

2025 Phenotype Phebruary Closing session

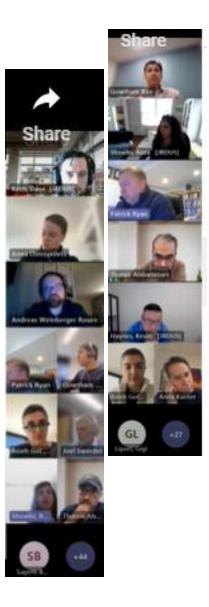


Following the footsteps of Dry January: Help 14 study leads and the teams create and evaluate all cohort definitions for their studies



2025 Phenotype Phebruary team & study leads

Patrick Ryan Gowtham Rao Azza Shoaibi **Anna Ostropolets** Lana Shubinsky **Chris Mecoli Oleg Zhuk** Vlad Korsik Maria Khitrun Bohdan Khilchevskyi **Chen Yanover** Tetiana Skugarevskaya **Seng Chan You Cindy Cai Chang Hoon Ha Kevin Haynes Asieh Golozar Michelle Hribar**





Phenotype Phebruary 2025: what we achieved

- **3** Atlas and CD demos, **20+** sessions
- Clinical descriptions for 13 studies written
- **63** cohort definitions re-used from PL/os
- 165 cohort definitions built
- **118** cohort definitions built and publicly shared
- 2 cohort diagnostics run on results.ohdsi.org
- 40+ collaborators reviewed literature, built cohorts, reviewed cohort diagnostics or attended calls

\times	Phenotype Phebruary 2025 Calendar						
Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturda	у
Week 1: Clinical Descriptions & Prior Work							
		Community call kick off & clinical description demos	Office hours	Study team's complete clinical description	Office hours		
	2	3	4	5	6	7	8
Week 2: Building Concept Sets and Logic							
		Community call update & Atlas demos	Office hours	Study team's complete 1st drafts of cohorts in Atlas			
	9	10	11	12	13	14	15
Week 3: CohortDiagnostics & Iterations							
		Community call update & CD demos	Office hours	Study team's complete cohort review and iterations	Office hours		
	16	17	18		20	21	22
Week 4: Finalize Phenotypes & Tools for Evaluation							
		Community call update & evaluation demos	Office hours	Study team's wrap up cohorts and document evaluation	Office hours	All cohor done	
	23	24	25	26	27	28	

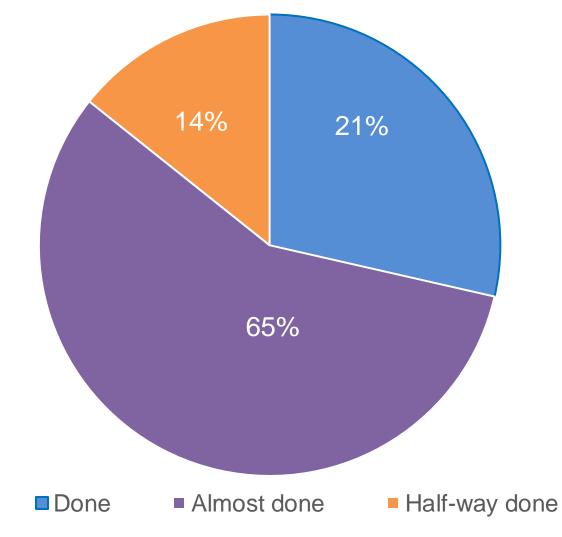
Phenotype Phebruary content location: Phenotype development and evaluation WG -> Files -> Phenotype Phebruary 2025

https://results.ohdsi.org/app/26_PhenotypePhebruary 2025



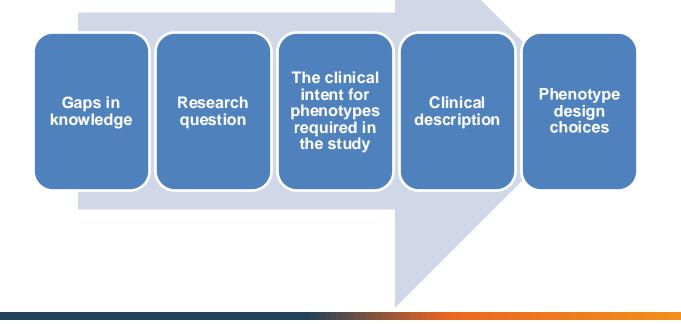
Phenotype Phebruary 2025: what we achieved

