nature genetics

Article

https://doi.org/10.1038/s41588-023-01372-4

Inference of chronic obstructive pulmonary disease with deep learning on raw spirograms identifies new genetic loci and improves risk models

In the format provided by the authors and unedited

Supplementary Notes

ML-based COPD enriched in lung tissue

Utilizing S-LDSC to perform tissue and cell-type specific analysis, we observed that fetal lung and smooth-muscle are the relevant tissues for ML-based COPD (Supplementary Tables **11** and **12**). These tissues and cells are similar to previous work [S1], further indicating that our ML-based COPD is a valid COPD phenotype and the improvement observed in number of additional hits/loci are not due to capturing other non-COPD phenotypes with high heritability (e.g., height). Furthermore, we observed that colon smooth muscle (H3K4me1; $P = 5 \times 10^{-10}$) and fetal lung (H3K4me1; $P = 2 \times 10^{-9}$) are the relevant tissues for ML-based COPD GWAS conditional on FEV_1/FVC (Supplementary Table **13**). Similar to S-LDSC analysis, GARFIELD [S2] indicated that the fetal lung has the largest enrichment (Supplementary Figure **6**) and the conditional GWAS of ML-based COPD on FEV_1/FVC was enriched in fetal lung and embryonic lung (Supplementary Figure **7**). Lastly, to understand the effect of cis-regulatory interactions, we applied GREAT [S3] to ML-based COPD GWS loci. The ML-based loci were significantly enriched for 82 ontology terms, primarily development and morphogenesis-related. Of particular note, GREAT results were enriched for respiratory and cardiovascular system development and morphology terms (Supplementary Table **14**).

PheWAS analysis of significant ML-based COPD hits

Phenome-wide association studies (PheWAS) are used to examine pleiotropic effects, which are particularly relevant when considering pharmacological interventions on implicated genes or pathways. We performed PheWAS for the 796 independent ML-based GWAS hits using 4,083 phenotypes in UKB and 2,803 phenotypes in FinnGen. We used a false discovery rate (FDR) of 5% to detect phenotype and variant pairs that are significant in our PheWAS (Supplementary Table **15**). Not surprisingly, most of the significant associations detected by PheWAS are related to different lung function measures, such as FVC, FEV_1, FEV_1/FVC, and PEF (Supplementary Table **16**). Similar to Sakornsakolpat et al. [S1], our PheWAS analysis identified association with body composition: Weight (131 hits), BMI (96 hits), and fat-free mass (89 hits). In addition, PheWAS detected multiple significant associations with blood counts: white blood cell count (85 hits), red blood cell counts (85

hits), haemoglobin concentration (85 hits), and platelet count (83 hits).

Conditional and DeepNull GWAS on ML-based COPD

Previous works [S8, S10] showed that many COPD hits are shared with FEV_1/FVC, FEV_1, FVC, and peak expiratory flow (PEF) hits. To ensure our ML-based COPD GWAS is not solely driven by FEV_1/FVC, we performed a secondary ML-based COPD GWAS conditioned on FEV_1/FVC (Supplementary Figure 8 illustrates the Manhattan plot and Supplementary Figure 9 illustrates the Q-Q plot). The SNP-heritability estimated from S-LDSC for conditional GWAS was 0.11 (s.e.m=0.01). The conditional GWAS identified 175 independent GWS hits at 117 independent GWS loci after merging hits within 250kb (Supplementary Table 21 and Table 22). Although the conditional analysis ensures that our ML-based COPD GWAS is not solely driven by FEV_1/FVC, it does not rule out cases where FEV_1/FVC has a non-linear effect on ML-based COPD. We utilized DeepNull [S26] to account for possible non-linear relationships between age, sex, and FEV_1/FVC. The ML-based COPD GWAS using DeepNull (Extended Data Fig 4 illustrates the Manhattan plot and Supplementary Figure 11 illustrates the Q-Q plot) identified 181 independent GWS at 129 independent GWS loci after merging hits within 250kb (Supplementary Tables 25 and 26). Thus, our ML-based COPD prediction captures a disease signal beyond FEV_1/FVC (Supplementary Table 27 and Table 28). Furthermore, we performed ML-based COPD GWAS conditioned on FEV_1/FVC, FEV_1, FVC, and PEF, observing a SNP-heritability of 0.04 (s.e.m = 0.00) and 41 independent GWS hits at 31 independent GWS loci (Supplementary Figure 10, Supplementary Table 23 and Table 24).

Permutation tests to validate ML-based phenotyping

To ensure that our result is not influenced by some bias introduced by our ML-based phenotyping procedure, we trained the ResNet18 model using permuted medical-record-based COPD labels and observed that the model predicts almost the same value for all individuals (0.0384 ± 0.0000), which matches the disease prevalence in the training data. In other words, the model cannot detect any patterns from inputs to the permuted labels and falls back on the best guess of prevalence for the probability of having COPD as ML-based COPD is agnostic to covariates such as age and height. Furthermore, we ran a GWAS on a permuted version of the original ML-based COPD phenotype

and observed that this permuted GWAS has SNP-heritability of zero (0.00 ± 0.01) and that no GWS variants were detected. Lastly, to ensure that the relationship between ML-based COPD and covariates is maintained when permuting GWAS inputs, we performed an additional experiment that shuffled our ML-based COPD phenotype and all GWAS covariates together. Note that, assuming low stochasticity in the model training process, this experimental setup is equivalent to jointly permuting spirometry curves, labels, and covariates across individuals, rerunning model training, and performing a GWAS. In this experiment as well no GWS variants were detected.

ML-based COPD identifies COPD better than proxy-GOLD for genetic discovery

First, we binarized the ML-based COPD risk into case/control labels with 50% prevalence (Methods) and compared a GWAS on this phenotype (hereafter "binarized ML-based COPD") with GWAS performed on medical-record-based COPD labels. We observed that binarized ML-based COPD has a higher significance level for all hits and the absolute magnitude of binarized ML-based COPD is larger than the raw label equivalents for all hits (Supplementary Figure **14**). We compared our binarized ML-based COPD with Sakornsakolpat et al. [S8] where we observed that our binarized ML-based COPD variants are more significant (Extended Data Fig **7**a) than Sakornsakolpat et al. [S8] while having the same effect size estimates ($R^2 = 0.91$; Extended Data Fig **7**b). Furthermore, when binarizing ML-based COPD so that disease prevalence matches Sakornsakolpat et al. [S8] (prevalence = 13.86%), the prevalence-matched binarized ML-based COPD has better power. Finally, when comparing a binarized ML-based COPD where we match the prevalence to proxy-GOLD, we observed that binarized ML-based COPD outperforms proxy-GOLD on all metrics including replicating previously known COPD hits (Extened Data Fig **8** and Figure **15**). Thus, even when ML-based COPD is considered as a binary trait, we show increased power (Supplementary Table **8**).

Replication of ML-based COPD novel loci in three independent datasets

A GWAS on the ML-based liability score identifies 265 COPD risk loci in addition to 91 previously known COPD loci with respect to to Sakornsakolpat et al. [S8] and GWAS catalog entries (as of 2022-07-09) for COPD, emphysema, chronic bronchitis. Out of 265 COPD-specific loci, 221 independently re-identified as associated with COPD or COPD-related lung function as follows. We observed that 101 out of 265 is statistically significant in a previous COPD GWAS [S8] after Bonferroni correction.

Also, 198 out of 265 are previously known FEV_1 or FEV_1/FVC loci with respect to [S10] and GWAS catalog entries (Supplementary Table **31** and Extended Data Fig **3**). From the remaining 67 novel loci out of 265, which are not previously known loci for COPD or COPD-related lung function, 23 is statistically significant in a previous COPD GWAS [S8] after Bonferroni correction that includes UKB samples. Furthermore, we analyzed three additional studies that do not include UKB samples to further quantify the replication status of these 67 loci. These three datasets are GBMI (Global Biobank Meta-analysis Initiative) [S27] excluding the UKB samples, SpiroMeta [S28], and ICGC (International COPD Genetics Consortium) [S7]. We defined two replication strategies: First, we defined *supportive* replication as consistent effect size direction between our ML-based COPD and the three comparators. The ICGC and GBMI GWAS are based on a COPD phenotype; thus, we expect their effect size signs to match our ML-based COPD. SpiroMeta phenotypes, on the other hand, capture lung function, so we expect their effect size signs to be the opposite of our ML-based COPD signs. Second, we defined *strict* replication as consistent effect size direction in any study with Bonferroni-corrected $P < 0.1$ (one-sided) for that study. We observed that 38/67 loci have *supportive* replication where the chance of this happening randomly is extremely small ($P \leq 2 \times 10^{-16}$). In addition, we observed that 6/67 have *strict* replication and, when relaxing the strict replication p-value from Bonferroni-corrected to nominal $P < 0.1$, 27/67 loci replicate (Supplementary Table **31** and Extended Data Fig **3**).

Tissue/Cell-type specific enrichment analysis of ML-based COPD hits

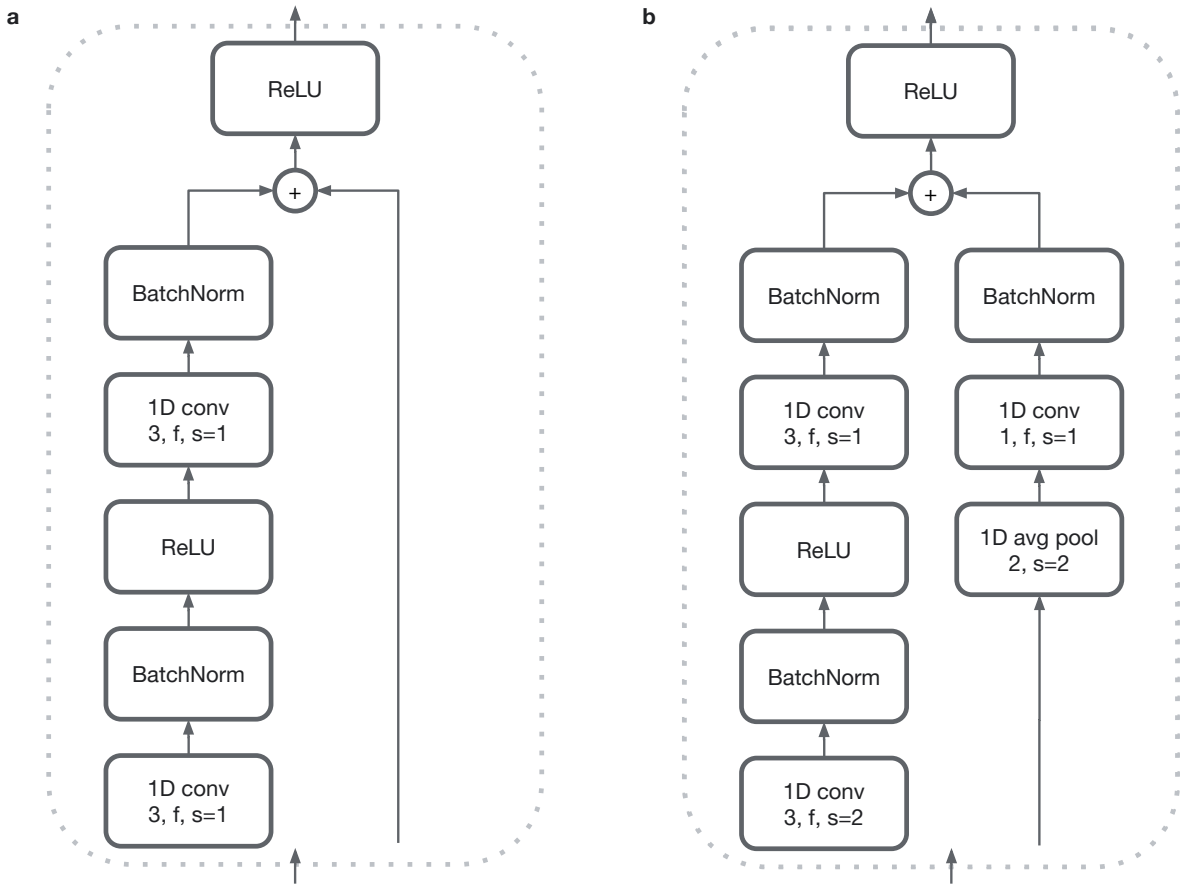
We utilized two methods to perform tissue/cell-type specific enrichment analysis. First, we utilized the tissues specific analysis in S-LDSC [S4, S5] where we utilized 53 baseline version 1 annotations (see URLs), "Multi_tissue_gene_expr" (includes both GTEx [S6] and Franke lab data [S7, S8]) and "Multi_tissue_chromatin" (includes both Roadmap [S9, S10] and EN-TEX data). In the case of gene expression, we utilize 53 tissues or cell types created by Finucane et al.[S5] while Franke lab data consists of 152 tissues or cell types. In the case of chromatin data, Roadmap [S9, S10] has 397 cell-type- or tissue-specific annotations while EN-TEX data has 93 cell-type- or tissue-specific annotations. As recommended by the S-LDSC authors, we used the -log (p-value) of regression coefficient (τ) as the metric to pick the specific tissue or cell-type. Second, we utilize GARFIELD [S2] to perform tissue-specific analysis where we utilized 424 DNase I hypersensitive site hotspot

annotations provided by the GARFIELD authors [S2] and we used the default parameters.

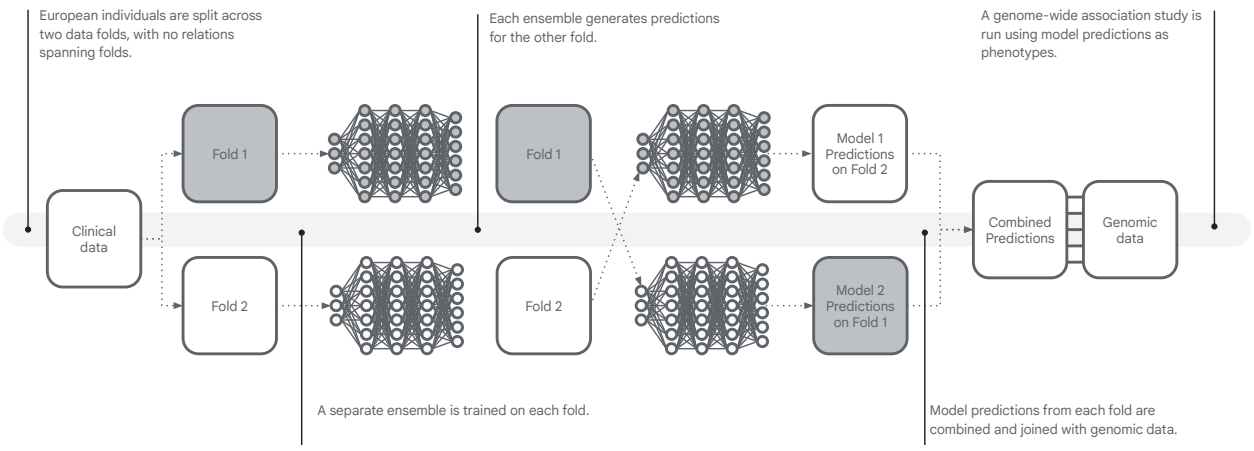
Functional analyses with GREAT

We utilized GREAT v4.0.4 [S3] on the human GRCh37 assembly to perform functional enrichment analysis of ML-based COPD risk loci. The default "basal+extension" region-gene association rule was used with 5 kb upstream, 1 kb downstream, 1000 kb extension, and curated regulatory domains included. GREAT analyzes enrichment of terms drawn from multiple data sources including Gene Ontology Biological Process (GOBP), the Mammalian Phenotype Ontology for phenotypes induced by a single gene knockout (MP1KO), and the Human Phenotype Ontology (HP). We considered terms to be statistically significant if the Bonferroni-corrected P-values for both the region-based and gene-based tests were ≤ 0.05.

