epidemiology and Genetics,
199 Brno, Czech Republic.

200 ⁸³ Chaim Sheba Medical Center, The Susanne Levy Gertner Oncogenetics Unit, Ramat Gan,
201 Israel.

202 ⁸⁴ Tel Aviv University, Sackler Faculty of Medicine, Ramat Aviv, Israel.

203 ⁸⁵ Jonsson Comprehensive Cancer Centre, UCLA, Schools of Medicine and Public Health,
204 Division of Cancer Prevention & Control Research, Los Angeles, CA, USA.

- 205 ⁸⁶ Dana-Farber Cancer Institute, Cancer Risk and Prevention Clinic, Boston, MA, USA.
- 206 ⁸⁷ University of Florence, Department of Experimental and Clinical Biomedical Sciences 'Mario
207 Serio', Medical Genetics Unit, Florence, Italy.
- 208 ⁸⁸ Cancer Council Victoria, Cancer Epidemiology Division, Melbourne, Victoria, Australia.
- 209 ⁸⁹ The University of Melbourne, Centre for Epidemiology and Biostatistics, Melbourne School
210 of Population and Global Health, Melbourne, Victoria, Australia.
- 211 ⁹⁰ Monash University, Precision Medicine, School of Clinical Sciences at Monash Health,
212 Clayton, Victoria, Australia.
- 213 ⁹¹ University of Kansas Medical Center, Department of Pathology and Laboratory Medicine,
214 Kansas City, KS, USA.
- 215 ⁹² Cedars-Sinai Medical Center, Samuel Oschin Comprehensive Cancer Institute, Cancer
216 Prevention and Genetics Program, Los Angeles, CA, USA.
- 217 ⁹³ National Cancer Institute, Clinical Genetics Branch, Division of Cancer Epidemiology and
218 Genetics, Bethesda, MD, USA.
- 219 ⁹⁴ Pomeranian Medical University, Department of Genetics and Pathology, Szczecin, Poland.
- 220 ⁹⁵ QIMR Berghofer Medical Research Institute, Population Health Department, Brisbane,
221 Queensland, Australia.
- 222 ⁹⁶ Faculty of Medicine and University Hospital Cologne, University of Cologne, Center for
223 Integrated Oncology (CIO), Cologne, Germany.
- 224 ⁹⁷ University of Southern California, Department of Preventive Medicine, Keck School of
225 Medicine, Los Angeles, CA, USA.
- 226 ⁹⁸ Karolinska Institutet, Institute of Environmental Medicine, Stockholm, Sweden.
- 227 ⁹⁹ German Cancer Research Center (DKFZ), Molecular Genetics of Breast Cancer, Heidelberg,
228 Germany.
- 229 ¹⁰⁰ Rigshospitalet, Copenhagen University Hospital, Department of Clinical Genetics,
230 Copenhagen, Denmark.
- 231 ¹⁰¹ Fred Hutchinson Cancer Research Center, Program in Epidemiology, Division of Public
232 Health Sciences, Seattle, WA, USA.
- 233 ¹⁰² University of Washington, Department of Epidemiology, Seattle, WA, USA.
- 234 ¹⁰³ National University of Singapore and National University Health System, Saw Swee Hock
235 School of Public Health, Singapore, Singapore.
- 236 ¹⁰⁴ National University Health System, Department of Surgery, Singapore, Singapore.
- 237 ¹⁰⁵ Humboldt-Universität zu Berlin, and Berlin Institute of Health, Department for Gynecology
238 with the Center for Oncologic Surgery Charitǝ Campus Virchow-Klinikum, Charitǝ Ā
239 Universitǝtsmedizin Berlin, corporate member of Freie Universitǝt Berlin, Berlin, Germany.

- 240 ¹⁰⁶ University of Texas MD Anderson Cancer Center, Department of Epidemiology, Houston,
241 TX, USA.
- 242 ¹⁰⁷ Danish Cancer Society Research Center, Department of Virus, Lifestyle and Genes,
243 Copenhagen, Denmark.
- 244 ¹⁰⁸ University of Copenhagen, Molecular Unit, Department of Pathology, Herlev Hospital,
245 Copenhagen, Denmark.
- 246 ¹⁰⁹ University of Copenhagen, Department of Gynaecology, Rigshospitalet, Copenhagen,
247 Denmark.
- 248 ¹¹⁰ Roswell Park Cancer Institute, Center For Immunotherapy, Buffalo, NY, USA.
- 249 ¹¹¹ BC Cancer, Vancouver General Hospital, and University of British Columbia, British
250 Columbia's Ovarian Cancer Research (OVCARE) Program, Vancouver, BC, Canada.
- 251 ¹¹² University of British Columbia, Department of Pathology and Laboratory Medicine,
252 Vancouver, BC, Canada.
- 253 ¹¹³ University of British Columbia, Department of Obstetrics and Gynecology, Vancouver, BC,
254 Canada.
- 255 ¹¹⁴ BC Cancer Research Centre, Department of Molecular Oncology, Vancouver, BC, Canada.
- 256 ¹¹⁵ N.N. Petrov Institute of Oncology, St. Petersburg, Russia.
- 257 ¹¹⁶ Coordinating center: The Netherlands Cancer Institute, The Hereditary Breast and Ovarian
258 Cancer Research Group Netherlands (HEBON), Amsterdam, The Netherlands.
- 259 ¹¹⁷ Lombardi Comprehensive Cancer Center, Georgetown University, Washington, DC, USA.
- 260 ¹¹⁸ Pomeranian Medical University, Independent Laboratory of Molecular Biology and Genetic
261 Diagnostics, Szczecin, Poland.
- 262 ¹¹⁹ Peter MacCallum Cancer Center, Parkville Familial Cancer Centre, Melbourne, Victoria,
263 Australia.
- 264 ¹²⁰ Vilnius University Hospital Santariskiu Clinics, Hematology, oncology and transfusion
265 medicine center, Dept. of Molecular and Regenerative Medicine, Vilnius, Lithuania.
- 266 ¹²¹ State Research Institute Centre for Innovative Medicine, Vilnius, Lithuania.
- 267 ¹²² Landspítali University Hospital, Department of Oncology, Reykjavik, Iceland.
- 268 ¹²³ Stanford University School of Medicine, Department of Epidemiology & Population Health,
269 Stanford, CA, USA.
- 270 ¹²⁴ Stanford Cancer Institute, Stanford University School of Medicine, Department of Medicine,
271 Division of Oncology, Stanford, CA, USA.
- 272 ¹²⁵ The Institute of Cancer Research, Division of Genetics and Epidemiology, London, UK.

- 273 ¹²⁶ Seoul National University College of Medicine, Department of Preventive Medicine, Seoul,
274 Korea.
- 275 ¹²⁷ Seoul National University Graduate School, Department of Biomedical Sciences, Seoul,
276 Korea.
- 277 ¹²⁸ Seoul National University, Cancer Research Institute, Seoul, Korea.
- 278 ¹²⁹ University of California at Los Angeles, David Geffen School of Medicine, Department of
279 Obstetrics and Gynecology, Los Angeles, CA, USA.
- 280 ¹³⁰ UC Davis Medical Center, Department of Pathology and Laboratory Medicine, Sacramento,
281 CA, USA.
- 282 ¹³¹ Medical University of South Carolina, Hollings Cancer Center, Charleston, SC, USA.
- 283 ¹³² Saint Petersburg State University, Saint Petersburg, Russia.
- 284 ¹³³ Sungkyunkwan University School of Medicine, Department of Obstetrics and Gynecology,
285 Samsung Medical Center, Seoul, Korea.
- 286 ¹³⁴ City of Hope Clinical Cancer Genetics Community Research Network, Duarte, CA, USA.
- 287 ¹³⁵ Cancer Genetics Centre, Hong Kong Hereditary Breast Cancer Family Registry, Happy
288 Valley, Hong Kong.
- 289 ¹³⁶ The University of Hong Kong, Department of Surgery, Pok Fu Lam, Hong Kong.
- 290 ¹³⁷ Hong Kong Sanatorium and Hospital, Department of Surgery, Happy Valley, Hong Kong.
- 291 ¹³⁸ VIB Center for Cancer Biology, Leuven, Belgium.
- 292 ¹³⁹ University of Leuven, Laboratory for Translational Genetics, Department of Human
293 Genetics, Leuven, Belgium.
- 294 ¹⁴⁰ Mayo Clinic, Department of Health Sciences Research, Division of Biomedical Statistics and
295 Informatics, Rochester, MN, USA.
- 296 ¹⁴¹ ONCOBELL-IDIBELL-IGTP, Catalan Institute of Oncology, CIBERONC, Hereditary
297 Cancer Program, Barcelona, Spain.
- 298 ¹⁴² BC Cancer, Cancer Control Research, Vancouver, BC, Canada.
- 299 ¹⁴³ Inserm U900, Genetic Epidemiology of Cancer team, Paris, France.
- 300 ¹⁴⁴ NYU Langone Medical Center, Gynecologic Oncology, Laura and Isaac Pearlmuter Cancer
301 Center, New York, NY, USA.
- 302 ¹⁴⁵ Genome Institute of Singapore, Human Genetics Division, Singapore, Singapore.
- 303 ¹⁴⁶ University of Texas MD Anderson Cancer Center, Department of Gynecologic Oncology
304 and Clinical Cancer Genetics Program, Houston, TX, USA.

