

1 **Association between Fuchs Endothelial Corneal Dystrophy, Diabetes Mellitus and**
2 **Multimorbidity**

3

4 Cari L. Nealon, OD,^{1,#} Christopher W. Halladay, MS,^{2,#} Bryan R. Gorman, OD,^{3,4,#} Piana
5 Simpson, OD,¹ David P. Roncone, OD,¹ Rachael L. Canania, OD,¹ Scott A. Anthony, OD,¹
6 Lea R. Sawicki Rogers, OD,⁵ Jenna N. Leber, OD,⁵ Jacquelyn M. Dougherty, OD,⁵ Jessica
7 N. Cooke Bailey, PhD, PhD,^{6,7,8} Dana C. Crawford, PhD,^{6,7,8} Jack M. Sullivan, MD, PhD,^{5,9,10}
8 Anat Galor, MD, MSPH,^{11,12} Wen-Chih Wu, MD,¹³ Paul B. Greenberg, MD,^{14,15} the Million
9 Veteran Program, Jonathan H. Lass, MD,^{16,17} Sudha K. Iyengar, PhD,^{6,7,8,5} Neal S. Peachey,
10 PhD^{8,18,19,5}

11

12 **Author affiliations:**

13 ¹ Eye Clinic, VA Northeast Ohio Healthcare System, Cleveland, Ohio, USA

14 ² Center of Innovation in Long Term Services and Supports, Providence VA Medical Center,
15 Providence, Rhode Island, USA

16 ³ VA Cooperative Studies Program, VA Boston Healthcare System, Boston, Massachusetts

17 ⁴ Booz Allen Hamilton, McLean, Virginia, USA

18 ⁵ Ophthalmology Section, VA Western NY Health Care System, Buffalo, New York, USA

19 ⁶ Cleveland Institute for Computational Biology, Case Western Reserve University School of
20 Medicine, Cleveland, Ohio, USA

21 ⁷ Department of Population & Quantitative Health Sciences, Case Western Reserve
22 University School of Medicine, Cleveland, Ohio, USA

23 ⁸ Research Service, VA Northeast Ohio Healthcare System, Cleveland, Ohio, USA

24 ⁹ Research Service, VA Western NY Health Care System, Buffalo, New York, USA

25 ¹⁰ Department of Ophthalmology (Ross Eye Institute), University at Buffalo-SUNY, Buffalo,
26 New York, USA

27 ¹¹ Miami Veterans Affairs Medical Center, Miami, Florida, USA

28 ¹² Bascom Palmer Eye Institute, University of Miami, Miami, Florida, USA

29 ¹³ Cardiology Section, Medical Service, Providence VA Medical Center, Providence, Rhode
30 Island, USA

31 ¹⁴ Ophthalmology Section, Providence VA Medical Center, Providence, Rhode Island, USA

32 ¹⁵ Division of Ophthalmology, Alpert Medical School, Brown University, Providence, Rhode
33 Island, USA

34 ¹⁶ Department of Ophthalmology & Visual Sciences, Case Western Reserve University,
35 Cleveland, Ohio, USA

36 ¹⁷ University Hospitals Eye Institute, Cleveland, Ohio, USA

37 ¹⁸ Cole Eye Institute, Cleveland Clinic Foundation, Cleveland, Ohio, USA

38 ¹⁹ Department of Ophthalmology, Cleveland Clinic Lerner College of Medicine of Case
39 Western Reserve University, Cleveland, Ohio, USA

40 # These authors contributed equally to this work

41 § These authors jointly supervised this work

42

43 Correspondence:

44 Neal S. Peachey, Research Service (151W), VA Northeast Ohio Healthcare System, 10701
45 East Boulevard, Cleveland OH 44106, USA (neal.peachey@va.gov; 216-791-3800 x64661);
46 Sudha K. Iyengar, Case Western Reserve University School of Medicine, 2103 Cornell Road,
47 Cleveland, OH 44106, USA (ski@case.edu; 216-368-4388)

48

49 Conflict of interest: none

50

51 Key words: Corneal dystrophy, diabetes, comorbidity, Fuchs endothelial corneal dystrophy

52

53 Supported by the VA Office of Research & Development (I01 BX003364), the Cleveland
54 Institute for Computational Biology, NIH Core Grants (P30 EY025585, P30 EY011373), the
55 Clinical and Translational Science Collaborative of Cleveland (UL1TR002548) from the
56 National Center for Advancing Translational Sciences (NCATS) component of the NIH and
57 NIH Roadmap for Medical Research, and unrestricted grants from Research to Prevent
58 Blindness to Case Western Reserve University, Cleveland Clinic Lerner College of Medicine
59 of Case Western Reserve University, and the University of Buffalo. This work was also
60 supported by the Harper-Inglis Memorial for Eye Research, The Peierls Foundation, and That
61 Man May See.

62

63 **Abstract**

64 **Purpose:** To assess risk for demographic variables and other health conditions that are
65 associated with Fuchs endothelial corneal dystrophy (FECD).

66 **Methods:** We developed a FECD case-control algorithm based on structured EHR data and
67 accuracy confirmed by individual review of charts at three VA Medical Centers. This algorithm
68 was applied to the Department of Veterans Affairs Million Veteran Program cohort from whom
69 sex, genetic ancestry, comorbidities, diagnostic phecodes and laboratory values were
70 extracted. Single and multiple variable logistic regression models were used to determine the
71 association of these risk factors with FECD diagnosis.

72 **Results:** Being a FECD case was associated with female sex, European genetic ancestry,
73 and a greater number of comorbidities. Of 1417 diagnostic phecodes evaluated, 213 had a
74 significant association with FECD, falling in both ocular and non-ocular conditions, including
75 diabetes mellitus (DM). Five of 69 laboratory values were associated with FECD, with the
76 direction of change for four being consistent with DM. Insulin dependency and type 1 DM
77 raised risk to a greater degree than type 2 DM, like other microvascular diabetic
78 complications.

79 **Conclusions:** Female gender, European ancestry and multimorbidity increased FECD risk.
80 Endocrine/metabolic clinic encounter codes as well as altered patterns of laboratory values
81 support DM increasing FECD risk. Our results evoke a threshold model in which the FECD
82 phenotype is intensified by DM and potentially other health conditions that alter corneal
83 physiology. DM may modify FECD onset and encourage progression among susceptible
84 individuals, suggesting that optimizing glucose control may be an effective preventative for
85 FECD.

86

87 Running Head: Fuchs Endothelial Corneal Dystrophy

88 Longitudinal electronic health records (EHRs) represent an amazing opportunity to identify
89 previously unappreciated or unrecognized relationships between complex conditions. In the
90 present study, we focused on Fuchs endothelial corneal dystrophy (FECD) a common
91 corneal dystrophy and a leading indication for corneal transplantation in the United States
92 (US).¹⁻⁴ FECD is an important healthcare challenge with an estimated prevalence in the adult
93 population of 4%.⁵ FECD is characterized by bilateral endothelial cell pleomorphism and
94 polymegathism, decreased endothelial cell density, guttae, and corneal edema.^{1,6} Two
95 subtypes are recognized based on the age of onset. The early subtype, linked to mutations in
96 the *COL8A2* gene, is rare and considered to have a more severe disease course.^{7,8} The
97 more common late subtype has an onset during the second or third decade becoming
98 symptomatic in the fifth to sixth decade⁶ with well-established links to mutations in *TCF4* and
99 *SLC4A11*.⁹

100 The etiology of FECD involves a complex interaction between incompletely penetrant
101 genetic factors that increase risk and biological and environmental factors that contribute to
102 disease manifestation.⁴ The profile of genetic risk related to FECD has become clearer
103 thanks to recent large-scale genetic analyses.¹⁰ Biological factors such as estrogen levels
104 and oxidative stress, and environmental factors such as UV light exposure, have been
105 implicated in FECD.^{4,11} This increased understanding of FECD pathogenesis prompted the
106 present study, wherein we reasoned that additional genetic, non-genetic biological and
107 environmental factors that relate to FECD might also be involved in other disorders, and thus
108 lead to an increased incidence in patients with FECD. This possibility has not been a major
109 research focus but is supported from two reports. Xu and colleagues¹² analyzed a U.S.
110 Medicare database to demonstrate that patients with FECD were at increased risk for breast,
111 ovarian, thyroid and several cutaneous malignancies, while also being at decreased risk for
112 lung and prostate cancer. Chang and colleagues¹³ analyzed the Taiwan Longitudinal Health
113 Insurance Database 2000 and noted that patients with ocular allergic conditions were more
114 likely to be diagnosed with FECD. These reports motivate our effort to comprehensively
115 analyze the relationship between FECD and other disease phenotypes. To achieve this goal,
116 we used the US Department of Veterans Affairs (VA) sponsored Million Veteran Program
117 (MVP), an ongoing observational cohort and mega-biobank, to explore whether other disease
118 phenotypes are associated with FECD. MVP integrates data from EHRs, health surveys, and
119 genomic data,^{14,15} with more than 825,000 Veterans successfully enrolled across the VA