Phenotyping for ~85% of studies is done or almost done

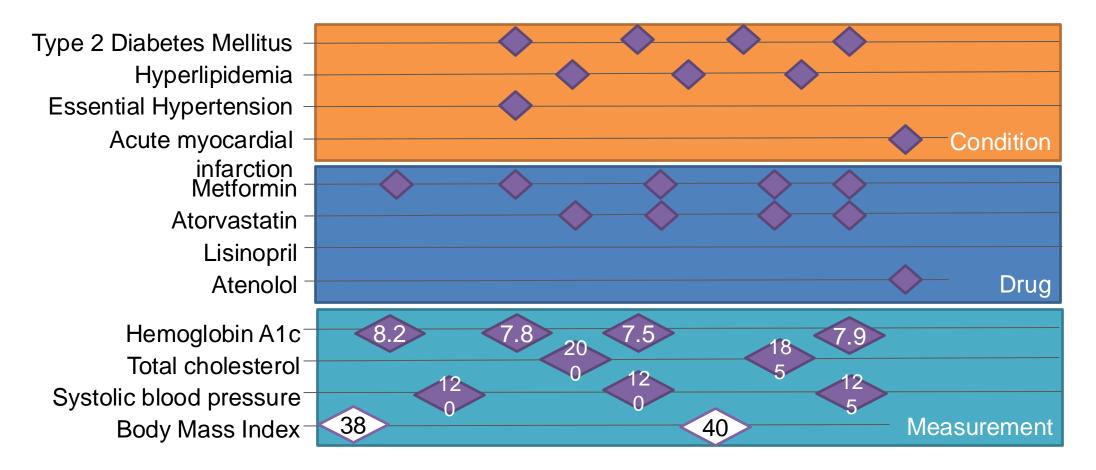




• Explicitly specifying clinical ideas that one wants to study (e.g., through clinical description) is crucial



What we *HAVE*? Observational data for a single person

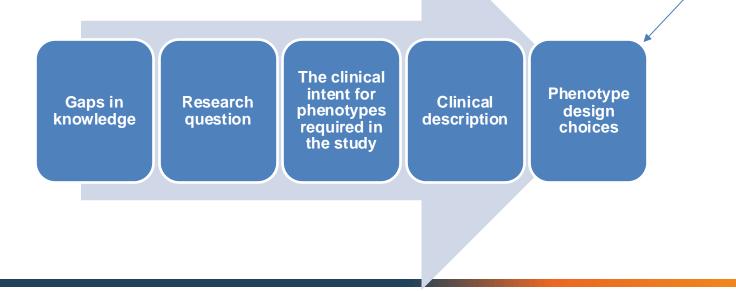


Observation Time

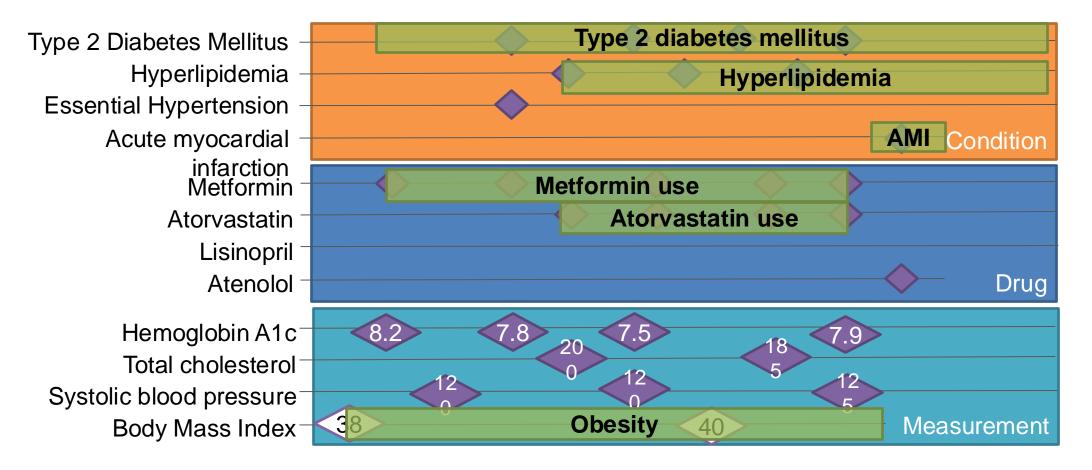


 We rely on our clinical knowledge when asking for specific patient populations. We should also listen to the data to see what it can support