- 305 ¹⁴⁷ Magee-Womens Hospital, University of Pittsburgh School of Medicine, Pittsburgh, PA,
306 USA.
- 307 ¹⁴⁸ Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, Unit of Medical Genetics,
308 Department of Medical Oncology and Hematology, Milan, Italy.
- 309 ¹⁴⁹ Duke University Hospital, Department of Surgery, Durham, NC, USA.
- 310 ¹⁵⁰ University of Hawaii Cancer Center, Cancer Epidemiology Program, Honolulu, HI, USA.
- 311 ¹⁵¹ Aichi Cancer Center Research Institute, Division of Cancer Epidemiology and Prevention,
312 Nagoya, Japan.
- 313 ¹⁵² Nagoya University Graduate School of Medicine, Division of Cancer Epidemiology,
314 Nagoya, Japan.
- 315 ¹⁵³ Samuel Lunenfeld Research Institute, Public Health Ontario, Toronto, ON, Canada.
- 316 ¹⁵⁴ Imperial College London, Division of Cancer and Ovarian Cancer Action Research Centre,
317 Department Surgery & Cancer, London, UK.
- 318 ¹⁵⁵ University of Glasgow, Institute of Cancer Sciences, Glasgow, UK.
- 319 ¹⁵⁶ University College London, MRC Clinical Trials Unit at UCL, Institute of Clinical Trials &
320 Methodology, London, UK.
- 321 ¹⁵⁷ Roswell Park Cancer Institute, NRG Oncology, Statistics and Data Management Center,
322 Buffalo, NY, USA.
- 323 ¹⁵⁸ Roswell Park Cancer Institute, Division of Cancer Prevention and Control, Buffalo, NY,
324 USA.
- 325 ¹⁵⁹ Magee-Womens Research Institute and Hillman Cancer Center, Womens Cancer Research
326 Center, Pittsburgh, PA, USA.
- 327 ¹⁶⁰ University of Pittsburgh School of Medicine, Division of Gynecologic Oncology,
328 Department of Obstetrics, Gynecology and Reproductive Sciences, Pittsburgh, PA, USA.
- 329 ¹⁶¹ Oregon Health & Science University, Department of Obstetrics and Gynecology, Portland,
330 OR, USA.
- 331 ¹⁶² Oregon Health & Science University, Knight Cancer Institute, Portland, OR, USA.
- 332 ¹⁶³ University of Helsinki, Department of Obstetrics and Gynecology, Helsinki University
333 Hospital, Helsinki, Finland.
- 334 ¹⁶⁴ National Cancer Centre, Cancer Genetics Service, Singapore, Singapore.
- 335 ¹⁶⁵ Nanyang Technological University, Lee Kong Chian School of Medicine, Singapore,
336 Singapore.
- 337 ¹⁶⁶ Odense University Hospital, Department of Clinical Genetics, Odense C, Denmark.
- 338 ¹⁶⁷ Latvian Biomedical Research and Study Centre, Riga, Latvia.

- 339 ¹⁶⁸ Roswell Park Cancer Institute, Department of Gynecologic Oncology, Buffalo, NY, USA.
- 340 ¹⁶⁹ Memorial Sloan Kettering Cancer Center, Clinical Genetics Research Lab, Department of
341 Cancer Biology and Genetics, New York, NY, USA.
- 342 ¹⁷⁰ Memorial Sloan Kettering Cancer Center, Clinical Genetics Service, Department of
343 Medicine, New York, NY, USA.
- 344 ¹⁷¹ National Institute of Oncology, Department of Molecular Genetics, Budapest, Hungary.
- 345 ¹⁷² University Hospitals Leuven, Division of Gynecologic Oncology, Department of Obstetrics
346 and Gynaecology and Leuven Cancer Institute, Leuven, Belgium.
- 347 ¹⁷³ The University of Chicago, Center for Clinical Cancer Genetics, Chicago, IL, USA.
- 348 ¹⁷⁴ Memorial Sloan-Kettering Cancer Center, Department of Epidemiology and Biostatistics,
349 New York, NY, USA.
- 350 ¹⁷⁵ Centro de Investigación en Red de Enfermedades Raras (CIBERER), Madrid, Spain.
- 351 ¹⁷⁶ QIMR Berghofer Medical Research Institute, Department of Genetics and Computational
352 Biology, Brisbane, Queensland, Australia.
- 353 ¹⁷⁷ Aalborg University Hospital, Molecular Diagnostics, Aalborg, Denmark.
- 354 ¹⁷⁸ Aalborg University Hospital, Clinical Cancer Research Center, Aalborg, Denmark.
- 355 ¹⁷⁹ Aalborg University, Department of Clinical Medicine, Aalborg, Denmark.
- 356 ¹⁸⁰ Portuguese Oncology Institute, Department of Genetics, Porto, Portugal.
- 357 ¹⁸¹ Moffitt Cancer Center, Department of Cancer Epidemiology, Tampa, FL, USA.
- 358 ¹⁸² IFOM - the FIRC Institute of Molecular Oncology, Genome Diagnostics Program, Milan,
359 Italy.
- 360 ¹⁸³ Bashkir State University, Department of Genetics and Fundamental Medicine, Ufa, Russia.
- 361 ¹⁸⁴ Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Unit of Molecular Bases of Genetic
362 Risk and Genetic Testing, Department of Research, Milan, Italy.
- 363 ¹⁸⁵ Karolinska Institutet, Clinical Genetics, Stockholm, Sweden.
- 364 ¹⁸⁶ Duke University Hospital, Department of Gynecologic Oncology, Durham, NC, USA.
- 365 ¹⁸⁷ Yale School of Public Health, Chronic Disease Epidemiology, New Haven, CT, USA.
- 366 ¹⁸⁸ Fox Chase Cancer Center, Population Studies Facility, Philadelphia, PA, USA.
- 367 ¹⁸⁹ National Institute of Environmental Health Sciences, NIH, Epidemiology Branch, Research
368 Triangle Park, NC, USA.
- 369 ¹⁹⁰ Fundación Pública Galega Medicina Xenómica, Santiago De Compostela, Spain.

- 370 ¹⁹¹ Instituto de Investigación Sanitaria de Santiago de Compostela, Santiago De Compostela,
371 Spain.
- 372 ¹⁹² Centre Hospitalier Universitaire de Québec – Université Laval Research Center, Genomics
373 Center, Québec City, QC, Canada.
- 374 ¹⁹³ Faculty of Medicine and University Hospital Cologne, University of Cologne, Center for
375 Molecular Medicine Cologne (CMMC), Cologne, Germany.
- 376 ¹⁹⁴ Hebei Medical University, Fourth Hospital, Department of Obstetrics and Gynaecology,
377 Shijiazhuang, China.
- 378 ¹⁹⁵ Icahn School of Medicine at Mount Sinai, Department of Population Health Science and
379 Policy, New York, NY, USA.
- 380 ¹⁹⁶ Icahn School of Medicine at Mount Sinai, Department of Genetics and Genomic Sciences,
381 New York, NY, USA.
- 382 ¹⁹⁷ Centre Hospitalier Universitaire de Québec-Université Laval Research Center, Genomic
383 Center, Québec City, QC, Canada.
- 384 ¹⁹⁸ Medical University of Vienna, Dept of OB/GYN and Comprehensive Cancer Center, Vienna,
385 Austria.
- 386 ¹⁹⁹ University of Cambridge, Department of Public Health and Primary Care, Cambridge, UK.
- 387 ²⁰⁰ The University of Melbourne, Department of Clinical Pathology, Melbourne, Victoria,
388 Australia.
- 389 ²⁰¹ Royal Alexandra Hospital, Department of Obstetrics and Gynecology, Division of
390 Gynecologic Oncology, Edmonton, Alberta, Canada.
- 391 ²⁰² Institut Curie, Service de Génétique, Paris, France.
- 392 ²⁰³ Université Paris Descartes, Paris, France.
- 393 ²⁰⁴ University of South Florida, Epidemiology Center, College of Medicine, Tampa, FL, USA.
- 394 ²⁰⁵ The Institute of Cancer Research, Division of Breast Cancer Research, London, UK.
- 395 ²⁰⁶ Medical University of Vienna, Dept of OB/GYN and Comprehensive Cancer Center, Vienna,
396 Austria.
- 397 ²⁰⁷ University of Porto, Biomedical Sciences Institute (ICBAS), Porto, Portugal.
- 398 ²⁰⁸ Cancer Research Malaysia, Breast Cancer Research Programme, Subang Jaya, Selangor,
399 Malaysia.
- 400 ²⁰⁹ University of Malaya, Department of Surgery, Faculty of Medicine, Kuala Lumpur,
401 Malaysia.
- 402 ²¹⁰ Brigham and Women's Hospital and Harvard Medical School, Obstetrics and Gynecology
403 Epidemiology Center, Boston, MA, USA.

- 404 ²¹¹ Columbia University, Department of Epidemiology, Mailman School of Public Health, New
405 York, NY, USA.
- 406 ²¹² Magee-Womens Hospital, University of Pittsburgh School of Medicine, Department of
407 Medicine, Pittsburgh, PA, USA.
- 408 ²¹³ McGill University, Program in Cancer Genetics, Departments of Human Genetics and
409 Oncology, Montréal, QC, Canada.
- 410 ²¹⁴ University of Cambridge, Department of Medical Genetics, Cambridge, UK.
- 411 ²¹⁵ Dartmouth College, Geisel School of Medicine, Hanover, NH, USA.
- 412 ²¹⁶ The Ohio State University, Department of Cancer Biology and Genetics, Columbus, OH,
413 USA.
- 414 ²¹⁷ Pontificia Universidad Javeriana, Institute of Human Genetics, Bogota, Colombia.
- 415 ²¹⁸ University of Oxford, Cancer Epidemiology Unit, Oxford, UK.
- 416 ²¹⁹ Beth Israel Deaconess Medical Center, Department of Medical Oncology, Boston, MA,
417 USA.
- 418 ²²⁰ Harvard T.H. Chan School of Public Health, Department of Epidemiology, Boston, MA,
419 USA.
- 420 ²²¹ University Medical Center Groningen, University Groningen, Department of Genetics,
421 Groningen, The Netherlands.
- 422 ²²² University of Pretoria, Department of Genetics, Arcadia, South Africa.
- 423 ²²³ Centro de Investigación en Red de Enfermedades Raras (CIBERER), Madrid, Spain.
- 424 ²²⁴ Fundación Pública Galega de Medicina Xenómica, Santiago de Compostela, Spain.
- 425 ²²⁵ Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Complejo
426 Hospitalario Universitario de Santiago, SERGAS, Santiago de Compostela, Spain.
- 427 ²²⁶ Vanderbilt University Medical Center, Division of Quantitative Sciences, Department of
428 Obstetrics and Gynecology, Department of Biomedical Sciences, Women's Health Research,
429 Nashville, TN, USA.
- 430 ²²⁷ Duke Cancer Institute, Cancer Control and Population Sciences, Durham, NC, USA.
- 431 ²²⁸ Duke University Hospital, Department of Community and Family Medicine, Durham, NC,
432 USA.
- 433 ²²⁹ National Institute of Environmental Health Sciences, NIH, Biostatistics and Computational
434 Biology Branch, Research Triangle Park, NC, USA.
- 435 ²³⁰ City of Hope, Clinical Cancer Genomics, Duarte, CA, USA.
- 436 ²³¹ Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

437 ²³² Stanford University School of Medicine, Department of Biomedical Data Science, Stanford,
438 CA, USA.

439 ²³³ Uppsala University, Department of Surgical Sciences, Uppsala, Sweden.

440 ²³⁴ University of Malaya, Department of Obstetrics and Gynaecology, University of Malaya
441 Medical Centre, Kuala Lumpur, Malaysia.

442 ²³⁵ Hebei Medical University, Fourth Hospital, Department of Molecular Biology, Shijiazhuang,
443 China.

444 ²³⁶ National Centre for Scientific Research 'Demokritos', Molecular Diagnostics Laboratory,
445 INRASTES, Athens, Greece.

446 ²³⁷ Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Cedars-
447 Sinai Medical Centre, Department of Obstetrics and Gynecology, Los Angeles, CA, USA.

