Supplemental Text

S1: Clinical Examination

To better contextualize the questions chosen for the minimum question set, we provided the final panel to clinicians to judge its feasibility. Clinicians provided feasibility ratings for 61 PXS and five PSS items. Agreement was highest for the implementation fields most relevant to question-set reduction. Reviewers assigned the same feasibility tier for 59 of 66 items (89.4%), and 62 of 66 items (93.9%) were rated within one feasibility tier. Capture-mode agreement was similarly high, with exact agreement for 55 of 65 valid item pairs (84.6%). Agreement was also high for respondent/staff burden (52 of 65, 80.0%), reflecting a shared view that most candidate items could be collected with low incremental burden. Agreement was lower for pilot recommendation (34 of 65, 52.3%) and clinical actionability, indicating that these fields captured more subjective judgments about wording, setting specificity, and how results would be used in care. Overall, 33 of 66 items were recommended by both reviewers for inclusion as written, and 52 of 66 were recommended by both reviewers for inclusion either as written or with wording/workflow revision. No item was marked for deprioritization by either reviewer. Among the PXS items, 47 of 61 were recommended by both reviewers for inclusion as written or with revision, while 14 had at least one unsure or pilot-only recommendation. Items requiring adjudication were concentrated among constructs with ambiguous wording, derived variables, home dampness/water damage constructs, sleep-related constructs, and occupational or hobby exposures that may require examples or specialty-specific exposure modules.

Supplemental Figures

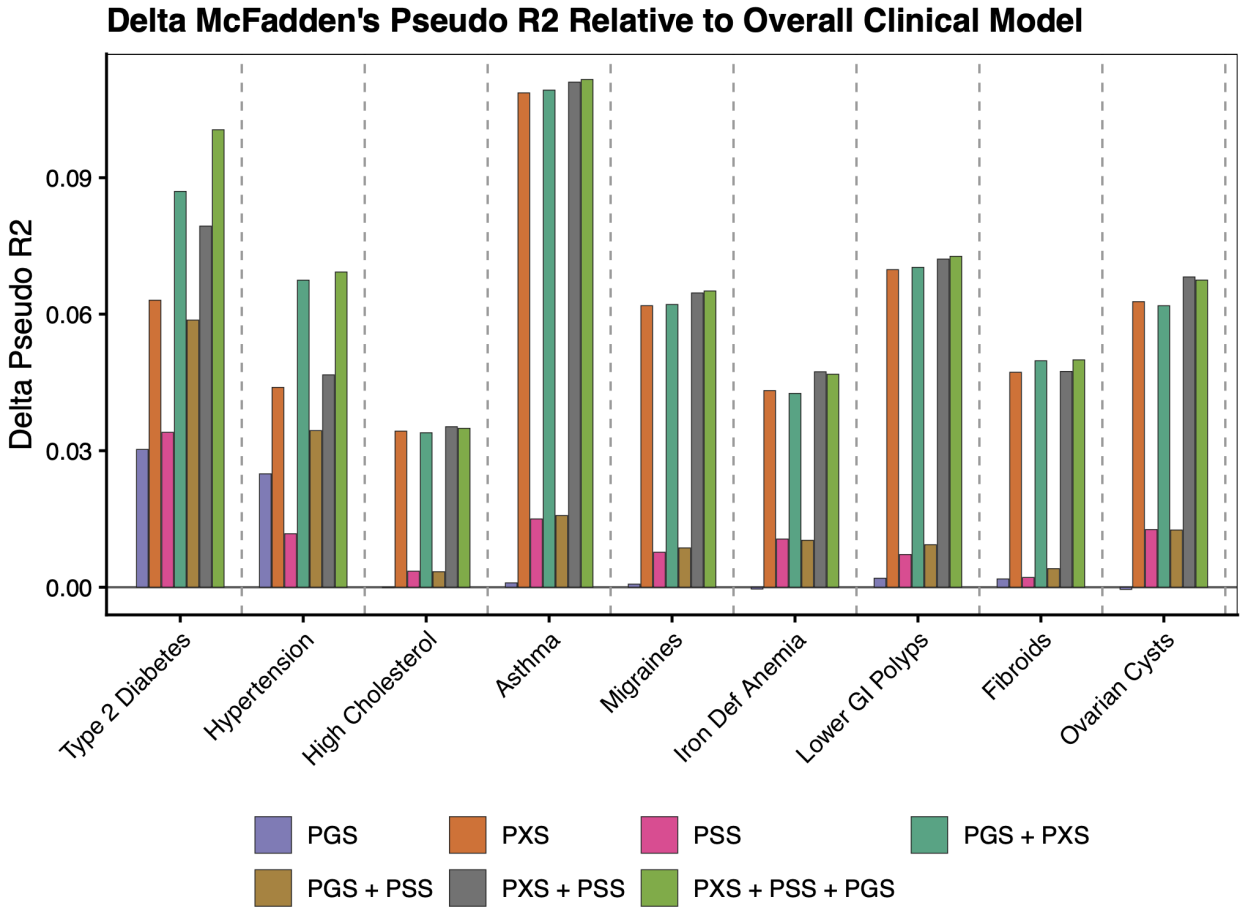

Figure S1: Disease-specific McFadden's Pseudo R^2. The X axis is the disease of interest, and the Y axis is ΔMcFadden's pseudo R^2 compared to age and sex alone.