120 healthcare system (www.mvp.va.gov). MVP has substantial ethnic diversity that parallels the
121 US population providing research opportunities across multiple ethnic groups.¹⁵

122 Before embarking on this analysis, we first addressed an important challenge in using
123 EHR-linked biobanks to accurately identify cases and controls through computable or
124 electronic phenotyping. Computable phenotyping utilizes diagnostic, procedure, and office
125 visit codes as well as prescription and/or laboratory tests to identify cases and controls for the
126 disease of interest.¹⁶ Rules-based algorithms have been successfully used within MVP to
127 identify other ocular diseases including age-related macular degeneration¹⁷ and primary open
128 angle glaucoma.¹⁸ To support our analysis, we developed a FECD phenotype algorithm
129 based on International Classification of Diseases Clinical Modification (ICD-CM) codes and
130 Current Procedural Terminology (CPT) codes to accurately identify FECD cases and controls
131 within the VA Computerized Patient Record System (CPRS).¹⁹

132 We used our FECD case-control algorithm to conduct a phenotype wide association
133 study (PheWAS) for FECD. Although PheWAS was originally developed to understand
134 genetic pleiotropy,²⁰ the use of PheWAS has been broadened to identify relationships
135 between a disease of interest and other conditions.^{21,22} Herein we used the PheWAS
136 approach to identify relationships between FECD and other disease phenotypes. Considering
137 prior reports that FECD incidence may be higher in females^{4,23,24} and may vary between
138 racial groups^{25,26} we incorporated these factors into our analysis and noted that female sex
139 and European ancestry were associated with FECD. Our main finding was that diagnostic
140 codes for multiple ocular and non-ocular conditions were elevated in FECD cases. Further,
141 these associations and laboratory value alterations identified in FECD indicate that diabetes
142 mellitus (DM) may contribute to the development of FECD.

143

144 **Materials & Methods**

145 **Study Participants**

146 The MVP provides access to health, genetic and lifestyle information for U.S. Veterans who
147 provided informed consent (<https://www.research.va.gov/mvp/>). When this study was
148 conducted, MVP had provided data for 658,582 Veterans. Local chart reviews were
149 conducted at the Eye Clinics of three VA Medical Centers (VAMCs; VA Northeast Ohio
150 Healthcare System (Cleveland, OH), Providence VA Medical Center (Providence, RI), and
151 VA Western New York Healthcare System (Buffalo, NY). This study adhered to the tenets of

152 the Declaration of Helsinki and was approved by the VA Office of Research & Development
153 Central Institutional Review Board (VA CIRB 18-41).

154

155 **Developing a Case-Control Algorithm for FECD**

156 We designed a rules-based algorithm utilizing structured EHR data (ICD-9-CM, ICD-10-CM,
157 and CPT codes) to identify FECD cases and controls within CPRS at each VAMC. Cases
158 greater than 18 years old were identified based on the presence of two FECD ICD-9-CM or
159 ICD-10-CM codes along with the absence of other ICD-9-CM or ICD-10-CM codes for
160 corneal conditions or complicated cataract surgery that could confound the diagnosis.
161 Controls were identified by excluding ICD-9-CM or ICD-10-CM codes for FECD, confounding
162 corneal conditions, and complicated cataract surgery in Veterans 65 years and older. The
163 algorithm was refined over five iterations; in each iteration, we used the algorithm to identify
164 up to 25 cases and 25 controls, in both non-Hispanic white (NHW) and non-Hispanic black
165 (NHB) Veterans and evaluated algorithm performance in those patients with chart reviews.
166 Race and ethnicity were based on the consensus self-reported identify maintained by the VA
167 Information Resource Center.

168 Chart review by Cleveland VAMC optometrists allowed for our FECD case-control
169 algorithm to be tested and refined until the final version was determined. To be confirmed as
170 a case, Veterans must have documented slit lamp findings of bilateral corneal endothelial
171 pathology. These structural signs were accompanied by FECD-relevant descriptors: bilateral
172 corneal guttae, a “beaten metal appearance”, stromal corneal edema, bullae, and bullous
173 keratopathy.^{1,6} Case exclusion criteria included prior complicated ocular surgeries or traumas
174 causing secondary corneal endothelial disease. Cases of postoperative intraocular surgery
175 with pseudophakic corneal edema, pseudophakic bullous keratopathy, or corneal
176 decompensation were included if FECD was documented pre-operatively. Clinicians were
177 instructed to use best clinical judgment and available clinical documentation to determine
178 inclusion status.

179 To be confirmed as controls, Veterans must have no signs of FECD or any other
180 corneal endothelial disease. Finalized inclusion criteria codes included H18.51 for ICD-10-CM
181 and 371.57 for ICD-9-CM. Exclusion criteria codes are displayed in **Supplementary Table 1**.

182 **Figure 1** presents an overview of the process used to evaluate algorithm performance.
183 We first calculated positive and negative predictive values (PPV, NPV) for the Cleveland

184 VAMC site, in parallel for NHB and NHW Veterans. Once we achieved >85% PPV and NPV
185 at the Cleveland VAMC, the algorithm was also evaluated by Eye Clinic staff at the Buffalo
186 and Providence VAMCs.

187

188 **Analyses Within the Million Veteran Program Biobank**

189 We applied the final FECD case-control algorithm to the available MVP cohort of 658,582
190 Veterans with available genetic data. Following identification of cases and controls, data on
191 sex and current age were extracted from medical records or the MVP baseline lifestyle
192 survey. Using participant genotypes²⁷ we classified individuals by genetic ancestry into
193 admixed African (AFR), Asian (ASN), European (EUR), and admixed American (HIS)
194 populations via the HARE (harmonized ancestry and race/ethnicity) machine learning
195 algorithm.²⁸

196 We used single and multiple variable logistic regression to understand the impact of
197 sex, genetic ancestry (coded as a categorical variable) and Charlson Comorbidity Index (CCI)
198 on FECD. Since our algorithm used age to select cases and controls, we could not test the
199 effect of age, but did use this variable to adjust for confounding. In our multiple regression
200 model, we also examined comorbidity in cases and controls using the CCI.²⁹ Based on prior
201 reports that US adults have an average of 2.2 comorbid conditions,^{30,31} we compared
202 Veterans without comorbid conditions to those with either 1-3 or 4 or more comorbidities,
203 coded as categorical variables.

204

205 **Association of FECD Diagnosis with other Phenotypes and Laboratory Values**

206 Phecodes combine ICD codes into clinically meaningful phenotypes^{32,33} and have traditionally
207 been used to perform PheWAS in the context of genetic risk factors.²⁰ In our study we used
208 phecodes to understand phenotypic relationships between FECD and multiple conditions. We
209 also determined if relationships existed between FECD and laboratory values. Phenotype
210 and laboratory analyses were both implemented in the R PheWAS package.³⁴ For each ICD
211 code, cases were defined as at least 2 counts of the code on separate days, and controls
212 were 0 counts. Subjects with 1 count were excluded. Additionally, we applied phenotype-
213 based control exclusion (i.e., exclusion of controls that were cases in similar phenotypes) and
214 we excluded females in male-specific phenotypes and vice versa. We only considered
215 phenotypes with 100 or more cases. For laboratory values, we used the median value in

216 regression analyses, and required that everyone had at least two lab measurements to be
217 included. Labs with measurements in at least 50 individuals were considered. All ICD code-
218 based phenotypes were tested with logistic regression, and all lab-based phenotypes were
219 tested with linear regression, adjusting for age, age-squared, sex (for non-sex specific
220 phenotypes), and 20 Principal Components of genome-wide genetic markers representing
221 population structure in both.

222

223 **Results**

224 **Developing an Accurate FECD Case-Control Algorithm**

225 Initial algorithm development and refinement via chart review was completed at the Cleveland
226 VAMC Eye Clinic. At each iteration, the algorithm identified 50 putative cases (25 NHW, 25
227 NHB) and 50 putative controls (25 NHW, 25 NHB) whose charts were then individually
228 reviewed. The first four iterations identified errors which were used to refine the algorithm.
229 For example, during early algorithm development, multiple putative cases were identified as
230 epithelial basement membrane dystrophy which affects the corneal epithelium rather than the
231 endothelium which is involved in FECD. Control accuracy was also affected by coding errors.
232 For example, some putative controls had clear changes to the corneal endothelium (guttae)
233 but this was not coded in the medical record.