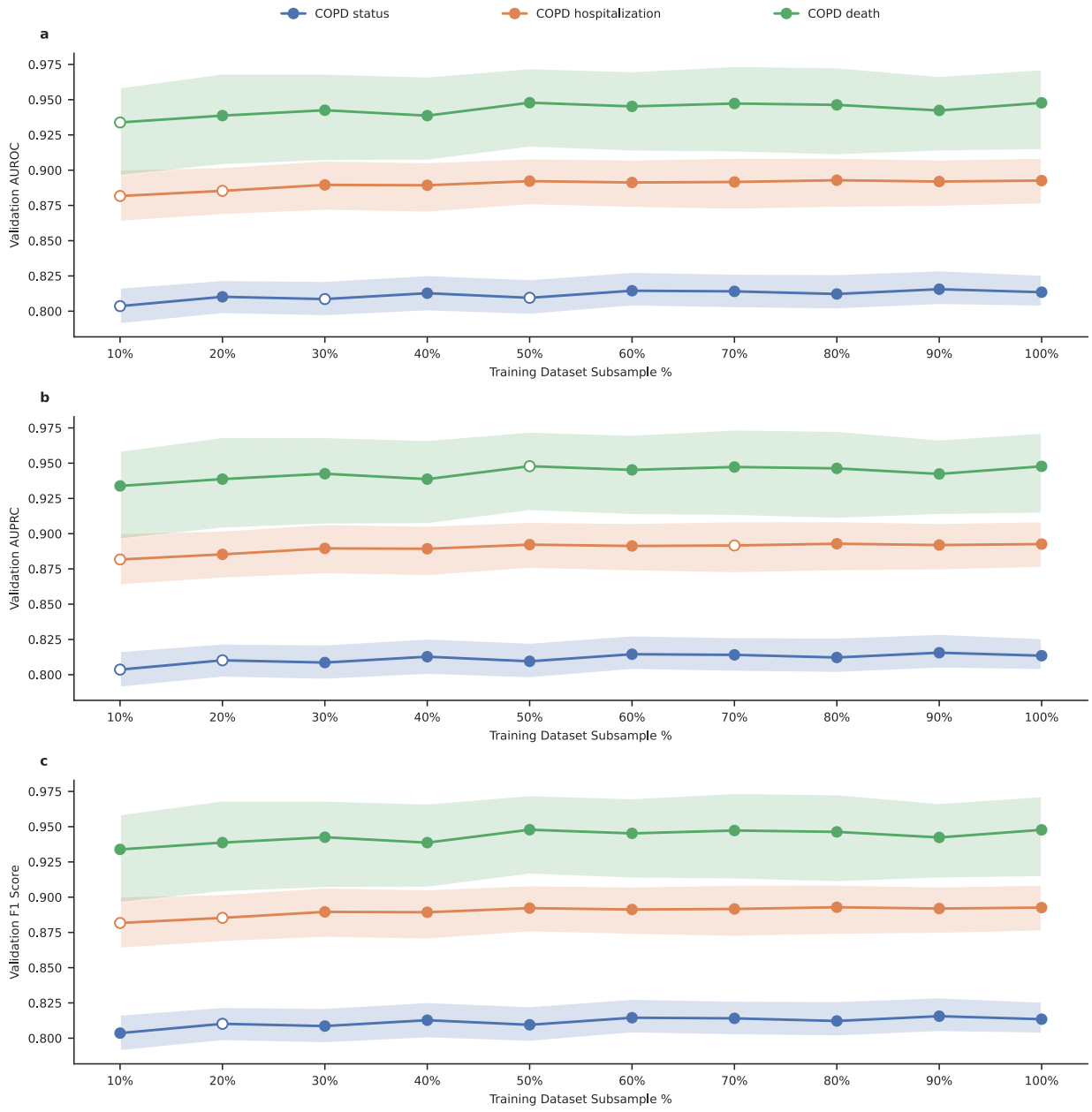
Supplementary Figures



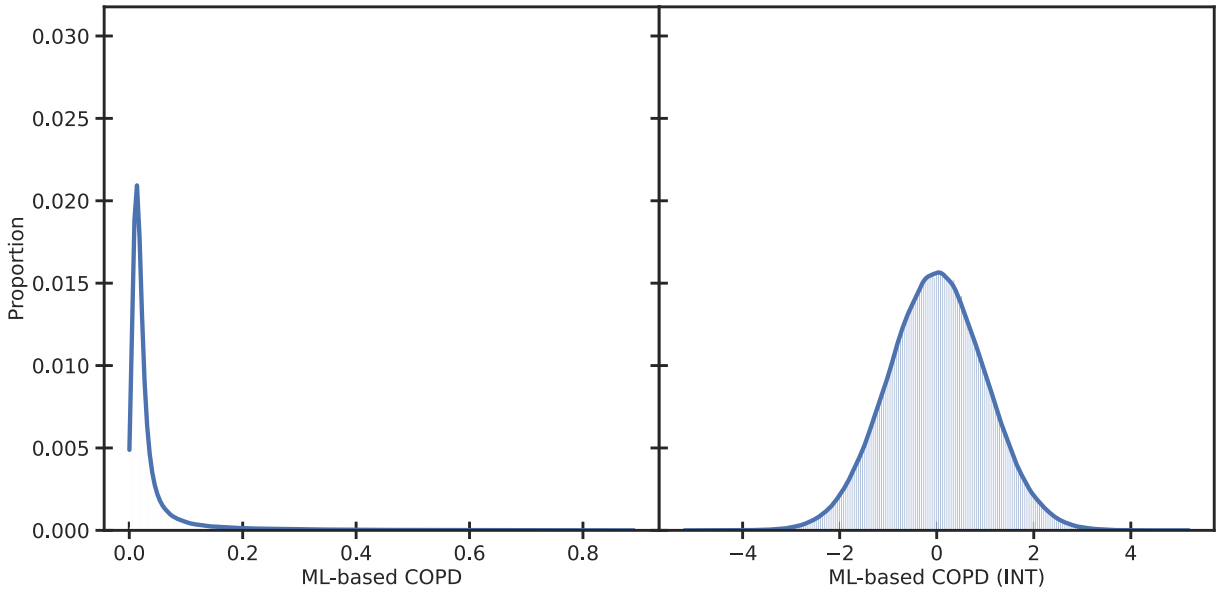
Supplementary Fig. 1: An overview of the one-dimension ResNet18-D residual and downsample layers. a) A basic residual layer [S11]. b) A downsampling residual layer following ResNet18-D architecture modifications [S12].



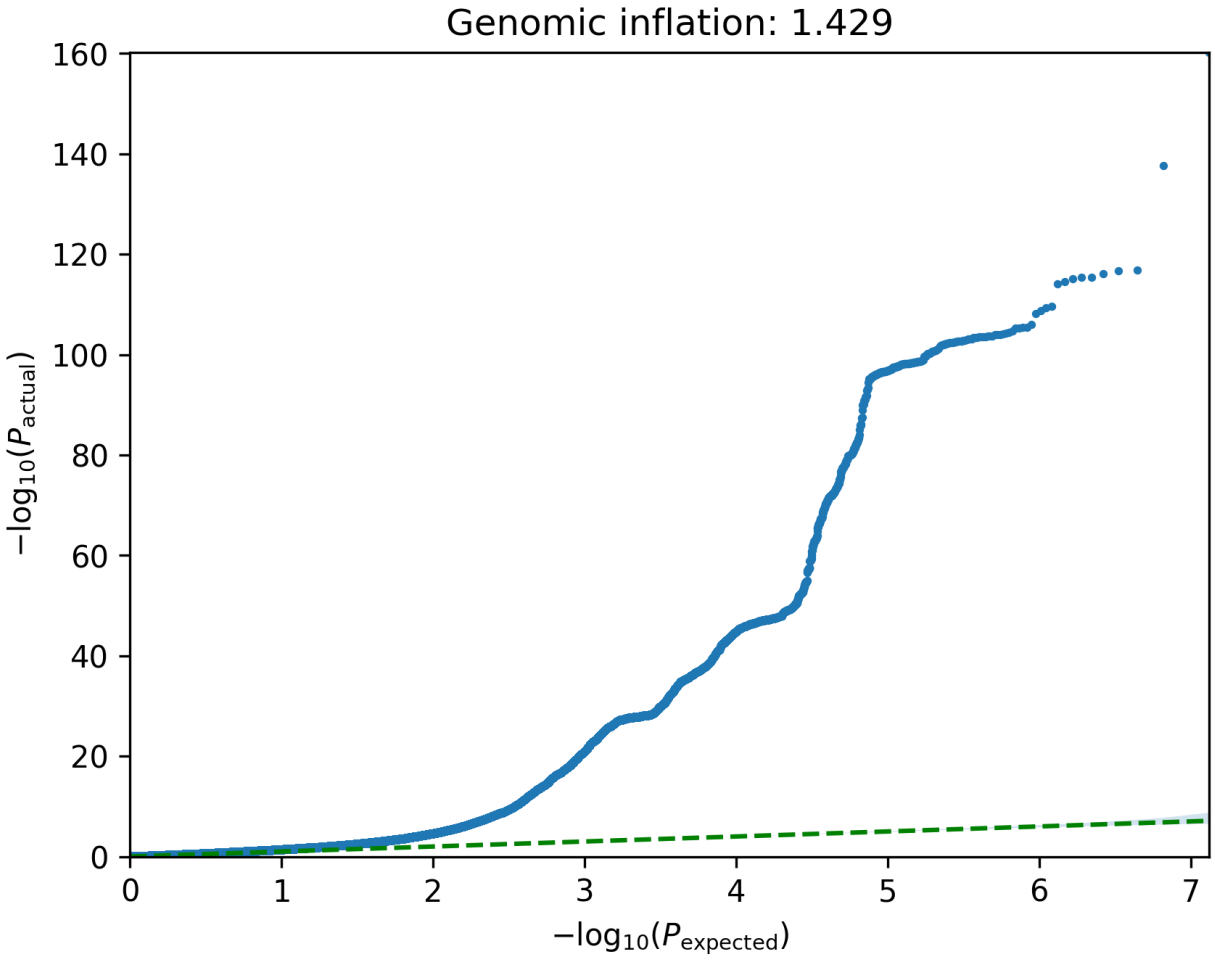
Supplementary Fig. 2: An overview of the cross-fold training and application process.



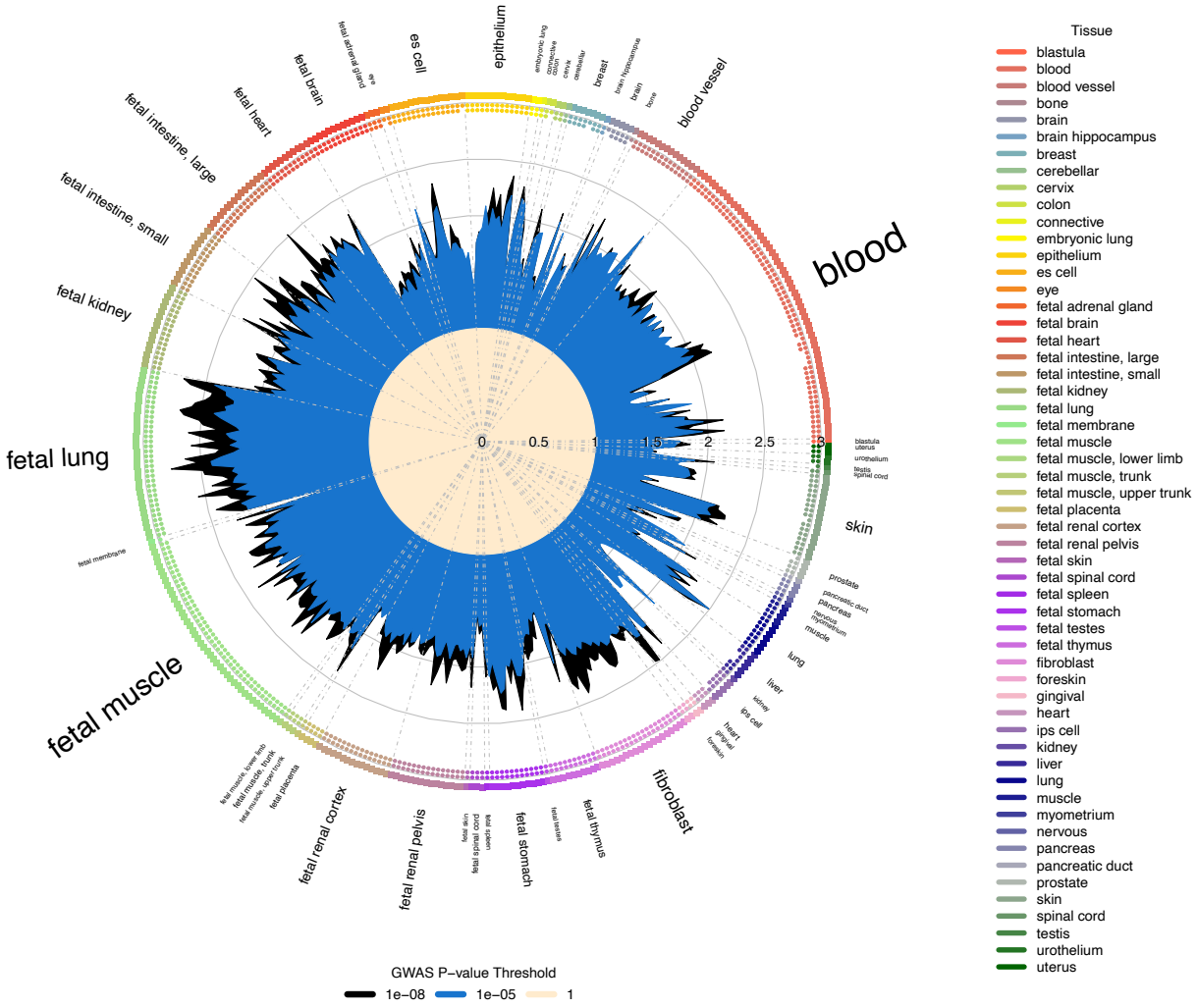
Supplementary Fig. 3: An ablation study assessing the impact of training dataset size on ResNet18 model performance. Ten separate ResNet18 models were trained on subsampled versions of the full training dataset (n=259,748) and then evaluated across the COPD status, hospitalization, and death tasks using the full validation dataset. Each subsampled dataset is a valid subset of all larger datasets. For example, all samples in the 40% dataset are also contained in the 50-100% datasets. Each point in the figure denotes the average metric over 100 bootstrapping samples while error bars denote the associated 95% confidence intervals. Unfilled points represent metrics with statistically significant differences compared to the corresponding 100% model under paired boostrapping. a) A comparison of validation AUROC across dataset sizes. b) A comparison of validation AUPRC across dataset sizes. c) A comparison of validation F1 scores across dataset sizes.