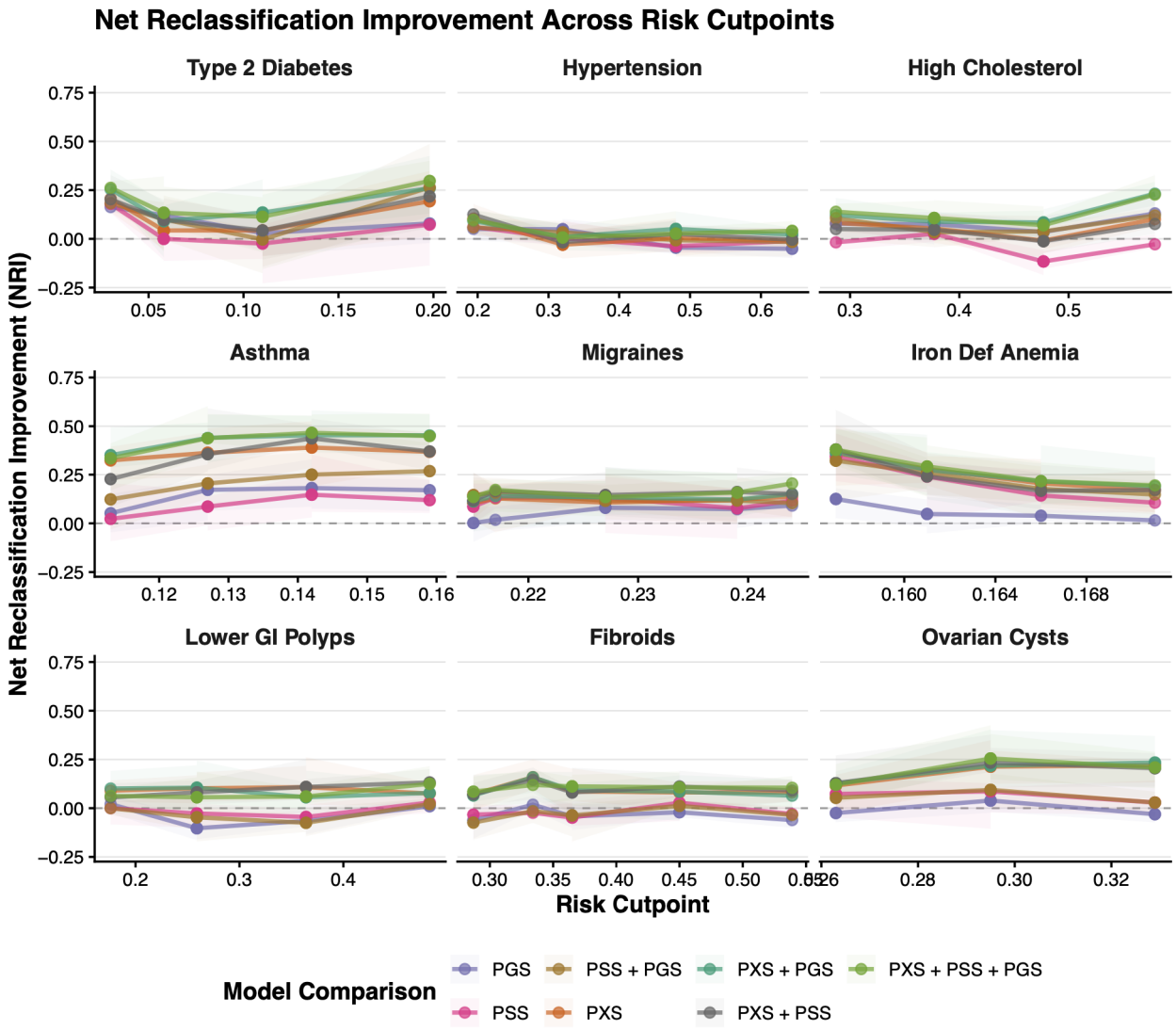

Figure S2: Disease-specific Net Reclassification Index (NRI). NRI across phenotypes and models. The Y axis is the NRI for each model at the X-axis predicted risk cutpoint.

