

Identification of sub-groups of chronic obstructive pulmonary disease patients using clustering analysis of small airway histopathology

Exploratory airway phenotyping • clinical prediction modelling • fairness

Urbi Roychowdhury

The University of Manchester | urbi.roychowdhury@student.manchester.ac.uk

*Supervisor: David A Jenkins | Division of Informatics,
Imaging & Data Sciences*

How do we find useful structure in messy clinical data; without mistaking patterns for proof?

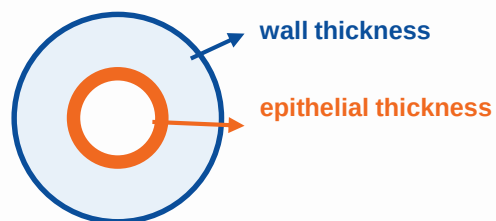


My Leading Approach

A common theme: how do we know when a pattern or prediction is genuinely reliable?

1 COPD PHENOTYPING

Can repeated small-airway measurements reveal recurring patient patterns?



Exploratory clustering → clinical profiling → sensitivity checks

2 PREDICTION + FAIRNESS

Can a clinical risk model stay trustworthy across time and patient groups?

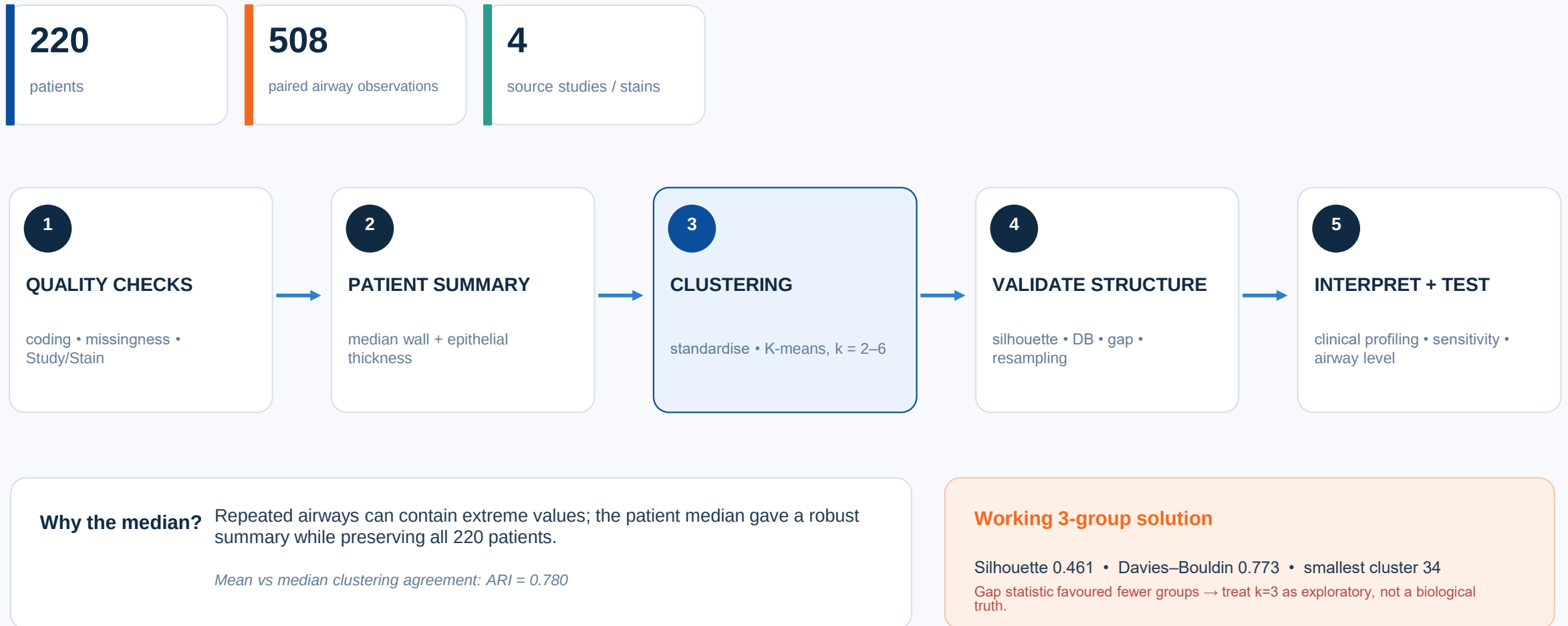


Reproduce → validate → stratify by sex → inspect change over time

COMMON THREAD • stability • uncertainty • data quality • cautious interpretation

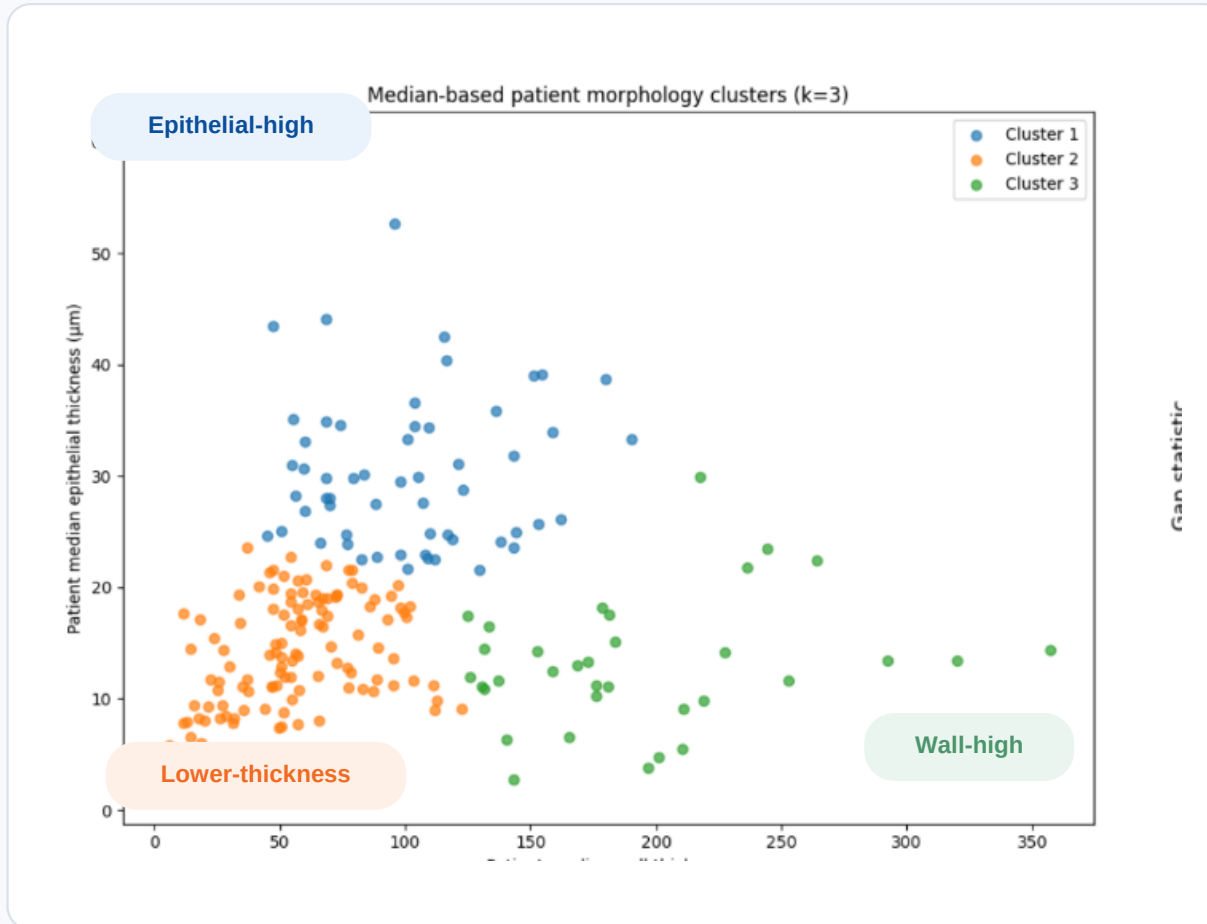
From repeated airway measurements to patient-level morphology

I built the analysis in stages so each step answered a different question



Three morphology patterns — useful as hypotheses, not “three COPD subtypes”

The most interesting result was not just that clusters existed, but how differently they behaved clinically.



WALL-HIGH GROUP

96.8% COPD • 74.2% current smokers

An interesting candidate pattern for follow-up — not evidence of causation or diagnostic value.

LOWER-THICKNESS GROUP

Lung-function variables reorganised ~26% of patients

Pack-years changed only 2.2%. This suggests clinical heterogeneity within a seemingly similar morphology region.