234 For the fifth and final iteration, review of the cases identified six errors in which
235 postoperative complications made it difficult to determine if FECD was present before
236 surgical intervention, yielding a PPV of 88%. Review of the 50 controls identified three errors
237 where guttae were present, but a FECD ICD-9-CM or ICD-10-CM was not recorded in the
238 chart, yielding a NPV of 94%.

239 We next evaluated algorithm performance at the Buffalo VAMC and Providence VAMC
240 Eye Clinics. At Buffalo, 42 FECD cases (25 NHW, 17 NHB) and 50 controls (25 NHW, 25
241 NHB) were identified by the algorithm. After chart review by eye care providers, a PPV of
242 78% and NPV of 98% were determined. At Providence, 26 FECD cases (25 NHW, 1 NHB)
243 and 49 (24 NHW, 25 NHB) controls were reviewed by eye care providers. The PPV was 77%
244 while the NPV was 100%.

245 Combined across all three VA sites, the FECD algorithm performed with an 82% PPV
246 and 97% NPV (**Table 1**). The PPV was higher in NHB Veterans (93%) than in NHW Veterans
247 (76%). NPV was similar between NHB and NHW Veterans at 98% and 96%, respectively.

248

249 **FECD Risk Factors include Female Sex and European Ancestry**

250 When applied to the genotyped MVP cohort, the algorithm identified 2,663 cases and
251 146,659 controls (**Table 2**), exceeding the sizes of the case and control cohorts analyzed in
252 the largest reported genome-wide association study for FECD.¹⁰ While this seems like a
253 large drop-off from the original MVP cohort of 658,582, a major factor is the requirement that
254 controls be 65 years of age or older and presence of an eye exam. Most cases and controls
255 were male, reflecting the historical prevalence of males in the U.S. military and that >90% of
256 MVP enrollees are male.³⁵ Stratified by ancestry, our algorithm identified 2,264 EUR, 315
257 AFR, 80 HIS, and 4 ASN cases, partially reflecting the relative ancestry proportions in MVP
258 (~70% EUR overall).

259 Prior reports indicate the frequency of FECD is elevated in females as compared to
260 males, although the estimate of the magnitude of this difference varies^{4,23,24} and has not been
261 examined in diverse populations. We examined the frequency of females vs. males among
262 cases and noted that the frequency of FECD varied across race and was much higher in
263 females (**Table 2**). Single variable logistic regression (**Supplementary Table 2**) showed that
264 FECD cases had a 5.6X greater odds of being female (OR=5.60; CI: 4.93-6.33; $p = 1.07 \times 10^{-160}$).
265 Single variable analysis of genetic ancestry as a categorical variable (EUR vs AFR, EUR
266 vs HIS, and EUR vs ASN) showed that individuals of EUR ancestry were at greatest risk for
267 FECD (**Supplementary Table 2**). The mean (\pm S.D.) age of FECD cases was significantly
268 younger (69.52 ± 10.04 years) than controls (71.68 ± 6.05 years) (**Fig. 2**). As our algorithm
269 implemented different age criteria for FECD cases and controls, we used logistic regression
270 to adjust for this age difference in subsequent multivariable logistic regression models (**Table**
271 **3**) but did not formally test for age differences. We incorporated the CCI in the model to
272 determine whether the non-ocular conditions as ascertained by higher scores demonstrated
273 association with FECD.

274 Although women comprise only ~9% of MVP enrollees, we found a much higher risk
275 for females after adjusting for age, and ancestry group (OR=4.60; CI: 4.03 – 5.22; $p < 0.001$).
276 In fact, female sex was the most significant effect in our analyses. Non-European ancestry
277 (AFR, ASN, HIS) was associated with reduced risk (OR=0.81, 95% CI: 0.71 - 0.91, $p < 0.001$;
278 OR=0.18, 95% CI: 0.06 - 0.42, $p < 0.001$; OR=0.53, 95% CI: 0.42 – 0.66, $p < 0.001$,
279 respectively).

280

281 **FECD is associated with multiple other disease conditions**

282 The CCI predicts long-term mortality for multiple medical conditions and can be used to
283 identify the impact of overall health on diagnostic and prognostic factors.³⁶ After adjusting for
284 other covariates (age, sex and ancestry group), we found that cases with FECD were 1.21
285 times more likely to have 1-3 comorbid conditions identified via the CCI, compared with a
286 control without FECD (95% CI: 1.06 - 1.39, $p = 0.005$). A similar analysis of an increasing
287 number of comorbid conditions, 4 or greater, elevated the odds ratio to 1.40 (95% CI: 1.11 -
288 1.75; $p = 0.003$).

289 We used a PheWAS to better understand the nature of the elevated number of
290 comorbidities noted in our more numerous Veterans of European ancestry. Of the total
291 number of 1,814 diagnostic phecodes, 1,417 met our criteria of including at least 100 cases.
292 Of these, 213 (15.0%) were more prevalent in FECD cases than controls after Bonferroni
293 correction (**Supplementary Table 3**). Significant associations were noted in virtually every
294 phecode group (**Figure 3; Table 4**). The phegroup '**Sense Organs**' had the highest number
295 and rate of significant phecodes (51 out of 118; 43.22%); the five most significant phecodes
296 were for corneal dystrophy (***cornea replaced by transplant, corneal dystrophy, corneal***
297 ***opacity and other disorders, Fuchs dystrophy, corneal edema***). Other highly significant
298 phecodes were ***inflammation of the eye*** (OR=2.25; 95% CI: 2.03 – 2.49; $p < 2.59 \times 10^{-51}$),
299 ***dry eyes*** (OR=1.95; 95% CI: 1.78 – 2.15; $p < 1.24 \times 10^{-44}$), and ***uveitis, noninfectious or***
300 ***NOS*** (OR=4.43; 95% CI: 3.60 – 5.39; $p < 8.86 \times 10^{-34}$). The phegroup '**Dermatologic**' also
301 had a high rate of significant phecodes (19 out of 73; 26.0%) and these included ***corns and***
302 ***callosities*** (OR=1.56; 95% CI: 1.36 – 1.77; $p < 1.99 \times 10^{-10}$), ***other hypertrophic and***
303 ***atrophic conditions*** (OR=1.53; 95% CI: 1.36 – 1.72; $p < 5.78 \times 10^{-12}$) and ***disturbances of***
304 ***skin sensation*** (OR=1.54; 95% CI: 1.34 – 1.76; $p < 2.59 \times 10^{-9}$). The '**Musculoskeletal**'
305 phegroup had 16 out of 109 (14.7%) phecodes that were significantly higher in cases than
306 controls, and these included a number related to arthritis (***peripheral enthesopathies and***
307 ***allied syndromes*** (OR=1.45; 95% CI: 1.32 – 1.59; $p < 8.86 \times 10^{-14}$), ***osteoarthritis NOS***
308 (OR=1.35; 95% CI: 1.23 – 1.48; $p < 1.58 \times 10^{-10}$), and abnormal wear of spinal disks
309 (***spondylosis and allied disorders*** (OR=1.42; 95% CI: 1.27 – 1.58; $p < 1.68 \times 10^{-9}$),
310 ***spondylosis without myelopathy*** (OR=1.40; 95% CI: 1.25 – 1.57; $p < 1.58 \times 10^{-8}$).

311 Within the '**Endocrine/Metabolic**' phegroup, we noted a number of phecodes related
312 to DM. For example, the odds of having the **insulin pump user** phecode was 2.2 times
313 higher among FECD cases versus controls (95% CI: 1.82 – 2.68; $p < 2.75 \times 10^{-13}$). Phecodes
314 for **Type 1 diabetes** (OR=2.01; 95% CI: 1.69 – 2.37; $p < 1.18 \times 10^{-13}$) and **Type 2 diabetes**
315 (OR=1.25; 95% CI: 1.15 – 1.37; $p < 4.55 \times 10^{-7}$) were both more prevalent in FECD cases,
316 along with a number of phecodes related to associations of Type 1 or Type 2 DM and specific
317 systems (neurological, ophthalmic, renal). The phecode **obesity** was also more prevalent in
318 FECD cases (OR=1.34; 95% CI: 1.22 – 1.47; $p < 1.73 \times 10^{-10}$) and this may be related to the
319 multiple '**Digestive**' (15 out of 146; 10.3%) and '**Respiratory**' (12 out of 77; 15.6%) phecodes
320 that that were significantly more prevalent in FECD cases than controls.

321 The only phegroups where FECD cases were not distinguished from controls were
322 '**pregnancy**' (0 out of 1; 0%), '**congenital anomalies**' (1 out of 30; 3.3%) and '**neoplasms**'
323 (3 out of 113; 2.7%). Disorders of pregnancy and congenital anomalies have relative low
324 relevance to the adult Veteran population that we examined.