448 ²³⁸ 1032 Royal Pass Road, Tampa, FL, USA.

449 ²³⁹ University of NSW Sydney, School of Women's and Children's Health, Faculty of Medicine,
450 Sydney, New South Wales, Australia.

451 ²⁴⁰ University of NSW Sydney, Adult Cancer Program, Lowy Cancer Research Centre, Sydney,
452 New South Wales, Australia.

453 ²⁴¹ University of Michigan School of Public Health, Department of Epidemiology, Ann Arbor,
454 MI, USA.

455 ²⁴² University of Southern California Norris Comprehensive Cancer Center, Department of
456 Preventive Medicine, Keck School of Medicine, Los Angeles, CA, USA.

457 ²⁴³ Mayo Clinic, Department of Health Science Research, Division of Epidemiology, Rochester,
458 MN, USA.

459 ²⁴⁴ Emory University, Department of Epidemiology, Rollins School of Public Health, Atlanta,
460 GA, USA.

461

462 *These authors had equal contributions.

463 Corresponding Author: pp10001@medschl.cam.ac.uk

464

465 **Running Title: Polygenic Risk Modelling for Prediction of Epithelial Ovarian**
466 **Cancer Risk**

467

468 **Abstract**

469 Polygenic risk scores (PRS) for epithelial ovarian cancer (EOC) have the potential to
470 improve risk stratification. Joint estimation of Single Nucleotide Polymorphism (SNP)
471 effects in models could improve predictive performance over standard approaches of
472 PRS construction. Here, we implemented computationally-efficient, penalized, logistic
473 regression models (lasso, elastic net, stepwise) to individual level genotype data and a
474 Bayesian framework with continuous shrinkage, “select and shrink for summary
475 statistics” (S4), to summary level data for epithelial non-mucinous ovarian cancer risk
476 prediction. We developed the models in a dataset consisting of 23,564 non-mucinous
477 EOC cases and 40,138 controls participating in the Ovarian Cancer Association
478 Consortium (OCAC) and validated the best models in three populations of different
479 ancestries: prospective data from 198,101 women of European ancestry; 7,669 women
480 of East Asian ancestry; 1,072 women of African ancestry, and in 18,915 *BRCA1* and
481 12,337 *BRCA2* pathogenic variant carriers of European ancestry. In the external
482 validation data, the model with the strongest association for non-mucinous EOC risk
483 derived from the OCAC model development data was the S4 model (27,240 SNPs) with
484 odds ratios (OR) of 1.38(95%CI:1.28–1.48,AUC:0.588) per unit standard deviation, in
485 women of European ancestry; 1.14(95%CI:1.08–1.19,AUC:0.538) in women of East
486 Asian ancestry; 1.38(95%CI:1.21-1.58,AUC:0.593) in women of African ancestry;
487 hazard ratios of 1.37(95%CI:1.30–1.44,AUC:0.592) in *BRCA1* pathogenic variant
488 carriers and 1.51(95%CI:1.36-1.67,AUC:0.624) in *BRCA2* pathogenic variant carriers.
489 Incorporation of the S4 PRS in risk prediction models for ovarian cancer may have
490 clinical utility in ovarian cancer prevention programs.

491

492 **INTRODUCTION**

493 Rare mutations in known high and moderate penetrance susceptibility genes (*BRCA1*,
494 *BRCA2*, *BRIP1*, *PALB2*, *RAD51C*, *RAD51D* and the mis-match repair genes) account
495 for about 40 percent of the inherited component of EOC disease risk (1,2). Genome
496 wide association studies (GWAS), reviewed in Kar et. al. and Jones et. al. (1,3), have
497 identified 39 common (minor allele frequency [MAF] > 0.05) susceptibility variants
498 which together explain about 6% of the heritability of EOC. Polygenic risk scores
499 (PRS) provide an opportunity for refined risk stratification in the general population as
500 well as in carriers of rare moderate or high risk alleles (4,5).

501 A PRS is calculated as the weighted sum of the number of risk alleles carried for a
502 specified set of genetic variants. The best approach to identify the set of variants and
503 their weights in order to optimize the predictive power of a PRS is unknown. A
504 common approach involves selecting a set of variants that reach a threshold for
505 association based on the p-value for each variant with or without clumping and pruning
506 to remove highly correlated variants (6,7). More complex prediction models, based on
507 machine learning approaches that do not assume variant independence have also been
508 used to construct PRS for complex traits in humans (8,9). To date, these methods have
509 produced only modest gains in predictive power for highly polygenic phenotypes (8,10).
510 Penalized regression approaches such as the lasso, elastic net and the adaptive lasso
511 have also been used for the joint estimation of variant effects using individual level data
512 for large data sets (11). While they have the potential advantage of improving
513 performance, the major drawback of these methods is the high computational burden
514 required to fit the models (11,12).

515 In this study, we present a novel implementation of computationally-efficient PRS
516 models using two approaches: 1) penalized regression models including the lasso,
517 elastic net and minimax concave penalty, for use when individual genotype data are
518 available; and 2) a Bayesian regression model with continuous shrinkage priors on
519 variant effect sizes, for use in broader settings where summary statistics are available,
520 hereafter referred to as the “select and shrink with summary statistics” (S4) method. We
521 compare these models with two commonly used methods, stepwise regression with p-
522 value thresholding and LDpred.

523 **MATERIALS (SUBJECTS) AND METHODS**

524 *EOC Histotypes*

525 EOC is a highly heterogeneous phenotype with five major histotypes for invasive
526 disease – high-grade serous, low-grade serous, endometrioid, clear cell and mucinous
527 histotype. The mucinous histotype is the least common and its origin is the most
528 controversial with up to 60% of diagnosed cases of mucinous ovarian cancer often
529 being misdiagnosed metastasis from non-ovarian sites (13). Recent molecular analyses
530 have concluded that most primary invasive mucinous cases are not extra-ovarian
531 metastases (14). However, accurate diagnosis relies on expert histopathology and
532 immuno-histochemical profiling (15), which remains a challenge in clinical practice and
533 can be an issue in different cohorts from different time periods. Therefore, in this study,
534 we performed PRS modelling and association testing for all cases of invasive EOC,
535 excluding mucinous cases, hereafter referred to as non-mucinous EOC.

536 ***Model Development Study Population***

537 We used genotype data from 23,564 invasive non-mucinous EOC cases and 40,138
538 controls with >80% European ancestry from 63 case-control studies included in the
539 Ovarian Cancer Association Consortium (OCAC) for model development. The study
540 protocol was approved by the institutional review boards of the Brigham and Women's
541 Hospital and Harvard T.H. Chan School of Public Health, and those of participating
542 registries as required. The distribution of cases by histotype was high-grade serous
543 (13,609), low-grade serous (2,749), endometrioid (2,877), clear cell (1,427), and others
544 (2,902). All mucinous EOC histotypes (2,587) were excluded. Sample collection,
545 genotyping and quality control have been previously described (16). Genotype data
546 were imputed to the Haplotype Reference Consortium reference panel on the Michigan
547 Imputation server, using 470,825 SNPs that passed quality control. Of the 32 million
548 SNPs imputed, 10 million had imputation $r^2 > 0.3$ and were included in this analysis.

549 ***Model Validation Study Population***

550 **UK Biobank Population**

551 We validated the best-fitting PRS models developed in the OCAC data in 657 prevalent
552 and incident cases of invasive EOC (346 serous, 98 endometrioid, 51 clear cell and 162
553 other) and 198,101 female controls of European ancestry from the UK Biobank. As with
554 the model development data, all mucinous histotypes (166) were excluded. Samples
555 were genotyped using either the Affymetrix UK BiLEVE Axiom Array or Affymetrix
556 UK Biobank Axiom Array (which share 95% marker content), and then imputed to a
557 combination of the Haplotype Reference Consortium, the 1000Genomes phase 3 and
558 the UK10K reference panels (17). We restricted analysis to genetically confirmed

559 females of European/white British ancestry. We excluded individuals if they were
560 outliers for heterozygosity, had low genotyping call rate <95%, had sex chromosome
561 aneuploidy, or if they were duplicates (cryptic or intended) (16). All SNPs selected in
562 the model development phase were available in the UK Biobank.

563 **Non-European Ancestry Population**

564 We investigated transferability of the best-fitting PRS models to populations of non-
565 European ancestry using genotype data from females of East Asian and African
566 ancestries that had been genotyped as part of the OCAC OncoArray Project (18,19).
567 Women of East Asian ancestry - 2,841 non-mucinous invasive EOC (1,960 high-grade
568 serous, 136 low-grade serous, 400 endometrioid, 271 clear cell, 74 other histotypes)
569 and 4,828 controls - were identified using a criterion of >80% Asian ancestry. This
570 group included samples collected from population-based studies in China, Japan, Korea,
571 Malaysia as well as samples from studies conducted in the US, Europe and Australia.
572 Details of these data have been previously described (18). Similarly, women of African
573 ancestry - 368 cases of non-mucinous invasive EOC (261 high-grade serous, 35 low-
574 grade serous, 47 endometrioid, 7 clear cell, 53 other histotypes) and 704 controls,
575 mainly from studies conducted in the US, were identified using a criterion of >80%
576 African ancestry as described previously (19).

577 **BRCA1/BRCA2 Pathogenic Variant Carrier Population**

578 We also assessed the performance of the best-fitting PRS models in women of European
579 ancestry (>80% European ancestry) with the pathogenic *BRCA1* and *BRCA2* variants
580 from the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA). We used
581 genotype data from 18,915 *BRCA1* (2,053 invasive EOC cases – 712 serous, 115

582 endometrioid, 9 clear cell, 1217 unknown/other) and 12,337 *BRCA2* (717 invasive EOC
583 cases – 26 serous, 4 endometrioid, 1 clear cell, 686 unknown/other) pathogenic variant
584 carriers from 63 studies contributing to CIMBA for independent model validation.
585 Details of the study population and sample collection have been described previously
586 (16). Genotyping, data quality control measures, intercontinental ancestry assessment
587 and imputation to the HRC reference panel are as described for the OCAC study
588 population.

589 **PRS from Meta-analysis of Summary Statistics**

590 We leveraged the increase in sample size resulting from a meta-analysis of EOC risk
591 associations, using both the CIMBA and OCAC data described above, to explore
592 performance of PRS approaches based on summary statistics.

593 **Statistical Analysis**

594 ***Polygenic Risk Models***

595 For all PRS models, we created scores as linear functions of the allele dosage in the
596 general form $PRS_i = \sum_j^p x_{ij} \beta_j$ where genotypes are denoted as x (taking on the minor
597 allele dosages of 0, 1 and 2), with x_{ij} representing the i th individual for the j th SNP
598 (out of p SNPs) on an additive log scale and β_j represents the weight - the log of the
599 odds ratio - of the j th SNP. We used different approaches to select and derive the
600 optimal weights, β_j , in models as described below.