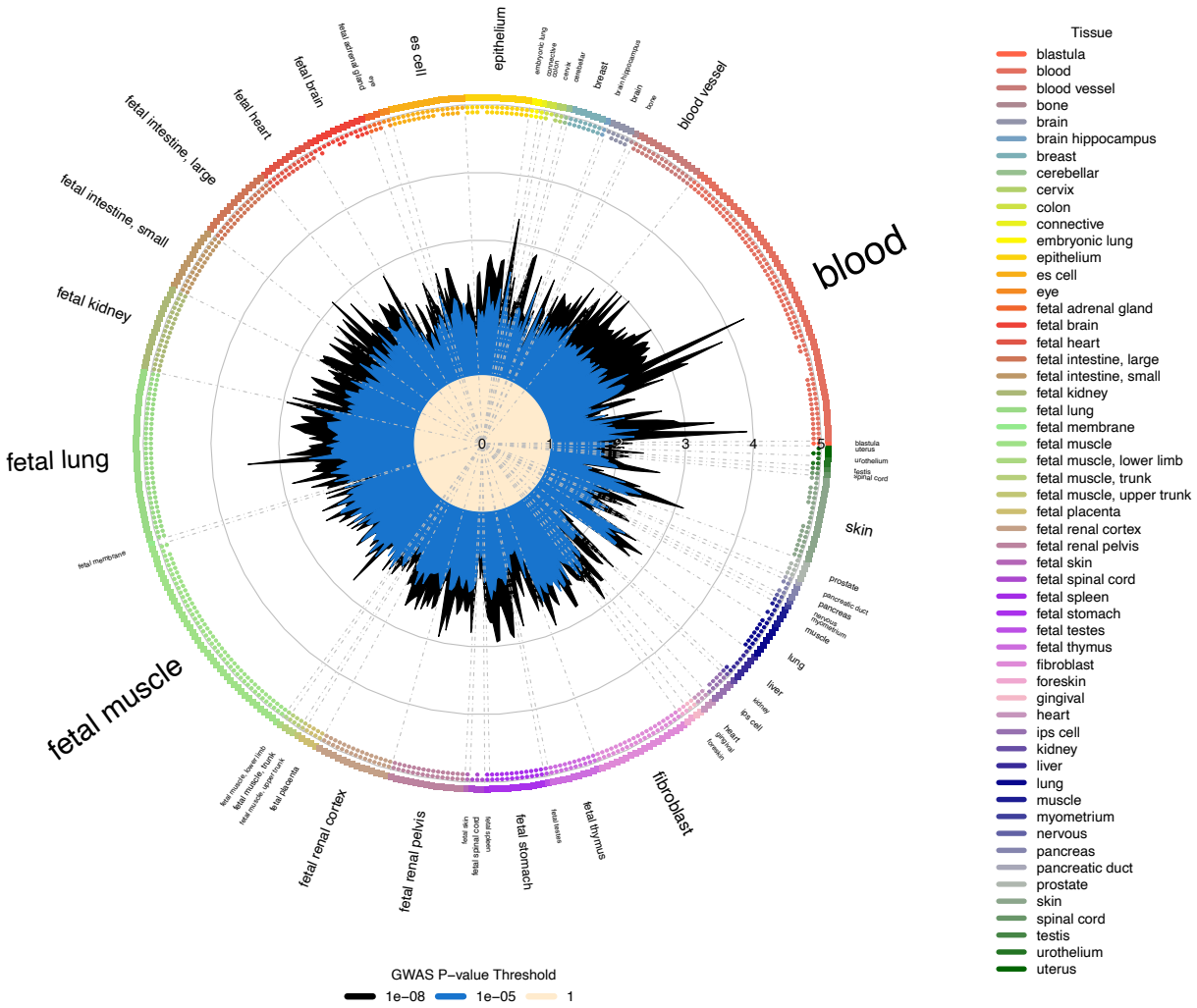
Supplementary Fig. 4: Histograms comparing our ML-based COPD liability score before and after an inverse normal transformation.



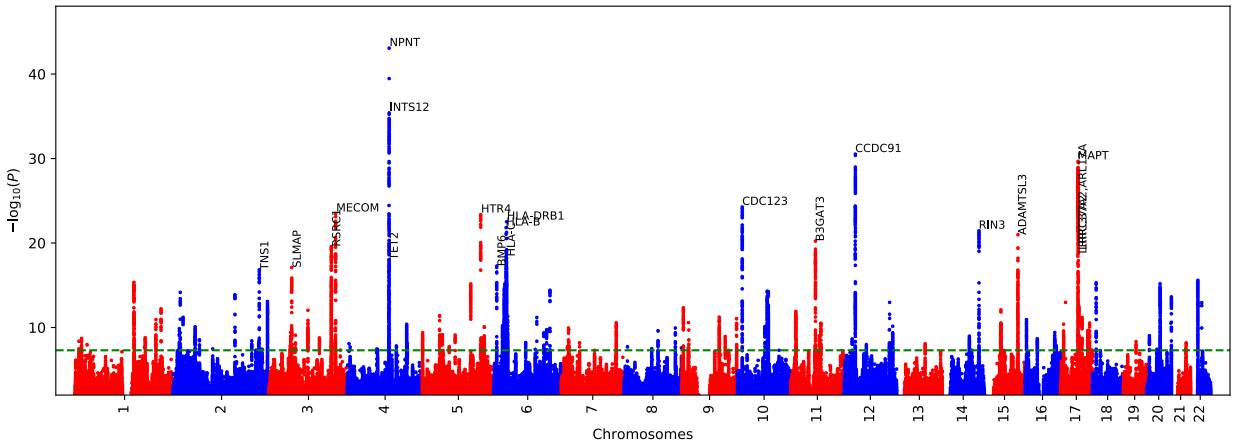
Supplementary Fig. 5: **QQplot of ML-based COPD GWAS.** The error band indicates the 95% confidence interval. We obtained P_{actual} from BOLT-LMM using a two-sided test.



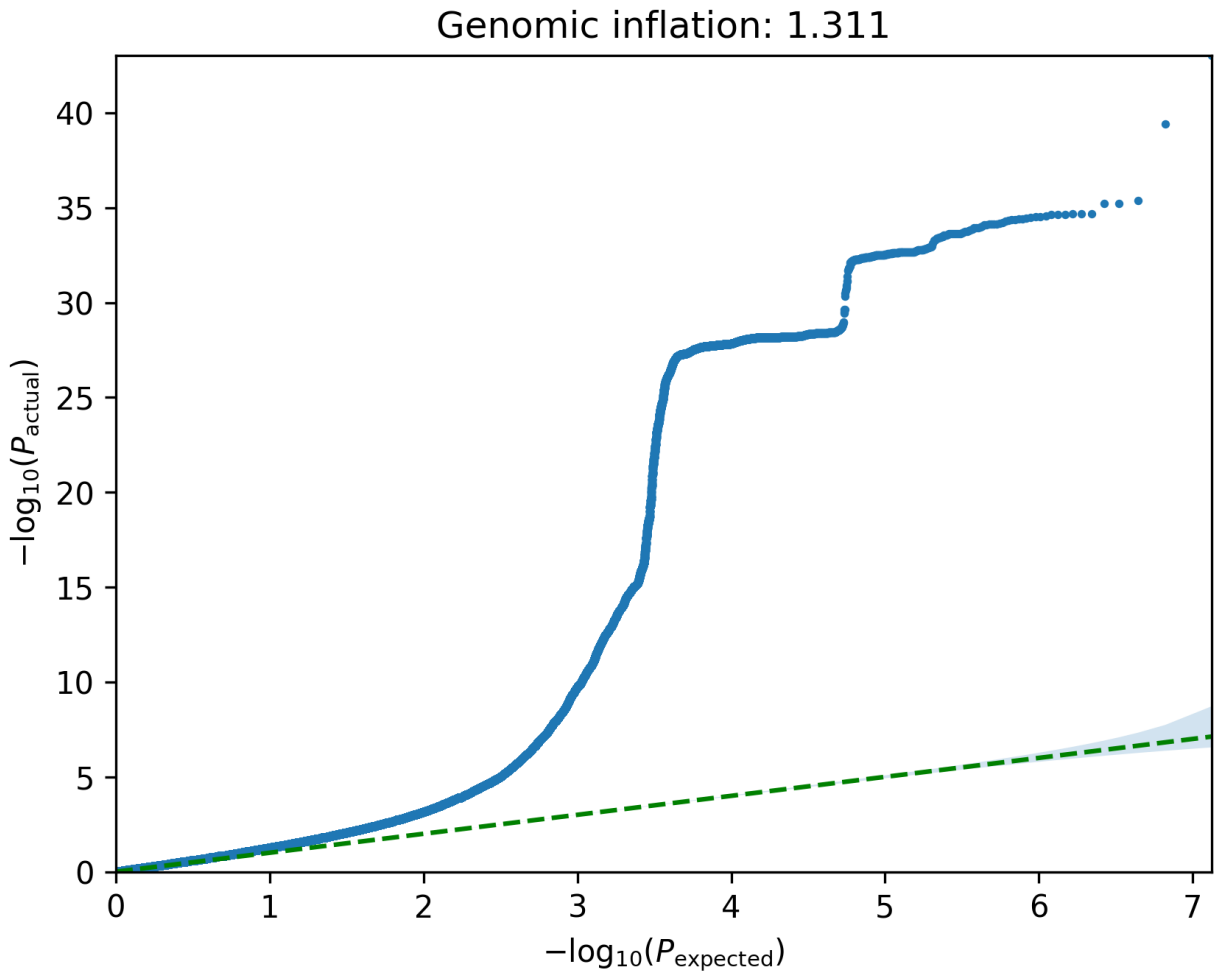
Supplementary Fig. 6: **Enrichment overlap of ML-based COPD GWAS with DNase I hotspots computed using GARFIELD.** Radial plot illustrates the enrichment (OR) in each cell type for different GWAS p-value thresholds ($P < 10^{-8}$ and 10^{-5}). In addition, the small dots on the outer side of the plot indicates enrichment significant level computed by GARFIELD for different significant level of 10^{-5}, 10^{-6}, 10^{-7}, and 10^{-8} in direction of outside to insider of plot. GWAS p-values were obtained from BOLT-LMM using a two-sided test.



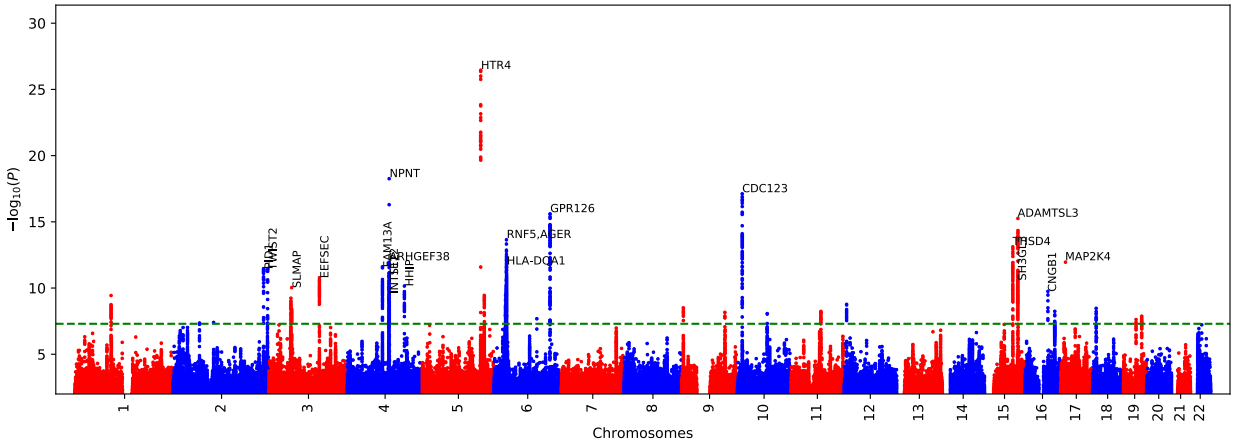
Supplementary Fig. 7: **Enrichment overlap of ML-based COPD conditional on FEV_1/FVC GWAS with DNase I hotspots computed using GARFIELD.** We observed the strongest enrichment in blood, fetal lung, and embryonic lung. GWAS p-values were obtained from BOLT-LMM using a two-sided test.



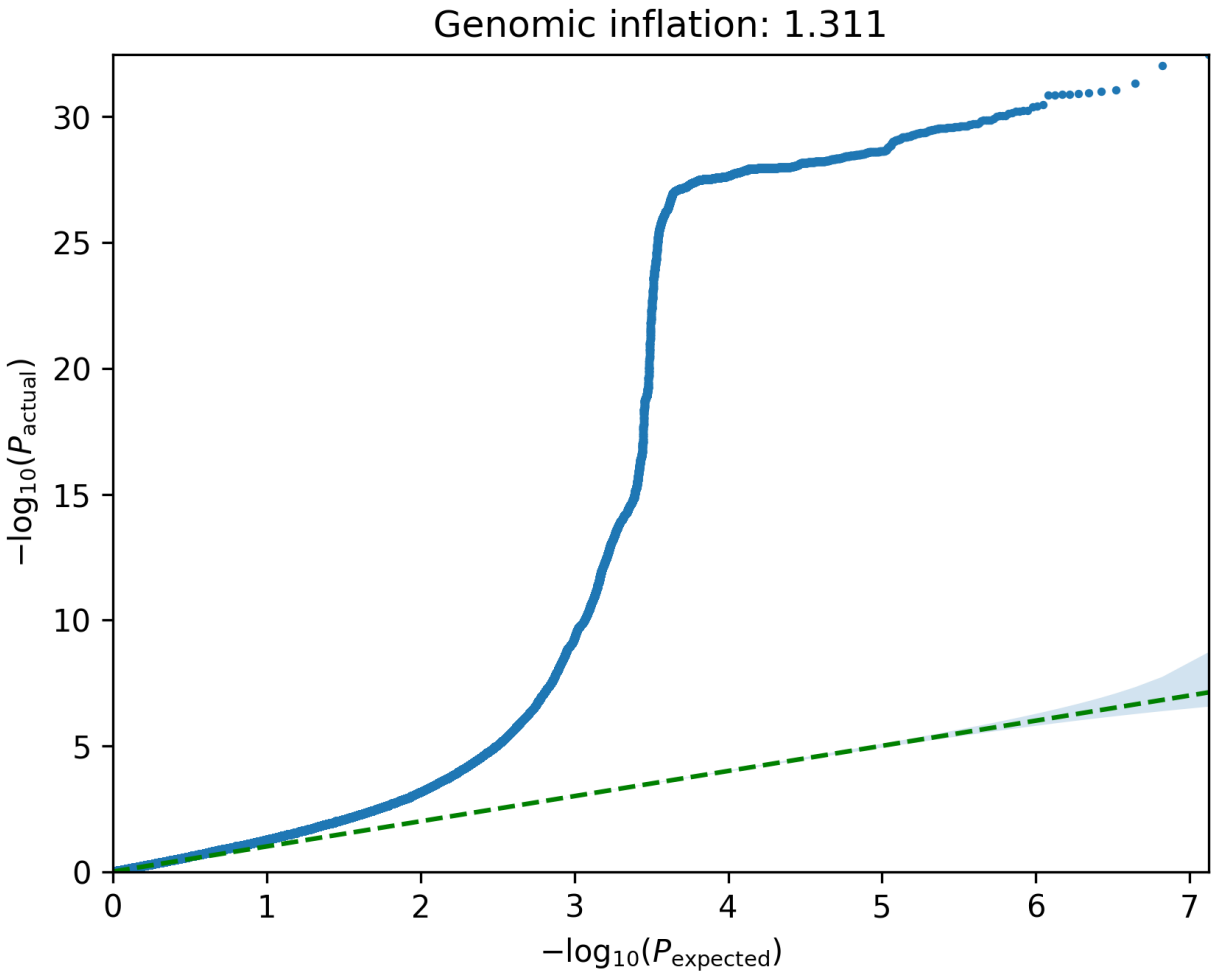
Supplementary Fig. 8: **ML-based COPD GWAS Manhattan plot conditional on FEV$_1$/FVC.** We obtained p-values from BOLT-LMM using a two-sided test.



