

1 Title

2 Mendelian randomisation analyses of eosinophils and other blood cell types in relation to lung
3 function and disease.

4 Authors

5 Anna L Guyatt,^{1*} Catherine John,¹ Alexander T Williams,¹ Nick Shrine,¹ Nicola Reeve,¹ SpiroMeta
6 consortium, Ian P Hall,² Louise V Wain,^{1,3} Nuala A Sheehan,¹ Frank Dudbridge,¹ Martin D Tobin^{*1,3}

7 Affiliations

- 8 1. Department of Health Sciences, University of Leicester, Leicester, UK
- 9 2. Division of Respiratory Medicine and NIHR-Nottingham Biomedical Research Centre,
10 University of Nottingham, Nottingham, UK
- 11 3. National Institute for Health Research, Leicester Respiratory Biomedical Research Centre,
12 Glenfield Hospital, Leicester, UK

14 Corresponding author statement

15 * = Corresponding Authors.

16 Corresponding author emails: anna.guyatt@leicester.ac.uk (Anna L Guyatt) mt47@le.ac.uk (Martin D
17 Tobin)

18 Abstract

19 **Background:** Eosinophils are granulocytes associated with airway inflammation in respiratory
20 disease. Eosinophil production and survival is controlled by interleukin-5: anti-interleukin-5 agents
21 reduce asthma and COPD exacerbation frequency, and response correlates with baseline eosinophil
22 counts. However, causal relationships between eosinophils and other respiratory phenotypes are
23 less studied.

24 **Methods:** We investigated causality between eosinophils and: lung function, acute exacerbations of
25 COPD (AECOPD), asthma-COPD overlap (ACO), moderate-to-severe asthma, and respiratory
26 infections. We performed Mendelian randomization (MR) using 151 genetic variants from genome-
27 wide association studies of blood eosinophil counts in UK Biobank/INTERVAL, and respiratory data
28 from UK Biobank, using MR methods relying on different assumptions for validity. Multivariable MR
29 using eight blood cell type exposures was performed for outcomes showing evidence of causation by
30 eosinophils.

31 **Findings:** There was evidence that higher eosinophils reduce FEV₁/FVC and FEV₁ (weighted median
32 estimator, SD change FEV₁/FVC per SD eosinophils: -0.054 [95%CI -0.078,-0.029]). There was also
33 evidence that eosinophils cause ACO (weighted median OR 1.44 [95%CI 1.19,1.74]), and asthma
34 (weighted median OR 1.50 [95%CI 1.23,1.83]). Multivariable MR for FEV₁/FVC, FEV₁, ACO and asthma
35 suggested that eosinophils were the cell type with the most important effect. Causal estimates of
36 individual variants were heterogeneous, which may arise from pleiotropy.

37 **Interpretation:** We found evidence that eosinophils reduce lung function, and increase ACO and
38 asthma risk, on average over the set of genetic variants studied. Eosinophils appear to be causal
39 determinants of fixed airflow obstruction among individuals with features of both asthma and COPD.

40 **Funding:** Wellcome, BHF, MRC, BBSRC CASE studentship with GSK, GSK/BLF.

41 **Research in Context**

42 **Evidence before this study**

43 High eosinophils are observed in some individuals with asthma and COPD. Anti-IL5 and anti-IL5R
44 drugs, such as mepolizumab and benralizumab, respectively, are currently licensed for treatment of
45 severe eosinophilic asthma, and response correlates with baseline eosinophil levels. Clinical trials of
46 anti-IL5 agents in COPD have reported reductions in blood and sputum eosinophil counts, but not
47 reproducible reductions in COPD exacerbations, and any clinical improvement in COPD has been
48 hypothesised to be smaller than that in asthma, and similarly related to the degree of eosinophilic
49 inflammation.

50 **Added value of this study**

51 We performed Mendelian randomization (MR) analyses, using genetic variants that predict blood
52 cell counts, to investigate causation between eosinophils and several quantitative lung function
53 traits and clinical respiratory outcomes, encompassing both fixed and reversible airflow obstruction.
54 Where there was evidence of causation by eosinophils, we explored whether other blood cell types
55 may have contributed to this association, since blood cells counts are correlated. Overall, our aim
56 was to provide a comprehensive assessment of the causal role of blood eosinophil counts in
57 respiratory health and disease.

58 **Implications of all available evidence**

59 Although clinical trials of anti-IL5 agents have not directly tested the causal role of eosinophils in
60 asthma and COPD, they are hypothesised to act via a reduction in eosinophil counts. We found
61 evidence for causality of eosinophils on FEV₁ and FEV₁/FVC (used in the diagnosis of COPD), on
62 asthma-COPD overlap, and on asthma. We did not find evidence for eosinophils causing COPD
63 exacerbations, but cannot rule out a small effect. Whilst the average effect of lowering eosinophils is
64 to improve certain respiratory phenotypes we noted heterogeneity in our causal estimates from
65 individual genetic variants, suggesting pleiotropic effects of SNPs affecting eosinophil levels. Taken
66 together with prior evidence, our findings suggest that clinical trials of eosinophil lowering agents
67 are warranted in patients with history of both asthma and COPD.

68 Introduction

69 Eosinophils are proinflammatory granulocytes associated with symptom severity and exacerbation
70 frequency in patients with asthma and chronic obstructive pulmonary disease (COPD).²⁻⁴ The degree
71 of eosinophilia in these obstructive lung diseases is variable: whilst eosinophil inflammation brought
72 about by allergic sensitisation has been considered characteristic of asthma, not all patients with
73 asthma have eosinophilia.^{2,5} Moreover, whilst airway inflammation in COPD is typically mediated by
74 neutrophils, a subset of individuals with COPD have raised eosinophil counts.^{2,6}

75 The production and survival of eosinophils is regulated by interleukin-5 (IL-5), and anti-IL5 therapies
76 (e.g. mepolizumab, and the anti-IL5RA agent, benralizumab) are now licensed in many countries for
77 the treatment of severe eosinophilic asthma.⁷⁻¹³ The decision to treat asthma with these drugs is
78 currently based upon blood eosinophil count,² since post-hoc analyses of clinical trials stratified by
79 eosinophil levels have shown increased efficacy of mepolizumab for treating severe asthma in those
80 with higher baseline eosinophils.³ Results from Mendelian randomization (MR) analyses have also
81 provided evidence for a causal role of eosinophils in asthma.¹⁴ MR analyses use genetic variants as
82 instrumental variables (IVs) to investigate causality between exposure and outcome, and under
83 certain assumptions may obviate some of the problems of traditional observational epidemiology
84 (reverse causation and confounding), permitting causal inference.