601 ***Penalized logistic regression models: the lasso, elastic net and minimax concave***
602 ***penalty***

603 A penalized logistic regression model for a set of SNPs aims to identify a set of
604 regression coefficients that minimize the regularized loss function given by

$$plr(x; \lambda, \kappa) = \begin{cases} x - \lambda \text{sign}(x)/(1 - \kappa) & \text{if } |x| < \lambda/\kappa \text{ and } |(x)| > \lambda \\ x & \text{if } |x| \geq \frac{\lambda}{\kappa} \\ 0 & \text{if } |(x)| < \lambda \end{cases}$$

605 where x is the effect estimate of a SNP, λ is the tuning parameter and κ is the threshold
606 (penalty) for different regularization paths. λ and κ are parameters that need to be
607 chosen during model development to optimize performance. The lasso, elastic net,
608 minimax concave penalty (MCP), and p-value thresholds are instances of the function
609 with different κ values. We minimized the winner's curse effect on inflated effect
610 estimates for rare SNPs by penalizing rarer SNPs more heavily than common SNPs.
611 Details are provided in the Supplementary Methods.

612 We used a two-stage approach to reduce computational burden without a corresponding
613 loss in predictive power. The first stage was a SNP selection stage using a sliding
614 windows approach, with 5.5Mb data blocks and a 500kb overlap between blocks. SNP
615 selection was performed for each block and selected SNPs were collated. Single SNP
616 association analyses were then run, and all SNPs with a χ^2 test statistic of less than 2.25
617 were excluded. Penalized regression models were applied to the remaining SNPs
618 using λ values of 3.0 and κ values of 0.0, 0.2, 0.4, 0.6, 0.8 and 1.0. SNPs selected in any
619 of these models were included in subsequent analyses.

620 In the second stage, we fit penalized regression models to the training dataset with λ
621 values ranging from 3.0 to 5.5 in increments of 0.1 iterated over κ values from -3.0 to 1
622 in increments of 0.1. The lasso model ($\kappa = 0$) for each value of λ was fitted first, to
623 obtain a unique maximum. From the fitted maximum the κ value was changed, and the
624 model refitted.

625 We applied this two-stage approach with five-fold cross-validation (**Figure 1**). The
626 variants and their weights from the two-stage penalized logistic regression modelling in
627 the training data were used to calculate the area under the receiver operating
628 characteristic curve (AUC) in the test data. We repeated this process for each cross-
629 validation iteration to obtain a mean AUC for each combination of λ and κ .

630 Finally, we selected the tuning and threshold parameters from the lasso, elastic net and
631 minimax concave penalty models with the maximum mean cross-validated AUC and
632 fitted penalized logistic regression models with these parameters to the entire OCAC
633 dataset to obtain SNP weights for PRS scores.

634 *Stepwise logistic regression with variable P-value threshold*

635 This model is a general PLR model with $\kappa=1$. As with the other PLR models, we
636 investigated various values for λ values (corresponding to a variable P-value threshold
637 for including a SNP in the model). However, we observed that the implementation of
638 this model on individual level data was more difficult than for other κ values because
639 the model would sometimes converge to a local optimum rather than the global
640 optimum. Therefore, we applied an approximate conditional and joint association
641 analysis using summary level statistics correcting for estimated LD between SNPs,

642 using a reference panel of 5,000 individual level genotype OCAC data as described in
643 Yang et.al (20). Details are provided in the Supplementary Methods.

644 *LDPred*

645 LDPred is a Bayesian approach that adjusts GWAS summary statistics for the effects of
646 LD by shrinking the posterior mean effect size of each genetic marker based on a point-
647 normal prior on the effect sizes and LD information from an external reference panel.
648 We derived seven candidate polygenic risk scores assuming the fractions of associated
649 variants were 0.001, 0.003, 0.01, 0.03, 0.1, 0.3 and 1.0 respectively (21) using an LD
650 reference panel of 503 samples of European ancestries from the 1000 Genomes phase 3
651 release and effect estimates from the genome wide association analysis on the OCAC
652 model development data.

653 *Select and shrink using summary statistics (S4)*

654 The S4 algorithm is similar to the PRS-CS algorithm (22) which is a Bayesian approach
655 that uses summary statistics and between-SNP correlation data from a reference panel to
656 generate the PRS scores by placing a continuous shrinkage prior on effect sizes. We
657 adapted this method with penalization of rarer SNPs by correcting for the standard
658 deviation resulting in the selection of fewer SNPs. To implement this algorithm, we
659 varied three parameters, a , b , ϕ , which together control the degree of shrinkage of effect
660 estimates. Φ , the overall shrinkage parameter, is influenced by values of a which
661 control shrinkage of effect estimates around 0 and b which control shrinkage of larger
662 effect estimates. Smaller values of a result in more severe shrinkage of effect estimates
663 than larger values. Conversely, smaller values of b produce less severe shrinkage than
664 larger values.

665 We generated summary statistics for each cross-validation training set and selected the
666 parameters that gave the best results on average from the cross-validation and applied
667 these to the set of summary statistics for the complete OCAC data set to obtain the final
668 set of weights.

669 ***PRS based on meta-analysis of OCAC-CIMBA summary statistics***

670 We conducted a meta-analysis of the EOC associations in *BRCA1* variant carriers,
671 *BRCA2* variant carriers and the participants participating in OCAC using previously
672 described methodological approaches (16). Additional details are provided in the
673 Supplementary Methods. We constructed two PRS models using results from the
674 OCAC-CIMBA meta-analysis: the Select and Shrink (OCAC-CIMBA) PRS and the
675 Stepwise (OCAC-CIMBA) PRS. To construct the Select and Shrink (OCAC-CIMBA)
676 PRS, we applied the a , b and ϕ parameters from the Select and Shrink model described
677 above to the summary statistics from the meta-analysis to obtain a different set of SNPs
678 and weights. We generated the Stepwise (OCAC-CIMBA) PRS by using histotype-
679 specific results from the meta-analysis. We selected all SNPs that were genome-wide
680 significant at nominal thresholds ($p < 5 \times 10^{-8}$), along with any independent signals in the
681 same region with $p < 10^{-5}$ from the histotype specific analyses for low-grade serous,
682 high-grade serous, endometrioid, clear cell ovarian cancer and non-mucinous invasive
683 EOC.

684 ***Polygenic risk score performance.***

685 The best lasso, elastic net, stepwise and S4 models from the model development stage
686 were validated using two independent data sources: the UK Biobank data and
687 *BRCA1/BRCA2* pathogenic variant carriers from the CIMBA. In the UK Biobank data,

688 we evaluated discriminatory performance of the models using the AUC and examined
689 the association between standardized PRS and risk of non-mucinous EOC using
690 logistic regression analysis. For the CIMBA data, we assessed associations for each
691 version of the PRS and invasive non-mucinous EOC risk using weighted Cox regression
692 methods previously described (5). PRSs in the CIMBA data were scaled to the same
693 PRS standard deviations as the OCAC data, meaning that per standard deviation hazard
694 ratios estimated on CIMBA data are comparable to PRS associations in the OCAC and
695 UK Biobank data. The regression models were adjusted for birth cohort (<1920, 1920-
696 1929, 1930-1939, 1940-1949, \geq 1950) and the first four ancestry informative principal
697 components (calculated separately by iCOGS/OncoArray genotyping array) and
698 stratified by Ashkenazi Jewish ancestry and country. Absolute risks by PRS percentiles
699 adjusting for competing risks of mortality from other causes were calculated as
700 described in the Supplementary Material.

701 *Transferability of PRS scores to non-European Ancestry*

702 We implemented two straightforward approaches to disentangle the role of ancestry on
703 polygenic risk scoring. We selected homogenous ancestral samples by using a high cut-
704 off criterion of 80% ancestry and we standardized the polygenic risk scores by mean-
705 centering within each population. These approaches led to a more uniform distribution
706 of polygenic risk scores within each ancestral population. Further adjustments using
707 principal components of ancestry did not attenuate risk estimates.

708 **Data Availability**

709 OncoArray germline genotype data for the OCAC studies have been deposited at the
710 European Genome-phenome Archive (EGA; <https://ega-archive.org/>), which is hosted

711 by the EBI and the CRG, under accession EGAS00001002305. Summary results are
712 available from the Ovarian Cancer Association Consortium
713 (<http://ocac.ccge.medschl.cam.ac.uk/>). A subset of the OncoArray germline genotype
714 data for the CIMBA studies will be made publically available through the database of
715 Genotypes and Phenotypes (dbGaP) under accession phs001321.v1.p1. The complete
716 data set will not be made publically available because of restraints imposed by the ethics
717 committees of individual studies; requests for further data can be made to the Data
718 Access Coordination Committee (<http://cimba.ccge.medschl.cam.ac.uk/>)

719 **Ethics Statement**

720 All study participants provided written informed consent and participated in research or
721 clinical studies at the host institute under ethically approved protocols. The studies and
722 their approving institutes are listed in the Supplementary Material (Ethics Statement)

723 **Results**

724 *Model development*

725 For models based on individual level genotype data, the elastic net model had the best
726 predictive accuracy (model parameters: $\lambda=3.3$, $\kappa=-2.2$, AUC=0.586) Predictive accuracy
727 for the lasso model ($\lambda=3.3$, AUC= 0.583) was slightly lower (**Table 1**). The optimal
728 value of λ obtained from regularization paths for the MCP model was 3.3. Further
729 reductions in the degree of penalization for the MCP models did not improve prediction
730 accuracy. Therefore, the best MCP model was equivalent to the lasso model. For
731 models based on summary statistics, the best-fitting S4 model had the best performance
732 ($a=2.75$, $b=2$, $\phi=3e-6$, AUC=0.593), whereas the best LDPred model had the poorest

733 performance of the methods tested ($\rho=0.001$, AUC=0.552). The mean odds ratios per
 734 standard deviation are shown in **Table 1** along with the number of SNPs included in the
 735 final model when the models were built with the relevant parameters using the complete
 736 dataset. All SNPs selected and the associated weights for each model are provided in
 737 Supplementary Tables 1 – 6. Given the poorer performance of the LDpred model and
 738 the very large number of SNPs included in the final model it was not considered for
 739 further validation in other datasets.

