Expanded detection of early fibrotic phenotypes using lobar traction bronchiolectasis in lung cancer screening

Daryl O. Cheng [1,2] (0000-0001-9906-4298)
Daisuke Yamada [1] (0000-0003-1561-5908)
Mehran Azimbagirad [1] (0000-0001-6321-7643)
Amyn Bhamani [2] (0000-0002-8575-6119)
Ryoko Egashira [1] (0000-0001-7566-546X)
Robert Chapman [1] (0000-0003-0093-1086)
John McCabe [1,2] (0000-0002-1114-8276)
Shanshan Wang [1]
Jennifer Naftel [3]
Sarah Marguerie [3]
Tim Wallis [3]
Jonathan Dunbar [3]
Pardeep Vasudev [1]
SUMMIT consortium *
Arjun Nair [4] (0000-0001-9270-3771)
Daniel C. Alexander [5] (0000-0003-2439-350X)
Sam M. Janes [2] (0000-0002-6634-5939)
Mark G. Jones [3, 6, †] (0000-0001-6308-6014)
Brintha Selvarajah [2,#, †] (0000-0001-8342-9117)
Joseph Jacob [1,#, †] (0000-0002-8054-2293)

1- Satsuma Lab, Hawkes' Institute, University College London, UK
2- Lungs for Living Research Centre, University College London, UK
3- NIHR Southampton Biomedical Research Centre, University Hospitals Southampton, UK
4- Department of Radiology, University College London Hospitals NHS Trust, UK
5- Hawkes' Institute, University College London, UK

6- School of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK
* Full list in acknowledgements
† Joint senior authors
# Corresponding authors: j.jacob@ucl.ac.uk, brintha.hayes@ucl.ac.uk

Word count: 3496 / 3500
Figure count: 3; Table count: 3

This article has an online data supplement which is accessible at the Supplement Tab.

Artificial Intelligence Disclaimer: No artificial intelligence tools were used in writing this manuscript

Author contributions: Conception and design: JJ, DY, DC; Data acquisition: MA, SW, NJ, SM, TW, JD, PV; Analysis and interpretation: DC, JJ, BS, MGJ, DY, RE, MA; Drafting and review: DC, JJ, AN, DCA, SMJ, MGJ, RC, AB, JMcC.

At a Glance

Scientific Knowledge on the Subject: Interstitial lung abnormalities (ILA) are common incidental findings in lung cancer screening. ILAs may represent a precursor to interstitial lung disease, but have also been shown to have associations with reduced lung function, increased symptoms, and reduced overall survival. Existing classifications may underestimate the prevalence of clinically important ILA, especially in modern lung cancer screening populations with improved scan resolution.

What This Study Adds to the Field: A new ILA classification, the TBe score, was developed emphasizing traction bronchiolectasis, an imaging feature reflecting early interstitial damage. The study then investigated whether this earlier disease state associated with increased respiratory morbidity and reduced overall survival.

The TBe score was shown to identify 86% more participants in lung cancer screening with undiagnosed fibrotic interstitial lung disease compared to the existing American Thoracic Society (ATS) classification. These participants have higher rates of respiratory hospitalisation and increased risk of death. Despite higher sensitivity, participants identified by TBe had comparable disease severity to those identified by ATS criteria. In lung cancer screening settings, TBe can identify more participants than ATS criteria who may warrant further investigation and follow up.

Abstract

Rationale: Lung cancer screening regularly identifies participants with interstitial lung abnormalities (ILA). Existing classification methods may underestimate the prevalence of clinically relevant ILA phenotypes.

Objectives: Can a classification system for ILAs developed in a lung cancer screening setting identify clinically relevant phenotypes?

Methods: Classification criteria based on the presence and lobar extent of traction bronchiolectasis (TBe) were developed internally by expert consensus. Categories included: no ILA, non-fibrotic ILA (NF-ILA), fibrotic ILA (F-ILA), and undiagnosed fibrotic ILD (U-ILD). Interobserver agreement was calculated between two readers. Clinical characteristics, respiratory hospitalisations, and survival were compared between participants of different ILA grades.

Measurements and Main Results: 8,169 participants were included in the final analysis. TBe showed improved interobserver agreement compared to the American Thoracic Society (ATS) classification, identifying 344 participants (4%) with U-ILD, 86% more than the ATS classification. An additional 405 had F-ILA (5%) and 667 had NF-ILA (8%). Compared to participants without ILA, participants with U-ILD had a higher rate of respiratory hospitalisation (IRR=4.4, 95% CI 2.7–7.5, p<0.001) and increased risk of death (aHR=2.4, 95% CI 1.9–3.0, p<0.001). Increasing ILA grade was associated with higher modified Medical Research Council dyspnoea scores (OR=1.1, 95% CI 1.0–1.1, p=0.02).

Conclusions: In a lung cancer screening setting, an ILA scoring system focused on lobar traction bronchiolectasis identifies more high-risk participants and demonstrates improved interobserver concordance than the ATS classification. TBe identifies participants with a respiratory phenotype who may warrant further investigation and follow up.

Word count: 241 / 250

Introduction

Lung cancer screening (LCS) programs are expanding worldwide. The UK government anticipates that once fully rolled out, by 2030, the national screening program will enroll more than 300,000 participants every year, generating almost one million scans annually [1]. The overwhelming majority of participants will not have lung cancer diagnosed. But the low dose computed tomography (LDCT) scan they receive may contain incidental findings such as interstitial lung abnormalities (ILA).