REPEATED AIRWAYS

42.3% of patients with repeated samples spanned >1 airway cluster

Yet dominant airway pattern agreed with the median-based group for 96% of unambiguous patients.

Counterbalance: after age, gender, pack-years and Study, morphology did not clearly add independent linear explanatory value for FEV1/FVC or FEV1 % predicted.

Rebuilding a cardiac-surgery risk model taught me what “validation” really means

A clinically useful model is not just fitted once : it has to remain calibrated, discriminative and reproducible.



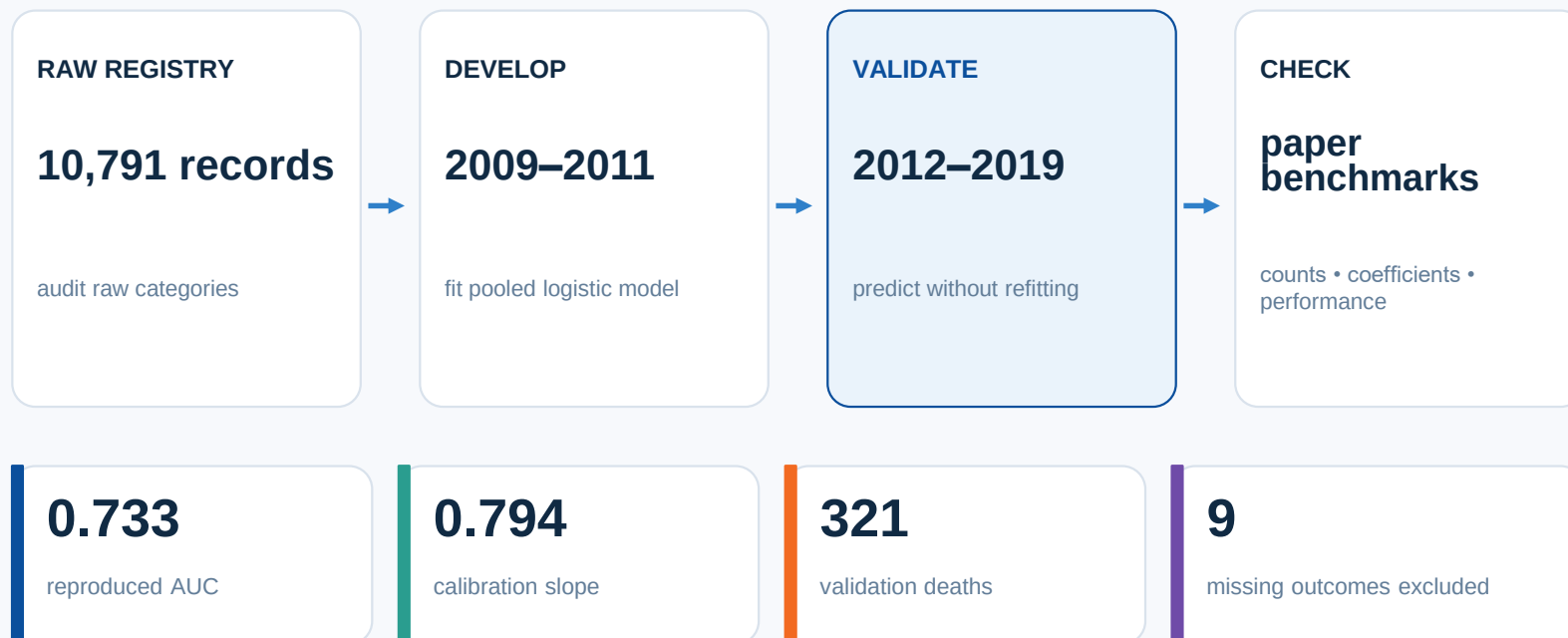
Jenkins et al.
(MEDINFO 2023)

NACSA cardiac-surgery
data • Wythenshawe
Hospital • 2009–2019

Published study compared four
approaches:

- Fixed logistic
- Yearly recalibration
- Varying coefficient
- Continually updating Bayesian