325 In our PheWAS for Veterans of AFR ancestry, forty out of 942 (4.25%) phecodes were
326 more prevalent in FECD cases than controls. While this analysis had less power due to the
327 smaller sample size, many of the phecodes that were significant reflected our EUR analysis,
328 including **Fuchs' dystrophy, corneal dystrophy, corneal opacity and other disorders of**
329 **cornea, corneal edema, dry eyes, and osteoarthritis**. Both Type 1 and Type 2 DM were
330 more frequent in AFR cases, attaining nominal significance that did not withstand multiple
331 testing correction, but displaying the same direction of effect. In sum, the effects were
332 consistent across groups, but lacked adequate sample sizes in AFR.

333 While more work is needed to fully understand this higher rate of non-ocular phecodes
334 in FECD cases than controls, our analysis of laboratory data were consistent with an
335 important role for DM. Out of the 69 available measures, five were significantly altered in
336 FECD cases of EUR ancestry after Bonferroni correction (**Supplementary Table 4**). For four
337 of these, the change direction was consistent with DM. FECD cases had elevated levels of
338 hemoglobin A1C (OR=1.11; CI: 1.06 – 1.16; $p < 2.47 \times 10^{-6}$) and serum glucose (OR=1.08; CI:
339 1.03 – 1.13; $p = 5.54 \times 10^{-4}$) while hematocrit (OR=0.90; CI: 0.86 – 0.94; $p < 1.57 \times 10^{-6}$) and
340 hemoglobin (OR=0.90; CI: 0.86 – 0.94; $p < 2.60 \times 10^{-6}$) were decreased. Changes in
341 bicarbonate levels had the strongest association, showing elevation in FECD cases
342 (OR=1.15; CI: 1.10 – 1.20; $p < 7.31 \times 10^{-10}$), but the cause of this pH increase in blood and

343 relationship to FECD is unknown. Hemoglobin A1C and serum glucose were also elevated in
344 FECD cases of AFR ancestry, but these changes did not reach statistical significance due to
345 the smaller sample size. Effect estimates were consistent with results from the EUR ancestry
346 analysis.

347 In summary, the most important demographic risk factors for FECD are female sex
348 and European ancestry. Further, we note that FECD risk increases with increasing multi-
349 morbidity and multiple systemic conditions. The associations with phecodes related to
350 endocrine/metabolic conditions including DM are supported by the significant laboratory
351 values that are associated with FECD.

352

353 **Discussion**

354 PheWAS has traditionally been used to identify pleiotropy, where a genetic trait is associated
355 with multiple conditions.²⁰ In this study, we applied the PheWAS approach to the MVP
356 biobank to comprehensively investigate whether FECD may be associated with other
357 conditions. We noted that Veterans with a CCI score of 0 (zero comorbidities) were less
358 likely to have a diagnosis of FECD as compared to those with at least one comorbidity. To
359 explore this observation, we examined the association between FECD status and the entire
360 library of phecodes. In comparison to controls, we noted that FECD cases had higher rates of
361 many conditions, including endocrine/metabolic, infectious disease, neurological, and
362 musculoskeletal phegroups and indicating that as a group FECD cases have many more
363 healthcare challenges than controls. Among these, a primary association converged on DM,
364 where we noted that FECD cases had elevated rates of DM-related phecodes and laboratory
365 profiles consistent with DM.

366 What underlies the many associations noted between FECD and other conditions? We
367 favor a model in which these other conditions render the cornea more vulnerable to
368 developing the FECD phenotype. FECD genetic risk factors are incompletely penetrant, so
369 an additional factor is thought to be needed to induce the disease phenotype. A number of
370 such factors that contribute to the FECD phenotype have been identified, including estrogen,
371 oxidative stress, and exposure to UV light.^{4,11} Our results indicate that DM may be a
372 particularly important factor that contributes to FECD etiology. This concept is consistent with
373 a broad literature. For example, DM is known to alter corneal physiology³⁷⁻⁴⁰ leading to
374 complications that increase keratopathy. These include abnormal tear secretion perhaps

375 secondary to reduced goblet cell density, decreased corneal sensitivity, delayed corneal re-
376 epithelialization, and altered corneal morphology, function and biomechanics.^{37,41-46}
377 Consistent with these observations, the Cornea Preservation Time Study documented
378 elevated risk for graft failure for eyes that received a cornea transplant from donors with
379 DM.⁴⁷ The mechanism underlying the association between DM and FECD may involve TCF4.
380 Expansion of a common trinucleotide repeat within the *TCF4* locus is strongly linked to
381 FECD,⁴⁸ and *TCF4* had the strongest signal in a large-scale genome wide analysis of
382 FECD.¹⁰ TCF4 (also referred to as E2-2) is involved in the regulation of insulin genes⁴⁹ and is
383 required to maintain the fate of mature plasmacytoid dendritic cells,⁵⁰ a cell type which has
384 been implicated in Type 1 DM.⁵¹

385 We also noted that the link between FECD and DM was stronger for the earlier onset
386 Type 1 than the Type 2 form, consistent with the relationship between DM and microvascular
387 complications of the retina, kidney, peripheral nervous system and other systems. It will be
388 interesting to determine whether the presence of DM results in a more severe FECD
389 phenotype in patients with comparable risk factors, and whether DM also impacts other
390 corneal dystrophies. It will also be interesting to determine, for example, if FECD is reduced
391 in prevalence or severity in patients with well-managed DM and whether the beneficial effects
392 of topical insulin observed in animal models³⁸ will alleviate FECD severity in patients with DM.

393 Diabetes mellitus may also be central to many of the associations between FECD and
394 other phecodes. For example, anemia can be induced by the shorter life span and greater
395 fragility of red blood cells in DM⁵²⁻⁵⁶ In addition, DM has been linked to multiple dermatologic
396 conditions, as well as enthesopathy, osteoarthritis and spinal disk disorders.⁵⁷⁻⁵⁹ In
397 comparison, while DM increases the risk of gout, this may reflect other comorbid conditions
398 and not a direct effect of DM.⁶⁰

399 We found that FECD had higher rates of other ocular conditions including cataracts
400 and dry eye. Both are more common in DM,³⁷ suggestive of a central role for DM in the
401 etiology of FECD. However, we cannot rule out the possibility that the elevated rate of
402 diagnostic codes for these and potentially other ocular conditions may simply reflect these
403 patients undergoing more intensive ophthalmic care. Further, the association with cataract
404 may reflect that cataract surgery alone can induce a phenotype involving damage to the
405 corneal endothelium and resembling FECD.⁶¹⁻⁶⁵ The association with dry eye may also reflect
406 that FECD symptoms related to corneal pain and/or sensitivity inspire clinicians to add codes

407 related to dry eye to the medical record. Prior studies have reported that patients with FECD
408 are more likely to have glaucoma.⁶⁶

409 We cannot exclude the possibility that pleiotropy underlies some of the FECD
410 associations that we observed. A more complete understanding of the genetic factors that
411 contribute to FECD and conditions associated with FECD will be needed to understand the
412 relative roles played by pleiotropy versus the physiological impacts of diabetes and potentially
413 other disorders on the cornea in FECD etiology.

414 FECD has been reported to be more common in women than in men.^{4,24,25} While
415 >90% of U.S. Veterans are male, 1.80% of EUR males (=2,057/111,935) and 7.72%
416 (=207/2,473) of EUR females were classified as cases, consistent with excess risk to
417 females. Using our case-control algorithm, we observed a 1.78% prevalence of FECD in the
418 MVP (2,663/149,322). This is somewhat lower than prior reports that FECD prevalence is
419 approximately 4%.^{5,64} A lower estimate of FECD prevalence could reflect that MVP enrollees
420 are >90% male³⁵ and that males have a lower occurrence of FECD. A lower estimate of
421 FECD prevalence could also reflect that our FECD algorithm (a) eliminated individuals with a
422 single positive code and (b) included many exclusion codes that could potentially co-exist
423 with FECD. Our estimate of FECD prevalence is, however, much higher than an analysis of
424 data from a national managed-care network which found that the prevalence of all corneal
425 dystrophies was less than 0.1%.⁶⁸ This study did not have an age requirement and therefore
426 included data from young patients not yet at risk for FECD. Additional studies of FECD
427 prevalence in large healthcare databases should help to define a precise estimate of FECD
428 prevalence.