ILAs are a common finding in smoking cohorts with meta-analyses estimating a 7% prevalence in LCS populations [2–4]. Demographic factors such as advanced age and smoking history confer increased risk [5]. ILAs may represent a precursor to, or an early stage of interstitial lung disease (ILD), and a proportion are reported to progress to ILD at follow up [2]. They are also increasingly recognised as important clinical entities in their own right, given reported associations with reduced lung function, increased respiratory symptoms, and increased mortality [6, 7].

A Position Paper from the Fleischner Society proposed criteria for ILA classification [8] which with slight modifications have now been ratified in an Official American Thoracic Society (ATS) Clinical Statement [9]. The ATS definition of an ILA requires bilateral and non-dependent ground-glass, and reticular opacities involving >5% of a lung zone. Fibrotic ILAs comprise honeycombing and/or reticulation with traction bronchiectasis involving >5% of a lung zone and <5% of total lung volume by visual estimate. ILD for those with ILAs in the ATS statement included “honeycombing and/or reticulation with traction bronchiectasis involving >5% of total lung volume by visual estimate”, or the presence of a major fibrotic pattern, or diffuse ground glass

abnormalities. The presence of traction bronchiolectasis was explicitly excluded from criteria that determine the presence of fibrosis, as evidence for its prognostic importance was felt to be lacking.

A focus on advanced fibrotic features (honeycombing and traction bronchiectasis) when defining early disease, and the requirement of >5% lung involvement, may result in an underestimation of the prevalence of early fibrosing ILD. Traction bronchiolectasis, comprising dilatation of airways in the peripheral 2 cm of the lung, is an imaging pattern recognized by experts as indicating the presence of a fibrosing ILD [10], despite being omitted from the ATS Statement.

A prior pilot study evaluated the feasibility of applying traction bronchiolectasis based criteria on screen detected ILAs, using a lobar, rather than a zonal or lung volume percentage based approach. This showed good interobserver agreement, and strong associations with mortality [11, 12].

With the increasing adoption of LCS worldwide, there is growing urgency around improving radiological classification of ILAs to distinguish which participants need onward referral. Here, we describe a modified ILA classification system developed to identify both ILA and undiagnosed ILD based on the presence and lobar extent of traction bronchiolectasis and explore its clinical associations.

Methods

Study population

The SUMMIT Study was a longitudinal, prospective, cohort study investigating the deliverability of low dose CT screening for lung cancer in a high-risk population in London, United Kingdom (NCT03934866). Participants aged 55–77 were enrolled from April 2019 to March 2020. Ethical approval for the SUMMIT Study and ongoing analyses was obtained from an NHS Research Ethics Committee (17/LO/2004) and the NHS Health Research Authority's Confidentiality Advisory Group (18/CAG/0054).

Baseline LDCT scan data were used for visual assessment. Participants were first screened for eligibility by trial staff by telephone, after which they were invited to an in-person baseline lung health check appointment (LHC). Demographic, smoking, and medical history data were collected at the LHC, including: self-reported ethnicity, smoking status, smoking duration, smoking dose (pack years, PY), body mass index (BMI), self-reported chronic obstructive pulmonary disease (COPD) diagnosis, spirometry data, and modified Medical Research Council dyspnoea score (mMRC). Respiratory hospitalisations were defined as any admission to hospital from with either the primary or secondary diagnosis coded from International Classification of Diseases-10 as ‘Diseases of the Respiratory System’, using Hospital Episode Statistics from NHS England at year 1 and year 2. Spirometry data collection was discontinued from 19 March 2020, because of COVID-19 restrictions. Age, gender and socioeconomic deprivation measured by Indices of Multiple Deprivation (IMD) rank (described in Supplementary Materials) were collected from primary care records.

Participants with known ILD, as well as participants with lung cancer identified at baseline scan, were excluded from the analysis. Mortality data were collected through study linkage with the National Disease Registration Service from NHS Digital. Further details on the study protocols of the SUMMIT study have been published previously [13, 14].

External validation population

A study population from the Targeted Lung Health Check (TLHC) program in Southampton was used for validation of the clinical and survival findings from SUMMIT. The TLHC program, designed as a pilot for the roll out of lung cancer screening (LCS) in the UK, had eligibility criteria closely mirroring that for SUMMIT, enrolling participants aged 55–77, and meeting a 6-year lung cancer risk of ≥1.51% ($PLCO_{m2012}$ [15]). For the Southampton cohort, ethical approval was obtained from the London-Hampstead Research Ethics Committee (REC: 17/LO/2037).

Baseline LDCT scans in Southampton TLHC subjects referred to clinical services on the basis of interstitial change identified by local thoracic radiologists were selected for assessment. These CTs were rescored by a study radiologist (JJ) using the new ILA classification. The same clinical data as in SUMMIT were collected, except IMD measures which were not available.