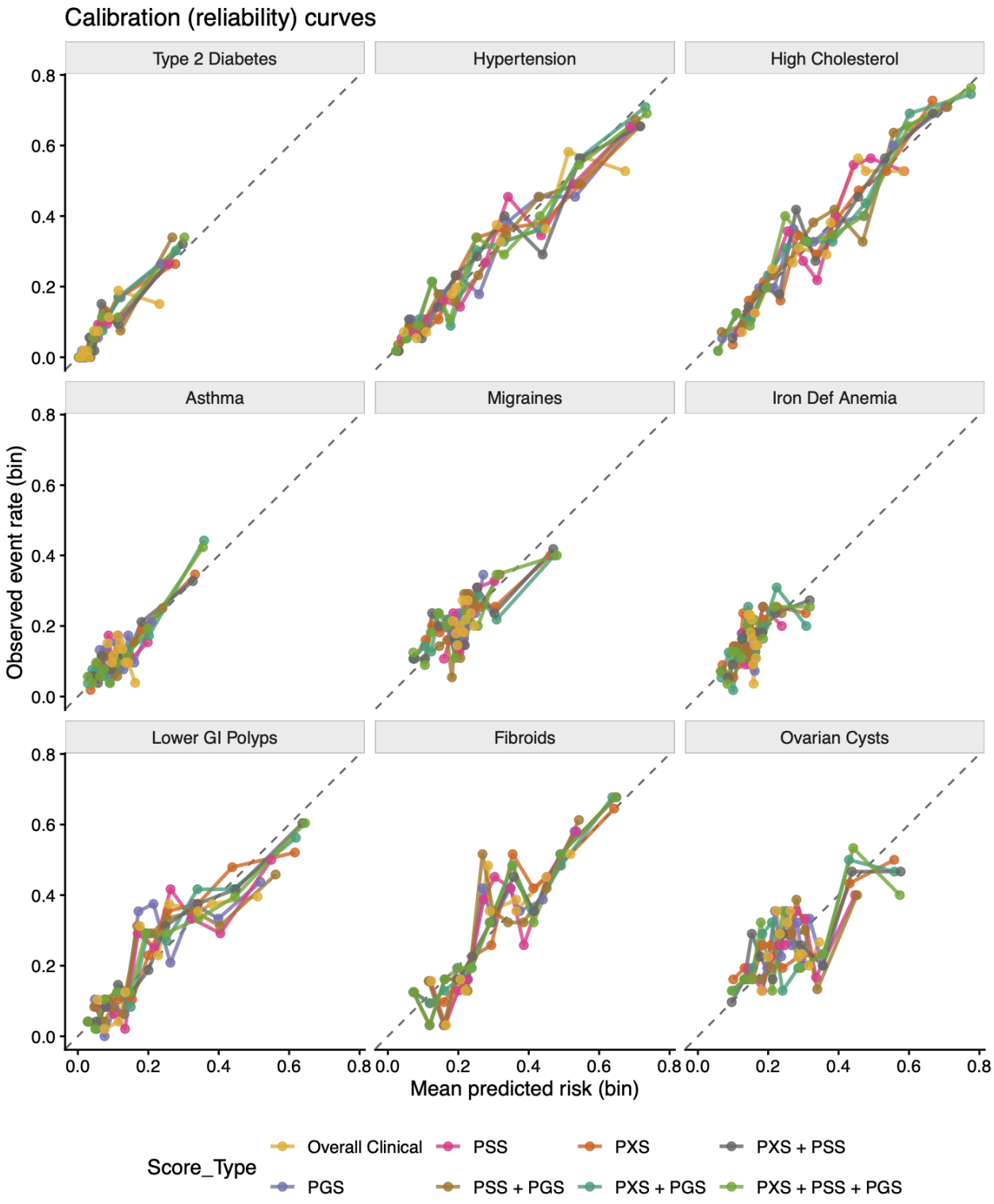

Figure S3: Disease-specific Calibration.
Calibration curves for each score combination and phenotype across models. The Y-axis is the observed event rate, and the X-axis is the mean predicted risk.