429 We further found that FECD frequency was higher in EUR Veterans than in other
430 ancestries, after adjustment for age and sex. This finding agrees with that of Eghrari et al.⁶⁹
431 who reported a modest reduction in the incidence of FECD in NHB as compared to NHW
432 patients. In comparison, Mahr and colleagues²⁵ did not observe a difference in FECD
433 incidence between NHW and NHB patients based on their analysis of U.S Medicare claims.
434 Any comparative increase in FECD incidence in NHW may reflect previously reported
435 average differences in corneal thickness across populations.⁷⁰

436

437 **Limitations and Strengths**

438 As MVP does not provide imaging data, this study was limited to clinical coding, billing, and
439 prescription data. We addressed this limitation by developing an algorithm to accurately
440 identify FECD cases and controls within CPRS. Through a local chart review, involving three
441 VAMC Eye Clinics, we were able to define a final algorithm that had high PPV (probability of
442 a case identified by the algorithm having FECD) and high NPV (probability of a control
443 identified by the algorithm not having FECD). Our algorithm thus provides the opportunity to
444 pursue FECD genetic studies within MVP and additional largescale biobanks.

445 In contrast to our algorithm for FECD, we did not use a similar approach to validate the
446 many phecodes that were examined in the PheWAS. These phecodes could suffer from
447 coding and diagnostic errors, although we would expect these to be random such errors will
448 reduce power by introducing noise.

449 This study was limited to the population served by the Veterans Health Administration.
450 While it is unclear that the Veteran and general population are markedly different, it will be
451 valuable to examine these associations in additional patient cohorts, including those including
452 a larger fraction of women enrollees.

453 The statistical model used to conduct our PheWAS and our analysis of laboratory
454 values adjusted for obvious confounders (age, sex, genetic ancestry). We cannot rule out
455 however, that some hidden residual confounding may remain.

456

457 **Conclusions**

458 By developing a rules-based algorithm to identify FECD and applying this to a large EHR-
459 based biobank, we provide new insights into the etiology of FECD. In the future it will be
460 important to better understand the relationship between FECD and DM. Insights regarding
461 this relationship may identify opportunities for slowing FECD progression. We anticipate that
462 our case-control algorithm will open the door for further FECD gene discovery.

463

464 **Acknowledgments**

465 We are grateful to the VINCI and GENISIS support teams, and to the MVP Core Statistical
466 Analysis team. This publication is solely the responsibility of the authors and does not

467 necessarily represent the official views of the NIH. This publication does not represent the
468 views of the Department of Veterans Affairs or the United States Government.

469

470 **References**

- 471 1. Vedana G, Villarreal Jr G, Jun AS. Fuchs endothelial corneal dystrophy: current
472 perspectives. *Clin Ophthalmol*. 2016;10:321-330.
- 473 2. Blitzer AL, Colby KA. Update on the surgical management of Fuchs Endothelial
474 Corneal Dystrophy. *Ophthalmol Ther*. 2020;9(4):757-765.
- 475 3. Eye Bank Association of America. Eye Banking Statistical Report. Eye Bank
476 Association of America, Washington, D.C. 2021.
- 477 4. Ong Tone S, Kocaba V, Bohm M, et al. Fuchs endothelial corneal dystrophy: the
478 vicious cycle of Fuchs pathogenesis. *Prog Ret Eye Res*. 2021;80:100863.
- 479 5. Rosenblum P, Stark WJ, Maumenee IH, Hirst LW, Maumenee AE. Hereditary Fuchs'
480 dystrophy. *Am J Ophthalmol*. 1980;90(4):455-462.
- 481 6. Nanda GG, Alone DP. Review: current understanding of the pathogenesis of Fuchs
482 endothelial corneal dystrophy. *Molec Vis*. 2019;25:295-310.
- 483 7. Biswas S, Munier FL, Yardley J, et al. Missense mutations in *COL8A2*, the gene
484 encoding the alpha2 chain of type VIII collagen, cause two forms of corneal endothelial
485 dystrophy. *Hum Mol Genet*. 2001;10(21):2415-2423.
- 486 8. Gottsch JD, Sundin OH, Liu SH, et al. Inheritance of a novel *COL8A2* mutation defines
487 a distinct early-onset subtype of Fuchs corneal dystrophy. *Invest Ophthalmol Vis Sci*.
488 2005;46(6):1934-1939.
- 489 9. Liu X, Zheng T, Zhao C, Zhang Y, Liu H, Wang L. Genetic mutations and molecular
490 mechanisms of Fuchs endothelial corneal dystrophy. *Eye Vis*. 2021;8(1):24.
- 491 10. Afshari NA, Igo RP Jr, Morris NJ, et al. Genome-wide association study identifies three
492 novel loci in Fuchs endothelial corneal dystrophy. *Nat Commun*. 2017;8:14898.
- 493 11. Liu C, Miyajima T, Melangath G, et al. Ultraviolet A light induces DNA damage and
494 estrogen-DNA adducts in Fuchs endothelial corneal dystrophy causing females to be
495 more affected. *Proc Natl Acad Sci USA*. 2020;117(1):573-583.
- 496 12. Xu TT, Baratz KH, Fautsch MP, Hodge DO, Mahr MA. Cancer risk in patients with
497 Fuchs endothelial corneal dystrophy. *Cornea*. 2022;41(9):1088-1093.

- 498 13. Chang Y-S, Ho C-H, Wang J-J, Tseng S-H, Jan R-L. The sociodemographic and risk
499 factors for Fuchs endothelial dystrophy: A nationwide, matched case-control study in
500 Taiwan. *J Pers Med.* 2022;12(2):305.
- 501 14. Gaziano JM, Concato J, Brophy M, et al. Million Veteran Program: a mega-biobank to
502 study genetic influences on health and disease. *J Clin Epidemiol.* 2016;70:214-223.
- 503 15. Nguyen X-MT, Quaden RM, Song RJ, et al. Baseline characterization and annual
504 trends of body mass index for a mega-biobank cohort of US Veterans 2011-2017. *J*
505 *Health Res Rev Dev Ctries.* 2018;5(2):98-107.
- 506 16. Pendergrass SA, Crawford DC. Using electronic health records to generate
507 phenotypes of research. *Curr Protoc Hum Genet.* 2019;100(1):e80.
- 508 17. Halladay CW, Hadi T, Anger MD, et al. Genetically guided algorithm development and
509 sample size optimization for age-related macular degeneration cases and controls in
510 electronic health records from the VA Million Veteran Program. *AMIA Jt Summits*
511 *Transl Sci Proc. AMIA Jt Summits Transl Sci.* 2019;2019:153-162.
- 512 18. Nealon CL, Halladay CW, Kinzy TG, et al. Development and evaluation of a rules-
513 based algorithm for primary open-angle glaucoma in the VA Million Veteran Program.
514 *Ophthalmic Epidemiol.* 2022;29(6):640-648.
- 515 19. Fihn SD, Francis J, Clancy C, et al. Insights from advanced analytics at the Veterans
516 Health Admin. *Health Aff.* 2014;33(7):1203-1211.
- 517 20. Denny JC, Bastarache L, Roden DM. Phenome-wide association studies as a tool to
518 advance precision medicine. *Annu Rev Genomics Hum Genet.* 2016;17:353-373.
- 519 21. Zhang T, Goodman M, Zhu F, et al. Phenome-wide examination of comorbidity burden
520 and multiple sclerosis disease severity. *Neurol Neuroimmunol Neuroinflamm.*
521 2021;7(6):e864.
- 522 22. Song RJ, Ho Y-L, Schubert P, et al. Phenome-wide association of 1809 phenotypes
523 and COVID-19 disease progression in the Veterans Health Administration Million
524 Veteran Program. *PLOS One* 2021;16(5):e0251651.
- 525 23. Yeh P, Colby K. Corneal endothelial dystrophies. In *The Cornea*, Foster CAD,
526 Dohlman C (Eds). (Lippincott Williams & Wilkins, Philadelphia, 2004).
- 527 24. Wilson SE, Bourne WM. Fuchs dystrophy. *Cornea.* 1988;7(1):2-18.