Definition of lobar traction bronchiolectasis score

We developed a four-point lobar TBe score based on the presence of traction bronchiolectasis. Traction bronchiolectasis was defined as dilated bronchioles in the peripheral 2cm of lung occurring within areas of reticulation and/or ground glass opacification (GGO). For ILA to be present, damage had to exist in more than one

lobe. No ILA therefore included GGO or reticulation in a maximum of one lobe. For the ILA categories, Non-fibrotic ILA (NF-ILA): included non-dependent GGO and/or reticulation in two or more lobes, with no traction bronchiolectasis. Fibrotic ILA (F-ILA) used the same criteria as NF-ILA but with traction bronchiolectasis evident in 1 or 2 lobes. Undiagnosed fibrotic ILD (U-ILD) again used the same criteria as NF-ILA but with traction bronchiolectasis present in 3 or more lobes (Figure 1). The lingula segment of the left upper lobe was considered a separate lobe.

Assessment of concordance

SUMMIT scans were split into two sets and scored independently by two expert thoracic radiologists (JJ:16 years' experience; DY: 10 years' experience). All ILA cases and a random selection of 10% of the no ILA cases scored by one radiologist were rescored by the other radiologist. The final reviewed dataset underwent consensus review. The scoring process is outlined in Supplementary Materials (Supplementary methods, Figure E2). Methods of analysis of concordance in subgroups are also outlined in Supplementary Materials.

Comparison with ATS Statement Classification

All cases were scored using the TBe score and the ATS Statement classifications (Supplementary Figure E2). All participants identified by ATS as ILD were compared with participants identified by TBe as U-ILD, excluding those already identified by ATS as ILD. We also compared those classified by TBe as U-ILD with those classified by ATS as F-ILA.

Furthermore, clinical characteristics were compared between those classified as F-ILA by both systems, and those classified as F-ILA by TBe only, those classified as

F-ILA by ATS only. Similarly, clinical characteristics were compared between those classified as ILD by both systems, those classified as U-ILD by TBe, and those classified by ILD by ATS. These comparisons are listed in Supplementary Table E10.

The ATS classification includes the use of symptom or spirometry-based criteria, in addition to imaging domains. Only pre-bronchodilator spirometry data were available, but this was used to complete an additional sensitivity analysis based on whether spirometry-based criteria to the ATS classification modified the findings compared to the imaging domains only. For this, the definitions of the TBe score and ATS classification as above were modified so that participants who met radiological criteria for NF-ILA or F-ILA and had pre-bronchodilator FVC <80% predicted were classified as U-ILD.

Statistical methods

Clinical characteristics

To investigate the relationship between the TBe score and clinical characteristics, the score was classed as an ordinal variable, ranging through no ILA, non-fibrotic ILA, fibrotic ILA and undiagnosed ILD. Predictors were selected *a priori* based on published literature of associations with ILA and ILD, as well as clinical plausibility. Details of feature selection are included in the Online Data Supplement. Proportional odds assumptions were assessed using the Brant test [16]. Variables which violated the proportional odds assumption were modeled with threshold-specific effects using a partial proportional odds model. The remaining variables were modelled with a single effect estimate. Potential collinearity between lung function measures were

assessed with correlation matrices, as well as variance inflation factors. Interobserver agreement was calculated using weighted and unweighted Cohen's kappa.

Respiratory hospitalisations

Poisson regression models were used to assess the relationship between baseline TBe classification and subsequent respiratory hospitalisations during follow up. Where significant overdispersion was present, negative binomial regression models were chosen. Models were adjusted for the clinical variables of age, gender, smoking status, smoking dose (PY), BMI, FEV_1, FVC, COPD diagnosis at baseline, as well as varying follow up times using a time offset. Incidence rate ratios (IRRs) were reported to quantify the relative increase in hospitalisation risk across TBe categories.

Mortality

The hazard of overall death was modelled using a Kaplan-Meier estimator, with log-rank tests for trend comparing survival distributions across ordered TBe classification groups, assessing the progressive relationship between TBe grade and survival. Multivariable survival analyses were performed using Cox proportional hazards models, and included age, gender, smoking status, smoking duration, and FEV_1, selected *a priori,* as described previously. Proportional hazards assumptions were assessed using Schoenfeld residuals for individual variables graphically and globally. Sample size calculation was performed to ensure that the data contained an adequate number of events per variable for reliable estimation [17].

Results

Clinical characteristics of study population

8,169 participants were included in the SUMMIT analysis. The median age was 66 years (interquartile range, IQR: 60–70), and 40% of participants were of female gender (Table 1, Supplementary Figure E1). Median follow up time was 5.0 years (IQR: 4.7–5.2), representing 38,460 person-years at risk. Using the TBe score, 344 participants were classed as undiagnosed ILD (4.2%), 405 with fibrotic ILA (5%), 667 with non-fibrotic ILA (8.2%), and the remaining 6753 with no ILA (83%).

Violations of proportional odds assumptions for age and FVC% required implementation of a partial proportional odds model to estimate associations with increasing TBe grade. TBe grade increase is defined as a progression of no ILA, NF-ILA, F-ILA and U-ILD. Odds ratios (OR) at the thresholds between grades were reported: threshold 1 represents the OR between no ILA and any ILA, threshold 2 the OR between NF-ILA and F-ILA, and threshold 3, the OR between F-ILA and U-ILD.

Every decade of increasing age was associated with greater odds of increased TBe grade (OR ranging from 2.4 to 3.3, $p<0.001$, Table 2). Male gender was associated with increased odds of a higher grade of TBe (OR = 1.5, 95% CI: 1.3–1.7, $p<0.001$), as was each 10 kg/m2 of increasing BMI (OR = 1.3, 95% CI: 1.2–1.5, $p<0.001$). Smoking status and smoking dose did not demonstrate associations with TBe grade.