85 In addition to asthma, blood eosinophil counts are associated with quantitative measures of lung
86 function in healthy populations (i.e. including individuals without asthma).¹⁵ However, causality of
87 these associations has yet to be established, with the only previous MR of lung function being of
88 small sample size, with imprecise effect estimates precluding confident inference.¹⁶ Moreover,
89 causal effects of eosinophils on other respiratory phenotypes, such as asthma-COPD overlap (ACO),
90 and respiratory infections are yet to be investigated. Diagnosis of COPD is made by spirometry if the
91 ratio of the forced expiratory volume in one second (FEV_1) to the forced vital capacity (FVC),
92 FEV_1/FVC , is <0.7 , with airflow obstruction being graded according to predicted values of FEV_1 .
93 Therefore, studying the role of eosinophils in determining quantitative lung function (i.e. lung
94 function measured as continuous traits) is a powerful way of understanding their role in the
95 development of fixed airflow obstruction such as is observed in COPD.^{17,18} Understanding the causal
96 relationship between eosinophils and fixed airflow obstruction is particularly pertinent given recent
97 interest in the potential use of mepolizumab in COPD.¹⁰⁻¹³

98 We undertook two-sample MR analyses using summary-level genome-wide association study
99 (GWAS) data to assess causality between eosinophils and respiratory traits and conditions
100 encompassing fixed and reversible airflow obstruction, using genetic variants associated with blood

101 eosinophil counts in the largest GWAS to date as IVs.¹⁴ First, we investigated a causal effect of
102 eosinophils on three quantitative lung function traits measured by spirometry, and four clinical
103 respiratory phenotypes (moderate-to-severe asthma, acute exacerbations of COPD [AECOPD], ACO,
104 and respiratory infections). We used three MR approaches that rely on different assumptions for
105 validity, and followed-up traits showing evidence of causality to assess whether the instrumental
106 variables affected lung function purely via eosinophil counts and not via other blood cell types.
107 Overall, our aim was to provide a comprehensive assessment of the causal role of blood eosinophil
108 counts in relation to respiratory health and disease.

109 **Methods**

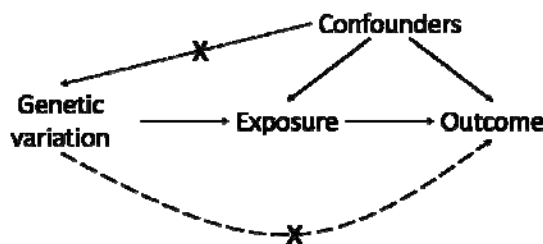
110 We assessed causal associations between eosinophils and other blood cell counts in relation to
111 multiple respiratory outcomes, using MR.^{1,19} Briefly, MR involves using genetic variants (here single-
112 nucleotide polymorphisms, SNPs), as instrumental variables (IVs) for an exposure of interest, in this
113 case eosinophil counts. This is done by comparing the magnitude of the effect of the SNPs on the
114 outcome to the effect of the SNPs on the exposure.^{1,19} All analyses reported here are two-sample

Box 1 Mendelian randomization: core assumptions

Assumptions

Mendelian randomization may be used to test for causality between an exposure (e.g. eosinophils) and outcome (e.g. FEV₁/FVC), if the following core assumptions hold:

- i) the genetic variation (single nucleotide polymorphisms in this work) used as instrumental variables (IVs) are associated with the exposure of interest;
- ii) the genetic variants are not associated with unobserved confounders of the exposure-outcome association (short dashed arrow). Genetic variants are allocated randomly at conception (Mendel's law of independent assortment) and so typically should not be associated with these confounding variables;
- iii) association between the genetic variants and the outcome is via the exposure, and not via an alternate pathway (i.e. there is no 'horizontal pleiotropy', see long dashed arrow). Whilst difficult to verify, reassurance that this assumption holds can be provided using biological knowledge of how the SNP functions, and by checking whether multiple MR methods, each relying on different assumptions for validity, give consistent results (known as triangulation).¹



115 MR analyses, since SNP-exposure and SNP-outcome associations were extracted from different (yet
116 overlapping²⁰) samples. The core assumptions of MR for inferring causality between an exposure and
117 outcome are described in **Box 1**. Additional assumptions required for accurate point estimation are
118 discussed in the **Extended Methods**, and elsewhere.²¹

119 All GWAS datasets analysed included participants from UK Biobank,²² incorporating other studies
120 where available. All GWAS included were of individuals of European ancestry. Datasets are
121 summarised here, and descriptions of covariate adjustments, and exposure-outcome GWAS overlap
122 are given in the **Extended Methods**.

123 **Exposure GWAS data sets (blood cell parameters)**

124 We used summary-level data from a previously published GWAS of blood cell counts¹⁴ in the initial
125 release of UK Biobank genetic data (around one third of all participants), plus the INTERVAL study
126 (eosinophil GWAS N=172,275, see **Extended Methods**).¹⁴ These included counts of eosinophils,
127 basophils, neutrophils, monocytes, lymphocytes, platelets, red blood cells and reticulocytes.

128 **Outcome GWAS data sets (respiratory outcomes)**

129 *Quantitative lung function GWASs*

130 We used published summary-level data from three GWAS of quantitative lung function (FEV₁, FVC
131 and FEV₁/FVC), undertaken in 400,102 individuals.¹⁸

132 *Clinical disease GWAS*

133 *Moderate-to-severe asthma*

134 We used a published GWAS of moderate-to-severe asthma comprising 5,135 cases and 25,675
135 controls, sampled from the initial release of UK Biobank, supplemented with external cases.²³

136 *Acute exacerbations of COPD*

137 We defined acute exacerbations of COPD (AECOPD) in UK Biobank; the eligible sample was restricted
138 to individuals with FEV₁/FVC < 0.7. From this subset, exacerbation cases (N=2,771) had an ICD-10 code
139 for AECOPD or a lower respiratory tract infection in the hospital episode statistics data provided by
140 UK Biobank (**Supplementary Table 1**). Controls (N=42,052) were people with COPD without a
141 recorded ICD-10 AECOPD code. Associations were adjusted for age (at recruitment), age², sex,
142 smoking status (ever/never), genotyping array and 10 ancestry principal components.

143 *Asthma-COPD overlap (ACO)*

144 We defined ACO cases in UK Biobank (N=8,068) as individuals with a self-report of a doctor diagnosis
145 of asthma, and both FEV₁/FVC < 0.7 and FEV₁ < 80% predicted at any study visit. Controls (N=40,360)
146 were selected in a ratio of approximately five controls to one case from participants reporting no
147 asthma or COPD, and with FEV₁ > 80% predicted and FEV₁/FVC > 0.7. Associations were adjusted for
148 age (at recruitment), sex, smoking status, and 10 ancestry principal components.

149 *Respiratory infections*

150 We defined respiratory tract infections requiring hospital admission in UK Biobank, using a wide
151 range of ICD-10 codes (**Supplementary Table 2**). Cases were individuals with one or more admission
152 for respiratory infections (N=19,459). Controls had no admissions for respiratory infections and were
153 selected in a ratio of approximately five controls to one case (n=101,438 after exclusions).
154 Associations were adjusted for age (at recruitment), age², sex, smoking status, genotyping array, and
155 10 ancestry principal components.