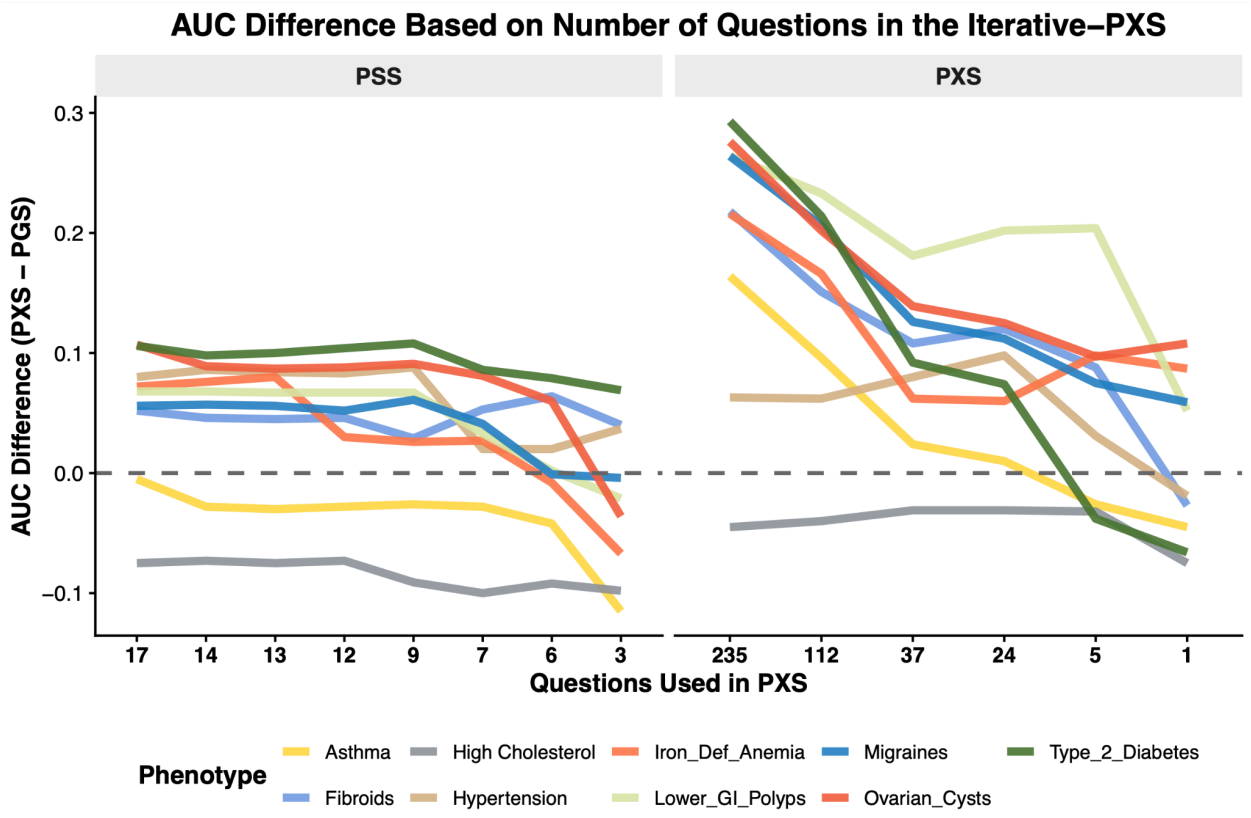

Figure S4: Iterative PXS Dropoff. The Y axis is the AUC difference between the PXS and the PGS for the specific disease, and the Y axis is the number of questions used to create that PXS, colored by disease.

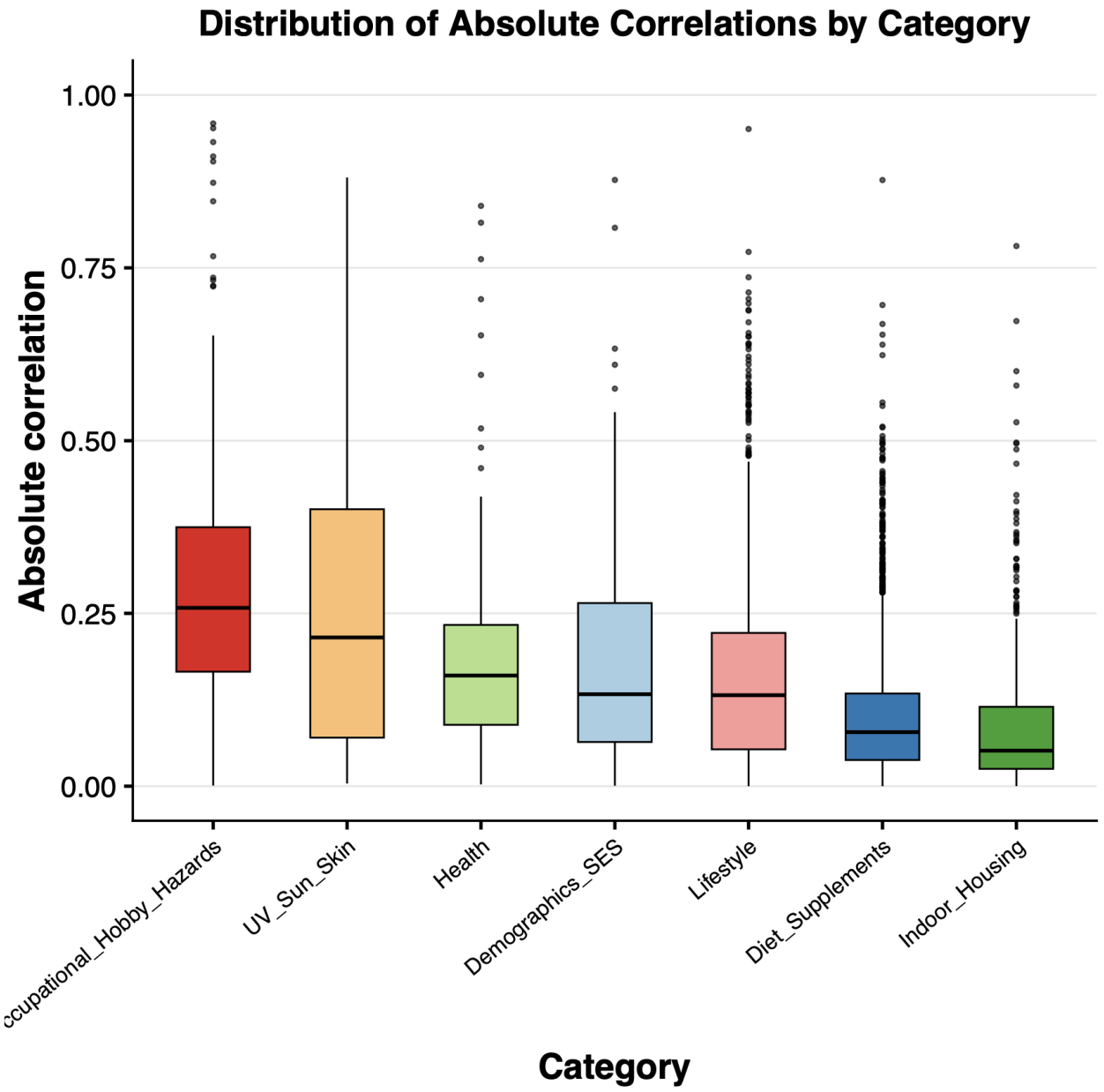

Figure S5: Exposure Correlations. The X-axis is the exposure category, and the Y-axis is the absolute value of the correlation per pair within the category.