740 **Table 1: Performance of different PRS models in five-fold cross validation of OCAC data**

Model	Number of SNPs*	Tuning parameter for best performance	AUC	OR per 1 SD of PRS	95% CI
a) Models based on individual level genotype data					
Lasso	1,403	$\lambda=3.3$	0.583	1.35	1.30 – 1.39
Elastic net	10,797	$\lambda=3.3, \kappa=-2.2$	0.586	1.36	1.31 – 1.40
MCP	1,403	$\lambda=3.3$	0.583	1.35	1.30 – 1.39
b) Models based on summary statistics					
LDpred	5,291,719	$\rho=0.001$	0.552	1.21	1.13 – 1.29
Stepwise	22	$\lambda=5.4$	0.572	1.30	1.26 – 1.34
Select and Shrink (OCAC)	27,240	$a=2.75, b=2, \phi=3e-6$	0.593	1.39	1.34 – 1.44

Abbreviations: AUC-Area Under the Receiver Operating Characteristic (ROC) Curve AUC); OR-Odds Ratio; SD-Standard deviation; PRS-Polygenic Risk Score; CI-Confidence Interval. NA-Not Applicable.

* Number of SNPs in PRS model run on full OCAC data set after selection of model parameters

741 ***Model validation in women of European ancestry: general population (UK Biobank)***
 742 ***and BRCA1/BRCA2 pathogenic variant carriers (CIMBA)***

743 The best AUC estimates derived from cross-validation are likely to be upwardly biased
 744 due to overfitting. Therefore, we used the UK Biobank data as an external validation
 745 dataset. Overall the PLR models performed slightly better in the UK Biobank data than
 746 the model development data (**Table 2**). Of the models developed using the OCAC
 747 model development data, the association between PRS and non-mucinous EOC was
 748 strongest with the Select and Shrink derived PRS (OR per unit SD=1.38, 95%CI:1.28–
 749 1.48) and slightly lower for the lasso PRS (OR per unit SD=1.37, 95%CI:1.27–1.48),

750 the elastic net PRS (OR per unit SD=1.36, 95%CI:1.26–1.47) and the stepwise PRS
 751 model (OR per unit SD=1.35, 95%CI:1.26–1.46). In *BRCA1* and *BRCA2* variant
 752 carriers, prediction accuracy was generally higher among *BRCA2* carriers than *BRCA1*
 753 carriers. Consistent with results from the general population in the UK Biobank, the
 754 Select and Shrink PRS model also had the strongest association and predictive accuracy
 755 for invasive EOC risk in both *BRCA1* (HR per unit SD=1.37, 95%CI:1.30–1.44,
 756 AUC=0.592) carriers and *BRCA2* carriers (HR per unit SD=1.51, 95%CI:1.36–1.67,
 757 AUC=0.624). The PRS models developed using the OCAC-CIMBA meta-analysis
 758 results had better discriminative ability in the UK Biobank than the PRS models
 759 developed using only OCAC data. Compared with the Select and Shrink model using
 760 only OCAC data, the Select and Shrink PRS model derived from the meta-analysis had
 761 fewer SNPs (n=18,007), a stronger association with invasive EOC risk (OR per unit
 762 SD=1.42, 95%CI:1.32–1.54) and better predictive accuracy (AUC=0.596). Similarly,
 763 the Stepwise model from the OCAC-CIMBA meta-analysis performed better than the
 764 Stepwise model from only OCAC data (OR per unit SD=1.39, 95%CI:1.29–1.50,
 765 AUC=0.595), but included more SNPs (n=36)

766 **Table 2: External validation of PRS models in European populations using data from UK**
 767 **Biobank and CIMBA**

Model (data set)	SNPs	UK Biobank			CIMBA <i>BRCA1</i> carriers†			CIMBA <i>BRCA2</i> carriers†		
		AUC	OR	95% CI	AUC	HR	95% CI	AUC	HR	95% CI
a) PRS models based on OCAC data										
Lasso (OCAC)	1,403	0.587	1.37	1.27 – 1.48	0.573	1.29	1.23 – 1.36	0.627	1.49	1.34 – 1.65
Elastic net (OCAC)	10,797	0.588	1.36	1.26 – 1.47	0.583	1.34	1.27 – 1.41	0.617	1.50	1.35 – 1.66
Stepwise (OCAC)	22	0.588	1.35	1.26 – 1.46	0.563	1.22	1.17 – 1.28	0.605	1.40	1.26 – 1.55
Select and shrink (OCAC)	27,240	0.588	1.38	1.28 - 1.48	0.592	1.37	1.30 – 1.44	0.624	1.51	1.36 – 1.67
b) PRS models based on meta-analysis of OCAC and CIMBA data										
Stepwise (OCAC-	36	0.595	1.39	1.29 - 1.50						

CIMBA)*

Select and shrink (OCAC-CIMBA)	18,007	0.596	1.42	1.32 – 1.54
-----------------------------------	--------	-------	------	-------------

Abbreviations: AUC-Area Under the Receiver Operating Characteristic Curve; OR-Odds Ratio; HR-Hazards ratio

*results in CIMBA are overfitted as the CIMBA data was used for model development

† HR are adjusted for birth cohort and the first four ancestry informative principal components. AUC were estimated from models fitting only the PRS as the independent variable.

768

769 Compared with women in the middle quintile, women in the top 95th percentile of the
770 lasso derived PRS model had 2.23-fold increased odds of non-mucinous EOC (**Table**
771 **3**). The observed distribution of the OR estimates was consistent with ORs obtained
772 from theoretical predicted values under the assumption that all SNPs interact
773 multiplicatively, especially for the lasso model (**Figure 3**), with all 95% confidence
774 intervals intersecting with the theoretical estimates for women of European ancestry.

Table 3: Association between polygenic risk scores and non-mucinous EOC by PRS percentiles and ancestry

Percentile	UK Biobank			East Asian			African		
	Controls (n)	Cases (n)	OR (95% CI)	Controls (n)	Cases (n)	OR (95% CI)	Controls (n)	Cases (n)	OR (95% CI)
a. Lasso									
0-5	9880	12	0.42 (0.22 – 0.72)	278	106	0.65 (0.51 – 0.83)	35	19	0.89 (0.47 – 1.65)
5-10	9870	24	0.83 (0.52 – 1.27)	271	112	0.71 (0.55 – 0.90)	41	13	0.52 (0.25 – 1.01)
10-20	19733	53	0.92 (0.66 – 1.27)	487	280	0.98 (0.82 – 1.18)	81	26	0.53 (0.31 – 0.88)
20-40	39468	104	0.90 (0.69 – 1.18)	993	541	0.93 (0.80 – 1.08)	154	60	0.64 (0.42 – 0.99)
40-60	39457	115	1.00	967	566	1.00	133	81	1.00
60-80	39425	147	1.28 (1.00 – 1.64)	941	593	1.08 (0.93 – 1.25)	136	78	0.94 (0.64 – 1.39)
80-90	19699	87	1.52 (1.14 – 2.00)	466	301	1.10 (0.92 – 1.32)	63	44	1.15 (0.71 – 1.84)
90-95	9842	51	1.78 (1.27 – 2.46)	214	169	1.35 (1.07 – 1.69)	34	20	0.97 (0.51 – 1.78)
95-100	9830	64	2.23 (1.64 – 3.02)	211	173	1.40 (1.12 – 1.76)	27	27	1.64 (0.90 – 3.00)
b. Elastic Net									
0-5	9876	17	0.67 (0.39 – 1.09)	277	107	0.72 (0.56 – 0.92)	35	19	0.90 (0.47 – 1.64)
5-10	9876	17	0.67 (0.39 – 1.09)	271	112	0.78 (0.61 – 0.99)	41	13	0.52 (0.25 – 1.01)
10-20	19740	45	0.89 (0.62 – 1.26)	497	270	1.02 (0.85 – 1.22)	81	26	0.53 (0.31 – 0.88)
20-40	39453	120	1.19 (0.91 – 1.55)	967	567	1.10 (0.95 – 1.28)	154	60	0.64 (0.42 – 0.96)
40-60	39471	101	1.00	1000	533	1.00	133	81	1.00
60-80	39413	159	1.58 (1.23 – 2.03)	926	608	1.23 (1.06 – 1.43)	136	78	0.94 (0.64 – 1.39)
80-90	19695	91	1.80 (1.36 – 2.40)	457	310	1.27 (1.06 – 1.52)	63	44	1.15 (0.71 – 1.84)
90-95	9841	52	2.07 (1.47 – 2.87)	226	157	1.30 (1.04 – 1.64)	34	20	0.97 (0.51 – 1.78)
95-100	9839	55	2.18 (1.56 – 3.02)	207	177	1.60 (1.28 – 2.01)	27	27	1.64 (0.90 – 3.00)
c. Stepwise									
0-5	9880	13	0.39 (0.21 – 0.67)	254	130	0.90 (0.71 – 1.14)	40	14	0.75 (0.37 – 1.44)

Percentile	UK Biobank			East Asian			African		
	Controls (n)	Cases (n)	OR (95% CI)	Controls (n)	Cases (n)	OR (95% CI)	Controls (n)	Cases (n)	OR (95% CI)
5-10	9874	19	0.57 (0.34 – 0.91)	268	115	0.76 (0.59 – 0.96)	43	11	0.55 (0.26 – 1.10)
10-20	19742	44	0.67 (0.47 – 0.93)	494	273	0.98 (0.81 – 1.17)	80	27	0.72 (0.42 – 1.21)
20-40	39470	102	0.77 (0.60 – 1.00)	970	564	1.03 (0.89 – 1.19)	142	72	1.09 (0.73 – 1.63)
40-60	39440	132	1.00	979	564	1.00	146	68	1.00
60-80	39414	158	1.20 (0.95 – 1.51)	951	583	1.08 (0.94 – 1.25)	130	84	1.39 (0.93 – 2.07)
80-90	19697	88	1.33 (1.02 – 1.75)	456	311	1.21 (1.01 – 1.44)	61	46	1.62 (1.00 – 2.61)
90-95	9853	41	1.24 (0.86 – 1.75)	236	147	1.10 (0.87 – 1.38)	35	19	1.17 (0.61 – 2.17)
95-100	9834	60	1.82 (1.33 – 2.46)	220	164	1.32 (1.04 – 1.65)	27	27	2.15 (1.17 – 3.95)
d. Select and Shrink									
0-5	9957	16	0.54 (0.31 – 0.89)	279	105	0.63 (0.49 – 0.81)	38	16	0.71 (0.36 – 1.33)
5-10	9888	15	0.51 (0.29 – 0.85)	254	129	0.85 (0.67- 1.08)	41	13	0.53 (0.26 – 1.03)
10-20	19812	51	0.87 (0.62 – 1.20)	489	278	0.96 (0.80-1.14)	81	26	0.54 (0.32 – 0.90)
20-40	39435	113	0.97 (0.75 – 1.25)	1013	521	0.86 (0.75 – 1.00)	156	58	0.62 (0.41 – 0.94)
40-60	39512	117	1.00	961	572	1.00	134	80	1.00
60-80	39316	158	1.36 (1.07 – 1.73)	950	584	1.03 (0.89 – 1.20)	137	77	0.94 (0.63 – 1.40)
80-90	19718	77	1.32 (0.98 – 1.76)	434	333	1.29 (1.08 – 1.54)	61	46	1.26 (0.79 – 2.02)
90-95	9791	45	1.55 (1.09 – 2.17)	233	150	1.08 (0.86 – 1.36)	30	24	1.34 (0.73 – 2.45)
95-100	9775	65	2.25 (1.65 – 3.03)	215	169	1.32 (1.05 – 1.66)	26	28	1.80 (0.99 – 3.31)

OR: Odds Ratio; CI: Confidence Interval.