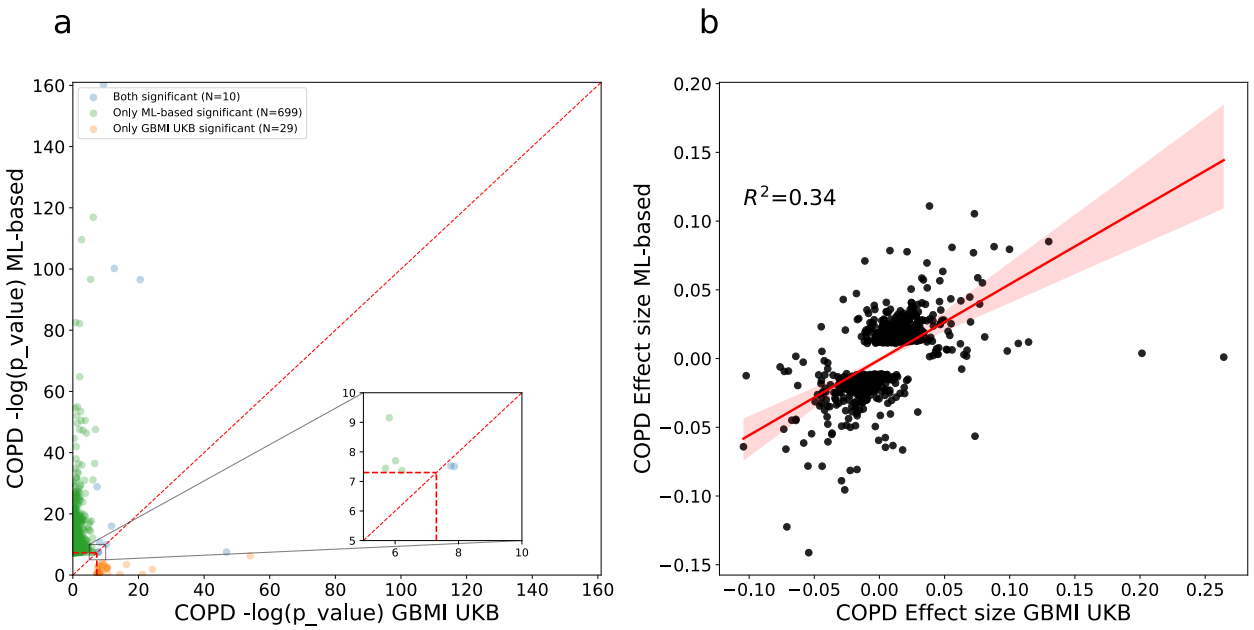
Supplementary Fig. 9: **QQplot of ML-based COPD GWAS conditional on FEV$_1$/FVC.** The error band indicates the 95% confidence interval. We obtained P_{actual} from BOLT-LMM using a two-sided test.



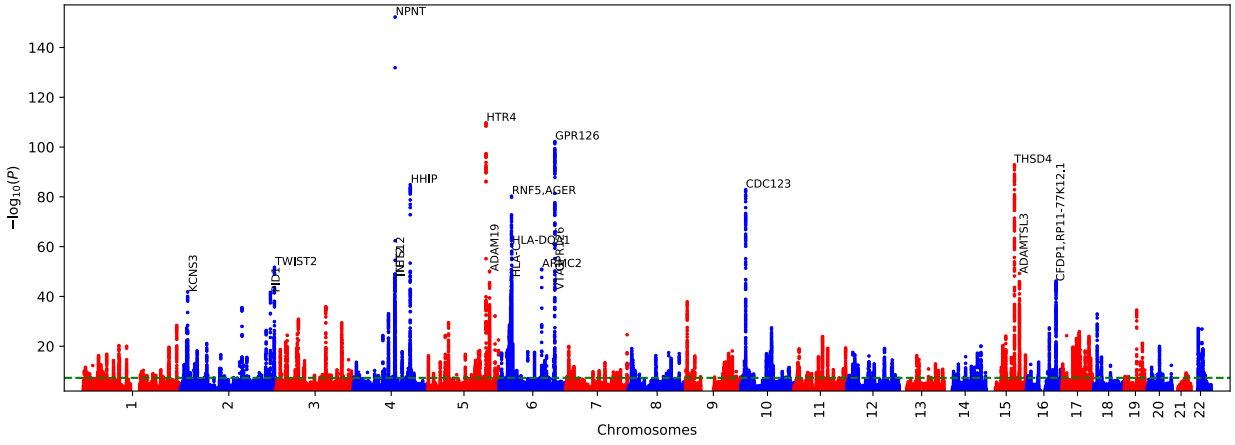
Supplementary Fig. 10: **ML-based COPD GWAS Manhattan plot conditional on $\mathbf{FEV_1/FVC}$, $\mathbf{FEV_1}$, FVC, and PEF.** We obtained p-values from BOLT-LMM using a two-sided test.



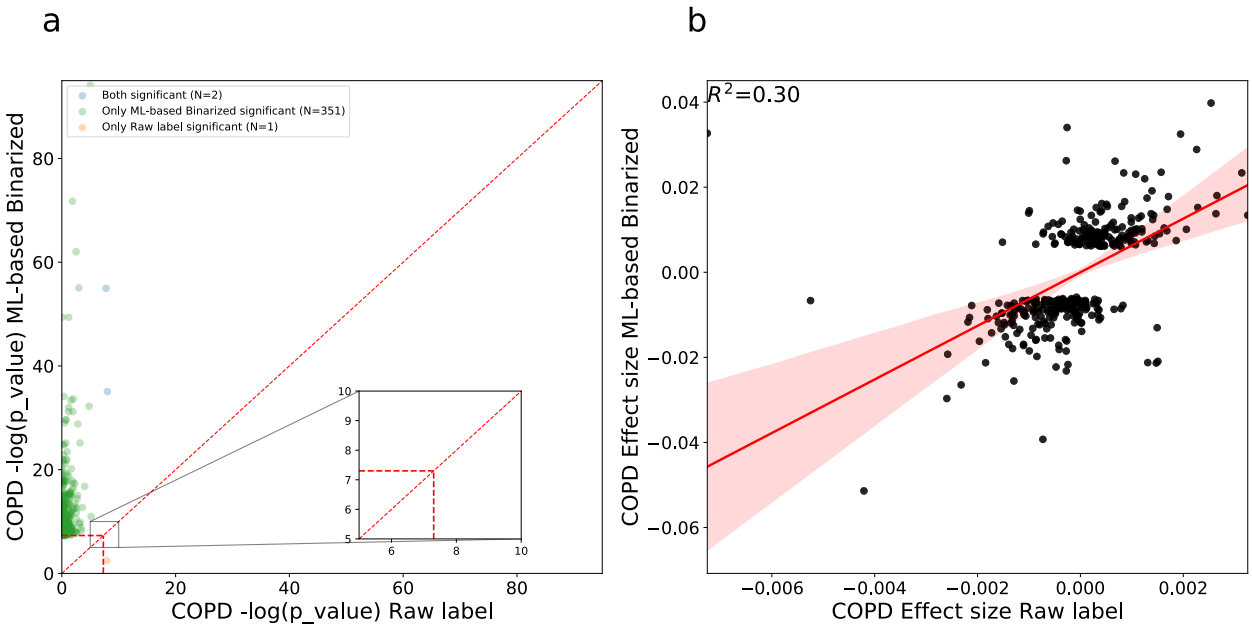
Supplementary Fig. 11: **QQplot of ML-based COPD GWAS via DeepNull.** The error band indicates the 95% confidence interval. We obtained P_{actual} from BOLT-LMM using a two-sided test.



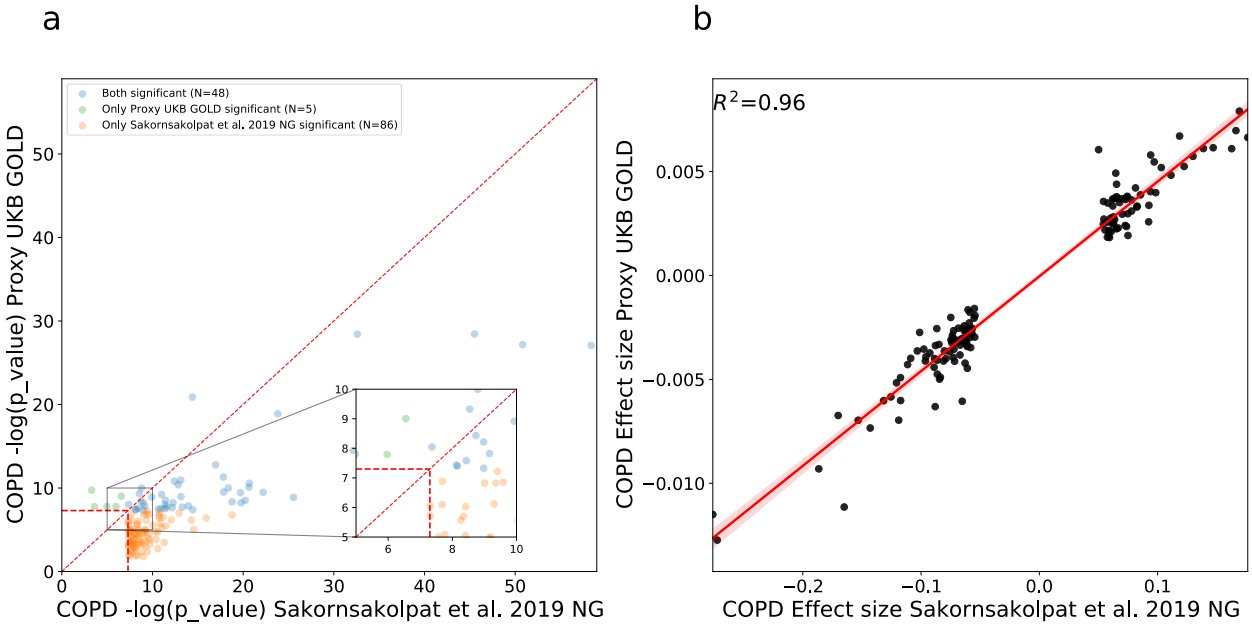
Supplementary Fig. 12: **Statistical power comparison of ML-based COPD with GBMI COPD excluding UKB.** a) The X-axis is the -log p-value of GBMI COPD excluding UKB. The Y-axis is the -log p-value of the ML-based COPD. Both p-values are computed using two-sided tests. The vertical and horizontal red line indicates the genome-wide significance level. The diagonal red line indicates the y=x. The orange dots indicate variants-in-hits that are significant for GBMI COPD excluding UKB but not significant for our ML-based COPD and green dots indicate variants-in-hits that are significant for our ML-based COPD but not significant for the GBMI COPD excluding UKB GWAS. b) Effect size correlation of ML-based COPD and GBMI COPD excluding UKB. The X-axis is the effect size of GBMI COPD excluding UKB for all GWS variants-in-hits and Y-axis is the effect size of our ML-based COPD. Light red band is the 95% confidence interval (e.g., band) of effect size correlation.



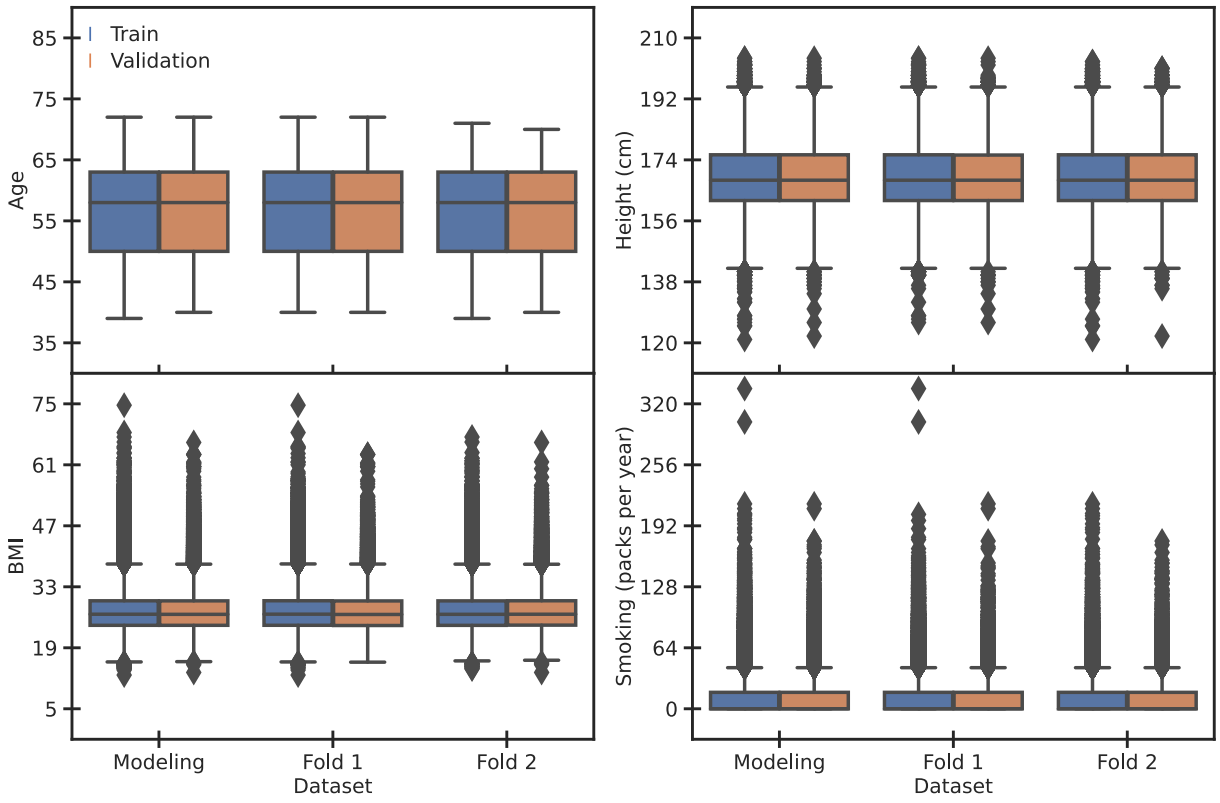
Supplementary Fig. 13: **ML-based COPD without MRB COPD case Manhattan plot.** We obtained p-values from BOLT-LMM using a two-sided test.