Breathlessness and spirometry

Increasing breathlessness as measured by the mMRC Dyspnoea Scale was associated with increasing TBe grade (OR=1.1, 95% CI: 1.0–1.2, p=0.02). FVC % predicted was not associated with increasing grade from no ILA to any ILA (OR = 1.0, 95% CI: 1.0–1.1, p=0.3). FVC was similarly not different between participants without ILA and NF-ILA, and participants with fibrotic patterns (F-ILA and U-ILD) (OR = 1.0, 95% CI: 0.9–1.0, p=0.3). However, lower FVC was associated with U-ILD (OR = 0.9, 95% CI: 0.9–1.0, p=0.005) compared to those with no ILA, NF-ILA and F-ILA. The absolute difference in FVC % predicated between participants without ILA (median 86%, IQR: 72–98%) and those with U-ILD (median 84%, IQR: 74–97%) was small.

Higher FEV_1/FVC ratios were associated with increasing TBe grade (OR=1.2, 95% CI: 1.1–1.2, p<0.001). FEV_1/FVC and FVC had a negative interaction (β = -0.053, standard error = 0.013, p<0.001).

Respiratory hospitalisations

Participants with no ILA had 14.8 respiratory hospitalisations per 1000 person-years, whilst participants with NF-ILA had 18.1, those with F-ILA had 17.4, and those with U-ILD had 35.9 (Table 3). In a negative binomial regression, participants with U-ILD had higher incidence rate ratio (IRR) for respiratory hospitalisations compared to those without ILA (IRR=4.4, 95% CI: 2.7–7.5, p<0.001). Participants with non-fibrotic ILA (IRR = 1.6, 95% CI: 1.0–2.5, p=0.03) and fibrotic ILA (IRR=1.9, 95% CI: 1.1–3.3, p=0.01) also had increased respiratory hospitalisation rates compared to those without ILA. Age, smoking dose, reduced FEV_1/FVC, reduced FVC, and COPD

diagnosis were also associated with increased respiratory hospitalisation (Supplementary Table E3).

Overall survival probability

Kaplan-Meier estimates showed decreased survival probability with increasing TBe grade (log-rank test for trend p<0.0001, Supplementary Figure E3). In Cox proportional hazards models, after adjustment for age, gender, BMI, smoking status, smoking dose (PY), COPD diagnosis at baseline, and FEV_1, TBe U-ILD was associated with an adjusted Hazard Ratio, aHR=2.4 (95% CI: 1.9–3.0, p<0.001, Table E2) compared to those with no ILA.

Classification performance

Interobserver concordance

2,282 cases which underwent consensus review were used to assess concordance (Supplementary Figure E2). The TBe classification had excellent interobserver agreement (weighted kappa = 0.90, 95% CI: 0.89–0.91). This was improved compared to the ATS criteria (weighted kappa = 0.80, 95% CI: 0.78–0.83). When comparing unweighted kappa statistics, metrics for the TBe classification (0.76 [95% CI: 0.74–0.79]) was higher compared to the ATS criteria (0.70 [95% CI: 0.67–0.73]). Both the new TBe classification and the ATS Society classification showed stability in concordance within the analyzed subgroups of age, gender, ethnicity, smoking, and socioeconomic deprivation as measured by IMD (Figure 2).

Comparison with ATS Classification

The participants in each TBe/ILA category were cross tabulated for comparison (Figure 3, Table E3). Of the 8,169 participants, 88% (7,227) were concordantly classified. 11% (916) participants had a TBe classification higher than that of the ATS classification.

ATS classified 2.6% (185) of participants as U-ILD, whereas TBe classified 4.2% of participants (344) as U-ILD. The majority of ATS U-ILD (164) were concordantly classified by TBe, whereas an additional 180 participants were classified as U-ILD by TBe only, meaning that TBe classified 86% more U-ILD cases compared to ATS. Participants who were additionally classified by TBe as U-ILD and had not been classified by ATS as U-ILD had similar survival to those classified by ATS as U-ILD (χ^2=1.7, p=0.2, Supplementary Figure E6).

Participants classified as F-ILA or U-ILD by the TBe score showed similar clinical characteristics to those classified as U-ILD by the ATS Classification (Supplementary Table E4 and E5). Only 26 participants (0.3%) had a TBe classification of lower grade than that of the ATS Classification, of which 21 were classified using ATS as ILD based on non-infective ground glass changes considered extensive enough to represent ILD.

Sensitivity analyses with spirometry criteria was performed, where participants with radiological NF-ILA and F-ILA with FVC <80% were classified as U-ILD. Similar findings were seen, where TBe identified 69% more participants as U-ILD compared to ATS criteria (271 vs 392 participants, Figure E11). Similarly, increasing dyspnoea remained associated increasing TBe grade (Table E16). The trend for increasing

respiratory hospitalizations with increasing TBe grade was also preserved (Table E17 and Table E18).