156 **Statistical methods**

157 *Univariable MR analyses of eosinophils and all respiratory traits and diseases*

158 We performed seven separate MR analyses of eosinophils on all outcomes, including three
159 quantitative lung function traits (FEV₁, FVC, and FEV₁/FVC); and four clinical disease phenotypes
160 (asthma, AECOPD, ACO and respiratory infections) utilising genetic IVs selected from¹⁸. The selection
161 of genetic IVs and the harmonisation of SNP-exposure and SNP-outcome datasets is detailed in the
162 **Extended Methods**. A total of 151 eosinophil IVs were used for this analysis. The primary analysis
163 used the inverse-variance weighted (IVW) MR method, and we checked sensitivity to condition (iii)
164 described in **Box 1** using sensitivity analyses, conducted using MR-Egger regression,²⁴ and the
165 weighted median estimator²⁵ (see **Extended Methods** for details, including the different
166 assumptions that the methods rely upon for validity). Further sensitivity analyses: i) restricted to
167 non-UKB FEV₁/FVC GWAS data, to assess sensitivity to sample overlap, and ii) restricted to FEV₁/FVC
168 GWAS data in UKB, excluding individuals with asthma.

169 *Multivariable MR analyses of multiple blood cell types and respiratory outcomes*

170 Since SNPs affecting eosinophils also affect other blood count types,¹⁴ we used multivariable MR in
171 order to estimate the influence of multiple cell types on respiratory outcomes, after conditioning on
172 the effects of the SNPs on other cell types. Multivariable MR analyses were run for respiratory
173 outcomes that showed evidence of eosinophil causation in the IVW MR analyses above, and that had
174 broadly consistent effect estimates in the weighted median and MR-Egger analyses. We also
175 performed an analysis of FEV₁/FVC in UKB (excluding asthma cases).

176 There were 1166 SNPs associated with at least one of eight blood count traits reported by Astle *et*
177 *al*¹⁴ at a genome-wide threshold. SNPs were LD clumped, and effect sizes were extracted from each
178 blood cell trait GWAS, and from the outcome GWASs. Effects for 318 clumped SNPs were
179 successfully harmonised, i.e. so effect sizes for the SNP-exposure and SNP-outcome effects
180 corresponded to the same allele (**Supplementary Table 3, Extended Methods**).

181 To implement multivariable MR, we used the `mv_multiple()` function of the ‘TwoSampleMR’ R
182 package.²⁶⁻²⁸

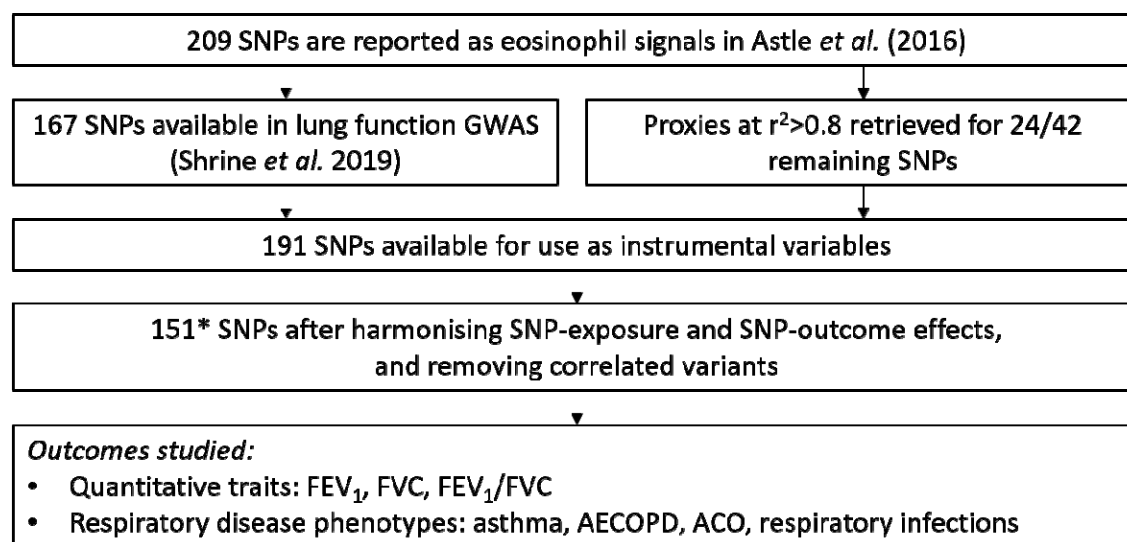
183 This analysis had two aims: i) to further investigate the possibility of horizontal pleiotropy affecting
184 the results of the eosinophil MR; and ii) to establish whether any other cell types besides eosinophils
185 could affect the respiratory outcomes studied.

186 Results

187 Univariable MR analyses of eosinophils and respiratory outcomes

188 There were 151 SNPs available for the univariable MR analyses, which were used in analyses of three
189 quantitative lung function traits (FEV₁, FVC, and FEV₁/FVC), and four respiratory disease phenotypes
190 (moderate-to-severe asthma, AECOPD, ACO and respiratory infections). Details of the selection of
191 SNP IVs is described in **Figure 1**.

192 **Figure 1** Selection of SNPs for univariable MR analyses of eosinophils and respiratory outcomes



193
194

195 Legend

196 Flowchart describing the analysis workflow for initial MR analyses of eosinophils. Of 209 SNPs
197 associated with eosinophil count, 167 were available in lung function GWASs (missingness is due to
198 some SpiroMeta studies not being imputed to the HRC panel).¹⁸ LD proxies at $r^2 > 0.8$ were retrieved
199 for 24/42 missing variants. Of the resulting 191 SNPs, 188 were successfully harmonised between the
200 SNP-eosinophil and SNP-lung function data sets, and 151* remained after LD clumping at an r^2
201 threshold of 0.01. These 151 SNPs were used in analyses. *One SNP, rs9974367, was missing in the
202 moderate-severe asthma GWAS. AECOPD=acute exacerbation of COPD; ACO=asthma-COPD overlap.

203 Results are presented in **Figure 2**. Amongst the quantitative lung function traits, there was strongest
204 evidence for a causal effect of eosinophils on FEV₁/FVC (SD change in FEV₁/FVC per SD eosinophils,
205 IVW estimate=-0.049 [95% CI: -0.079, -0.020]), with a smaller effect on FEV₁ (IVW estimate=-0.028
206 [95% CI: -0.054, -0.003]). There was substantial heterogeneity of SNP-specific causal estimates for all

207 three traits, as evidenced by the large values of Cochran's Q statistic. However, heterogeneity was
208 likely due to balanced pleiotropy, since weighted median estimates were consistent with the IVW
209 estimates. Confidence intervals for the MR Egger estimates were broad, consistent with the low
210 power of this method, but there was no evidence of unbalanced pleiotropy, as evidenced by the p-
211 values for the intercept of the MR Egger intercept term. Scatterplots of SNP-outcome against SNP-
212 exposure effects for the three quantitative outcomes are given in **Supplementary Figure 1**.