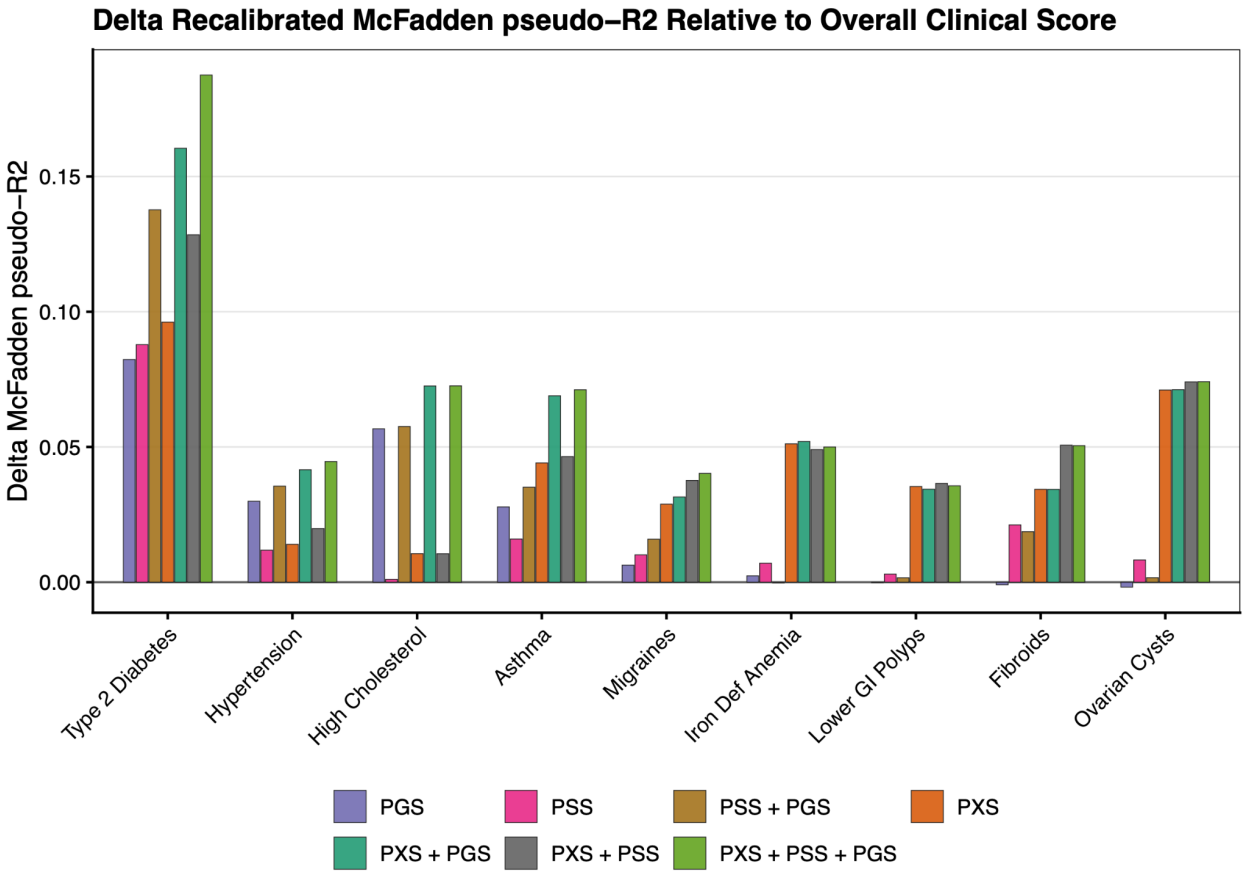

Figure S6: McFadden's R^2 for the Minimum Question Set. The X-axis is the disease of interest, and the Y-axis is the ΔMcFadden's pseudo R^2 compared to age and sex alone.