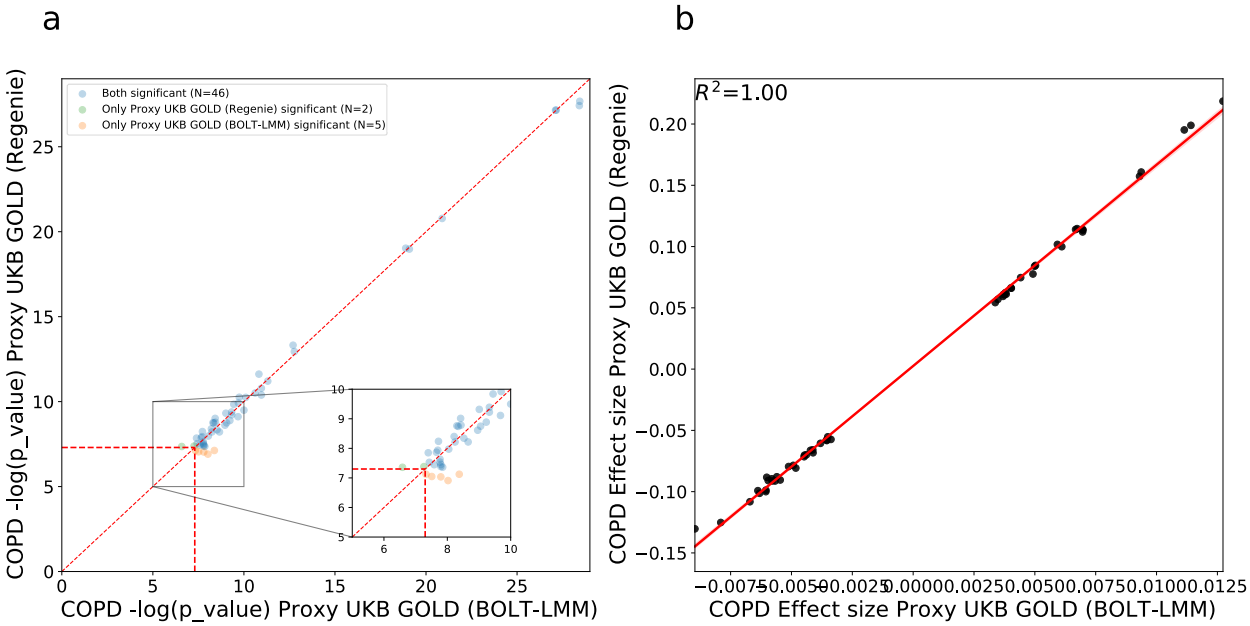
Supplementary Fig. 14: **Statistical power comparison of binarized ML-based COPD with medical-record-based COPD labels.** a) The X-axis is the -log p-value of medical-record-based COPD. The Y-axis is -log p-value of the binarized ML-based COPD. Both p-values are computed using two-sided tests. The vertical and horizontal red line indicates the genome-wide significance level. The diagonal red line indicates the y=x. The orange dots indicate variants-in-hits that are significant for medical-record-based COPD but not significant for our binarized ML-based COPD and green dots indicate variants-in-hits that are significant for our binarized ML-based COPD but not significant for the medical-record-based COPD label GWAS. b) Effect size correlation of binarized ML-based COPD and medical-record-based COPD GWAS. The X-axis is the effect size of medical-record-based label COPD for all GWS variants-in-hits and Y-axis is the effect size of our binarized ML-based COPD. Light red band is the 95% confidence interval (e.g., band) of effect size correlation.



Supplementary Fig. 15: **Statistical power comparison of proxy-GOLD with Sakornsakolpat et al.** a) The X-axis is the -log p-value of Sakornsakolpat et al. [S1]. The Y-axis is the -log p-value of proxy-GOLD. Both p-values are computed using two-sided tests. The vertical and horizontal red line indicates the genome-wide significance level. The diagonal red line indicates the y=x. The orange dots indicate variants-in-hits that are significant for Sakornsakolpat but not significant for proxy-GOLD and green dots indicate variants-in-hits that are significant for proxy-GOLD but not significant for Sakornsakolpat GWAS. b) Effect size correlation of proxy-GOLD and Sakornsakolpat GWAS. The X-axis is the effect size of Sakornsakolpat for all GWS variants-in-hits and Y-axis is the effect size of proxy-GOLD. Light red band is the 95% confidence interval (e.g., band) of effect size correlation.



Supplementary Fig. 16: **Covariates distribution across datasets.** We plot the histogram values of 4 continuous covariates (Age, Height, BMI, and Smoking per year) for the modeling (training $n = 259,746$, validation $n = 65,281$), Fold 1 (training $n = 128,739$, validation $n = 32,310$), and Fold 2 (training $n = 129,691$, validation $n = 32,637$) datasets. The whiskers extend between the largest and smallest values within 1.5 times the interquartile range. All values outside the whiskers are marked by points.



Supplementary Fig. 17: **Statistical power comparison of proxy-GOLD label using BOLT-LMM vs Regenie.** a) The X-axis is the -log p-value of proxy-GOLD via BOLT-LMM. The Y-axis is -log p-value of proxy-GOLD via Regenie. Both p-values are computed using two-sided tests. The vertical and horizontal red line indicates the genome-wide significance level. The diagonal red line indicates the y=x. The orange dots indicate variants-in-hits that are significant for BOLT-LMM but not significant for Regenie and green dots indicate variants that are significant for Regenie but not significant for BOLT-LMM. b) Effect size correlation of proxy-GOLD GWAS via BOLT-LMM and Regenie. The X-axis is the effect size of BOLT-LMM for all GWS hits and Y-axis is the effect size of Regenie obtained from proxy-GOLD COPD. Light red band is the 95% confidence interval (e.g., band) of effect size correlation.

Supplementary Tables

Label	Definition	Usage
Self report	Code 6 in field 6152 (medical conditions in touch screen questionnaires) or codes 1112, 1113, or 1472 in field 20002 (medical conditions in verbal interview).	Definition of medical-record-based COPD labels below used in training.
Primary Care	ICD-10 codes J41, J42, J43, or J44 in field 42040 (GP clinical event records), after Read v2 and v3 codes in the records mapped into ICD-10 codes.	Definition of medical-record-based COPD labels below used in training.
Training Hospitalization	Includes primary or secondary causes of hospitalization. ICD-9 codes 491, 492, or 496 in field 41271 (Diagnoses - ICD9), or ICD-10 codes J41, J42, J43, or J44 in field 41270 (Diagnoses - ICD10).	Definition of medical-record-based COPD labels below used in training.
Medical-record-based	If a COPD case in at least one of "self report", "primary care", and "training hospitalization" COPD labels.	Training of ML models.
Evaluation medical-record-based	Logical OR of "self report", "primary care", and "training hospitalization" labels only when all three sources exist for an individual.	Evaluation of ML models and GWAS hits (Results). Having all three sources increases the likelihood of a correct COPD label which is preferred for evaluation.
Future hospitalization	Only includes cases with COPD as primary cause of hospitalization after the spirometry test date. ICD-10 codes J41, J42, J43, or J44 in field 41234 (records in HES inpatient diagnoses dataset) after converting ICD-9 codes to ICD-10 codes.	Evaluation of ML models (Results).
Death	ICD-10 J41, J42, J43, or J44 codes in field 40001 (primary cause of death).	Evaluation of ML models and GWAS hits (Results).
Hospitalization	Similar to "future hospitalization" but also includes cases before the spirometry test date.	Evaluation of GWAS hits (Results).
Proxy-GOLD	Mirrors moderate or worse GOLD grading for a single blow without bronchodilation: $FEV_1/FVC < 0.7$ and FEV_1%predicted $< 80\%$.	Evaluating the noisiness of training labels, "medical-record-based" COPD. Evaluation of binarized ML-based COPD liability (Results).

Supplementary Table 1: **Definition of the binary COPD labels in UKB.** These labels were used for both training and evaluating ML models as well as running label-based GWAS.

See the attached Excel table.

Supplementary Table 2: **Overview of various model architectures.**

Model	Medical-record-based AUC	Hospitalization AUC	Death AUC
GOLD 2-4	0.6892 (0.6788–0.6991)	0.7607 (0.7350–0.7802)	0.8374 (0.7708–0.8978)
FEV_1/FVC Ratio	0.7785 (0.7652–0.7917)	0.8362 (0.8191–0.8566)	0.8871 (0.8047–0.9398)
FEV_1%predicted	0.7793 (0.7669–0.7909)	0.8712 (0.8517–0.8878)	0.9166 (0.8571–0.9654)
Spiro-metric LogReg	0.7976 (0.7852–0.8096)	0.8748 (0.8545–0.8924)	0.9348 (0.8981–0.9611)
Spiro-metric MLP	0.8023 (0.7903–0.8142)	0.8776 (0.8575–0.8958)	0.9379 (0.9086–0.9628)
Flow-volume MLP	0.8031 (0.7885–0.8146)	0.8808 (0.8604–0.9006)	0.9453 (0.9139–0.9720)
Flow-volume ResNet18	0.8161 (0.8055–0.8295)	0.8949 (0.8779–0.9106)	0.9490 (0.9207–0.9713)

Model	Medical-record-based AUPRC	Hospitalization AUPRC	Death AUPRC
GOLD 2-4	0.1447 (0.1328–0.1562)	0.0359 (0.0305–0.0410)	0.0053 (0.0034–0.0075)
FEV_1/FVC Ratio	0.2199 (0.1979–0.2364)	0.0833 (0.0663–0.1043)	0.0219 (0.0111–0.0374)
FEV_1%predicted	0.2100 (0.1888–0.2309)	0.0940 (0.0740–0.1144)	0.0292 (0.0147–0.0512)
Spiro-metric LogReg	0.2546 (0.2305–0.2768)	0.1078 (0.0862–0.1345)	0.0288 (0.0157–0.0508)
Spiro-metric MLP	0.2859 (0.2622–0.3112)	0.1343 (0.1042–0.1662)	0.0383 (0.0164–0.0774)
Flow-volume MLP	0.3160 (0.2896–0.3389)	0.1639 (0.1298–0.1976)	0.0426 (0.0193–0.0820)
Flow-volume ResNet18	0.3282 (0.3004–0.3527)	0.1777 (0.1403–0.2100)	0.0586 (0.0260–0.1223)

Model	Medical-record-based F1	Hospitalization F1	Death F1
GOLD 2-4	-	-	-
FEV_1/FVC Ratio	0.3067 (0.2892–0.3224)	0.1918 (0.1649–0.2170)	0.0490 (0.0046–0.0900)
FEV_1%predicted	0.2984 (0.2778–0.3199)	0.1901 (0.1583–0.2201)	0.0825 (0.0222–0.1544)
Spiro-metric LogReg	0.3435 (0.3217–0.3599)	0.2100 (0.1796–0.2352)	0.0567 (0.0136–0.1225)
Spiro-metric MLP	0.3536 (0.3291–0.3739)	0.2263 (0.1920–0.2556)	0.1158 (0.0425–0.2046)
Flow-volume MLP	0.3633 (0.3393–0.3852)	0.2603 (0.2191–0.2967)	0.1174 (0.0553–0.2151)
Flow-volume ResNet18	0.3793 (0.3565–0.3993)	0.2709 (0.2369–0.3040)	0.1205 (0.0374–0.2253)

Supplementary Table 3: **Comparison of model AUC, AUPRC, and F1 scores across evaluation medical-record-based COPD disease status, future COPD-related hospitalization, and COPD-related death.** GOLD 2-4, FEV_1/FVC Ratio, and FEV_1%predicted denote risk models based on standard spirometry metrics. The spirometry metric logistic regression and MLP models were trained to predict COPD status from only FEV_1/FVC ratio, FEV_1, FVC, and PEF while the flow-volume MLP and ResNet18 models utilized the entire flow-volume curve. We performed hyperparameter sweeps for each class of deep learning model, selecting the model that minimized the binary cross entropy loss over the modeling validation set. 95% confidential intervals were generated using bootstrapping ($n = 100$ bootstrapping trials; Methods).

Model	Medical-record-based AUC	Hospitalization AUC	Death AUC
GOLD 2-4	0.1269 (0.1156–0.1371)	0.1342 (0.1185–0.1527)	0.1116 (0.0596–0.1577)
FEV_1/FVC Ratio	0.0376 (0.0293–0.0460)	0.0587 (0.0446–0.0733)	0.0619 (0.0174–0.1177)
FEV_1%predicted	0.0368 (0.0276–0.0460)	0.0238 (0.0128–0.0398)	0.0324 (-0.0007–0.0768)
Spiro-metric LogReg	0.0185 (0.0112–0.0245)	0.0202 (0.0124–0.0277)	0.0141 (0.0017–0.0285)
Spiro-metric MLP	0.0138 (0.0075–0.0179)	0.0173 (0.0086–0.0258)	0.0110 (0.0015–0.0205)
Flow-volume MLP	0.0130 (0.0073–0.0185)	0.0142 (0.0078–0.0232)	0.0036 (-0.0048–0.0117)

Model	Medical-record-based AUPRC	Hospitalization AUPRC	Death AUPRC
GOLD 2-4	0.1836 (0.1636–0.2048)	0.1418 (0.1077–0.1711)	0.0533 (0.0212–0.1155)
FEV_1/FVC Ratio	0.1084 (0.0912–0.1252)	0.0944 (0.0704–0.1206)	0.0367 (0.0122–0.0823)
FEV_1%predicted	0.1182 (0.0940–0.1380)	0.0837 (0.0569–0.1096)	0.0294 (0.0043–0.0744)
Spiro-metric LogReg	0.0737 (0.0588–0.0914)	0.0699 (0.0504–0.0912)	0.0298 (0.0076-0.0738)
Spiro-metric MLP	0.0424 (0.0304–0.0548)	0.0434 (0.0275–0.0589)	0.0203 (-0.0003–0.0537)
Flow-volume MLP	0.0122 (0.0031–0.0193)	0.0138 (0.0020–0.0275)	0.0160 (0.0004–0.0389)

Model	Medical-record-based F1	Hospitalization F1	Death F1
GOLD 2-4	-	-	-
FEV_1/FVC Ratio	0.0725 (0.0587–0.0923)	0.0791 (0.0596–0.1024)	0.0715 (0.0005–0.1630)
FEV_1%predicted	0.0809 (0.0611–0.1017)	0.0809 (0.0545–0.1120)	0.0380 (-0.0449–0.1233)
Spiro-metric LogReg	0.0358 (0.0222–0.0514)	0.0609 (0.0404–0.0843)	0.0637 (-0.0053–0.1603)
Spiro-metric MLP	0.0257 (0.0120–0.0394)	0.0446 (0.0246–0.0704)	0.0047 (-0.1067–0.1027)
Flow-volume MLP	0.0160 (0.0042–0.0293)	0.0107 (-0.0147–0.0375)	0.0031 (-0.0782–0.0605)

Supplementary Table 4: **Paired bootstrapping between the flow-volume ResNet18 and other candidate COPD models' AUC, AUPRC, and F1 scores across evaluation medical-record-based COPD disease status, future COPD-related hospitalization, and COPD-related death.** These results show the relative significance level of improvement of the flow-volume ResNet18 model over other candidate models from Supplementary Table **3** using paired bootstrapping ($n = 100$ bootstrapping trials; Methods).