213 Amongst the respiratory disease phenotypes (also **Figure 2**), there was strong evidence for a causal
214 effect of eosinophils on asthma (OR per SD eosinophil count, IVW method=2.46 [95% CI: 1.98, 3.06]),
215 and ACO (IVW OR=1.86 [95% CI: 1.52, 2.27]). There was substantial heterogeneity of SNP-specific
216 causal estimates for these two traits, and the weighted median estimates were smaller in magnitude
217 than the IVW estimates (weighted median OR: 1.50 [95% CI: 1.23, 1.83] for asthma, and 1.44 [95%
218 CI: 1.19, 1.74] for ACO). Whilst the confidence intervals for the MR Egger estimates were still broad,
219 estimates were generally closer to the weighted median estimates, suggesting that the IVW
220 estimates were inflated due to the presence of horizontal pleiotropy, particularly for asthma.
221 Inflation for the asthma estimate is also more likely due to larger overlap between the SNP-exposure
222 and SNP-outcome datasets for this analysis (see **Extended Methods**). Scatterplots of SNP-outcome
223 against SNP-exposure effects for all four outcomes are given in **Supplementary Figure 2**, which
224 clearly demonstrate the heterogeneity in causal effect estimates. Conversely, there was no evidence
225 of association of eosinophil counts with AECOPD or respiratory infections. For both outcomes,
226 confidence intervals for all three MR estimates included the null, and all estimates were close to the
227 null.

228 See **Supplementary Table 4** for full results for all models.

229 *Sensitivity analysis to assess the effects of sample overlap for quantitative lung function traits*

230 UK Biobank featured in all GWAS datasets used, albeit the blood cell count GWAS and asthma GWAS
231 included only approximately one third of the UK Biobank genotype data.¹⁴ Since we had access to
232 data from the subset of the quantitative lung function GWASs that did not include UK Biobank, we
233 conducted sensitivity analyses to assess for the effect of sample overlap (see **Extended Methods**).
234 Results were generally consistent with the main results (e.g. SD change in FEV₁/FVC per SD
235 eosinophil count, IVW estimate=-0.041 [95% CI: -0.072, -0.009]; SD change FEV₁ per SD eosinophil
236 count=-0.043 [95%CI -0.077,-0.010]) (**Supplementary Table 5**).

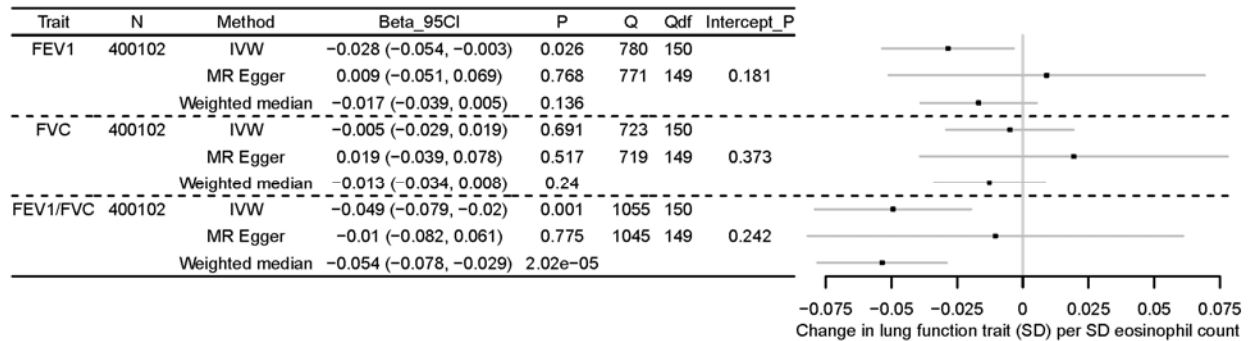
237 *Sensitivity analysis to assess the effect on FEV₁/FVC in individuals without asthma*

238 The causal effect of eosinophils on FEV₁/FVC was recalculated using data from UK Biobank, excluding
239 individuals with an asthma diagnosis. The effect size was attenuated compared to the main results,

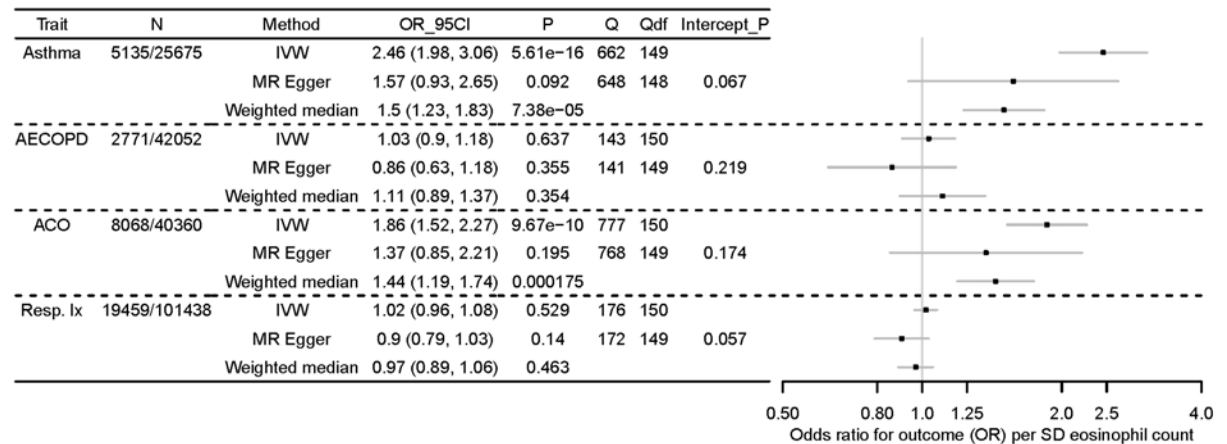
240 and the 95% confidence interval overlapped the null: IVW -0.013 [95% CI: -0.041, 0.015] (see

241 **Supplementary Table 6**).