Replication of findings in external data

Clinical characteristics and survival of patients classified using the TBe classification were similar between those in the development LCS dataset (SUMMIT) and the Southampton TLHC cohort (Table E8). In ordinal logistic modelling, increasing TBe grade was similarly associated with increasing age (OR=2.6, 95% CI 1.4–4.9, p=0.003), male gender (OR=4.3, 95% CI 2.3–8.3, p<0.001) and lower FVC (OR = 0.67, 95% CI 0.51–0.88, p=0.004; Table E9). Although no association with mMRC was found (OR=1.2, 95% CI 0.87–1.7, p=0.2), a reduced diffusing capacity for carbon monoxide (DLCO) was associated with increasing TBe grade (OR=0.69, 95% CI 0.57–0.82, p<0.001).

Participants with no ILA had the lowest rate of respiratory admission (14 admissions per 1000 person-years) compared to those with NF-ILA, F-ILA, and U-ILD (106, 38 and 55 admissions per 1000 person-years respectively, Table E10). In negative binomial regression models, participants with NF-ILA had a higher IRR of respiratory hospitalisation compared to those without ILA (IRR=4.4, 95% CI 1.1–22, p=0.05, Table E11), although the wide confidence interval points to a sparse data bias. No association was seen in those with F-ILA (IRR=1.3, 95% CI 0.23–7.5, p=0.8) and U-ILD (IRR=1.7, 95% CI 0.44–8.6, p=0.5).

The survival probability of participants identified as U-ILD was similar at 3 years between both LCS cohorts (SUMMIT, 0.88 [95% CI: 0.84–0.91], Southampton TLHC, 0.86 [95% CI: 0.79–0.94], Table E12).

Discussion

In this analysis, we describe a new ILA classification based on lobar traction bronchiolectasis (TBe score) developed in a large lung cancer screening cohort. Participants can be classified using our criteria with excellent interobserver concordance. Those with ILA and U-ILD are more breathless, have more respiratory hospitalisations, and a higher risk of death, compared to those without ILA. These findings have implications for lung cancer screening programmes to inform management pathways and recruitment into research studies for incidentally detected ILD.

To emphasise the detection of early disease phenotypes, the key imaging feature delineated in our TBe score was traction bronchiolectasis, an early manifestation of fibrosis. Traction bronchiolectasis has been linked to an increased risk of ILA progression to ILD, and its presence has been associated with genetic variants of *MUC5B*, which has a strong association with ILDs such as idiopathic pulmonary fibrosis (IPF) [19, 20]. Advances in CT scanner technology over the past 10 years have also allowed improved detection of traction bronchiolectasis when compared to the CT imaging acquired in the bioresources that largely informed the Fleischner Society and ATS ILA classification, such as the COPDGene study [21, 22].

Prior studies have demonstrated that the extent and progression of traction bronchiectasis and bronchiolectasis are associated with lung volume restriction, decline in gas exchange, and overall survival among participants with ILAs [23]. Hata *et al* [23] proposed the traction bronchiectasis index (TBI) as a severity-based tool for risk stratification in ILA populations. In contrast, our TBe score was designed as a complementary and categorical approach focused on the lobar distribution rather

than severity of traction bronchiolectasis. This was intended to enhance interpretability and applicability in lung cancer screening settings, where widespread subtle abnormalities are common and volume estimation can be challenging.

The recent ATS clinical statement [9] requires visual quantification of appropriate imaging features that involve >5% of a lung zone (for NF-ILA and F-ILA) and >5% of total lung volume for ILD. Although the ATS statement acknowledged that the 5% threshold of lung zone needs additional supporting data, no rationale was available to suggest an alternative. In contrast to the Fleischner and ATS ILA classification, the TBe score does not require assessment of lung volume percentages, which are challenging to estimate and replicate. Indeed, for the ATS definition of ILD, observer agreement is likely to be particularly challenging as it relies upon summing volumes of features which may exist in a number of different lobes. The ATS ILA classification also includes symptom and spirometry-based criteria as well as imaging domains. Whilst only pre-bronchodilator FVC was available (DLCO was unavailable) in our cohort, sensitivity analysis using both spirometry and imaging comparing TBe and ATS systems did not substantially change our results.

The prevalence of ILA or U-ILD in our cohort, whether classified by the TBe score (17%) or by the ATS classification (11%) was higher than previously reported. The difference in prevalence between the two classification systems is perhaps not surprising, because the TBe score did not require a minimum involvement of 5% of a lung zone for ILA classification. Indeed, removing the obstacle of having to be able to accurately determine lung volume involvement of ILA was a prime motivating factor for developing the TBe score. The higher prevalence of ILAs detected using our TBe score may mean more referrals and downstream investigations of patients

undergoing lung cancer screening. Nevertheless, our analyses demonstrate as a key outcome how the increased detection of ILA still identifies participants who are more breathless, have higher rates of respiratory hospitalization, and are at higher risk of death. This emphasizes that the higher TBe score prevalence represents the detection of more patients with symptoms and outcomes that an ILA classification should be finding.

In comparing the ILA or U-ILD prevalence in our cohort to previous reports in other populations, the most relevant classification is the ATS classification, since this clinical statement aligns most with the definitions used in the published literature. The total prevalence by the ATS classification of any ILA or U-ILD in our cohort was 11%, which is higher than the meta-analysed estimate in LCS cohorts of 7% [3]. When we interrogated the reported characteristics of the meta-analysed populations, we find that our cohort is older (by around 4.4 years) and more male (by around 2%). Since both age and gender are known associations with ILA, when we adjust the ATS classification prevalence in our SUMMIT population by the appropriate adjustment factors, we derive a prevalence of 6.9%. That is, if our SUMMIT population had the same age distribution and gender balance of those in the reported LCS meta-analysis, the ATS classification ILA prevalence would be 6.9%, very close to the reported 7% meta-analysis prevalence (Online Data Supplement Results). Importantly, participants classified as U-ILD by our TBe score who were deemed low risk by the ATS classification (non-fibrotic ILA) had similar survival than those who were concordantly classified as U-ILD, demonstrating the ability of the TBe score to better prognosticate their survival.