observe in the data



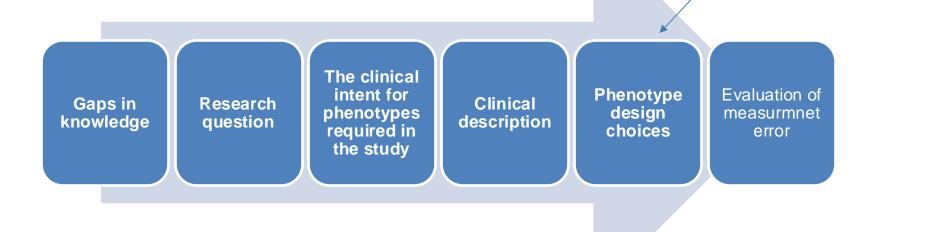
What we *WANT*? Longitudinal health status for a single person



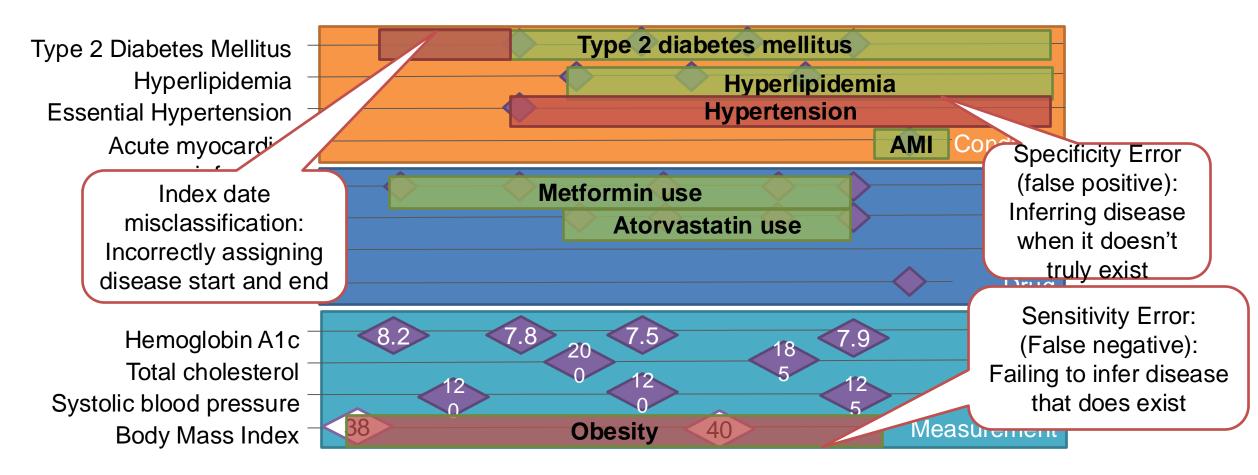
Observation Time



 In an absence of objective criteria, choices are subjective, and measurement error is not easily described or communicated
What we can reliably observe in the data



Potential errors from inference in disease phenotyping



Observation Time



 ~20% of the cohorts were re-used from the Phenotype Library/other sources. We could have spent 100+ hours less if all of them were available



Moving forward

- We can create Atlas shells for common cohort designs
- We can improve LLM prompts for clinical description
- We can create automated solutions for concept sets
- We can create a lightweight version of Cohort Diagnostics for easy exploration

$\left|\right\rangle$

Final note of celebration

2022: 80% of the phenotypes were done by 1 community leader

2023: 80% of the phenotypes were done by 5 prople 2024: 80% of the phenotypes were done by 20 people

2025: 80% of the phenotypes were done by 40 people

Building our community of phenotypers