242 **Figure 2 MR analyses of eosinophils (exposure) on three quantitative lung function traits (top) and**
 243 **four respiratory disease phenotypes (bottom), using 151 eosinophil-associated SNPs**



244
 245



246
 247

248 *Legend*

249 *Top: Results of MR analyses of eosinophil counts (exposure) on three quantitative lung function traits*
 250 *(outcome), FEV₁, FVC, and FEV₁/FVC. A forest plot of three estimates for each traits is shown (IVW,*
 251 *MR Egger, weighted median), along with the maximum sample size in the outcome GWAS (N), the*
 252 *effect size in SD change in outcome trait per SD eosinophil count, and 95% confidence interval, values*
 253 *for Cochran's Q statistic (Q) and the associated degrees of freedom (Q_df), and the P-value for the*
 254 *MR Egger intercept (Intercept_P). Boxes of the forest plot represent effect sizes, whiskers are 95%*
 255 *confidence intervals.*

256 *Bottom: Results of MR analyses of eosinophil counts (exposure) on four respiratory disease*
 257 *phenotypes (outcome), moderate-to-severe asthma, acute exacerbations of COPD (AECOPD),*

258 *asthma-COPD overlap (ACO), and respiratory infection (Resp. Ix). A forest plot of three estimates for*
259 *each traits is shown (IVW, MR Egger, weighted median), along with sample size in the outcome*
260 *GWAS for cases and controls, respectively (N), the effect size as odds ratio (OR) per SD eosinophil*
261 *count, and 95% confidence interval, values for Cochran's Q statistic (Q) and the associated degrees of*
262 *freedom (Q_df), and the P-value for the MR Egger intercept (Intercept_P). Boxes of the forest plot*
263 *represent odds ratios, whiskers are 95% confidence intervals. NB only 150/151 of the eosinophil SNPs*
264 *were available in the moderate-to-severe asthma GWAS.*

265 **Multivariable MR analyses of blood cell counts and respiratory outcomes**

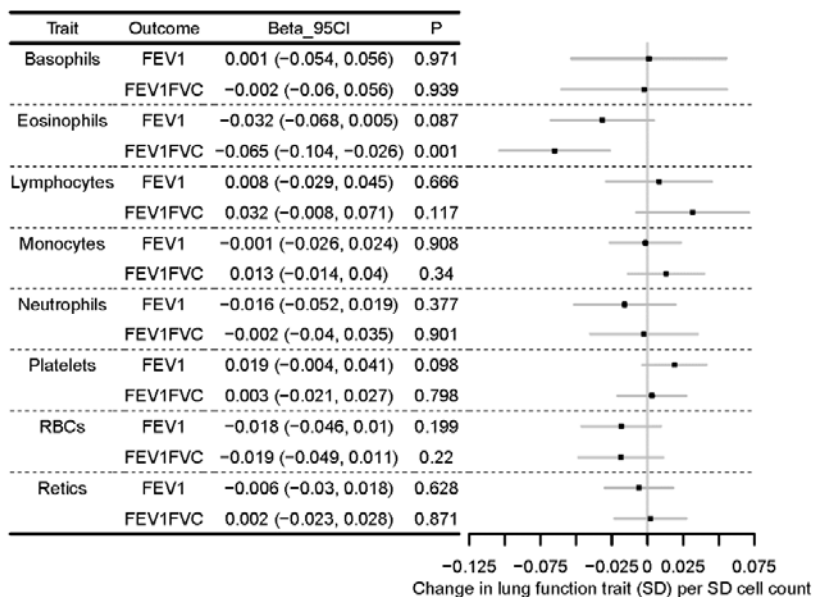
266 To further explore causality between blood cell parameters and FEV₁, FEV₁/FVC, moderate-to-severe
267 asthma and ACO, we carried out multivariable MR analyses, using eight cell types as exposures
268 (basophils, eosinophils, neutrophils, monocytes, lymphocytes, platelets, red blood cells and
269 reticulocytes).

270 The selection of SNP IVs for multivariable MR is described in the **Methods**, and **Supplementary Table**
271 **3**. Briefly, 1166 unique SNPs were associated with at least one of the eight cell types at $p < 8.31 \times 10^{-9}$
272 in Astle *et al.*, and were available in the outcome GWASs. After LD-clumping, 329 SNPs remained,
273 and after harmonising SNP-exposure and SNP-outcome effects, 318 SNPs remained. Estimated
274 variance explained (r^2) and instrument strength (F-statistics) by cell-type are also given in
275 **Supplementary Table 3**. We present these estimates as a guide of relative strength of sets of IVs, but
276 do not intend for them to be interpreted as measures of absolute strength, given the likelihood of
277 Winner's curse bias.²⁰

278 Multivariable MR results for FEV₁ and FEV₁/FVC are presented in **Figure 3**. Even after conditioning on
279 the effects of the SNPs on other cell types, eosinophils reduced lung function as measured by
280 FEV₁/FVC (multivariable estimate, SD change in FEV₁/FVC per SD eosinophils adjusted for other cell
281 types: -0.065 [95% CI: -0.104, -0.026]). The eosinophil point estimate for FEV₁ was consistent with
282 the univariable estimate (**Figure 2**), but results for eosinophils (-0.032 [95%CI: -0.068, 0.005]) and
283 other cell types were consistent with the null. When asthma cases were excluded from SNP-
284 FEV₁/FVC results, the eosinophil estimate attenuated, and confidence intervals overlapped the null (-
285 0.028 [95%CI: -0.069, 0.013]) consistent with the causal effect of eosinophils on lung function being
286 of greater magnitude in people with a history of asthma (**Supplementary Figure 3**).

287 Results of the multivariable MR analysis for the two respiratory disease outcomes (ACO and asthma)
288 are presented in **Figure 4**. There was an association of eosinophil count with both ACO (OR 1.95
289 [95% CI: 1.57, 2.42]) and asthma [OR 2.90 [95% CI: 2.31, 3.65]], after adjusting for the effects of the
290 SNPs on other cell types. Confidence intervals for other cell type estimates were consistent with the
291 null effect of 1, with the exception of neutrophils for ACO.

292 **Figure 3 Multivariable MR analyses of eight cell types and FEV₁ and FEV₁/FVC**