- 528 25. Mahr MA, Baratz KH, Hodge DO, Erie JC. Racial/ethnic differences in rates of
529 penetrating or endothelial keratoplasty for Fuchs endothelial corneal dystrophy among
530 US Medicare beneficiaries, 2014. *JAMA Ophthalmol.* 2016;134(10):1178-1180.
- 531 26. Eghrari AO, Vahedi S, Afshari NA, Riazuddin SA, Gottsch JD. CTG18.1 expansion in
532 TCF4 among African Americans with Fuchs corneal dystrophy. *Invest Ophthalmol Vis*
533 *Sci.* 2017;58(14):6046-6049.
- 534 27. Hunter-Zinck H, Shi Y, Li M, et al. Genotyping array design and data quality control in
535 the Million Veteran Program. *Amer J Hum Genet.* 2020;106(4):535-548.
- 536 28. Fang H, Hui Q, Lynch J, et al. Harmonizing genetic ancestry and self-identified
537 race/ethnicity in genome-wide association studies. *Amer J Hum Genet.*
538 2019;105(4):763-772.
- 539 29. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying
540 prognostic comorbidity in longitudinal studies: development and validation. *J Chronic*
541 *Dis.* 1987;40(5):373-383.
- 542 30. Akinyemiju T, Jha M, Moore JX, Pisu M. Disparities in the prevalence of comorbidities
543 among US adults by state Medicaid expansion status. *Prev Med.* 2016;88:196-202.
- 544 31. Akinyemiju T, Moore JX. Data on burden of comorbidities in the United States and
545 Medicaid expansion status. *Data Brief.* 2016;8:120-122.
- 546 32. Wei WQ, Bastarache L, Carroll RJ, et al. Evaluating phecodes, clinical classification
547 software, and ICD-9-CM codes for phenome-wide association studies in the electronic
548 health record. *PLOS One.* 2017;12(7):0175508.
- 549 33. Wu P Gifford A, Meng X, et al. Mapping ICD-10 and ICD-10-CM codes to phecodes:
550 Workflow development and initial evaluation. *JMIR Med Inform.* 2019;7(4):e14325.
- 551 34. Carroll RJ, Bastarache L, Denny JC. R PheWAS: data analysis and plotting tools for
552 phenome-wide association studies in the R environment. *Bioinformatics.*
553 2014;30(16):1375-1376.
- 554 35. Harrington KM, Nguyen X-MT, Song RJ, et al. Gender differences in demographic and
555 health characteristics of the Million Veteran Program cohort. *Womens Health Issues*
556 2019;29 Suppl 1(Suppl 1):S56-S66.
- 557 36. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson comorbidity index: A
558 critical review of clinimetric properties. *Psychother Psychosom.* 2022;91(1):8-35.

- 559 37. Skarbez K, Priestley Y, Hoepf M, Koevary SB. Comprehensive review of the effects of
560 diabetes on ocular health. *Expert Rev Ophthalmol*. 2010;5(4):557-577.
- 561 38. Ljubimov AJ. Diabetic complications in the cornea. *Vision Res*. 2017;139:138-152.
- 562 39. Kaltenece A, Ferdousi M, Azmi A, Marshall A, Soran H, Malik RA. Keratocyte density
563 is reduced and related to corneal nerve damage in diabetic neuropathy. *Invest*
564 *Ophthalmol Vis Sci*. 2018;59(8):3584-3590.
- 565 40. Shah R, Amador C, Tormanen K, et al. Systemic diseases and the cornea. *Exp Eye*
566 *Res*. 2021;204:108455.
- 567 41. Lass JH, Spurney RV, Dutt RM, et al. A morphologic and fluorophotometric analysis of
568 the corneal endothelium in type I diabetes mellitus and cystic fibrosis. *Am J*
569 *Ophthalmol*. 1985;100(6):783-788.
- 570 42. Weston BC, Bourne WM, Polse KA, Hodge DO. Corneal hydration control in diabetes
571 mellitus. *Invest Ophthalmol Vis Sci*. 1995;36(3):586-595.
- 572 43. Larsson LI, Bourne WM, Pach JM, Brubaker RF. Structure and function of the corneal
573 endothelium in diabetes mellitus type I and type II. *Arch Ophthalmol*. 1996;114(1):9-
574 14.
- 575 44. Ziadi M, Moiroux P, d'Athis P, et al. Assessment of induced corneal hypoxia in diabetic
576 patients. *Cornea*. 2002;21(5):453-457.
- 577 45. Liaboe CA, Aldrich BT, Carter PC, et al. Assessing the impact of diabetes mellitus on
578 donor corneal endothelial cell density. *Cornea*. 2017;36(5):561e566.
- 579 46. Ramm L, Herber R, Spoerl E, Pillunat LE, Terai N. Factors influencing corneal
580 biomechanics in diabetes mellitus. *Cornea*. 2020;39(5):552-557.
- 581 47. Terry MA, Aldave AJ, Szczotka-Flynn LB, et al. Donor, recipient, and operative factors
582 associated with graft success in the Cornea Preservation Time Study. *Ophthalmology*.
583 2018;125(11):1700-1709.
- 584 48. Wieben ED, Aleff RA, Tosakulwong N, Butz ML, Highsmith WE, Edwards AO, Baratz
585 KH. A common trinucleotide repeat expansion within the transcription factor 4 (TCF4,
586 E2-2) gene predicts Fuchs corneal dystrophy. *PLoS One*. 2012;7(11):e49083.
- 587 49. Vierra CA, Nelson C. The pan basic helix-loop-helix proteins are required for insulin
588 gene expression. *Molec Endocrinol*. 1995;9(1):64-71.

- 589 50. Ghosh HS, Cisse B, Bunin A, Lewis KL, Reizis B. Continuous expression of the
590 transcription factor E2-2 maintains the cell fate of mature plasmacytoid dendritic cells.
591 *Immunity*. 2010;33(6):905-916.
- 592 51. Ye Y, Gaugler B, Mohty M, Malard F. Plasmacytoid dendritic cell biology and its role in
593 immune-related diseases. *Clin Trans Immunol*. 2020;9(5):e1139.
- 594 52. Craig KJ, Williams JD, Riley SG, Smith H, Owens DR, Worthing D, Cavill I, et al.
595 Anemia and diabetes in the absence of nephropathy. *Diabetes Care*. 2005;28(5):1118-
596 1123.
- 597 53. El-Achkar TM, Ohmit SE, McCullough PA, Crook ED, Brown WW, Grimm R, Bakris
598 GL, et al. Higher prevalence of anemia with diabetes mellitus in moderate kidney
599 insufficiency: the kidney early evaluation program. *Kidney Int*. 2005;67(4):1483-1488.
- 600 54. Ford ES, Cowie CC, Li C, Handelsman Y, Bloomgarden ZT. Iron-deficiency anemia,
601 non-iron-deficiency anemia and HbA1c among adults in the US. *J Diabetes*.
602 2011;3(1):67-73.
- 603 55. Lippi G, Mercadanti M, Aloe R, Targher G. Erythrocyte mechanical fragility is
604 increased in patients with type 2 diabetes. *Eur J Intern Med*. 2012;23(2):150-153.
- 605 56. Knychala MA, da Silva Garrote-Filho M, da Silva BB, de Oliveira SN, Luz SY,
606 Rodrigues MOM, Penha-Silva N. Red cell distribution width and erythrocyte osmotic
607 stability in type 2 diabetes mellitus. *J Cell Mol Med*. 2021;25(5):2505-2516.
- 608 57. Piva SR, Susko AM, Khoja SS, Josbeno DA, Fitzgeralds GK, Toledo FGS. Links
609 between osteoarthritis and diabetes: Implications for management from a physical
610 activity perspective. *Clin Geriatr Med*. 2015;31(1):67-87.
- 611 58. Ursini F, Arturi F, Nicolosi K, et al. Plantar fascia enthesopathy is highly prevalent in
612 diabetic patients without peripheral neuropathy and correlates with retinopathy and
613 impaired kidney function. *PLoS One*. 2017;12(3):e0174529.
- 614 59. Park C-H, Min K-B, Min J-Y, Kim DH, Seo KM, Kim D-K. Strong association of type 2
615 diabetes with degenerative lumbar spine disorders. *Scientific Rep*. 2021;11(21):16472.
- 616 60. Wijnands JMA, van Durme DM, Ghislaine P, Driessen JHM, Boonen A, Klop C,
617 Leufkens B, Cooper C, Stehouwer CDA, de Vries F. Individuals with Type 2 diabetes
618 mellitus are at an increased risk of gout but this is not due to diabetes. A population-
619 based cohort study. *Medicine*. 2015;94(32):e1358.