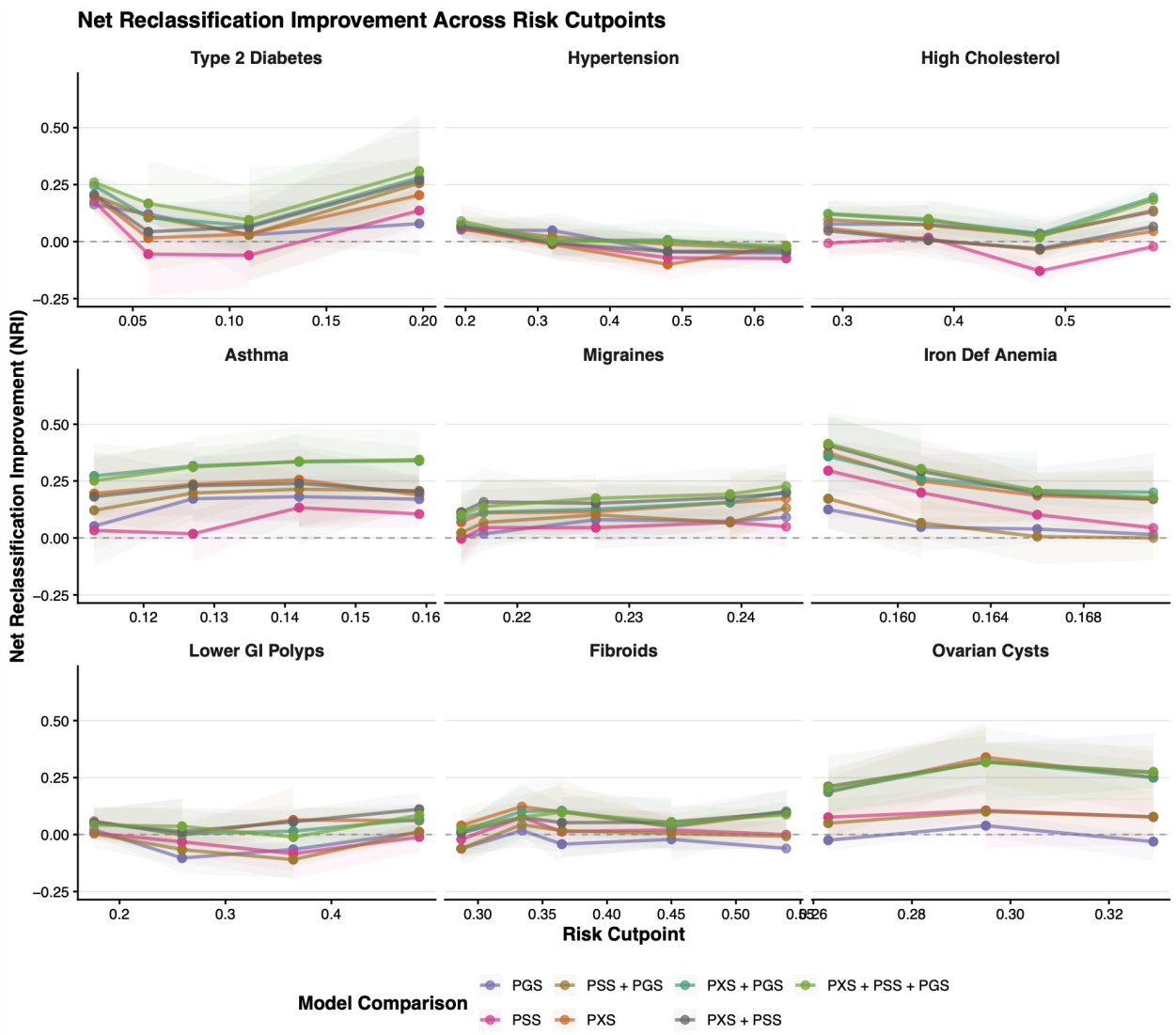

Figure S7: Net Reclassification Index (NRI) for the Minimum Question Set. NRI across phenotypes and models for the minimum question set. The Y-axis is the NRI for each model at the X-axis predicted risk cutpoint.