Model	Spearman R	Pearson R	Pearson R Squared
FEV_1/FVC Ratio	0.1744 (0.1674–0.1815)	0.2187 (0.2072–0.2310)	0.0479 (0.0429–0.0533)
FEV_1%predicted	0.1767 (0.1685–0.1843)	0.1938 (0.1850–0.2047)	0.0376 (0.0342–0.0419)
Spiro-metric LogReg	0.1905 (0.1843–0.1983)	0.2758 (0.2554–0.3026)	0.0762 (0.0652–0.0915)
Spiro-metric MLP	0.1964 (0.1907–0.2033)	0.3158 (0.2960–0.3361)	0.0998 (0.0876–0.1130)
Flow-volume MLP	0.1988 (0.1928–0.2062)	0.3366 (0.3177–0.3593)	0.1134 (0.1009–0.1291)
Flow-volume ResNet18	0.2037 (0.1975–0.2114)	0.3498 (0.3318–0.3703)	0.1225 (0.1101–0.1371)

Supplementary Table 5: **Comparison of model COPD risk with an individual's number of exacerbatory events.**

Model	Spearman R	Pearson R	Pearson R Squared
FEV_1/FVC Ratio	0.0292 (0.0258–0.0335)	0.1311 (0.1212–0.1421)	0.0746 (0.0665–0.0853)
FEV_1%predicted	0.0269 (0.0222–0.0323)	0.1560 (0.1410–0.1745)	0.0849 (0.0742–0.0985)
Spiro-metric LogReg	0.0131 (0.0099–0.0156)	0.0740 (0.0626–0.0848)	0.0463 (0.0386–0.0533)
Spiro-metric MLP	0.0072 (0.0049–0.0093)	0.0340 (0.0254–0.0413)	0.0227 (0.0169–0.0284)
Flow-volume MLP	0.0049 (0.0028–0.0073)	0.0132 (0.0078–0.0193)	0.0091 (0.0053–0.0134)

Supplementary Table 6: **Comparison of model COPD risk with an individual's number of exacerbatory events.** These results are similar to Supplementary Table 5 while to compute the significant level improvement between two models we used the paired bootstrapping ($n = 100$ bootstrapping trials).

See the attached Excel table.

Supplementary Table 7: **Comparison of cross-fold models.**

Phenotype	Prevalence	#Hits	#Loci	SNP-heritability	Notes
ML-based COPD	NA	796	356	0.2028 (0.0104)	
Proxy-GOLD	7.22%	45	30	0.0414 (0.0039)	
COPD (Medical-record-based)	4.44%	2	2	0.0112 (0.0029)	
COPD (Sakornsakolpat et al. 2019)	13.86%	117	84	0.0686 (0.0058)	
FEV_1 UKB (Shrine et al. 2019)	NA	406	244	0.1799 (0.0073)	
FVC UKB (Shrine et al. 2019)	NA	342	228	0.1779 (0.0063)	
FEV_1/FVC UKB (Shrine et al. 2019)	NA	669	328	0.2035 (0.01)	
Binarized ML-based COPD Sakornsakolpat et al. 2019 prevalence	13.86%	142	89	0.0726 (0.0054)	We fixed the prevalence of ML-based COPD to be the same as Sakornsakolpat et al.
Binarized ML-based COPD Proxy-GOLD prevalence	7.22%	56	37	0.0506 (0.0041)	We fixed the prevalence of ML-based COPD to be the same as Proxy-GOLD

Supplementary Table 8: **Comparison of different model GWAS results and prevalence.**

See the attached Excel table.

Supplementary Table 9: **ML-based COPD GWS hits.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, GWAS p-value. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 10: **ML-based COPD GWS loci.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, GWAS p-value. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 11: **Cell-type specific chromatin marks identify disease-relevant cell-type for ML-based COPD**. We performed cell-type specific chromatin heritability enrichment analysis via S-LDSC. We obtained the cell-type specific chromatin from S-LDSC website on BROAD. The second column is the regression coefficient corresponding to the cell type specific annotation and third column is the corresponding standard mean error. The last column is the p-value from a one-sided test that the regression coefficient is greater than zero with no adjustment for multiple-hypothesis tests. It is recommended by the authors [S5] that significant level of regression coefficient (e.g., p-value column) should be used to identify disease-relevant cell-type.

See the attached Excel table.

Supplementary Table 12: **S-LDSC enrichments of ML-based COPD for cell-type specific chromatin marks.** In addition to detecting the COPD-relevant cell-type, we compute the enrichment of ML-based COPD in cell-type specific chromatin marks. `Prop._SNPs` is the fraction of SNPs that is covered by each cell-type specific annotation. `Prop._h2` is the fraction of ML-based COPD that is captured by each cell-type specific annotation. `Enrichment` is defined as the $\frac{\texttt{Prop.\_h2}}{\texttt{Prop.\_SNPs}}$. `Enrichment_std_error` is the standard error of enrichment computed by S-LDSC via block-jackknife.

See the attached Excel table.

Supplementary Table 13: **Cell-type specific chromatin marks identify disease-relevant cell-type for ML-based COPD conditional on $\mathbf{FEV_1/FVC}$** . We performed cell-type specific chromatin heritability enrichment analysis via S-LDSC.

Term	Description	Region P	Gene P	Num regions
GO:0072359	circulatory system development	4.49e-05	2.62e-12	78
GO:0007507	heart development	1.37e-04	1.61e-10	56
MP:0005385	cardiovascular system phenotype	1.49e-04	1.35e-09	137
HP:0001626	Abnormality of the cardiovascular system	7.61e-04	7.53e-04	71
MP:0002127	abnormal cardiovascular system morphology	8.33e-04	2.82e-08	111
MP:0001544	abnormal cardiovascular system physiology	1.29e-03	6.25e-09	97
MP:0001175	abnormal lung morphology	5.76e-03	1.67e-05	47
GO:0003007	heart morphogenesis	5.80e-03	4.21e-07	34
MP:0001176	abnormal lung development	2.34e-02	1.33e-03	21
MP:0003115	abnormal respiratory system development	2.44e-02	3.32e-04	23
HP:0030680	Abnormality of cardiovascular system morphology	3.33e-02	4.65e-03	52
MP:0002132	abnormal respiratory system morphology	3.90e-02	1.77e-05	60
GO:0003205	cardiac chamber development	4.71e-02	1.09e-06	27
GO:0003279	cardiac septum development	4.78e-02	8.32e-07	22

Supplementary Table 14: **Cardiovascular and respiratory term enrichments of the ML-based COPD loci.** Enrichments were computed using GREAT [S3] with default parameters. The 82 total terms significant at Bonferroni-corrected P-value 0.05 by both the two-sided region-based binomial and two-sided gene-based hypergeometric tests were filtered to those with description that matched the regular expression 'cardiac|cardio|cardial|heart|circulatory|respir|lung|pulmon'.

See the attached Excel table.

Supplementary Table 15: **PheWAS of ML-based COPD significant hits.** We extracted the statistical test for all phenotypes in Neal lab and FinnGen of all ML-based COPD significant hits (796 independent GWS hits). In the case of FinnGen, variants position used the GRCh38-hg38 build, we used USCS liftover to match the variants. VID is the variant ID created by `chr:bp:ref:alt` where `chr` is the chromosome, `bp` is the variant position, `ref` is the reference allele, and `alt` is the alternative allele. P is the p-value obtrained from Neal lab and FinnGen GWAS and FDR is the computed false discovery rate.

See the attached Excel table.

Supplementary Table 16: **Summary of ML-based COPD PheWAS based on phenotypes**. We sorted phenotypes in Neal lab and FinnGen based on number significant hits obtained from ML-based COPD. We used the FDR $< 5\%$ to consider significant.

See the attached Excel table.

Supplementary Table 17: **Novel ML-based COPD GWS hits.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, GWAS p-value. We obtained p-values from BOLT-LMM using a two-sided test. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 18: **ML-based COPD GWS hits compared to other GWAS.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P_ML-based COPD, ML-based COPD GWAS P-value; P_DeepNull, ML-based COPD GWAS P-value via DeepNull; P_FEV1/FVC, FEV1/FVC GWAS P-value; P_GOLD 2-4, GOLD 2-4 GWAS P-value; P_Sakornsakolpat 2019; COPD GWAS P-value obtained from Sakornsakolpat 2019 [S1]. We obtained p-values from BOLT-LMM using a two-sided test.GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 19: **Novel ML-based COPD GWS loci.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, GWAS p-value. We obtained p-values from BOLT-LMM using a two-sided test. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 20: **ML-based COPD GWS Loci compared to other GWAS.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P_ML-based COPD, ML-based COPD GWAS P-value; P_DeepNull, ML-based COPD GWAS P-value via DeepNull; P_FEV1/FVC, FEV1/FVC GWAS P-value; P_GOLD 2-4, GOLD 2-4 GWAS P-value; P_Sakornsakolpat 2019; COPD GWAS P-value obtained from Sakornsakolpat 2019 [S1].

See the attached Excel table.

Supplementary Table 21: **ML-based COPD GWS hits conditional on FEV_1/FVC.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, ML-based COPD GWAS P-value. We obtained p-values from BOLT-LMM using a two-sided test. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 22: **ML-based COPD GWS loci conditional on FEV_1/FVC, FEV_1, FVC, and PEF.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, ML-based COPD GWAS P-value. We obtained p-values from BOLT-LMM using a two-sided test. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 23: **ML-based COPD GWS hits conditional on FEV_1/FVC, FEV1, FVC, and PEF.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, ML-based COPD GWAS P-value. We obtained p-values from BOLT-LMM using a two-sided test. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 24: **ML-based COPD GWS loci conditional on FEV_1/FVC**. CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, ML-based COPD GWAS P-value. We obtained p-values from BOLT-LMM using a two-sided test. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 25: **ML-based COPD GWS hits via DeepNull.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, ML-based COPD GWAS P-value. We obtained p-values from BOLT-LMM using a two-sided test. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 26: **ML-based COPD GWS loci via DeepNull.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, ML-based COPD GWAS P-value. We obtained p-values from BOLT-LMM using a two-sided test. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 27: **ML-based COPD GWS DeepNull hits not significant in the FEV_1/FVC ratio GWAS.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, ML-based COPD GWAS P-value. We obtained p-values from BOLT-LMM using a two-sided test. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

See the attached Excel table.

Supplementary Table 28: **ML-based COPD GWS DeepNull loci not significant in the FEV_1/FVC ratio GWAS.** CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; SRC, imputed or genotyped variant; INFO, imputation INFO score (set to 1 for genotyped variants); P, ML-based COPD GWAS P-value. We obtained p-values from BOLT-LMM using a two-sided test. GENE_CONTEXT, genomic context of the variant. Notation for gene context:

- Overlapping gene(s):
  - `[A]`: variant overlaps gene `A`
  - `[A,B]`: variant overlaps genes `A` and `B`
- Downstream genes:
  - `[]A`: variant position is $0 < p \leq 10^3$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^3 < p \leq 10^4$ bp upstream of closest downstream gene `A`
  - `[]-A`: variant position is $10^4 < p \leq 10^5$ bp upstream of closest downstream gene `A`
  - `[]--A`: variant position is $10^5 < p \leq 10^6$ bp upstream of closest downstream gene `A`
  - `[]`: closest downstream gene is further than 10^6 bp
- Upstream genes: mirrors downstream gene notation, e.g., `B-[]` means variant position is $10^3 < p \leq 10^4$ bp downstream of closest gene `B`.

Phenotype 1	Phenotype 2	r_g
ML-based COPD	COPD (Sakornsakolpat et al. 2019)	0.8985 (0.0736)
ML-based COPD	COPD (GBMI excluding UKB)	0.4845 (0.0513)
ML-based COPD	COPD (Hobbs et al. 2017)	0.8226 (0.0823)
MRB COPD	COPD (Hobbs et al. 2017)	0.7487 (0.1716)
ML-based COPD (No MRB Cases)	COPD (Hobbs et al. 2017)	0.8267 (0.0782)
ML-based COPD (No MRB Cases)	COPD (GBMI excluding UKB)	0.4395 (0.0468)
Binarized ML-base COPD (Proxy-GOLD prevalence)	COPD (Sakornsakolpat et al. 2019)	0.943 (0.0854)
ML-based COPD	FEV_1 UKB (Shrine et al. 2019)	-0.7031 (0.0441)
ML-based COPD	FEV_1/FVC UKB (Shrine et al. 2019)	-0.7837 (0.0491)
ML-based COPD	FVC UKB (Shrine et al. 2019)	-0.3440 (0.0261)
ML-based COPD	PEF UKB (Shrine et al. 2019)	-0.5757 (0.0385)
ML-based COPD	Height UKB	0.0580 (0.0214)
ML-based COPD	Asthma UKB	0.3980 (0.037)
COPD (Sakornsakolpat et al. 2019)	Asthma UKB	0.4815 (0.0531)
COPD (Sakornsakolpat et al. 2019)	Height UKB	0.1289 (0.0329)

Supplementary Table 29: **COPD genetic correlation with existing GWAS.** The genetic correlation between two phenotypes is computed using S-LDSC.

See the attached Excel table.

Supplementary Table 30: **Comparison of p-value for ML-based COPD with Hobbs et al. 2017 COPD [S13] for 22 previously detected COPD loci in Hobbs et al. 2017 [S13].**VID is the variant ID created by `chr:bp:ref:alt` where `chr` is the chromosome, `bp` is the variant position, `ref` is the reference allele, and `alt` is the alternative allele. HOBBES-2017-P is the p-value obtained from Hobbs et al. 2017 [S13] full summary statistics, ML-BASED-COPD-P is the p-value of GWAS on ML-based COPD risk, ML-BASED-COPD-NO-CASE-P is the p-value of GWAS on ML-based COPD risk restricted to control samples. We obtained p-values from BOLT-LMM using a two-sided test.

See the attached Excel table.

Supplementary Table 31: **ML-based COPD novel loci replication in GBMI, ICGC, and SpiroMeta.** We analyzed three studies that do not include UK Biobank samples to quantify loci replication. These three datasets are GBMI (Global Biobank Meta-analysis Initiative) [S14], SpiroMeta [S15], and ICGC (International COPD Genetics Consortium) [S13]. CHR, chromosome; POS, base-pair variant position; EA, effect allele; NEA, non-effect allele; BETA, estimated effect size; SE, standard error; ICGC-EA, ICGC-BETA, ICGC-P, GMBI-P, SpiroMeta-FVC-EA, SpiroMeta-FVC-BETA, SpiroMeta-FVC-P, SpiroMeta-FEV1-EA, SpiroMeta-FEV1-BETA, SpiroMeta-FEV1-P, SpiroMeta-Ratio-EA, SpiroMeta-Ratio-BETA, SpiroMeta-Ratio-P, Strict Replication, and Supportive Replication. All obtained p-values from previous studies used a two-sided test. We defined *supportive* replication as consistent effect size direction across studies. The ICGC and GBMI GWAS are based on a COPD phenotype; thus, we expect their effect size signs to match our ML-based COPD. SpiroMeta phenotypes, on the other hand, capture lung function, so we expect their effect size signs to be the opposite of our ML-based COPD signs. We defined *strict* replication as consistent effect size direction in any study with Bonferroni correction of $P < 0.1$ (one-sided) for that study.