- 620 61. Walkow T, Anders N, Klebe S. Endothelial cell loss after phacoemulsification: relation
621 to preoperative and intraoperative parameters. *J Cataract Refract Surg*.
622 2000;26(5):727–732.
- 623 62. Seitzman GD, Gottsch JD, Stark WJ. Cataract surgery in patients with Fuchs' corneal
624 dystrophy: expanding recommendations for cataract surgery without simultaneous
625 keratoplasty. *Ophthalmology*. 2005;112(3):441-446.
- 626 63. Güell JL, El Husseiny MA, Manero F, Gris O, Elies D. Historical review and update of
627 surgical treatment for corneal endothelial diseases. *Ophthalmol Ther*. 2014;3(1-2):1-
628 15.
- 629 64. van Cleynenbreugel H, Remeijer L, Hillenaar T. Cataract surgery in patients with
630 Fuchs' endothelial corneal dystrophy: when to consider a triple procedure.
631 *Ophthalmology*. 2014;121(2):445-453.
- 632 65. Yong WWD, Chai HC, Shen L, Manotosh R, Anna Tan WT. Comparing outcomes of
633 phacoemulsification with femtosecond laser-assisted cataract surgery in patients with
634 Fuchs endothelial dystrophy. *Am J Ophthalmol*. 2018;196:173-180.
- 635 66. Nagarsheth M, Singh A, Schmotzer B, et al. Relationship between Fuchs endothelial
636 corneal dystrophy severity and glaucoma and/or ocular hypertension. *Arch*
637 *Ophthalmol*. 2012;130(11):1384-1388.
- 638 67. Elhali H, Azizi B, Jurkunas UV. Fuchs endothelial corneal dystrophy. *The Ocular*
639 *Surface* 2010 8(4):173-184.
- 640 68. Musch DC, Niziol LM, Stein JD, Kamyar RM, Sugar A. Prevalence of corneal
641 dystrophies in the United States: estimates from claims data. *Invest Ophthalmol Vis*
642 *Sci*. 2011;52(9):6959-6963.
- 643 69. Eghrari AO, Vahedi S, Afshari NA, Riazuddin SA, Gottsch JD. CTG18.1 Expansion in
644 TCF4 among African Americans with Fuchs' Corneal Dystrophy. *Invest Ophthalmol Vis*
645 *Sci*. 2017;58(14):6046-6049.
- 646 70. McGlumphy EJ, Margo JA, Haidara M, Brown CH, Hoover CK, Munir WM. Predictive
647 value of corneal donor demographics on endothelial cell density. *Cornea*.
648 2018;37(9):1159-1162.
- 649

650 **Figure Legends**

651 **Figure 1:** Overview of the algorithm refinement process and application to MVP. FECD
652 case-control algorithms were evaluated and revised at the Cleveland VAMC Eye Clinic. Upon
653 reaching >85% PPV and >95% NPV, algorithm performance was confirmed at Buffalo and
654 Providence VAMC Eye Clinics, and then applied to the MVP.

655
656 **Figure 2:** Age distribution of FECD cases (N=2,663) (a) and controls (N=138,944) (b). (c)
657 Box-plot representation.

658
659 **Figure 3:** Phenome-wide association scan of EHR-derived phenotypes (main panel) and
660 normalized laboratory values (inset) with FECD. Upwards-pointing arrows indicate positive
661 association ($OR > 1$) and downwards-pointing arrows indicate negative association ($OR < 1$).
662 Red line indicates Bonferroni-corrected significance threshold. Phenotypes highly similar to
663 FECD are not plotted. Labs shown are those significant after multiple hypothesis testing
664 correction.

665

666 **Table 1:** Summary of Cases and Controls reviewed in final FECD algorithm.

		Algorithm Identified			
		FECD cases (female)	Controls (female)	PPV	NPV
Non-Hispanic Black	Cleveland	25 (4)	25 (0)	96%	96%
	Providence	1 (0)	25 (0)	100%	100%
	Buffalo	17 (3)	25 (1)	88%	100%
	Combined	43 (7)	75 (1)	93%	98%
Non-Hispanic White	Cleveland	25 (2)	25 (0)	80%	92%
	Providence	25 (0)	24 (0)	76%	100%
	Buffalo	25 (2)	25 (0)	72%	96%
	Combined	75 (4)	74 (0)	76%	96%
Total		118 (11)	149 (1)	82%	97%

667 FECD: Fuchs endothelial corneal dystrophy

668 PPV: positive predictive value

669 NPV: negative predictive value

670

671

672 **Table 2:** Genotyped MVP participants who meet the criteria for case with FECD or control.

Genetic ancestry	Cases all / male / female	Controls all / male / female	% FECD all / male / female
EUR	2,264 / 2,057 / 207	114,408 / 111,935 / 2,473	1.94 / 1.80 / 7.72
AFR	315 / 225 / 90	16,367 / 15,907 / 460	1.89 / 1.39 / 16.36
HIS	80 / 75 / 5	7,105 / 6,962 / 143	1.11 / 1.07 / 3.38
ASN	4 / 4 / 0	1,064 / 1,035 / 29	0.37 / 0.38 / not done
Total	2,663 / 2,361 / 302	138,944 / 135,839 / 3,105	1.88 / 1.71 / 8.86

673 MVP: Million Veteran Program

674 FECD: Fuchs endothelial corneal dystrophy

675 EUR: European ancestry

676 AFR: African ancestry

677 ASN: Asian ancestry

678 HIS: admixed American ancestry

679

680

681 **Table 3:** Multivariable logistic regression examining various factors influencing risk to FECD
682 (N = 141,607).

Variable	OR	95% CI	P value
Age	0.92	0.92, 0.93	< 0.001
Sex	4.60	4.03, 5.22	< 0.001
Ancestry – AFR vs EUR	0.81	0.71, 0.91	< 0.001
Ancestry – HIS vs EUR	0.53	0.42, 0.66	< 0.001
Ancestry – ASN vs EUR	0.18	0.06, 0.42	< 0.001
CCI Cat 1-3	1.21	1.06, 1.40	0.005
CCI Cat 4+	1.40	1.11, 1.75	0.003

683 FECD: Fuchs endothelial corneal dystrophy

684 AFR: admixed African ancestry

685 EUR: European ancestry

686 HIS: admixed American ancestry (Hispanic/Latino)

687 ASN: Asian ancestry

688 CCI: Charlson Comorbidity Index

689 OR: odds ratio

690 CI: confidence interval

691

692 **Table 4:** Summary of PheWAS organized by phecode groups. For each phecode group, the
 693 total number of phecodes that included at least 100 cases and those that achieved statistical
 694 significance after Bonferroni correction are noted.

Phecode group	Total phecodes	Significant phecodes	% significant
circulatory system	17	156	10.90
congenital anomalies	1	30	3.33
dermatologic	19	73	26.03
digestive	15	146	10.27
endocrine/metabolic	21	122	17.21
genitourinary	11	106	10.38
hematopoietic	3	47	6.38
infectious diseases	8	50	16.00
injuries & poisonings	8	91	8.79
mental disorders	8	67	11.94
musculoskeletal	16	109	14.68
neoplasms	2	113	1.77
neurological	11	73	15.07
pregnancy complications	0	1	0.00
respiratory	12	77	15.58
sense organs	51	118	43.22
symptoms	10	38	26.32
TOTAL	213	305	15.03