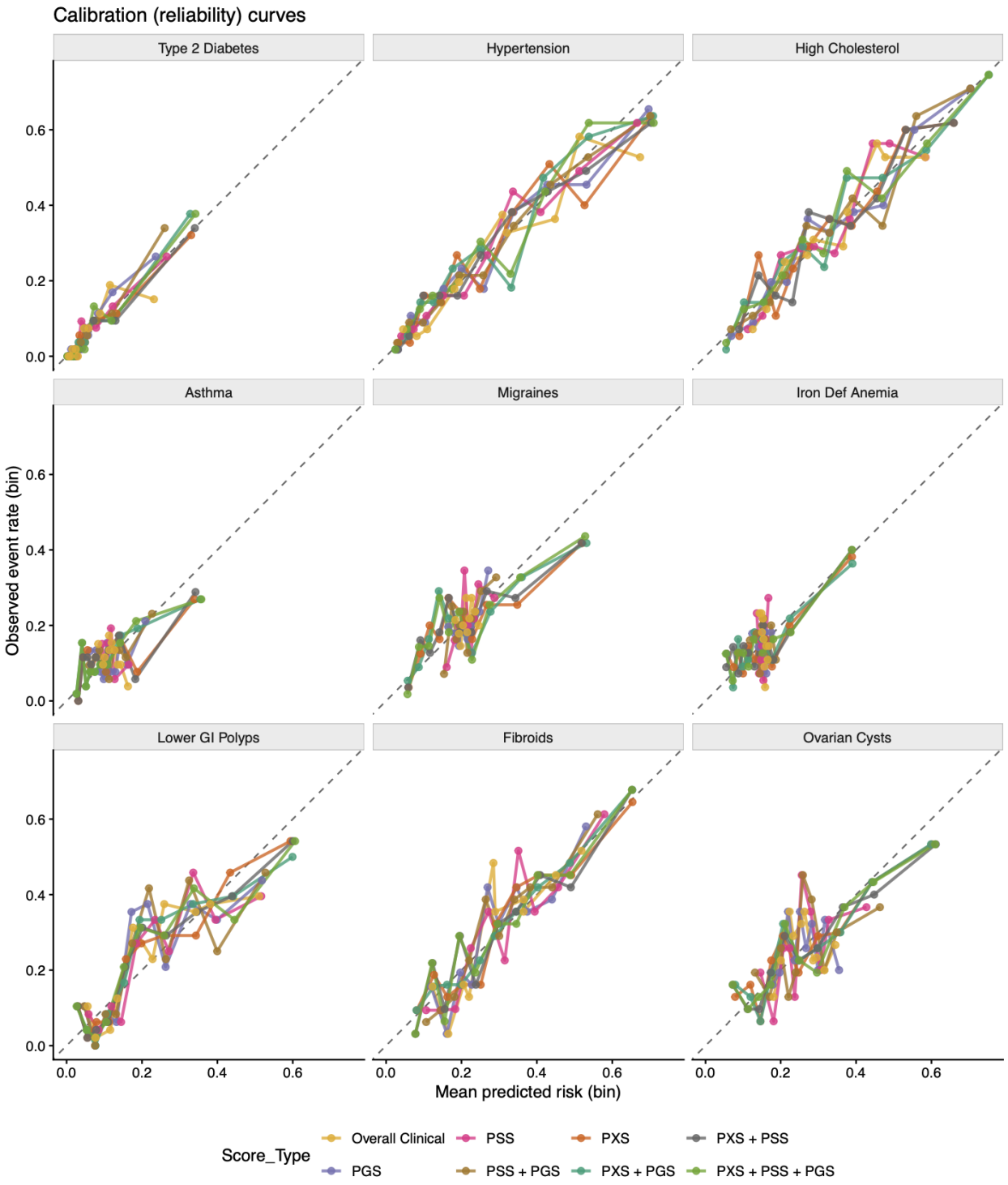

Figure S8: Calibration for the Minimum Question Set. Calibration curves for each score combination and phenotype across models. The Y-axis is the observed event rate, and the X-axis is the mean predicted risk.

Figure S9: Example Clinical Workflow. This illustrative workflow clarifies the intended clinical use of the PXS and PSS. A patient completes the 66-question assessment through the patient portal, tablet-based check-in, or nurse-led intake before a primary care visit. The electronic health record computes disease-specific PXS/PSS and flags elevated scores. For example, a 45-year-old patient without a prior asthma diagnosis but with a high asthma PXS and low PGS would not be diagnosed based on the score alone; rather, the score would prompt focused questioning about household dampness, irritant exposures, and occupational triggers, consideration of confirmatory evaluation, and targeted counseling or remediation support. In this workflow, PXS and PSS function as structured prompts for targeted follow-up rather than stand-alone diagnostic tools.

Supplemental Tables

Category	n	Features
UV_Sun_Skin	3	Sunscreen Use Frequency Sunglasses Use Frequency Skin Reaction Without Sunscreen
Health	7	History of Measles Number of Sleep Problems History of Mumps Bacterial Food Poisoning Genital Herpes German Measles (Rubella) Bacterial Pneumonia
Occupational_Hobby_Hazards	8	Cleaning Chemical Exposure Mercury Exposure X-Ray Exposure Acrylic Exposure Glue Exposure Ink Exposure Inkjet Printer Exposure Chemical-Resistant Gloves Use
Lifestyle	10	Sleep Disrupted by Pain Moderate Exercise Frequency Time Spent Walking Days Napped Nighttime Bathroom Use Snoring During Sleep Unusual Hyperactivity Time Spent Sitting Lifetime 100 Cigarettes Smoked Time Spent Standing
Indoor_Housing	11	Home Heating Month Start HEPA Filter Use Electric Heating Water Damage Repair Living Near Dry Cleaners Fireplace Use Home Water Filter Number of Bathrooms Home Dampness Home Mold Presence Vacuum Use
Diet_Supplements	22	Cream Consumption Grape Consumption Onion Consumption Cooked Spinach Consumption Ham Consumption Tortilla Consumption Low-Calorie Soda Consumption Low-Calorie Decaf Soda Consumption Coffee Consumption Water Consumption Mayonnaise Consumption Ready-to-Eat Meals Last Meal Time (AM/PM) Non-Fat Cheese Consumption Grapefruit Consumption

		Deli Meat Consumption **Peanut Consumption** **Splenda Use** **Iron Supplement Use** **Derived Iron Intake** **Fast Food Consumption** **Buffet Restaurant Consumption**

Table S1: Minimum Question Set for PXS

Category	n	Features
Demographics_SES	**5**	**Father's Highest Education** **Mother's Highest Education** **Participant's Highest Education** **Permanent Employment Status** **Home Type**

Table S2: Minimum Question Set for PSS

Phenotype	Usual diagnostic framework	Existing clinical assessment or risk context	What disease-specific PXS/PSS could add (future implementation frame)
High cholesterol	Diagnosed from history/physical plus lipid testing; high cholesterol is usually identified on a lipid panel rather than from symptoms[1].	Primary prevention decisions already use lipid values and atherosclerotic cardiovascular disease risk assessment[2].	A disease-specific PXS could help flag patients who may benefit from earlier lipid testing or a more structured exposure counseling workflow when routine history is limited. It should not substitute for measured lipid values.
Hypertension	Diagnosis depends on accurate blood-pressure measurement and out-of-office confirmation when hypertension is suspected[3].	Current care is stage-based and risk-based, with home monitoring and treatment intensification tied to blood-pressure level and overall cardiovascular risk[3].	A disease-specific PXS could help identify patients who merit home blood-pressure monitoring or a more structured exposure/lifestyle review before treatment escalation.
Type 2 diabetes	Diagnosis is laboratory-based, using A1C or plasma glucose criteria[4].	Current assessment centers on glycemic testing, prediabetes surveillance, and cardiometabolic risk reduction[4].	A disease-specific PXS could help prioritize earlier metabolic testing or prevention-focused counseling in patients with modifiable exposure profiles; it should not replace laboratory diagnosis.
Iron-deficient anemia	Diagnosis is confirmed with blood counts and iron studies; ferritin is used to diagnose iron deficiency, followed by	Standard care focuses on confirming anemia/iron deficiency and identifying etiology, often including gastrointestinal evaluation[5].	A disease-specific PXS could support a more structured exposure, diet, and environmental history alongside standard etiologic workup. It