Ground Truth	PRS	AUROC	AUPRC	Top Decile Prevalence	R
Eval. MRB COPD	ResNet18	0.550 (0.541–0.560)	0.088 (0.084–0.092)	0.097 (0.089–0.104)	0.046 (0.037–0.056)
Eval. MRB COPD	MRB COPD	0.517 (0.509–0.528)	0.079 (0.076–0.082)	0.081 (0.076–0.086)	0.018 (0.010–0.027)
Eval. MRB COPD	Sakornsakolpat	0.538 (0.529–0.548)	0.085 (0.081–0.089)	0.095 (0.086–0.103)	0.036 (0.028–0.045)
Eval. MRB COPD	FEV_1/FVC	0.540 (0.531–0.551)	0.085 (0.082–0.090)	0.095 (0.088–0.104)	0.037 (0.028–0.047)
Hospitalization	ResNet18	0.564 (0.549–0.577)	0.023 (0.021–0.025)	0.025 (0.022–0.028)	0.029 (0.022–0.036)
Hospitalization	MRB COPD	0.514 (0.504–0.526)	0.018 (0.017–0.020)	0.020 (0.018–0.021)	0.006 (0.001–0.013)
Hospitalization	Sakornsakolpat	0.551 (0.537–0.565)	0.022 (0.020–0.024)	0.025 (0.022–0.028)	0.024 (0.018–0.031)
Hospitalization	FEV_1/FVC	0.560 (0.546–0.573)	0.023 (0.021–0.025)	0.026 (0.023–0.029)	0.028 (0.021–0.034)
Death	ResNet18	0.598 (0.557–0.632)	0.004 (0.003–0.005)	0.004 (0.003–0.005)	0.016 (0.009–0.023)
Death	MRB COPD	0.503 (0.473–0.537)	0.002 (0.002–0.003)	0.002 (0.002–0.003)	0.001 (-0.004–0.007)
Death	Sakornsakolpat	0.575 (0.533–0.606)	0.003 (0.002–0.004)	0.003 (0.002–0.005)	0.013 (0.006–0.019)
Death	FEV_1/FVC	0.599 (0.554–0.645)	0.004 (0.003–0.005)	0.004 (0.003–0.005)	0.016 (0.009–0.024)

Supplementary Table 32: **Comparison of PRSs in UKB.** The PRSs are defined based on the GWAS effect sizes of ML-based COPD, Medical-record-based (MRB) COPD, and Sakornsakolpat et al. [S1], and the FEV_1/FVC. The numbers in the parenthesis show the 95% confidence interval. R is the Pearson correlation between true phenotypic values and estimated PRS.

Ground Truth	PRS	AUROC	AUPRC	Top Decile Prevalence	R
COPD Status	ResNet18	0.615 (0.598–0.631)	0.632 (0.615–0.650)	0.682 (0.647–0.720)	0.205 (0.178–0.232)
COPD Status	MRB COPD	0.525 (0.511–0.538)	0.550 (0.536–0.565)	0.547 (0.517–0.577)	0.046 (0.022–0.066)
COPD Status	Sakornsakolpat	0.616 (0.599–0.630)	0.623 (0.602–0.640)	0.689 (0.655–0.721)	0.202 (0.173–0.227)
COPD Status	FEV_1/FVC	0.618 (0.602–0.632)	0.627 (0.608–0.644)	0.688 (0.644–0.723)	0.206 (0.182–0.230)
pctEmph_vida	ResNet18	NA	NA	NA	0.106 (0.079–0.127)
pctEmph_vida	MRB COPD	NA	NA	NA	0.043 (0.017–0.066)
pctEmph_vida	Sakornsakolpat	NA	NA	NA	0.140 (0.114–0.168)
pctEmph_vida	FEV_1/FVC	NA	NA	NA	0.138 (0.113–0.157)
Pi10_SRWA_vida	ResNet18	NA	NA	NA	0.051 (0.030–0.072)
Pi10_SRWA_vida	MRB COPD	NA	NA	NA	0.000 (-0.023–0.024)
Pi10_SRWA_vida	Sakornsakolpat	NA	NA	NA	0.021 (-0.005–0.042)
Pi10_SRWA_vida	FEV_1/FVC	NA	NA	NA	0.026 (-0.001–0.051)

Supplementary Table 33: **Comparison of PRSs in COPDGene.** The PRSs are defined based on the GWAS effect sizes of ML-based COPD, Medical-record-based (MRB) COPD, and Sakornsakolpat et al. [S1], and the ratio, FEV_1/FVC. The numbers in the parenthesis show the 95% confidence interval. The effected individuals are defined as the individuals with final GOLD stage 2, 3, and 4 post-QA. R is the Pearson correlation between true phenotypic values and estimated PRS.

Model	Medical-record-based AUC	Hospitalization AUC	Death AUC
FEV_1/FVC Ratio	0.6977 (0.6622–0.7286)	0.8225 (0.7542–0.8705)	0.9862 (0.9732–0.9953)
Flow-volume ResNet18	**0.7523 (0.7201–0.7890)**	**0.8705 (0.8262–0.9151)**	**0.9918 (0.9781–0.9992)**
Model	**Medical-record-based AUPRC**	**Hospitalization AUPRC**	**Death AUPRC**
FEV_1/FVC Ratio	0.0914 (0.0708–0.1145)	0.0383 (0.0230–0.0587)	0.0357 (0.0107–0.0888)
Flow-volume ResNet18	**0.1786 (0.1398–0.2190)**	**0.1031 (0.0616–0.1558)**	**0.2851 (0.0472–0.5777)**
Model	**Medical-record-based F1**	**Hospitalization F1**	**Death F1**
FEV_1/FVC Ratio	0.1866 (0.1491–0.2178)	0.0898 (0.0588–0.1228)	0.0507 (0.0077–0.1048)
Flow-volume ResNet18	**0.2400 (0.1969–0.2856)**	**0.1915 (0.1259–0.2532)**	0.1495 (0.0000–0.4369)

Supplementary Table 34: **Comparison of model AUC, AUPRC, and F1 scores across tasks in non-Europeans.** Performance metrics are calculated using non-European individuals with valid spirometry blows for medical-record-based COPD disease status ($n = 8,201$, prevalence $= 3.244\%$), future COPD-related hospitalization($n = 21,878$, prevalence $= 0.384\%$), and COPD-related death($n = 26,125$, prevalence$= 0.034\%$). FEV_1/FVC Ratio denotes a risk model based on standard spirometry metrics. The flow-volume ResNet18 model utilized the entire flow-volume curve. We performed hyperparameter sweeps for each class of deep learning model, selecting the model that minimized the binary cross entropy loss over the modeling validation set. 95% confidential intervals were generated using bootstrapping ($n = 100$ bootstrapping trials). Bold values denote that a model is statistically better than other models for the given task metric ($n = 100$ bootstrapping trials).

Codes	Description	MRB COPD
J41	Simple and mucopurulent chronic bronchitis	0.6%
J42	Unspecified chronic bronchitis	3.4%
J43	Emphysema	21.4%
J44	Other chronic obstructive pulmonary disease	92.7%

Supplementary Table 35: **COPD ICD10 codes.** MRB stands for medical-record-based COPD. Note each individual might have multiple codes, but a large fraction of affected individuals have a J44 or its equivalent in their medical records and do not have information about COPD subtypes.

Dataset	Split	n	MRB	Eval. MRB	GOLD	Hospitalization	Death
Modeling	Train	259746	0.0383	0.0473	0.0725	0.0075	0.0007
Modeling	Validation	65281	0.0388	0.0479	0.0720	0.0072	0.0007
Fold 1	Train	128739	0.0384	0.0486	0.0711	0.0076	0.0007
Fold 1	Validation	32310	0.0391	0.0477	0.0719	0.0073	0.0008
Fold 2	Train	129691	0.0379	0.0456	0.0734	0.0072	0.0007
Fold 2	Validation	32637	0.0381	0.0480	0.0714	0.0069	0.0006
PRS	Holdout	110739	0.0640	0.0748	-	0.0053	0.0021

Supplementary Table 36: **Prevalence of COPD cases across datasets.** MRB stands for medical-record-based. GOLD labels are omitted for the PRS holdout set since it contains individuals with invalid blows. n is the number of individuals in each dataset.

Dataset	Split	n	FEV_1	FVC	FEV_1/FVC	FEV_1%predicted
Modeling	Train	259746	2.84 ± 0.75	3.76 ± 0.95	0.75 ± 0.07	96.47 ± 16.96
Modeling	Validation	65281	2.84 ± 0.75	3.76 ± 0.94	0.75 ± 0.07	96.48 ± 16.90
Fold 1	Train	128739	2.84 ± 0.75	3.76 ± 0.95	0.75 ± 0.07	96.50 ± 16.84
Fold 1	Validation	32310	2.84 ± 0.75	3.76 ± 0.95	0.75 ± 0.07	96.53 ± 16.90
Fold 2	Train	129691	2.84 ± 0.75	3.76 ± 0.95	0.75 ± 0.07	96.47 ± 17.04
Fold 2	Validation	32637	2.84 ± 0.75	3.76 ± 0.94	0.75 ± 0.07	96.47 ± 16.86

Supplementary Table 37: **Mean spirometry metrics across datasets.** The PRS holdout set is omitted since it contains only individuals with invalid blows. n is the number of individuals in each dataset and FEV_1%predicted is the predicted FEV_1 using age, sex, and height as features.

Dataset	Split	n	Age	Sex	Height (cm)	BMI	Smoking (packs per year)	Smoker	Occasional Smoker
Modeling	Train	259746	56.60 ± 7.98	0.4616	168.80 ± 9.14	27.38 ± 4.68	9.69 ± 14.64	0.3262	0.1448
Modeling	Validation	65281	56.59 ± 7.95	0.4599	168.82 ± 9.11	27.35 ± 4.66	9.66 ± 14.73	0.3246	0.1464
Fold 1	Train	128739	56.61 ± 7.98	0.4622	168.82 ± 9.12	27.38 ± 4.66	9.66 ± 14.60	0.3267	0.1437
Fold 1	Validation	32310	56.54 ± 7.96	0.4572	168.80 ± 9.10	27.31 ± 4.65	9.73 ± 14.85	0.3239	0.1481
Fold 2	Train	129691	56.59 ± 7.97	0.4607	168.78 ± 9.15	27.37 ± 4.70	9.70 ± 14.65	0.3251	0.1461
Fold 2	Validation	32637	56.64 ± 7.94	0.4622	168.84 ± 9.11	27.40 ± 4.67	9.57 ± 14.58	0.3247	0.1452

Supplementary Table 38: **Covariates mean across datasets.**

See the attached Excel table.

Supplementary Table 39: **Overview of hyperparameters considered for model architectures.**

See the attached Excel table.

Supplementary Table 40: **Overview of the final hyperparameters used for each architecture.**

Pheno Name	UKB Data-Field
Age	21003
BMI	21001
Sex	31
Height	50
Current tobacco smoking	1239
Past tobacco smoking	1249
Smoking pack years	20161

Supplementary Table 41: **Set of non-PC covariates utilized in ML-based GWAS.** UKB Data-Field code for non-PC covariates.

Supplementary References

[S1] Phuwanat Sakornsakolpat, Dmitry Prokopenko, Maxime Lamontagne, Nicola F. Reeve, Anna L. Guyatt, Victoria E. Jackson, Nick Shrine, Dandi Qiao, Traci M. Bartz, Deog Kyeom Kim, Mi Kyeong Lee, Jeanne C. Latourelle, Xingnan Li, Jarrett D. Morrow, Ma'en Obeidat, Annah B. Wyss, Per Bakke, R. Graham Barr, Terri H. Beaty, Steven A. Belinsky, Guy G. Brusselle, James D. Crapo, Kim de Jong, Dawn L. DeMeo, Tasha E. Fingerlin, Sina A. Gharib, Amund Gulsvik, Ian P. Hall, John E. Hokanson, Woo Jin Kim, David A. Lomas, Stephanie J. London, Deborah A. Meyers, George T. O'Connor, Stephen I. Rennard, David A. Schwartz, Pawel Sliwinski, David Sparrow, David P. Strachan, Ruth Tal-Singer, Yohannes Tesfaigzi, Jørgen Vestbo, Judith M. Vonk, Jae-Joon Yim, Xiaobo Zhou, Yohan Bossé, Ani Manichaikul, Lies Lahousse, Edwin K. Silverman, H. Marike Boezen, Louise V. Wain, Martin D. Tobin, Brian D. Hobbs, and Michael H. Cho and. Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. *Nature Genetics*, 51(3): 494–505, February 2019. doi: 10.1038/s41588-018-0342-2. URL `https://doi.org/10.1038/s41588-018-0342-2`.

[S2] Valentina Iotchkova, Graham R. S. Ritchie, Matthias Geihs, Sandro Morganella, Josine L. Min, Klaudia Walter, Nicholas John Timpson, Ian Dunham, Ewan Birney, and Nicole Soranzo. GARFIELD classifies disease-relevant genomic features through integration of functional annotations with association signals. *Nature Genetics*, 51(2):343–353, January 2019. doi: 10.1038/s41588-018-0322-6. URL `https://doi.org/10.1038/s41588-018-0322-6`.

[S3] Cory Y McLean, Dave Bristor, Michael Hiller, Shoa L Clarke, Bruce T Schaar, Craig B Lowe, Aaron M Wenger, and Gill Bejerano. GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology*, 28(5):495–501, May 2010. doi: 10.1038/nbt.1630. URL `https://doi.org/10.1038/nbt.1630`.

[S4] Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Nick Patterson, Mark J Daly, Alkes L Price, and Benjamin M Neale. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, 47(3): 291–295, February 2015. doi: 10.1038/ng.3211. URL `https://doi.org/10.1038/ng.3211`.

[S5] Hilary K. Finucane, , Yakir A. Reshef, Verneri Anttila, Kamil Slowikowski, Alexander Gusev, Andrea Byrnes, Steven Gazal, Po-Ru Loh, Caleb Lareau, Noam Shoresh, Giulio Genovese, Arpiar Saunders, Evan Macosko, Samuela Pollack, John R. B. Perry, Jason D. Buenrostro, Bradley E. Bernstein, Soumya Raychaudhuri, Steven McCarroll, Benjamin M. Neale, and Alkes L. Price. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nature Genetics*, 50(4):621–629, April 2018. doi: 10.1038/s41588-018-0081-4. URL `https://doi.org/10.1038/s41588-018-0081-4`.