MY RECONSTRUCTION WORKFLOW

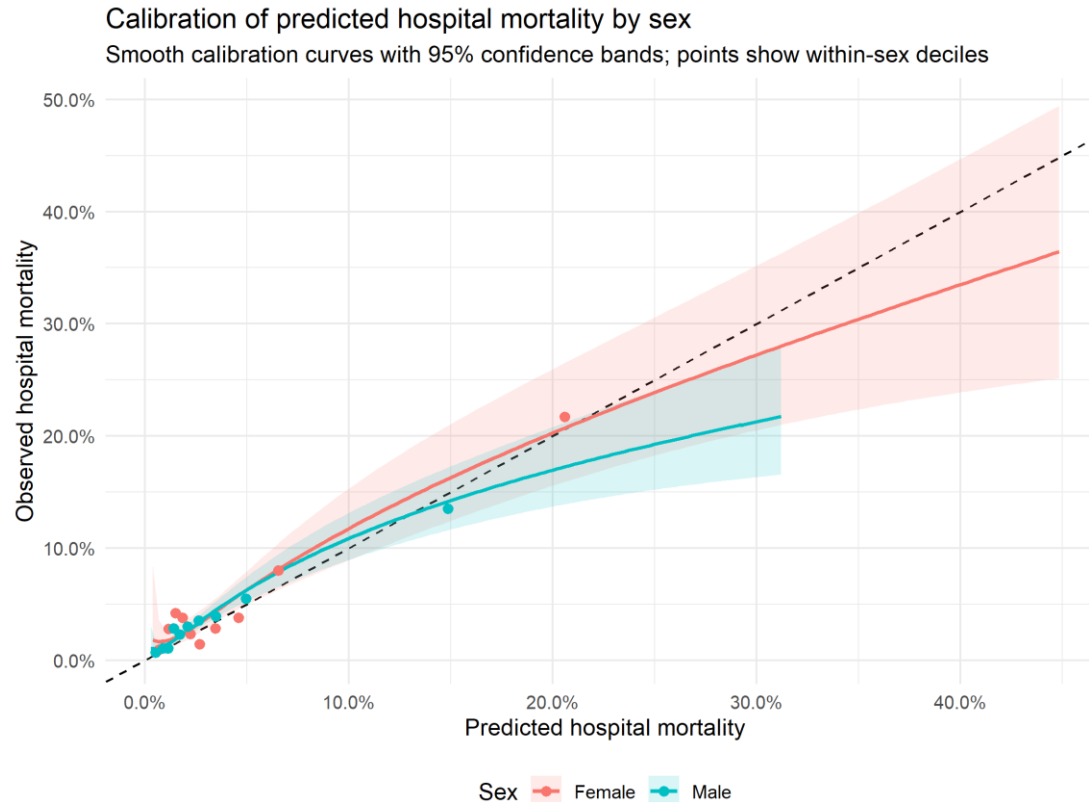


Reproduction lesson:

most baseline characteristics and coefficients matched closely, but absolute calibration did not match perfectly — so later fairness results were treated as an exploratory audit of the reconstructed model.

The pooled model performed similarly on average — but representation and time still mattered

Fairness was treated as a validation question: does the same model remain reliable in women and men?



VALIDATION COHORT

Women

27.4%

of validation cohort

Mortality 5.18%
AUC 0.732
O/E 1.144

Men

72.6%

of validation cohort

Mortality 3.74%
AUC 0.731
O/E 1.111

Aggregate result

No clear persistent sex-specific gap in calibration or discrimination.

But over time...

Calibration moved differently in some years; several female strata had few deaths, making estimates unstable.

Representation matters: the female-only refit had only 32 development deaths and became highly unstable (AUC 0.385). More targeted modelling cannot compensate for very limited subgroup information.

The internship changed how I think about “finding a result”

The strongest outcome was a more cautious, validation-first way of doing health data science research

1

PATTERNS ARE STARTING POINTS

Clustering can reveal useful structure, but internal metrics do not turn clusters into biological subtypes. The next question is what the pattern represents and whether it replicates.

2

DATA WORK IS PART OF THE SCIENCE

Raw coding, missingness, repeated observations, study effects and reference categories can change the answer. Reproduction checks should come before interpretation.

3

PERFORMANCE MUST BE MONITORED

A model can look similar across groups on average while still facing representation gaps, unstable subgroup-years or drift. Fairness is part of validation, not a one-off label.

NEXT STEPS

COPD

Validate morphology patterns in independently collected data with consistent sampling.

AIRWAYS

Model within-patient heterogeneity directly rather than reducing it to a single summary.

PREDICTION

Extend the sex-specific audit to the continually updating Bayesian model.

Thank you to my supervisors, the N8 CIR team and the University of Manchester!