389 *Absolute Risk of Developing Ovarian Cancer by PRS percentiles*

390 We estimated cumulative risk of EOC experienced between birth and the age of 80 within
391 PRS percentiles for women in the general population (**Figure 2**), by applying the odds ratio
392 from the PRS models to age-specific population incidence and mortality data for England in
393 2016. For *BRCA1* and *BRCA2* pathogenic variant carriers, we applied the estimated hazard
394 ratios from PRS models to age-specific incidence rates obtained from Kuchenbaecker et al.
395 (23).

396 For women in the general population, the estimated cumulative risks of EOC by age 80 for
397 women at the 99th centile of the PRS distribution were 2.24%, 2.18%, 2.54% and 2.81% for
398 the lasso, elastic net, stepwise and S4 models, respectively. In comparison, the absolute risks
399 of EOC by age 80 for women at the 1st centile were 0.76%, 0.78%, 0.64% and 0.56% for the
400 lasso, elastic net, stepwise and S4 models, respectively.

401 The absolute risks of developing EOC in *BRCA1* and *BRCA2* pathogenic variant carriers
402 were considerably higher than for women in the general population (Figures S1:Cumulative
403 risk of ovarian cancer risk in *BRCA1* carriers by polygenic risk score percentiles and S2:
404 Cumulative risk of ovarian cancer risk in *BRCA2* carriers by polygenic risk score
405 percentiles). The estimated absolute risk of developing ovarian cancer by age 80 for *BRCA1*
406 carriers at the 99th PRS centiles were 63.2%, 66.3%, 59.0% and 68.4% for the lasso, elastic
407 net, stepwise and S4 models, respectively. The corresponding absolute risks for women at the
408 1st PRS centile were 27.7%, 25.6%, 30.8% and 24.2%. Absolute risks of developing EOC
409 were lower for *BRCA2* carriers than *BRCA1* carriers, with absolute risks for women in the
410 99th centile being 36.3%, 36.3%, 33.0% and 36.9%; and absolute risks for women in the 1st
411 centile being 7.10%, 7.12%, 8.24% and 6.92% for the lasso, elastic net, stepwise and S4

412 models, respectively. Absolute risks for *BRCA1* and *BRCA2* carriers at the 10th and 90th
413 percentile are provided in Supplementary Table 7.

414

415 *PRS distribution and ancestry*

416 To investigate the transferability of the PLR derived PRS to other populations, we applied the
417 scores to women of African (N=1,072) and Asian (N=7,669) ancestry genotyped as part of
418 the OncoArray project. In general, the distributions of the raw polygenic scores were
419 dependent on both the statistical methods used in SNP selection and ancestral group. PRS
420 models that included more variants had less dispersion, such that the elastic net models had
421 the least between individual variation in all ancestral groups (standard deviation=0.15, 0.19
422 and 0.22 for individuals of Asian, African and European ancestries respectively), while the
423 distributions from the stepwise models were the most dispersed (standard deviation = 0.23,
424 0.27 and 0.30 for individuals of Asian, African and European ancestries respectively). As
425 expected, given the variation in variant frequencies by population, the distribution of
426 polygenic scores was significantly different across the three ancestral groups, with the least
427 dispersion among women of Asian ancestry and the most variation in women of European
428 ancestry. The difference in polygenic risk score distribution was minimized after correction
429 for ancestry by standardizing the PRS to have unit standard deviation using the control
430 subjects for each ancestral group. For comparison, we investigated the use of the first 20
431 principal components to correct for ancestry and we obtained similar results.

432 High PRSs were significantly associated with risk of non-mucinous EOC in both Asian and
433 African ancestries (**Table 4**), although the effects were weaker than in women of European
434 ancestry. For example, with the lasso model, the odds ratio (95% CI) per unit standard
435 deviation increment in polygenic score was estimated to be 1.16 (1.11–1.22) in women of

436 East Asian ancestry and 1.28 (1.13–1.45) in women of African ancestry compared to 1.37
 437 (1.27–1.48) in women of European ancestry (p for heterogeneity < 0.0001). Variability in
 438 effect sizes among ancestral groups was highest for the stepwise model ($I^2 = 92\%$) *versus*
 439 84% and 83% for elastic net and lasso derived polygenic scores respectively. The best
 440 discriminative model among women of East Asian and African ancestry were the Elastic net
 441 PRS (AUC=0.543) and the Select and Shrink PRS model derived from OCAC-CIMBA meta-
 442 analysis (AUC=0.596) respectively. Women of African ancestry in the top 5% of the PRS
 443 had about two-fold increased risk compared to women in the middle quintile (lasso
 444 OR:1.64,95%CI: 0.90–3.00; elastic net OR:1.64,95%CI:0.90–3.00; stepwise OR:2.15,
 445 95%CI:1.17–3.95; S4 OR:1.80, 95%CI:0.99–3.31). Effect estimates were smaller in women
 446 of East Asian ancestry with women in the top 5% of the PRS, having about a 1.5 fold
 447 increased risk compared to women in the middle quintile (lasso OR:1.40, 95%CI:1.12–1.76;
 448 elastic net OR:1.60, 95% CI:1.28–2.01; stepwise OR:1.32, 95%CI:1.04–1.65; S4 OR:1.32,
 449 95%CI:1.05–1.66).

450 **Table 4: External validation of PRS models in East Asian and African Populations**

Model	East Asian ancestries			African ancestries		
	AUC	OR	95% CI	AUC	OR	95% CI
Lasso	0.541	1.16	(1.11 – 1.22)	0.576	1.28	(1.13 – 1.45)
Elastic net	0.543	1.17	(1.12 – 1.23)	0.574	1.29	(1.14 – 1.47)
Stepwise (OCAC)	0.528	1.11	(1.06 – 1.16)	0.581	1.34	(1.18 – 1.52)
Select and shrink (OCAC)	0.538	1.14	(1.08 - 1.19)	0.593	1.38	(1.21 – 1.58)
Stepwise (OCAC-CIMBA)	0.542	1.17	(1.11 – 1.23)	0.594	1.37	(1.20 – 1.56)
Select and shrink (OCAC-CIMBA)	0.537	1.14	(1.08 – 1.19)	0.596	1.41	(1.23 – 1.61)

451

452 **Discussion**

453 Genetic risk profiling with polygenic risk scores has led to actionable outcomes for cancers
454 such as breast and prostate (24,25). Previous PRS scores for invasive EOC risk in the general
455 population and *BRCA1/BRCA2* pathogenic variant carriers have been based on genetic
456 variants for which an association with EOC risk had been established at nominal genome-
457 wide significance (4,5,26–28). Here, we explored the predictive performance of
458 computationally-efficient, penalized, regression methods in modelling joint SNP effects for
459 EOC risk prediction in diverse populations and compared them with common approaches. By
460 leveraging the correlation between SNPs which do not reach nominal genome-wide
461 thresholds and including them in PRS models, the polygenic risk scores derived from
462 penalized regression models in this analysis provide stronger evidence of association with
463 risk of non-mucinous EOC than previously published PRSs in both the general population
464 and in *BRCA1/BRCA2* pathogenic variant carriers.

465 Recently, Barnes et. al derived a PRS score using 22 SNPs that were significantly associated
466 with high-grade serous EOC risk in GWAS (PRS_{HGS}) to predict EOC risk in *BRCA1/BRCA2*
467 pathogenic variant carriers (5). To make effect estimates obtained in this analysis comparable
468 to the effect estimates obtained from the PRS_{HGS}, we standardized all PRSs using the
469 standard deviation from unaffected *BRCA1/BRCA2* carriers; all PRS models in this analysis
470 except the Stepwise (OCAC only) had higher effect estimates (5). However, the
471 corresponding AUCs were higher for the PRS_{HGS} model (0.604 for *BRCA1* carriers and 0.667
472 for *BRCA2* carriers), most likely as a result of inclusion of other predictors (birth cohort and
473 principal components) in the model. The AUC estimates for women in the general
474 population, as estimated from the UK Biobank, are slightly higher than estimates from

475 previously published PRS models for overall EOC risk by Wei et al (AUC=0.57) and Yang et
476 al (AUC=0.58) (26,28)

477 In theory, polygenic risk profiling has the potential for clinical utility, being the earliest
478 measurable contributor to risk which may lead to actionable outcomes. The level of risk
479 among women considered to have a high polygenic risk score, for example women in the 95th
480 percentile, for all of the models we considered approaches the same level of risk conferred by
481 pathogenic variants in moderate penetrance genes such as *FANCM* (RR=2.1, 95%CI=1.1–
482 3.9) and *PALB2* (RR=2.91 95%CI=1.40–6.04) (29,30). The inclusion of other risk factors
483 such as family history of ovarian cancer, presence of rare pathogenic variants, age at
484 menarche, oral contraceptive use, hormone replacement therapy, parity, and endometriosis in
485 combination with the PRS models could potentially improve risk stratification as has been
486 implemented in the CanRisk tool (www.canrisk.org), which currently uses a PRS model
487 based on 36 SNPs with the potential to use other PRS models (31,32).

488 An important consideration in the clinical utility of polygenic risk scores is the degree to
489 which results are applicable to diverse populations. We found that the discriminative ability
490 varied substantially by ancestral group. As expected, given that the model development
491 dataset consisted entirely of women of European ancestry, the models had greater
492 discriminative power in women of European ancestry, relative to women of African and East
493 Asian ancestry. We observed greater attenuation of discriminative ability in East Asian
494 populations than African populations. This finding is in contrast to what one would expect
495 given human demographic history, and results from genome wide association studies for
496 EOC (18,19,33,34). One possible explanation for this disparity is the small sample size and
497 imprecise effect estimates for women of African ancestry in this study, due to the larger
498 differences in allele frequency between this population and that of the cohort used to develop

499 the model. Although the model development data for this analysis was predominantly women
500 of European ancestry, the models developed using our approach performed substantially
501 better in women of African ancestry than a PRS model developed by combining 24 published
502 GWAS SNPs associated with non-mucinous EOC, for which the odds of EOC risk was 1.20
503 fold per standard deviation of PRS (19).