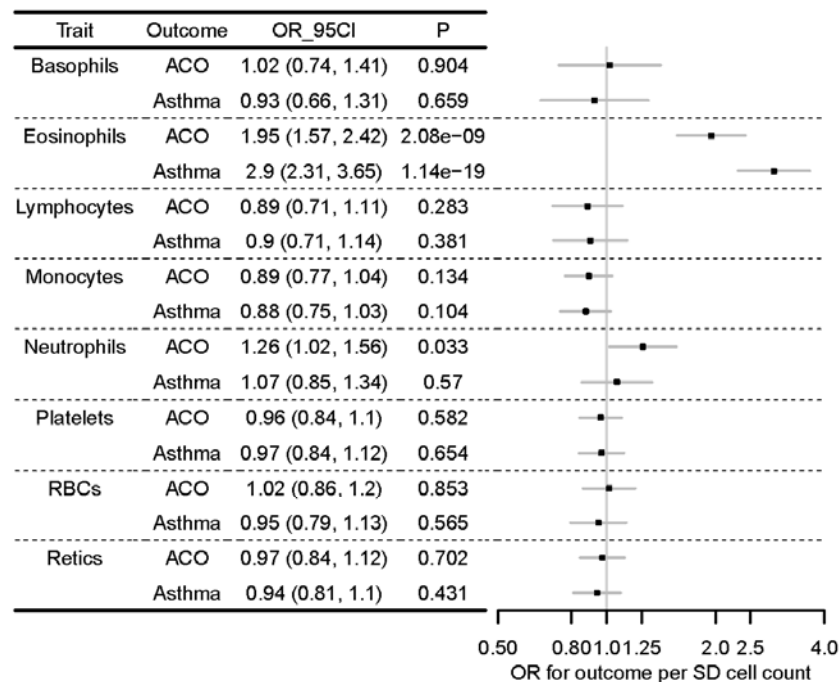
293

294 *Forest plot showing multivariable MR estimating the causal effect of multiple cell types on*
 295 *quantitative lung function outcomes, after conditioning on the effects of the SNPs on other cell types.*
 296 *Models were run for each of forced expiratory volume in 1 second (FEV₁) and the ratio of FEV₁ to*
 297 *forced vital capacity (FVC) separately, but effect sizes are shown next to one another for comparison.*
 298 *Effect sizes (beta, 95% confidence interval, 95CI) are in SD change in lung function outcome per SD*
 299 *cell count (adjusted for the effects of other cell types). Points of the forest plot represent effect size*
 300 *estimate; whiskers are 95% confidence intervals.*

301

302

303 **Figure 4 Multivariable MR analyses of eight cell types and two respiratory disease outcomes, ACO**
 304 **and asthma**



305

306 *Legend*

307 *Forest plot showing multivariable MR estimating the causal effect of multiple cell types on*
 308 *respiratory disease outcomes, after conditioning on the effects of the SNPs on other cell types.*
 309 *Models were run for each of ACO and asthma separately, but effect sizes are shown next to one*
 310 *another for comparison. Odds ratios (OR, 95% confidence interval, 95CI) are per SD cell count*
 311 *(adjusted for the effects of other cell types). Points of the forest plot represent odds ratios; whiskers*
 312 *are 95% confidence intervals.*

313 Discussion

314 Using Mendelian randomization, we found evidence for a causal effect of eosinophils in determining
315 FEV₁/FVC, FEV₁, ACO and asthma. For these four outcomes, we additionally performed multivariable
316 MR to investigate the causal role of a broader range of eight blood cell types, and found that
317 amongst these cell types eosinophils had the strongest evidence of an effect on the respiratory
318 outcomes studied.

319 To our knowledge, this is the largest MR study of eosinophils in relation to lung function, and the
320 first to investigate the relationship between eosinophils and AECOPD, ACO, and respiratory
321 infections. Whilst terminology of ACO has changed over time, the recognition that asthma and COPD
322 coexist in some patients has not changed,²⁹ and this is what our analysis aimed to capture.

323 A previous two-sample MR analysis of eosinophils in relation to asthma was undertaken by the
324 authors of the cell count GWAS used to derive IVs in the current study; these authors used asthma
325 GWAS data from the GABRIEL study.¹⁴ Our results support their conclusion that eosinophils are
326 causal determinants of asthma. We are aware of one other MR study of eosinophil counts in relation
327 to asthma, COPD, FEV₁ and FEV₁/FVC, conducted entirely within the LifeLines cohort (N=13,301, 5
328 SNPs used as IVs).¹⁶ In that study, confidence intervals for causal estimates of eosinophils were
329 consistent with the null, albeit point estimates were consistent with a harmful effect for FEV₁/FVC,
330 asthma and COPD. Here, we used the largest eosinophil count GWAS to date (N=172,275)¹⁴ to derive
331 IVs, and found evidence for causality of eosinophils in reducing FEV₁/FVC, the lung function trait in
332 which impairment is the key feature of COPD diagnosis, and FEV₁, the trait used to grade airflow
333 limitation in COPD. We did not find strong evidence of a causal effect of eosinophils on FVC, despite
334 observing negative point estimates. This is of interest because it is a reduction in FEV₁ (and therefore
335 in FEV₁/FVC) that is a hallmark of obstructive lung diseases, with effects on FVC being generally less
336 prominent. Our sensitivity analyses highlighted a larger estimated causal effect of eosinophils on
337 FEV₁/FVC among individuals with a history of asthma, with an attenuation of effect size estimates
338 seen when excluding this group. Together with the effect shown for ACO, our findings highlight the
339 importance of eosinophils on lung function and airflow obstruction in people with a history of
340 asthma.

341 Amongst the outcome traits showing evidence for causation by eosinophils, we observed substantial
342 heterogeneity in causal estimates for individual SNPs, i.e. different causal effects on the outcome
343 phenotypes may arise from different genetic variants influencing eosinophils. Accordingly, we
344 compared MR methods that rely on differing assumptions for validity (**Box 1**). When comparing the
345 IVW, MR-Egger and weighted median results, effect estimates for eosinophils on quantitative lung

346 function traits were broadly consistent, but point estimates for ACO and asthma attenuated
347 substantially when using the MR-Egger and weighted median approaches (albeit estimates were still
348 consistent with a harmful effect of raised eosinophils). This suggests there were SNPs with
349 pleiotropic associations, e.g. SNPs associated with asthma and ACO via pathways other than
350 eosinophils, which is a known challenge in MR studies (see also **Box 1**).

351 Since many of the SNPs used as IVs for eosinophil count are also associated with other cell counts,¹⁴
352 we performed multivariable MR to estimate the influence of multiple cell types simultaneously, after
353 conditioning on the effects of the SNPs on other cell types. Whilst such an analysis has previously
354 been undertaken for asthma,¹⁴ we also performed multivariable MR analyses for FEV₁, FEV₁/FVC,
355 asthma and ACO. Whilst we did not find substantial evidence for a harmful effect of neutrophils on
356 asthma, nor a protective effect of monocytes and lymphocytes, as reported by Astle *et al.*,¹⁴ the
357 effect directions in our multivariable MR analysis were consistent with those of this previous study
358 for neutrophils, monocytes and lymphocytes. In multivariable MR analyses, we observed a larger
359 effect of eosinophils on asthma than reported previously: this could be because our SNP-outcome
360 dataset was of moderate-to-severe asthma, and around half of the cases and the majority of
361 controls were also included in the exposure GWAS. Nevertheless, the univariable MR estimates from
362 MR-Egger regression and weighted median estimation were consistent with the previous estimate
363 reported for asthma in multivariable analysis by Astle *et al.*¹⁴

364 In comparison with the harmful effect of eosinophils on FEV₁/FVC, we did not observe a harmful
365 effect of neutrophils on FEV₁/FVC, despite the known relationship between neutrophilia and COPD.³⁰
366 This could be due to lower power of the neutrophil IVs compared to the eosinophil IVs
367 (**Supplementary Table 3**), but it is also possible that neutrophilic inflammation is more of a
368 consequence than a cause of reduced FEV₁/FVC, for example in response to infection,³¹ and
369 therefore may represent a marker of disease progression. There was very weak evidence for a
370 harmful effect of neutrophils on ACO. Whilst we did not find strong evidence for causality of
371 eosinophils on AECOPD and respiratory infections, point estimates were consistent with a harmful
372 effect on AECOPD, and may have been limited by power. The effects of anti-IL5 drugs that have been
373 attributed to the reduction of eosinophils have been noted to be smaller in acute exacerbations of
374 COPD compared to asthma.^{3,32}