Phenotype	Usual diagnostic framework	Existing clinical assessment or risk context	What disease-specific PXS/PSS could add (future implementation frame)
	evaluation for the underlying source[5].		should not replace laboratory tests.
Asthma	Diagnosis requires a history of variable respiratory symptoms plus objective evidence of variable expiratory airflow[6].	Trigger assessment is already part of care but is often unstructured and inconsistently captured in routine clinical workflows[6].	A disease-specific PXS could function as a pre-visit trigger inventory to prompt targeted questioning, spirometry, and environmental or occupational mitigation. A disease-specific PSS could flag structural constraints that affect control, access, or adherence.
Migraine	Migraine is primarily a clinical diagnosis based on history and neurologic examination; imaging is not routinely indicated when the presentation is typical and there are no red flags[7].	Management already relies on structured symptom history, trigger identification, and exclusion of secondary causes[7].	A disease-specific PXS could formalize trigger-focused history and counseling, improving how exposure patterns are captured and reused across visits without redefining diagnosis.
Fibroids	Fibroids may be identified on pelvic examination or pelvic ultrasound; hysteroscopy can further define location when needed[8].	Clinical significance depends on bleeding, pain, bulk symptoms, fertility goals, and anemia; asymptomatic fibroids may not require intervention[8].	A disease-specific PXS could help identify symptomatic patients who warrant more detailed exposure or lifestyle questioning alongside standard imaging and anemia evaluation. It should not be framed as a screening or diagnostic test for asymptomatic fibroids.
Ovarian cysts	Ovarian cysts are typically characterized with pelvic ultrasound, including size, shape, location, and whether contents are simple or complex[9].	Management is often triage and follow-up of an anatomic finding; many cysts are benign and can be observed[9].	This is best framed as a history or triage adjunct in symptomatic or recurrent cases rather than a screening tool for asymptomatic women.
Lower gastrointestinal polyps	Polyps are commonly ascertained through colorectal screening or endoscopic evaluation rather than symptom-based diagnosis[10].	Screening is already guideline-based in average-risk adults, so ascertainment partly reflects screening uptake and access[10].	A disease-specific PXS could prompt earlier shared decision-making about screening uptake, gastrointestinal follow-up, and focused exposure counseling. A disease-specific PSS may also capture structural barriers to screening.

Table S3: Contributions of PXS and PSS by Phenotype

Data component	Description	Number of participants
Health and Exposure Survey	Demographics, individual and family health history, environmental exposures, socioeconomic status, and lifestyle factors	9426
External Exposome Survey	Residential and occupational environmental exposures	3579
Internal Exposome Survey	Medication use, physical activity, stress, sleep, diet, genetics, and reproductive history	3034

Table S4: PEGS Data Components

Phenotype	Cases	Controls	N	Pct_Cases
High Cholesterol	949	1814	2763	0.343
Hypertension	799	1928	2727	0.293
Migraines	555	2198	2753	0.202
Lower GI Polyps	512	1867	2379	0.215
Iron Def Anemia	487	2263	2750	0.177
Fibroids	435	1120	1555	0.28
Ovarian Cysts	412	1122	1534	0.269
Asthma	334	2265	2599	0.129
Type 2 Diabetes	177	2490	2667	0.066

Table S5: Number of Cases for Each Phenotype

Supplemental References

1. Cholesterol test. https://www.mayoclinic.org/tests-procedures/cholesterol-test/about/pac-20384601.
2. 2026 Guideline on the Management of Dyslipidemia. *professional.heart.org* https://professional.heart.org/en/science-news/2026-guideline-on-the-management-of-dyslipidemia.
3. Top 10 Things to Know About the AHA/ACC High Blood Pressure Guideline. *www.heart.org* https://www.heart.org/en/health-topics/high-blood-pressure/the-facts-about-high-blood-pressure/high-bp-top-10 (2025).
4. American Diabetes Association Professional Practice Committee for Diabetes*. 2. Diagnosis and classification of diabetes: Standards of care in diabetes-2026. *Diabetes Care* **49**, S27–S49 (2026).
5. Short, M. W. & Domagalski, J. E. Iron deficiency anemia: evaluation and management. *Am. Fam. Physician* **87**, 98–104 (2013).
6. *Global Initiative for Asthma. Summary Guide for Asthma Management and Prevention*. (2025).
7. Overview of Migraine Diagnosis. *American Headache Society* https://americanheadachesociety.org/research/library/overview-of-migraine-diagnosis (2020).
8. Office of the Commissioner. Uterine fibroids. *U.S. Food and Drug Administration* https://www.fda.gov/consumers/womens-health-topics/uterine-fibroids (2025).
9. Ovarian Cysts. https://www.acog.org/womens-health/faqs/ovarian-cysts.
10. Colorectal Cancer: Screening. https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/colorectal-cancer-screening (2021).