504 Further refinements to our models, by exploring other penalty functions, may improve the
505 predictive value of the PRS. However, this approach may be complicated by difficulties that
506 arise due to the correlation structure between SNPs. Another option to optimizing the models
507 could be varying the penalization function based on prior knowledge. In genomic regions that
508 are known to have variants associated with EOC, one is more likely to find other risk-
509 associated variants. Therefore, varying the penalty function in these regions such that more
510 SNPs are selected into the model may improve the PRS. Finally, as more functional data
511 become available, modifying penalty functions to incorporate functional data may further
512 improve the PRS. Current approaches for incorporating functional annotation have resulted in
513 only modest gains in prediction accuracy for complex traits such as breast cancer, celiac
514 disease, type 2 diabetes and rheumatoid arthritis, much of which is attributed to the SNPs
515 selected in the models and not the functional annotation (35).

516 The UK Biobank, our model validation dataset for women in the general population, had a
517 small number of invasive EOC cases with a disproportionately high number of mucinous
518 cases (166 of the 823 invasive EOC cases or ~20%). Furthermore, cases of the serous
519 histotype could not be classified as either high-grade or low-grade. Therefore, we could not
520 investigate EOC histotype-specific polygenic scores. As the serous histotype is the most
521 common, it is possible that a high-grade serous EOC specific polygenic score may have
522 better predictive value than a non-mucinous polygenic score.

523 **Conclusion**

524 In conclusion, our results indicate that using the lasso model for individual level genotype
525 data and the S4 model for summary level data in polygenic risk score construction provide an
526 improvement in risk prediction for non-mucinous EOC over more common approaches. Our
527 approach overcomes the computational limitations in the use of penalized methods for large
528 scale genetic data, particularly in the presence of highly-correlated SNPs and the use of cross-
529 validation for parameter estimation is preferred. In practical terms, the polygenic risk score
530 provides sufficient discrimination, particularly for women of European ancestry, to be
531 considered for inclusion in risk prediction and prevention approaches for EOC in the future.
532 Further studies are required to optimize these polygenic risk scores in ancestrally diverse
533 populations and to validate their performance with the inclusion of other genetic and lifestyle
534 risk factors.

535 **Acknowledgements and Funding**

536 Full acknowledgement and funding details are provided in the Supplementary Material

537

538 **Conflicts of Interest**

539 Anna DeFazio has received a research grant from AstraZeneca, not directly related to the
540 content of this manuscript. Matthias W. Beckmann conducts research funded by Amgen,
541 Novartis and Pfizer. Peter A. Fashing conducts research funded by Amgen, Novartis and
542 Pfizer. He received Honoraria from Roche, Novartis and Pfizer. Allison W. Kurian reports
543 research funding to her institution from Myriad Genetics for an unrelated project. Usha
544 Menon owns stocks in Abcodia Ltd. Rachel A. Murphy is a consultant for Pharmavite. The
545 other authors declare no conflicts of interest.

546 **References**

- 547 1. Jones MR, Kamara D, Karlan BY, Pharoah PDP, Gayther SA. Genetic epidemiology of
548 ovarian cancer and prospects for polygenic risk prediction. *Gynecol Oncol.*
549 2017;147(3):705–13.
- 550 2. Lyra PCM, Rangel LB, Monteiro ANA. Functional Landscape of Common Variants
551 Associated with Susceptibility to Epithelial Ovarian Cancer. *Curr Epidemiol Rep.* 2020
552 Mar 1;7(1):49–57.
- 553 3. Kar SP, Berchuck A, Gayther SA, Goode EL, Moysich KB, Pearce CL, et al. Common
554 Genetic Variation and Susceptibility to Ovarian Cancer: Current Insights and Future
555 Directions. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored*
556 *Am Soc Prev Oncol.* 2018;27(4):395–404.
- 557 4. Kuchenbaecker KB, McGuffog L, Barrowdale D, Lee A, Soucy P, Dennis J, et al.
558 Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in
559 BRCA1 and BRCA2 Mutation Carriers. *J Natl Cancer Inst.* 2017 01;109(7).
- 560 5. Barnes DR, Rookus MA, McGuffog L, Leslie G, Mooij TM, Dennis J, et al. Polygenic
561 risk scores and breast and epithelial ovarian cancer risks for carriers of BRCA1 and
562 BRCA2 pathogenic variants. *Genet Med Off J Am Coll Med Genet.* 2020 Jul 15;
- 563 6. Wray NR, Goddard ME, Visscher PM. Prediction of individual genetic risk to disease
564 from genome-wide association studies. *Genome Res.* 2007 Oct;17(10):1520–8.
- 565 7. International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM,
566 O'Donovan MC, et al. Common polygenic variation contributes to risk of schizophrenia
567 and bipolar disorder. *Nature.* 2009 Aug 6;460(7256):748–52.
- 568 8. Abraham G, Kowalczyk A, Zobel J, Inouye M. Performance and robustness of
569 penalized and unpenalized methods for genetic prediction of complex human disease.
570 *Genet Epidemiol.* 2013 Feb;37(2):184–95.
- 571 9. Habier D, Fernando RL, Kizilkaya K, Garrick DJ. Extension of the bayesian alphabet
572 for genomic selection. *BMC Bioinformatics.* 2011 May 23;12:186.
- 573 10. Szymczak S, Biernacka JM, Cordell HJ, González-Recio O, König IR, Zhang H, et al.
574 Machine learning in genome-wide association studies. *Genet Epidemiol.* 2009;33 Suppl
575 1:S51-57.
- 576 11. Privé F, Aschard H, Blum MGB. Efficient Implementation of Penalized Regression for
577 Genetic Risk Prediction. *Genetics.* 2019;212(1):65–74.
- 578 12. Mak TSH, Porsch RM, Choi SW, Zhou X, Sham PC. Polygenic scores via penalized
579 regression on summary statistics. *Genet Epidemiol.* 2017;41(6):469–80.
- 580 13. Perren TJ. Mucinous epithelial ovarian carcinoma. *Ann Oncol.* 2016 Apr 1;27:i53–7.
- 581 14. Cheasley D, Wakefield MJ, Ryland GL, Allan PE, Alsop K, Amarasinghe KC, et al. The
582 molecular origin and taxonomy of mucinous ovarian carcinoma. *Nat Commun.* 2019
583 02;10(1):3935.

- 584 15. Meagher NS, Wang L, Rambau PF, Intermaggio MP, Huntsman DG, Wilkens LR, et al.
585 A combination of the immunohistochemical markers CK7 and SATB2 is highly
586 sensitive and specific for distinguishing primary ovarian mucinous tumors from
587 colorectal and appendiceal metastases. *Mod Pathol Off J U S Can Acad Pathol Inc.*
588 2019;32(12):1834–46.
- 589 16. Phelan CM, Kuchenbaecker KB, Tyrer JP, Kar SP, Lawrenson K, Winham SJ, et al.
590 Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian
591 cancer. *Nat Genet.* 2017 May;49(5):680–91.
- 592 17. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank
593 resource with deep phenotyping and genomic data. *Nature.* 2018;562(7726):203–9.
- 594 18. Lawrenson K, Song F, Hazelett DJ, Kar SP, Tyrer J, Phelan CM, et al. Genome-wide
595 association studies identify susceptibility loci for epithelial ovarian cancer in east Asian
596 women. *Gynecol Oncol.* 2019;153(2):343–55.
- 597 19. Manichaikul A, Peres LC, Wang X-Q, Barnard ME, Chyn D, Sheng X, et al.
598 Identification of novel epithelial ovarian cancer loci in women of African ancestry. *Int J*
599 *Cancer.* 2020 Jun 1;146(11):2987–98.
- 600 20. Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of ANthropometric
601 Traits (GIANT) Consortium, DIABetes Genetics Replication And Meta-analysis
602 (DIAGRAM) Consortium, et al. Conditional and joint multiple-SNP analysis of GWAS
603 summary statistics identifies additional variants influencing complex traits. *Nat Genet.*
604 2012 Mar 18;44(4):369–75, S1-3.
- 605 21. Vilhjálmsson BJ, Yang J, Finucane HK, Gusev A, Lindström S, Ripke S, et al.
606 Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *Am J*
607 *Hum Genet.* 2015 Oct 1;97(4):576–92.
- 608 22. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian
609 regression and continuous shrinkage priors. *Nat Commun.* 2019 16;10(1):1776.
- 610 23. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips K-A, Mooij TM, Roos-Blom M-J,
611 et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2
612 Mutation Carriers. *JAMA.* 2017 20;317(23):2402–16.
- 613 24. Mavaddat N, Michailidou K, Dennis J, Lush M, Fachal L, Lee A, et al. Polygenic Risk
614 Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am J Hum Genet.*
615 2019 03;104(1):21–34.
- 616 25. Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ, et al.
617 Association analyses of more than 140,000 men identify 63 new prostate cancer
618 susceptibility loci. *Nat Genet.* 2018;50(7):928–36.
- 619 26. Jia G, Lu Y, Wen W, Long J, Liu Y, Tao R, et al. Evaluating the Utility of Polygenic
620 Risk Scores in Identifying High-Risk Individuals for Eight Common Cancers. *JNCI*
621 *Cancer Spectr.* 2020 Jun;4(3):pkaa021.
- 622 27. Pearce CL, Stram DO, Ness RB, Stram DA, Roman LD, Templeman C, et al.
623 Population distribution of lifetime risk of ovarian cancer in the United States. *Cancer*

- 624 Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol.
625 2015 Apr;24(4):671–6.
- 626 28. Yang X, Leslie G, Gentry-Maharaj A, Ryan A, Intermaggio M, Lee A, et al. Evaluation
627 of polygenic risk scores for ovarian cancer risk prediction in a prospective cohort study.
628 J Med Genet. 2018;55(8):546–54.
- 629 29. Yang X, Leslie G, Doroszuk A, Schneider S, Allen J, Decker B, et al. Cancer Risks
630 Associated With Germline PALB2 Pathogenic Variants: An International Study of 524
631 Families. J Clin Oncol Off J Am Soc Clin Oncol. 2020 Mar 1;38(7):674–85.
- 632 30. Song H, Dicks EM, Tyrer J, Intermaggio M, Chenevix-Trench G, Bowtell DD, et al.
633 Population-based targeted sequencing of 54 candidate genes identifies PALB2 as a
634 susceptibility gene for high-grade serous ovarian cancer. J Med Genet. 2020 Jun 16;
- 635 31. Lee A, Mavaddat N, Wilcox AN, Cunningham AP, Carver T, Hartley S, et al.
636 BOADICEA: a comprehensive breast cancer risk prediction model incorporating
637 genetic and nongenetic risk factors. Genet Med. 2019 Aug;21(8):1708–18.
- 638 32. Welcome to CanRisk [Internet]. [cited 2020 Aug 31]. Available from:
639 <https://www.canrisk.org/>
- 640 33. Chen K, Ma H, Li L, Zang R, Wang C, Song F, et al. Genome-wide association study
641 identifies new susceptibility loci for epithelial ovarian cancer in Han Chinese women.
642 Nat Commun. 2014 Aug 19;5:4682.
- 643 34. Watkins WS, Rogers AR, Ostler CT, Wooding S, Bamshad MJ, Brassington A-ME, et
644 al. Genetic variation among world populations: inferences from 100 Alu insertion
645 polymorphisms. Genome Res. 2003 Jul;13(7):1607–18.
- 646 35. Hu Y, Lu Q, Powles R, Yao X, Yang C, Fang F, et al. Leveraging functional annotations
647 in genetic risk prediction for human complex diseases. PLoS Comput Biol. 2017
648 Jun;13(6):e1005589.
- 649

Figure captions

Figure 1: PRS model development using penalized regression and LDpred Bayesian approach

Figure 2: Cumulative risk of ovarian cancer between birth and age 80 by PRS percentiles and PRS models. Shown are the cumulative risk of ovarian cancer risk in UK women by polygenic risk score percentiles. The lasso (A) and elastic net (B) penalized regression models were applied to individual level genotype data, while the stepwise (C) and S4 (D) models were applied to summary level statistics.