Spirometry revealed a complex relationship with ILA grade, with interaction between FEV_1/FVC and FVC. This interaction suggests that restrictive spirometry patterns (with low FVC, and preserved or elevated FEV_1/FVC) are more strongly associated with higher ILA grade than obstructive patterns and underscores that the interpretation of the decreasing FVC with undiagnosed ILD depends upon FEV_1/FVC.

Our findings of increased risk of death remain consistent in an external validation lung cancer screening dataset also from the UK. DLCO data in this cohort, which were not available in our discovery cohort, show a strong association between increasing ILA grade and decreasing DLCO. This is consistent with previously published evidence that DLCO decrease may be one of the earliest spirometric associations of ILA [19], a finding supported by the fact that a reduced FVC was not seen until the final threshold of ILA grade in our discovery cohort.

A methodological strength of our analysis here is the use of partial proportional odds models. Whilst introducing complexity, this allows for the modeling of changing effect sizes between classes, reducing reliance on proportional odds assumptions which are often violated in clinical research. Subgroup analyses of our classification system for protected characteristics such as sex and ethnicity demonstrated that it achieved good stability, with no difference in interobserver concordance between groups.

Our study has some important limitations. Throughout the study, related hypotheses across multiple domains (hospitalisations, breathlessness, and mortality) were tested, as a validation of the construction of this validation system. Interpretation of confidence intervals and p-values should take this into context. Only two thoracic radiologists scored all the scans in this analysis, and more precise estimates of

interobserver concordance will require further validation with other readers. Furthermore, the same radiologists who developed the TBe system scored the scans, which could be a source of bias favoring the TBe system. The absence of cause-specific mortality data limits our ability to determine whether the excess mortality associated with ILA was attributable to respiratory causes such as fibrosis, or whether ILAs are associated with excess comorbidity. However, our finding that respiratory hospitalisations are increased in participants with ILA, with risk increasing with grade of ILA, suggests that the burden of clinical symptoms in these patients is indeed respiratory. The size of our external validation dataset was small, meaning that only unadjusted mortality findings were replicated, given the sample size limitations on the number of possible parameters, with no consistent effect for respiratory hospitalisations seen due to its limited power. There were minimal numbers of participants with non-fibrotic and fibrotic ILAs, meaning power was limited to detect effects for these groups. The selection of participants in our validation cohort, being those referred on to the ILA evaluation service, presents a potential source of selection bias, but likely one that underestimates potential effects. However, it was chosen because of its comparable geography and setting, as a lung cancer screening setting in the UK.

In conclusion, with the increasing adoption of lung cancer screening, there is a new urgency for simple, practical criteria to stratify the findings of ILA on low dose CT. Criteria which are simple to apply with radiological information only, have excellent interobserver concordance, and which identify relevant clinical phenotypes, have the potential to broaden the reach and impact of lung cancer screening programs.

Tables

Characteristic	Overall, n=8169 (100%)	No ILA, n=6753 (82.6%)	NF-ILA, n=667 (8.2%)	F-ILA, n=405 (5.0%)	U-ILD, n=344 (4.2%)
Age, y	66 (60, 70)	65 (60, 70)	67 (62, 72)	69 (64, 73)	70 (66, 74)
Gender, female	3,274 (40%)	2,794 (41%)	236 (35%)	143 (35%)	101 (29%)
Smoking, py	40 (32, 51)	40 (32, 51)	41 (33, 51)	41 (33, 51)	43 (34, 56)
Unknown	71	60	5	3	3
Smoking status, current	4,004 (49%)	3,409 (50%)	290 (43%)	155 (38%)	150 (44%)
Smoking duration, y	45 (40, 50)	45 (40, 50)	46 (41, 52)	46 (40, 51)	48 (43, 53)
BMI, kg/m²	27.3 (24.2, 30.7)	27.1 (24.1, 30.6)	27.7 (24.6, 31.4)	28.4 (25.4, 31.3)	28.0 (25.1, 31.1)
Ethnicity					
Asian	782 (9.6%)	660 (9.8%)	48 (7.2%)	37 (9.1%)	37 (11%)
Black	499 (6.1%)	431 (6.4%)	32 (4.8%)	15 (3.7%)	21 (6.1%)
Mixed	178 (2.2%)	155 (2.3%)	12 (1.8%)	8 (2.0%)	3 (0.9%)
Other	229 (2.8%)	208 (3.1%)	10 (1.5%)	7 (1.7%)	4 (1.2%)
White	6,481 (79%)	5,299 (78%)	565 (85%)	338 (83%)	279 (81%)
Baseline COPD diagnosis	1,956 (24%)	1,591 (24%)	168 (25%)	97 (24%)	100 (29%)
FEV_1, %pred	77 (63, 89)	77 (63, 89)	79 (68, 90)	80 (68, 90)	77 (64, 88)
Unknown	825	749	38	18	20