375 Key strengths of the study are that we carried out two-sample analyses with differing sensitivities to
376 underlying assumptions (including MR-Egger and weighted median analysis). We used the largest
377 GWAS dataset available for eosinophil count together with multiple large GWASs of lung function
378 and respiratory disease. In addition to the advantages of valid MR analyses in mitigating against

379 some key limitations of observational epidemiology (reverse causality and confounding), using
380 genetic proxies for eosinophils may provide a better estimate for long-term eosinophilia, which may
381 be useful, since one-off assays of cell counts may show a degree of lability.³³ Another strength of our
382 work is that we undertook multivariable MR analyses to investigate causality between multiple cell
383 types and the outcomes studied, whilst controlling for the effects of IVs that may have had
384 pleiotropic effects via other cell types.

385 We acknowledge several limitations. We did not have post-bronchodilator measures of spirometry
386 available. However, our criterion for defining asthma-COPD overlap required a % predicted FEV1 of
387 less than 80% (consistent with a spirometric definition of GOLD Stage 2-4 COPD); using the same
388 pre-bronchodilator spirometry definition of COPD we have previously shown a positive predictive
389 value of 98% for diagnosis of postbronchodilation-defined COPD.³⁴ Sample overlap between the
390 SNP-eosinophil and SNP-outcome datasets (all datasets included participants from UK Biobank)
391 could bias estimates towards the observational eosinophil-outcome association.²⁰ As a sensitivity
392 analysis, we repeated the univariable lung function MR analysis of eosinophils using SNP-lung
393 function results from the SpiroMeta consortium (i.e. omitting UK Biobank), and found a consistent
394 IVW estimate. Nevertheless, we acknowledge that our analyses of other outcomes (particularly the
395 asthma analysis) could be vulnerable to a degree of non-conservative bias.^{19,20} The results of our MR
396 analyses also use genome-wide results adjusted for covariates, and therefore may be susceptible to
397 collider bias.¹⁹ For the binary phenotypes of asthma, AECOPD and respiratory infections, there may
398 also have been bias due to misclassification in electronic healthcare records, which we would expect
399 to be non-differential, and therefore towards the null. There is also potential bias in the causal
400 estimates for binary outcomes due to non-collapsibility of the odds ratio.²¹ Our analyses do not
401 consider the possibility of non-linear effects between eosinophil counts and lung function, which
402 have been observed: the results presented are the average difference in lung function if eosinophils
403 were increased by the same amount for every individual in the population. Finally, we acknowledge
404 that the multivariable analyses may still be vulnerable to residual pleiotropy via pathways other than
405 the eight cell types studied. However, we argue that the sensitivity analyses undertaken together
406 support a causal role for eosinophils in the traits studied. The significant heterogeneity in SNP-
407 specific causal estimates may arise from pleiotropy. A specific genetic variant influencing eosinophils
408 may have a greater or lesser effect on these respiratory phenotypes than the direct effect of
409 eosinophils, possibly in the opposing direction.

410 Our work provides insight into the role of eosinophils in causing respiratory conditions characterised
411 by airflow obstruction. At present, treatment with the anti-IL5 agents mepolizumab and
412 benralizumab in asthma is initiated according to levels of eosinophil counts,⁹ yet it is possible that a

413 more proximal factor, potentially such as IL5 itself, may be an even better predictor of drug
414 response. Future work could seek therefore to identify whether particular pathways upstream of
415 eosinophil counts might help design better methods for deciding upon treatment initiation.

416 To conclude, we found new evidence for a causal effect of eosinophils on ACO, FEV₁, and FEV₁/FVC,
417 as well as reproducing a previous association with asthma. Taken together, our findings support
418 eosinophils being important causal determinants of airflow obstruction, and suggest that the
419 strongest causal effects of eosinophils are in people with a history of reversible airflow obstruction,
420 including the patient group with features of both asthma and COPD.

421 **URLs**

422 EBI GWAS catalog: www.ebi.ac.uk/gwas

423 rAggr: <http://raggr.usc.edu/>

424 'TwoSampleMR' R package documentation: <https://mrcieu.github.io/TwoSampleMR/>

425 **Author Contributions**

426 NS, ALG, ATW, CJ, NR, the SpiroMeta consortium, IPH, LVW and MDT were involved in data
427 collection, by generating GWAS data of the respiratory traits used in these analyses. NAS, FD, IPH
428 and MDT were involved in data interpretation. IPH and MDT were involved in study design. ALG was
429 involved in all aspects of the work, performed the analysis and wrote the first draft of the paper,
430 including all figures and tables. All authors critically reviewed the manuscript and contributed to
431 data interpretation and draft revision.

432 **Ethics committee approval**

433 This research used published summary data from published genome-wide association studies, plus
434 data from UK Biobank. This particular research was conducted using the UK Biobank Resource under
435 Application Number 648. UK Biobank has ethical approval from the UK National Health Service (NHS)
436 National Research Ethics Service (Ref 11/NW/0382).

437 **Role of the funding source**

438 ALG is funded by internal fellowships at the University of Leicester from the Wellcome Trust
439 Institutional Strategic Support Fund (204801/Z/16/Z) and the BHF Accelerator Award
440 (AA/18/3/34220). CJ holds a Medical Research Council Clinical Research Training Fellowship
441 (MR/P00167X/1). ATW is funded by a BBSRC CASE studentship with GSK. LVW holds a GSK/British
442 Lung Foundation Chair in Respiratory Research. MDT is supported by a Wellcome Trust Investigator

443 Award (WT202849/Z/16/Z). MDT and LVW have been supported by the MRC (MR/N011317/1). MDT
444 and IPH hold NIHR Senior Investigator Awards. The research was also supported by BREATHE — The
445 Health Data Research Hub for Respiratory Health (MC_PC_19004). The research was partially
446 supported by the NIHR Leicester Biomedical Research Centre and the NIHR Nottingham Biomedical
447 Research Centre; the views expressed are those of the author(s) and not necessarily those of the
448 NHS, the NIHR or the Department of Health. The funders had no role in the design of the Mendelian
449 randomization analyses. ALG and MDT were involved in all stages of study development and
450 delivery, and MDT had full access to all data in the study and final responsibility for the decision to
451 submit for publication.

452 Acknowledgements

453 We are grateful to Dr William J Astle for his advice. This research used the ALICE and SPECTRE High
454 Performance Computing Facilities at the University of Leicester. This analysis used summary data
455 available via the BREATHE Digital Innovation Hub hosted at SAIL databank
456 (<https://saildatabank.com/>).