695

Figure 1

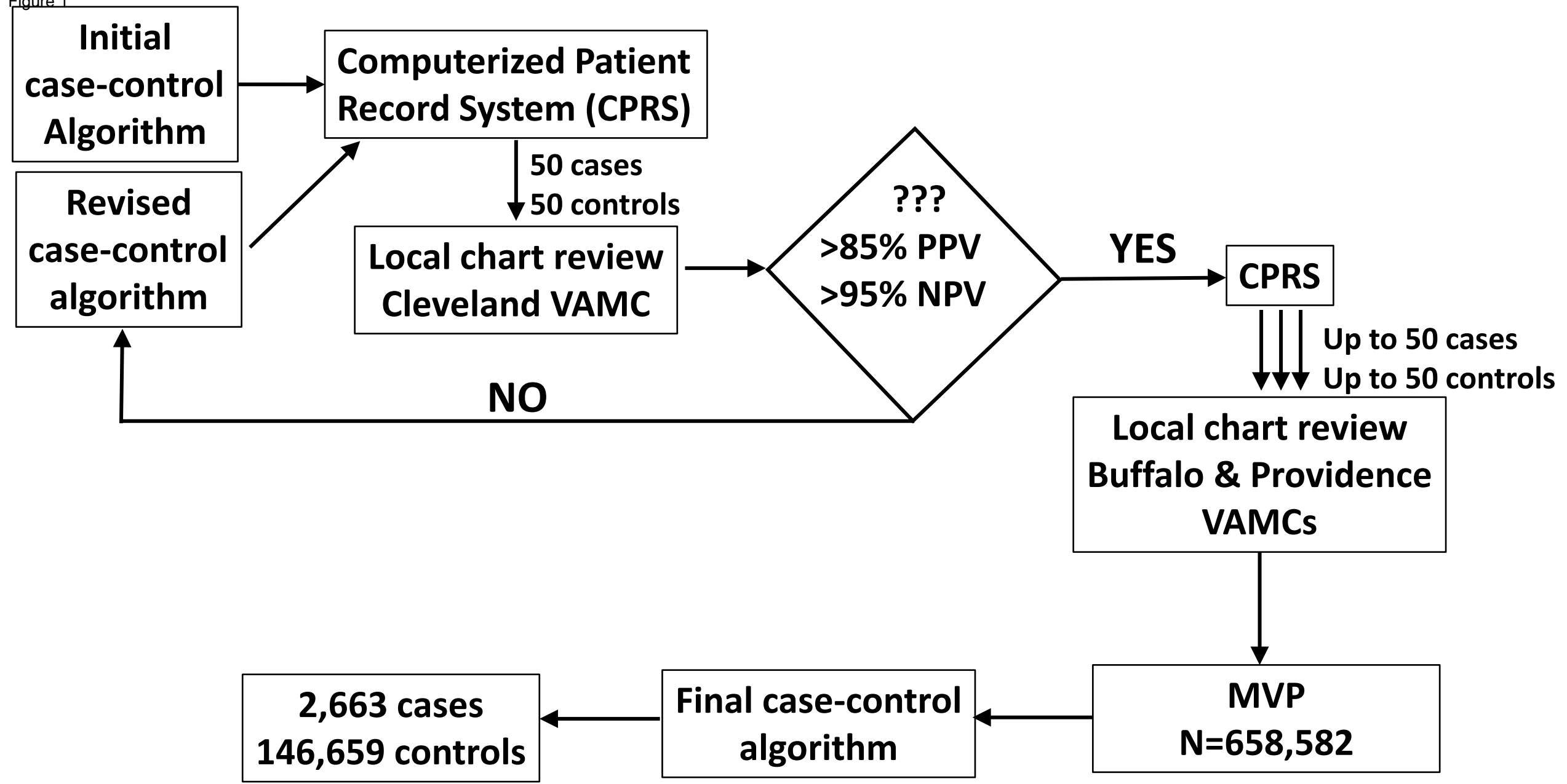


Figure 1: Overview of the algorithm refinement process and application to MVP. FECD case-control algorithms were evaluated and revised at the Cleveland VAMC. Upon reaching >85% PPV and >95% NPV, algorithm performance was confirmed at Buffalo and Providence VAMCs, and then applied to the MVP.

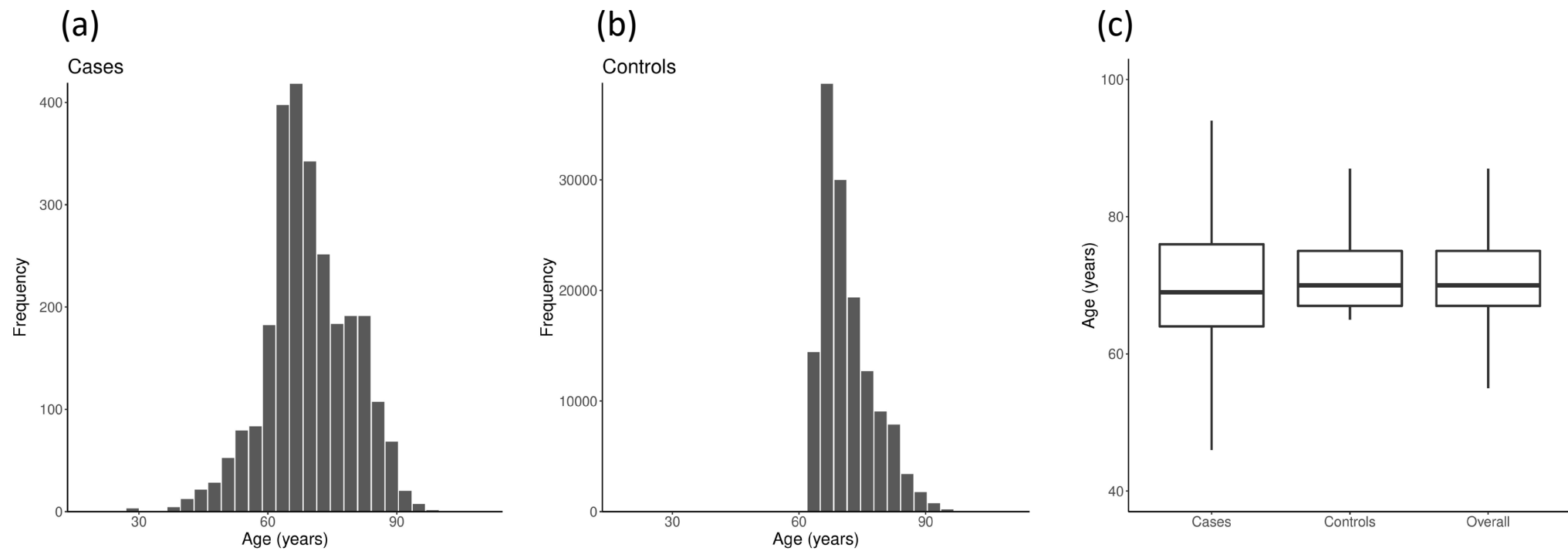


Figure 2: Age distribution of FECD cases (N=2,663) (a) and controls (N=138,944) (b). (c) Box-plot representation.

Figure 3

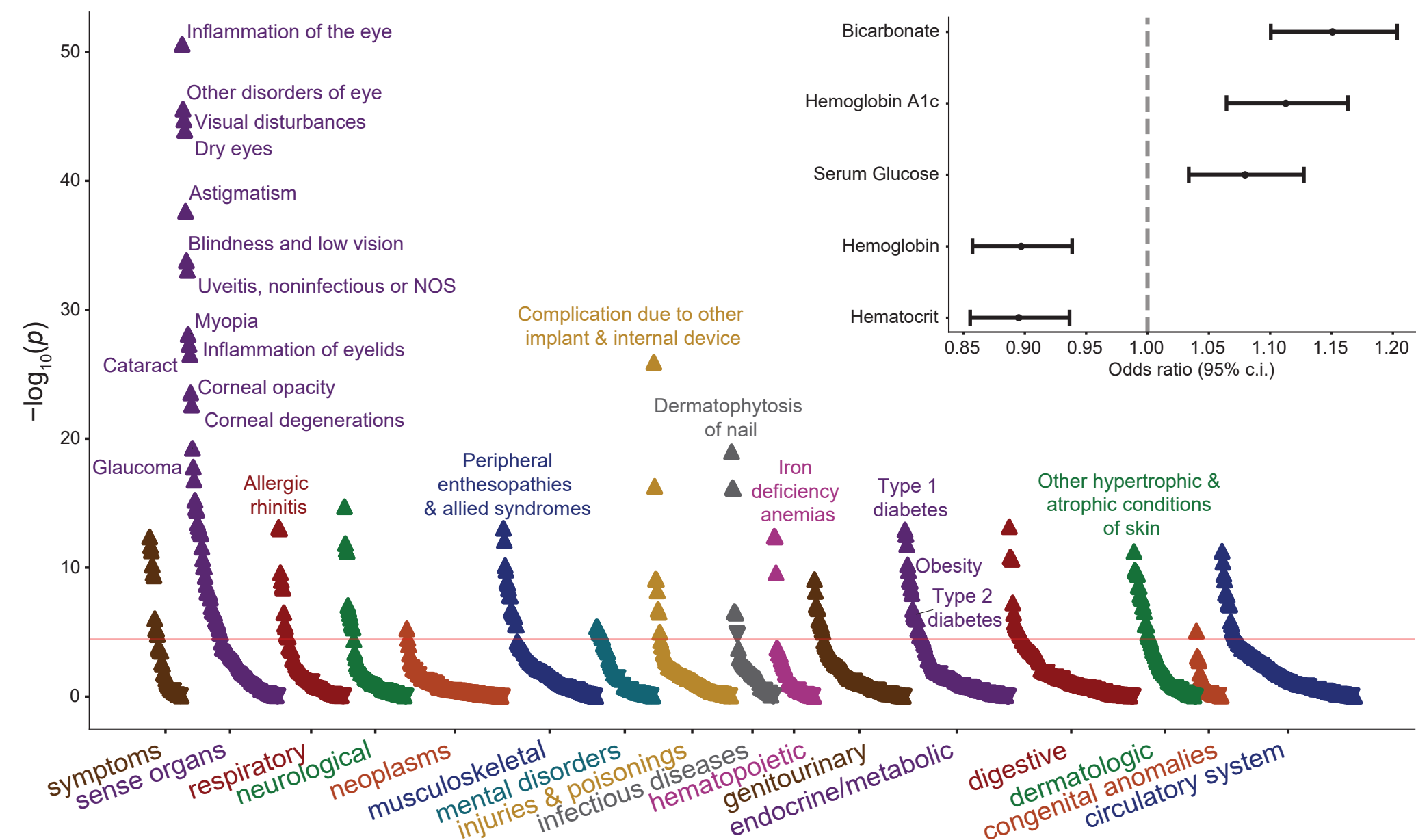


Figure 3. Phenome-wide association scan of EHR-derived phenotypes (main panel) and normalized laboratory values (inset) with FECD. Upwards-pointing arrows indicate positive association ($OR > 1$) and downwards-pointing arrows indicate negative association ($OR < 1$). Red line indicates Bonferroni-corrected significance threshold. Phenotypes highly similar to FECD are not plotted. Labs shown are those significant after multiple hypothesis testing correction.