FVC, %pred	87 (75, 98)	87 (75, 98)	89 (78, 100)	89 (79, 99)	85 (73, 98)
Unknown	825	749	38	18	20
FEV_1/FVC	70 (62, 76)	70 (62, 75)	69 (62, 75)	70 (63, 76)	70 (63, 77)
Unknown	825	749	38	18	20
mMRC Dyspnea Score					
0	2,645 (33%)	2,220 (33%)	214 (33%)	120 (30%)	91 (27%)
1	3,539 (44%)	2,919 (44%)	286 (44%)	181 (46%)	153 (46%)
≥2	1,846 (23%)	1,513 (23%)	149 (23%)	93 (24%)	91 (27%)
Unknown	139	101	18	11	9
IMD rank quintile					
1	2,620 (32%)	2,156 (32%)	213 (32%)	134 (34%)	117 (34%)
2	2,420 (30%)	2,024 (30%)	198 (30%)	98 (25%)	100 (29%)
3	1,443 (18%)	1,178 (18%)	121 (18%)	82 (21%)	62 (18%)
4	1,201 (15%)	997 (15%)	93 (14%)	62 (16%)	49 (14%)
5	402 (5.0%)	326 (4.9%)	37 (5.6%)	23 (5.8%)	16 (4.7%)
Unknown	83	72	5	6	0
Follow-up time, years	4.98 (4.73, 5.19)	4.98 (4.73, 5.19)	4.97 (4.72, 5.19)	5.00 (4.77, 5.19)	4.87 (4.54, 5.11)
Unknown	82	68	3	3	8
Total respiratory admissions	619	473	59	34	53
Death during follow-up	727 (9.0%)	533 (8.0%)	76 (11%)	37 (9.2%)	81 (24%)

Unknown	82	68	3	3	8

Table 1. Demographics and clinical characteristics of study participants. NF-ILA, non-fibrotic ILA; F-ILA, fibrotic ILA; U-ILD, undiagnosed ILD. Continuous variables presented as median (IQR); categorical variables presented as n (%).

	OR (95% CI)	p-value	OR for ILA grades*
Age, 10 y	2.4 (2.1–2.6)	<0.001	Any ILA vs None
	2.8 (2.5–3.3)	<0.001	NF-ILA/U-ILD vs NF-ILA/None
	3.3 (2.8–4.1)	<0.001	U-ILD vs All Others
Gender, male	1.5 (1.3–1.7)	<0.001	All
Smoking Status	0.9 (0.8–1.0)	0.168	All
Smoking dose, 10 py	1.0 (1.0–1.0)	0.808	All
BMI, 10 kg/m²	1.3 (1.2–1.5)	<0.001	All
FVC, 10 %pred†	1.0 (1.0–1.1)	0.332	Any ILA vs None
	1.0 (0.9–1.0)	0.342	NF-ILA/U-ILD vs NF-ILA/None
	0.9 (0.9–1.0)	0.005	U-ILD vs All Others
FEV_1/FVC‡	1.2 (1.1–1.2)	<0.001	All
mMRC Dyspnoea Scale	1.1 (1.0–1.2)	0.009	All
FVC %pred × FEV_1/FVC§	0.9 (0.9–1.0)	<0.001	All

*** Multiple** ORs **for variables violating proportional odds assumption**

† FVC % predicted centered around 100%

‡ FEV_1/FVC ratio centered around 0.75

§ Interaction term between FVC % predicted and FEV_1/FVC ratio

Table 2. Partial proportional odds model of increasing ILA grade. Thresholds reported for variables with non-proportional odds. Where parameters violate the proportional odds assumption, specific odds ratios (OR) given for comparisons between specific grades. BMI, body-mass-index; FEV_1, forced expiratory volume in first second; FVC, forced vital capacity; mMRC, modified Medical Research Council Dyspnoea scale.

ILA category (N)	Respiratory Hospitalizations				Overall Survival			
	Total respiratory admissions	Respiratory admissions rate (95% CI)*	IRR (95% CI)†	p-value	Deaths	Person-years at risk	HR (95% CI)‡	p-value
No ILA (6685)	473	14.83 (13.5–16.2)	—	—	533	31,895	—	—
NF-ILA (664)	59	18.81 (14.3–24.3)	1.61 (1.04–2.49)	0.025	76	3,136	1.34 (1.05–1.70)	0.020
F-ILA (402)	34	17.42 (12.1–24.3)	1.89 (1.10–3.25)	0.012	37	1,952	0.95 (0.68–1.34)	0.8
U-ILD (336)	53	35.88 (26.9–46.9)	4.44 (2.67–7.48)	<0.001	81	1,477	2.37 (1.87–3.01)	<0.001

*** Rate of respiratory admissions expressed per 1000 person-years (95% CI).**

† IRR = Incidence Rate Ratio. Estimates from negative binomial regression model, adjusted for age, gender, smoking status, smoking dose, BMI, FEV_1/FVC, FVC % pred, and baseline COPD diagnosis.

‡ HR = Hazard Ratio. Estimates from Cox proportional hazards model, adjusted for age, gender, smoking status, smoking dose, BMI, FVC % pred, and baseline COPD diagnosis.