Figure 3: Association between the PLR PRS models and non-mucinous ovarian cancer by PRS percentiles. Shown are estimated odds ratios (OR) and confidence intervals for women of European ancestry by percentiles of polygenic risk scores derived from lasso (A), elastic net (B), stepwise (C) and S4 (D) models relative to the middle quintile.

Figure 1: PRS model development using penalized regression and LDpred Bayesian approach.

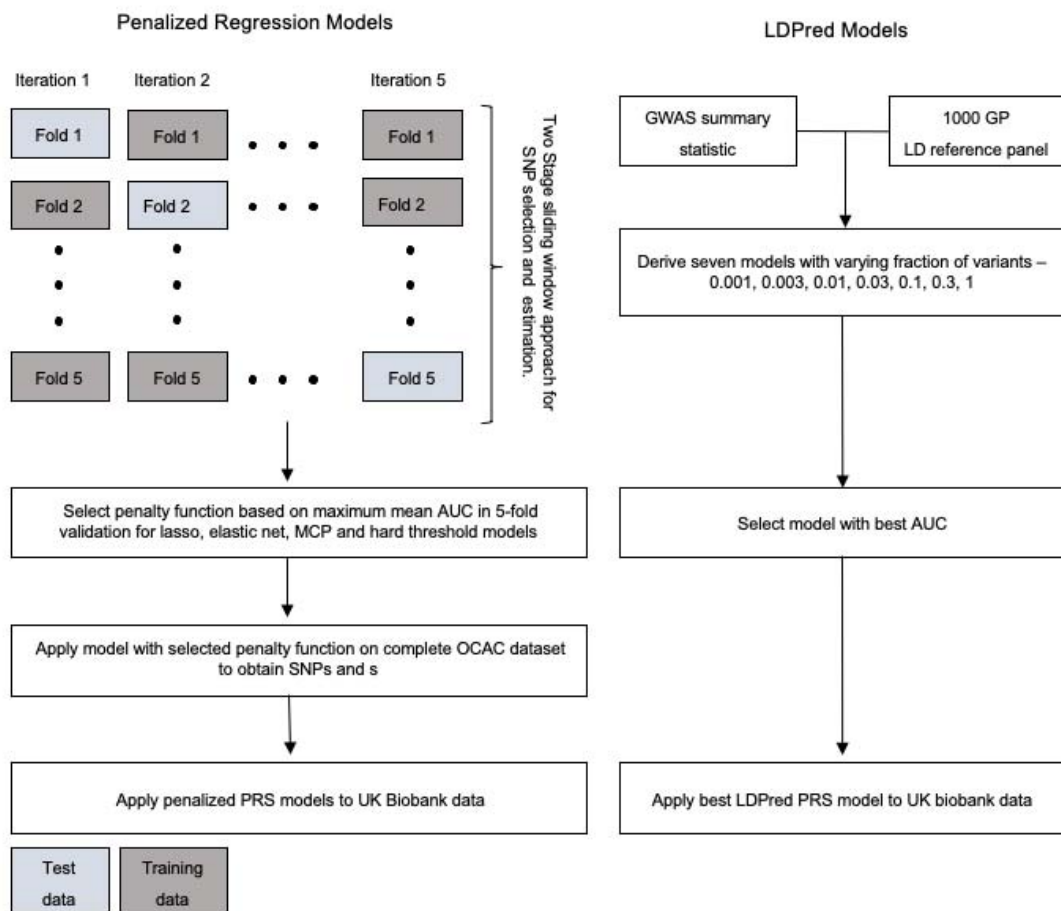


Figure 2: Cumulative risk of ovarian cancer between birth and age 80 by PRS percentiles and PRS models

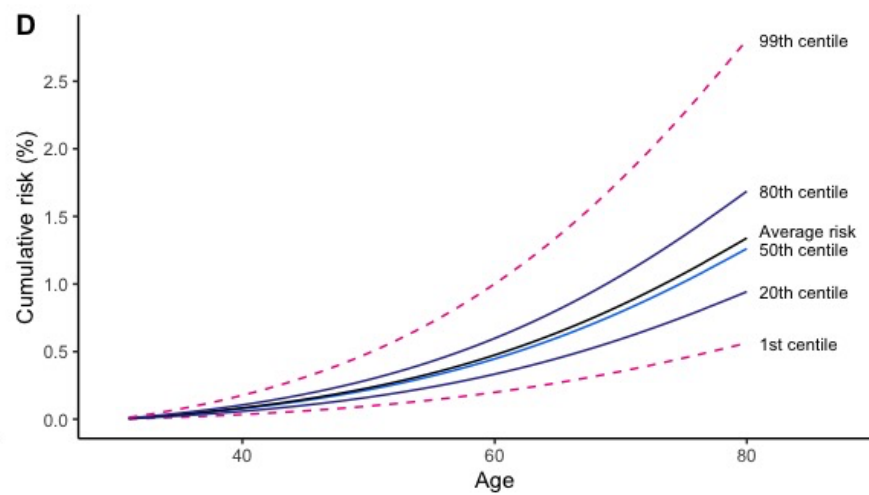
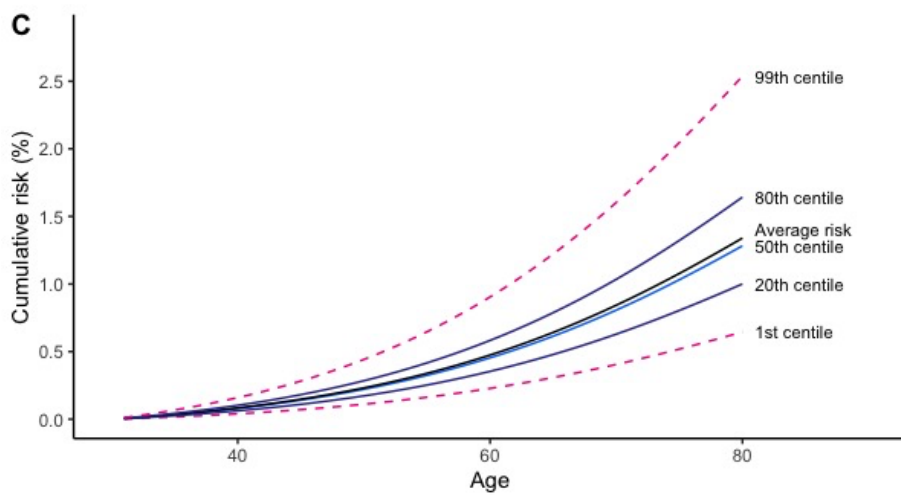
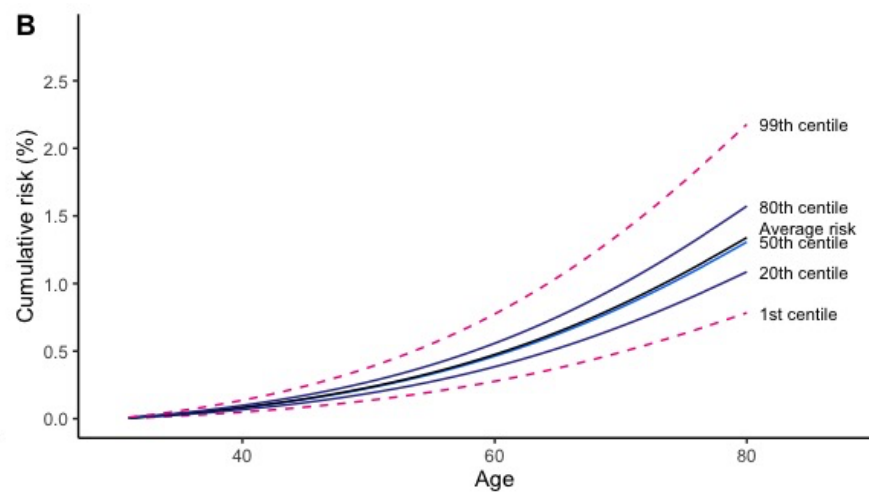
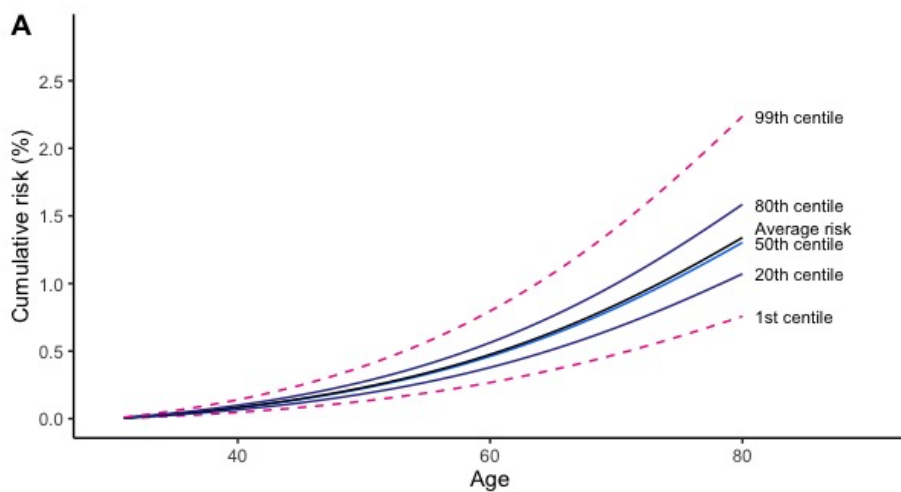


Figure 3: Association between the PLR PRS models and non-mucinous ovarian cancer by PRS percentiles