[S6] Kristin G. Ardlie, David S. Deluca, Ayellet V. Segrè, Timothy J. Sullivan, Taylor R. Young, Ellen T. Gelfand, Casandra A. Trowbridge, Julian B. Maller, Taru Tukiainen, Monkol Lek, Lucas D. Ward, Pouya Kheradpour, Benjamin Iriarte, Yan Meng, Cameron D. Palmer, Tõnu Esko, Wendy Winckler, Joel N. Hirschhorn, Manolis Kellis, Daniel G. MacArthur, Gad Getz, Andrey A. Shabalin, Gen Li, Yi-Hui Zhou, Andrew B. Nobel, Ivan Rusyn, Fred A. Wright, Tuuli Lappalainen, Pedro G. Ferreira, Halit Ongen, Manuel A. Rivas, Alexis Battle, Sara Mostafavi, Jean Monlong, Michael Sammeth, Marta Mele, Ferran Reverter, Jakob M. Goldmann, Daphne Koller, Roderic Guigó, Mark I. McCarthy, Emmanouil T. Dermitzakis, Eric R. Gamazon, Hae Kyung Im, Anuar Konkashbaev, Dan L. Nicolae, Nancy J. Cox, Timothée Flutre, Xiaoquan Wen, Matthew Stephens, Jonathan K. Pritchard, Zhidong Tu, Bin Zhang, Tao Huang, Quan Long, Luan Lin, Jialiang Yang, Jun Zhu, Jun Liu, Amanda Brown, Bernadette Mestichelli, Denee Tidwell, Edmund Lo, Mike Salvatore, Saboor Shad, Jeffrey A. Thomas, John T. Lonsdale, Michael T. Moser, Bryan M. Gillard, Ellen Karasik, Kimberly Ramsey, Christopher Choi, Barbara A. Foster, John Syron, Johnell Fleming, Harold Magazine, Rick Hasz, Gary D. Walters, Jason P. Bridge, Mark Miklos, Susan Sullivan, Laura K. Barker, Heather M. Traino, Maghboeba Mosavel, Laura A. Siminoff, Dana R. Valley, Daniel C. Rohrer, Scott D. Jewell, Philip A. Branton, Leslie H. Sobin, Mary Barcus, Liqun Qi, Jeffrey McLean, Pushpa Hariharan, Ki Sung Um, Shenpei Wu, David Tabor, Charles Shive, Anna M. Smith, Stephen A. Buia, Anita H. Undale, Karna L. Robinson, Nancy Roche, Kimberly M. Valentino, Angela Britton, Robin Burges, Debra Bradbury, Kenneth W. Hambright, John Seleski, Greg E. Korzeniewski, Kenyon Erickson, Yvonne Marcus, Jorge Tejada, Mehran Taherian, Chunrong Lu, Margaret Basile, Deborah C. Mash, Simona Volpi, Jeffery P. Struewing, Gary F. Temple,

Joy Boyer, Deborah Colantuoni, Roger Little, Susan Koester, Latarsha J. Carithers, Helen M. Moore, Ping Guan, Carolyn Compton, Sherilyn J. Sawyer, Joanne P. Demchok, Jimmie B. Vaught, Chana A. Rabiner, Nicole C. Lockhart, Kristin G. Ardlie, Gad Getz, Fred A. Wright, Manolis Kellis, Simona Volpi, and Emmanouil T. Dermitzakis. The genotype-tissue expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*, 348(6235):648–660, May 2015. doi: 10.1126/science.1262110. URL `https://doi.org/10.1126/science.1262110`.

[S7] Tune H. Pers, , Juha M. Karjalainen, Yingleong Chan, Harm-Jan Westra, Andrew R. Wood, Jian Yang, Julian C. Lui, Sailaja Vedantam, Stefan Gustafsson, Tonu Esko, Tim Frayling, Elizabeth K. Speliotes, Michael Boehnke, Soumya Raychaudhuri, Rudolf S. N. Fehrmann, Joel N. Hirschhorn, and Lude Franke. Biological interpretation of genome-wide association studies using predicted gene functions. *Nature Communications*, 6(1), January 2015. doi: 10.1038/ncomms6890. URL `https://doi.org/10.1038/ncomms6890`.

[S8] Rudolf S N Fehrmann, Juha M Karjalainen, Małgorzata Krajewska, Harm-Jan Westra, David Maloney, Anton Simeonov, Tune H Pers, Joel N Hirschhorn, Ritsert C Jansen, Erik A Schultes, Herman H H B M van Haagen, Elisabeth G E de Vries, Gerard J te Meerman, Cisca Wijmenga, Marcel A T M van Vugt, and Lude Franke. Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nature Genetics*, 47(2):115–125, January 2015. doi: 10.1038/ng.3173. URL `https://doi.org/10.1038/ng.3173`.

[S9] The ENCODE (ENCyclopedia of DNA elements) project. *Science*, 306(5696):636–640, October 2004. doi: 10.1126/science.1105136. URL `https://doi.org/10.1126/science.1105136`.

[S10] Anshul Kundaje, , Wouter Meuleman, Jason Ernst, Misha Bilenky, Angela Yen, Alireza Heravi-Moussavi, Pouya Kheradpour, Zhizhuo Zhang, Jianrong Wang, Michael J. Ziller, Viren Amin, John W. Whitaker, Matthew D. Schultz, Lucas D. Ward, Abhishek Sarkar, Gerald Quon, Richard S. Sandstrom, Matthew L. Eaton, Yi-Chieh Wu, Andreas R. Pfenning, Xinchen Wang, Melina Claussnitzer, Yaping Liu, Cristian Coarfa, R. Alan Harris, Noam Shoresh, Charles B. Epstein, Elizabeta Gjoneska, Danny Leung, Wei Xie, R. David Hawkins, Ryan Lister, Chibo Hong, Philippe Gascard, Andrew J. Mungall, Richard Moore, Eric Chuah, Angela Tam, Theresa K. Canfield, R. Scott Hansen, Rajinder Kaul, Peter J. Sabo, Mukul S.

Bansal, Annaick Carles, Jesse R. Dixon, Kai-How Farh, Soheil Feizi, Rosa Karlic, Ah-Ram Kim, Ashwinikumar Kulkarni, Daofeng Li, Rebecca Lowdon, GiNell Elliott, Tim R. Mercer, Shane J. Neph, Vitor Onuchic, Paz Polak, Nisha Rajagopal, Pradipta Ray, Richard C. Sallari, Kyle T. Siebenthall, Nicholas A. Sinnott-Armstrong, Michael Stevens, Robert E. Thurman, Jie Wu, Bo Zhang, Xin Zhou, Arthur E. Beaudet, Laurie A. Boyer, Philip L. De Jager, Peggy J. Farnham, Susan J. Fisher, David Haussler, Steven J. M. Jones, Wei Li, Marco A. Marra, Michael T. McManus, Shamil Sunyaev, James A. Thomson, Thea D. Tlsty, Li-Huei Tsai, Wei Wang, Robert A. Waterland, Michael Q. Zhang, Lisa H. Chadwick, Bradley E. Bernstein, Joseph F. Costello, Joseph R. Ecker, Martin Hirst, Alexander Meissner, Aleksandar Milosavljevic, Bing Ren, John A. Stamatoyannopoulos, Ting Wang, and Manolis Kellis. Integrative analysis of 111 reference human epigenomes. *Nature*, 518(7539):317–330, February 2015. doi: 10.1038/nature14248. URL `https://doi.org/10.1038/nature14248`.

[S11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[S12] Tong He, Zhi Zhang, Hang Zhang, Zhongyue Zhang, Junyuan Xie, and Mu Li. Bag of tricks for image classification with convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 558–567, 2019.

[S13] Brian D Hobbs, Kim de Jong, Maxime Lamontagne, Yohan Bossé, Nick Shrine, María Soler Artigas, Louise V Wain, Ian P Hall, Victoria E Jackson, Annah B Wyss, Stephanie J London, Kari E North, Nora Franceschini, David P Strachan, Terri H Beaty, John E Hokanson, James D Crapo, Peter J Castaldi, Robert P Chase, Traci M Bartz, Susan R Heckbert, Bruce M Psaty, Sina A Gharib, Pieter Zanen, Jan W Lammers, Matthijs Oudkerk, H J Groen, Nicholas Locantore, Ruth Tal-Singer, Stephen I Rennard, Jørgen Vestbo, Wim Timens, Peter D Paré, Jeanne C Latourelle, Josée Dupuis, George T O'Connor, Jemma B Wilk, Woo Jin Kim, Mi Kyeong Lee, Yeon-Mok Oh, Judith M Vonk, Harry J de Koning, Shuguang Leng, Steven A Belinsky, Yohannes Tesfaigzi, Ani Manichaikul, Xin-Qun Wang, Stephen S Rich, R Graham Barr, David Sparrow, Augusto A Litonjua, Per Bakke, Amund Gulsvik, Lies Lahousse, Guy G Brusselle, Bruno H Stricker, André G Uitterlinden, Elizabeth J Ampleford,

Eugene R Bleecker, Prescott G Woodruff, Deborah A Meyers, Dandi Qiao, David A Lomas, Jae-Joon Yim, Deog Kyeom Kim, Iwona Hawrylkiewicz, Pawel Sliwinski, Megan Hardin, Tasha E Fingerlin, David A Schwartz, Dirkje S Postma, William MacNee, Martin D Tobin, Edwin K Silverman, H Marike Boezen, and Michael H Cho. Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nature Genetics*, 49(3):426–432, February 2017. doi: 10.1038/ng.3752. URL `https://doi.org/10.1038/ng.3752`.

[S14] Wei Zhou, Masahiro Kanai, Kuan-Han H. Wu, Humaira Rasheed, Kristin Tsuo, Jibril B. Hirbo, Ying Wang, Arjun Bhattacharya, Huiling Zhao, Shinichi Namba, Ida Surakka, Brooke N. Wolford, Valeria Lo Faro, Esteban A. Lopera-Maya, Kristi Läll, Marie-Julie Favé, Juulia J. Partanen, Sinéad B. Chapman, Juha Karjalainen, Mitja Kurki, Mutaamba Maasha, Ben M. Brumpton, Sameer Chavan, Tzu-Ting Chen, Michelle Daya, Yi Ding, Yen-Chen A. Feng, Lindsay A. Guare, Christopher R. Gignoux, Sarah E. Graham, Whitney E. Hornsby, Nathan Ingold, Said I. Ismail, Ruth Johnson, Triin Laisk, Kuang Lin, Jun Lv, Iona Y. Millwood, Sonia Moreno-Grau, Kisung Nam, Priit Palta, Anita Pandit, Michael H. Preuss, Chadi Saad, Shefali Setia-Verma, Unnur Thorsteinsdottir, Jasmina Uzunovic, Anurag Verma, Matthew Zawistowski, Xue Zhong, Nahla Afifi, Kawthar M. Al-Dabhani, Asma Al Thani, Yuki Bradford, Archie Campbell, Kristy Crooks, Geertruida H. de Bock, Scott M. Damrauer, Nicholas J. Douville, Sarah Finer, Lars G. Fritsche, Eleni Fthenou, Gilberto Gonzalez-Arroyo, Christopher J. Griffiths, Yu Guo, Karen A. Hunt, Alexander Ioannidis, Nomdo M. Jansonius, Takahiro Konuma, Ming Ta Michael Lee, Arturo Lopez-Pineda, Yuta Matsuda, Riccardo E. Marioni, Babak Moatamed, Marco A. Nava-Aguilar, Kensuke Numakura, Snehal Patil, Nicholas Rafaels, Anne Richmond, Agustin Rojas-Muñoz, Jonathan A. Shortt, Peter Straub, Ran Tao, Brett Vanderwerff, Manvi Vernekar, Yogasudha Veturi, Kathleen C. Barnes, Marike Boezen, Zhengming Chen, Chia-Yen Chen, Judy Cho, George Davey Smith, Hilary K. Finucane, Lude Franke, Eric R. Gamazon, Andrea Ganna, Tom R. Gaunt, Tian Ge, Hailiang Huang, Jennifer Huffman, Nicholas Katsanis, Jukka T. Koskela, Clara Lajonchere, Matthew H. Law, Liming Li, Cecilia M. Lindgren, Ruth J.F. Loos, Stuart MacGregor, Koichi Matsuda, Catherine M. Olsen, David J. Porteous, Jordan A. Shavit, Harold Snieder, Tomohiro Takano, Richard C. Trembath,

Judith M. Vonk, David C. Whiteman, Stephen J. Wicks, Cisca Wijmenga, John Wright, Jie Zheng, Xiang Zhou, Philip Awadalla, Michael Boehnke, Carlos D. Bustamante, Nancy J. Cox, Segun Fatumo, Daniel H. Geschwind, Caroline Hayward, Kristian Hveem, Eimear E. Kenny, Seunggeun Lee, Yen-Feng Lin, Hamdi Mbarek, Reedik Mägi, Hilary C. Martin, Sarah E. Medland, Yukinori Okada, Aarno V. Palotie, Bogdan Pasaniuc, Daniel J. Rader, Marylyn D. Ritchie, Serena Sanna, Jordan W. Smoller, Kari Stefansson, David A. van Heel, Robin G. Walters, Sebastian Zöllner, Alicia R. Martin, Cristen J. Willer, Mark J. Daly, and Benjamin M. Neale. Global biobank meta-analysis initiative: Powering genetic discovery across human disease. *Cell Genomics*, 2(10):100192, October 2022. doi: 10.1016/j.xgen.2022.100192. URL `https://doi.org/10.1016/j.xgen.2022.100192`.

[S15] María Soler Artigas, , Louise V. Wain, Suzanne Miller, Abdul Kader Kheirallah, Jennifer E. Huffman, Ioanna Ntalla, Nick Shrine, Ma'en Obeidat, Holly Trochet, Wendy L. McArdle, Alexessander Couto Alves, Jennie Hui, Jing Hua Zhao, Peter K. Joshi, Alexander Teumer, Eva Albrecht, Medea Imboden, Rajesh Rawal, Lorna M. Lopez, Jonathan Marten, Stefan Enroth, Ida Surakka, Ozren Polasek, Leo-Pekka Lyytikäinen, Raquel Granell, Pirro G. Hysi, Claudia Flexeder, Anubha Mahajan, John Beilby, Yohan Bossé, Corry-Anke Brandsma, Harry Campbell, Christian Gieger, Sven Gläser, Juan R. González, Harald Grallert, Chris J. Hammond, Sarah E. Harris, Anna-Liisa Hartikainen, Markku Heliövaara, John Henderson, Lynne Hocking, Momoko Horikoshi, Nina Hutri-Kähönen, Erik Ingelsson, Åsa Johansson, John P. Kemp, Ivana Kolcic, Ashish Kumar, Lars Lind, Erik Melén, Arthur W. Musk, Pau Navarro, David C. Nickle, Sandosh Padmanabhan, Olli T. Raitakari, Janina S. Ried, Samuli Ripatti, Holger Schulz, Robert A. Scott, Don D. Sin, John M. Starr, Ana Viñuela, Henry Völzke, Sarah H. Wild, Alan F. Wright, Tatijana Zemunik, Deborah L. Jarvis, Tim D. Spector, David M. Evans, Terho Lehtimäki, Veronique Vitart, Mika Kähönen, Ulf Gyllensten, Igor Rudan, Ian J. Deary, Stefan Karrasch, Nicole M. Probst-Hensch, Joachim Heinrich, Beate Stubbe, James F. Wilson, Nicholas J. Wareham, Alan L. James, Andrew P. Morris, Marjo-Riitta Jarvelin, Caroline Hayward, Ian Sayers, David P. Strachan, Ian P. Hall, and Martin D. Tobin. Sixteen new lung function signals identified through 1000 genomes project reference panel imputation. *Nature Communications*, 6(1), December 2015. doi: 10.1038/ncomms9658.

URL https://doi.org/10.1038/ncomms9658.