457 Declaration of Interests

458 MDT and LVW receive funding from GSK for collaborative research projects outside of the submitted
459 work. IPH has funded research collaborations with GSK, Boehringer Ingelheim and Orion.

460 Data sharing statement

461 This manuscript uses summary-level GWAS data. Lung function GWAS summary statistics as well as
462 the blood cell GWAS summary statistics are available from the EBI GWAS catalog
463 (<https://www.ebi.ac.uk/gwas/>). Other GWAS summary statistics are available from the authors on
464 request.

465 References

- 466 1. Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal
467 inference in epidemiological studies. *Hum Mol Genet* **23**, R89-98 (2014).
- 468 2. George, L. & Brightling, C.E. Eosinophilic airway inflammation: role in asthma and chronic
469 obstructive pulmonary disease. *Thorax* **71**, 34-51 (2016).
- 470 3. Ortega, H.G. *et al.* Severe eosinophilic asthma treated with mepolizumab stratified by
471 baseline eosinophil thresholds: a secondary analysis of the DREAM and MENSA studies.
472 *Lancet Respir Med* **4**, 549-556 (2016).
- 473 4. Pavord, I.D. *et al.* Mepolizumab for Eosinophilic Chronic Obstructive Pulmonary Disease. *N*
474 *Engl J Med* **377**, 1613-1629 (2017).
- 475 5. Eltboli, O. & Brightling, C.E. Eosinophils as diagnostic tools in chronic lung disease. *Expert Rev*
476 *Respir Med* **7**, 33-42 (2013).
- 477 6. Saetta, M. *et al.* Airway eosinophilia in chronic bronchitis during exacerbations. *Am J Respir*
478 *Crit Care Med* **150**, 1646-52 (1994).
- 479 7. Pavord, I.D. *et al.* Mepolizumab for severe eosinophilic asthma (DREAM): a multicentre,
480 double-blind, placebo-controlled trial. *Lancet* **380**, 651-9 (2012).
- 481 8. Haldar, P. *et al.* Mepolizumab and exacerbations of refractory eosinophilic asthma. *N Engl J*
482 *Med* **360**, 973-84 (2009).

- 483 9. National Institute for Health and Care Excellence. Mepolizumab for treating severe
484 refractory eosinophilic asthma. Technology appraisal guidance [TA431]. (2017).
- 485 10. U.S. Food and Drug Administration (FDA). Drug Trials Snapshots: NUCAA. (2016).
- 486 11. U.S. Food and Drug Administration (FDA). Drug Trials Snapshots: FASENRA. (2017).
- 487 12. European Medicines Agency. Nucala (mepolizumab). (European Medicines Agency,
488 Amsterdam, 2019).
- 489 13. European Medicines Agency. Fasenra (benralizumab). (European Medicines Agency,
490 Amsterdam, 2019).
- 491 14. Astle, W.J. *et al.* The Allelic Landscape of Human Blood Cell Trait Variation and Links to
492 Common Complex Disease. *Cell* **167**, 1415-1429 e19 (2016).
- 493 15. Hancox, R.J., Pavord, I.D. & Sears, M.R. Associations between blood eosinophils and decline
494 in lung function among adults with and without asthma. *Eur Respir J* **51**(2018).
- 495 16. Amini, M. *et al.* Blood Eosinophil Count and Metabolic, Cardiac and Pulmonary Outcomes: A
496 Mendelian Randomization Study. *Twin Res Hum Genet* **21**, 89-100 (2018).
- 497 17. Sakornsakolpat, P. *et al.* Genetic landscape of chronic obstructive pulmonary disease
498 identifies heterogeneous cell-type and phenotype associations. *Nat Genet* **51**, 494-505
499 (2019).
- 500 18. Shrine, N. *et al.* New genetic signals for lung function highlight pathways and chronic
501 obstructive pulmonary disease associations across multiple ancestries. *Nat Genet* **51**, 481-
502 493 (2019).
- 503 19. Davies, N.M., Holmes, M.V. & Davey Smith, G. Reading Mendelian randomisation studies: a
504 guide, glossary, and checklist for clinicians. *BMJ* **362**, k601 (2018).
- 505 20. Burgess, S., Davies, N.M. & Thompson, S.G. Bias due to participant overlap in two-sample
506 Mendelian randomization. *Genet Epidemiol* **40**, 597-608 (2016).
- 507 21. Sheehan, N.A. & Didelez, V. Epidemiology, genetic epidemiology and Mendelian
508 randomisation: more need than ever to attend to detail. *Hum Genet* **139**, 121-136 (2020).
- 509 22. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature*
510 **562**, 203-209 (2018).
- 511 23. Shrine, N. *et al.* Moderate-to-severe asthma in individuals of European ancestry: a genome-
512 wide association study. *Lancet Respir Med* **7**, 20-34 (2019).
- 513 24. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid
514 instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*
515 **44**, 512-25 (2015).

- 516 25. Bowden, J., Davey Smith, G., Haycock, P.C. & Burgess, S. Consistent Estimation in Mendelian
517 Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet*
518 *Epidemiol* **40**, 304-14 (2016).
- 519 26. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the
520 human phenome. *eLife* **7**, e34408 (2018).
- 521 27. Burgess, S., Dudbridge, F. & Thompson, S.G. Re: “Multivariable Mendelian Randomization:
522 The Use of Pleiotropic Genetic Variants to Estimate Causal Effects”. *American Journal of*
523 *Epidemiology* **181**, 290-291 (2015).
- 524 28. Burgess, S. & Thompson, S.G. Multivariable Mendelian randomization: the use of pleiotropic
525 genetic variants to estimate causal effects. *Am J Epidemiol* **181**, 251-60 (2015).
- 526 29. Global Initiative for Chronic Obstructive Lung Disease, I. Global Strategy for the Diagnosis,
527 Management, and Prevention of Chronic Obstructive Pulmonary Disease (2020 Report).
528 (2020).
- 529 30. Quint, J.K. & Wedzicha, J.A. The neutrophil in chronic obstructive pulmonary disease. *J*
530 *Allergy Clin Immunol* **119**, 1065-71 (2007).
- 531 31. Jasper, A.E., McIver, W.J., Sapey, E. & Walton, G.M. Understanding the role of neutrophils in
532 chronic inflammatory airway disease. *F1000Res* **8**(2019).
- 533 32. Brightling, C. & Greening, N. Airway inflammation in COPD: progress to precision medicine.
534 *Eur Respir J* **54**(2019).
- 535 33. Shah, A.D. *et al.* White cell count in the normal range and short-term and long-term
536 mortality: international comparisons of electronic health record cohorts in England and New
537 Zealand. *BMJ Open* **7**(2017).
- 538 34. Soler Artigas, M. *et al.* Genome-wide association and large-scale follow up identifies 16 new
539 loci influencing lung function. *Nat Genet* **43**, 1082-90 (2011).

540

541