Table 3. Respiratory hospitalization numbers, rates, and incidence rate ratios by ILA category in the SUMMIT cohort as estimated by multivariate negative binomial

regression. CI, confidence interval; IRR, incidence rate ratio. HR, Hazard Ratio; NF-ILA, Non-fibrotic ILA; F-ILA, Fibrotic ILA; U-ILD, Undiagnosed ILD.

Figures

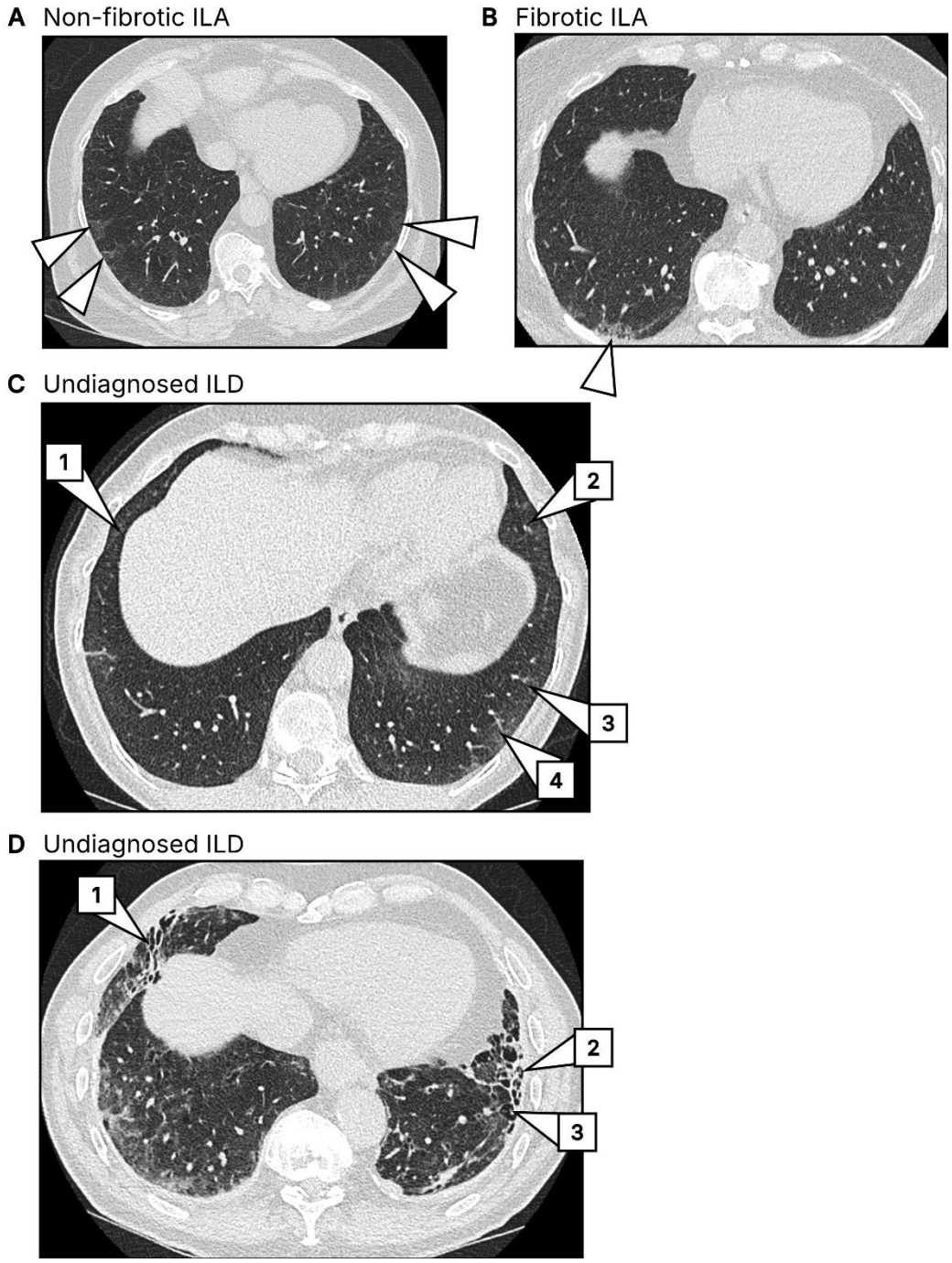


Figure 1. Axial CT images of cases classified by TBe score as A) non-fibrotic ILA (NF-ILA), B) Fibrotic ILA (F-ILA), and C and D) undiagnosed ILD (U-ILD). In panel A, arrows point to bilateral foci of ground glass change. In panel B, the enlarged image shows a focus of traction bronchiolectasis (arrowhead) occurring on the background

of dense lung parenchyma. Panel C, shows a case where the American Thoracic Society (ATS) criteria showed discrepant reads whereby one radiologist designated this as no ILA and a second radiologist designated this non-fibrotic ILA. The consensus ATS read was non-fibrotic ILA. The TBe classification was UILD based on bronchiole dilatation within dense lung parenchyma in the right middle (arrow 1), lingula (arrow 2) and left lower lobes (arrow 3 and 4).Panel D shows a case concordantly classified as UILD by the TBe score and ILD by the ATS criteria with traction bronchiectasis and bronchiolectasis in the right middle (arrow 1), lingula (arrow 2) and left lower lobes (arrow 3) occurring on the background of ground glass opacities and reticulation.

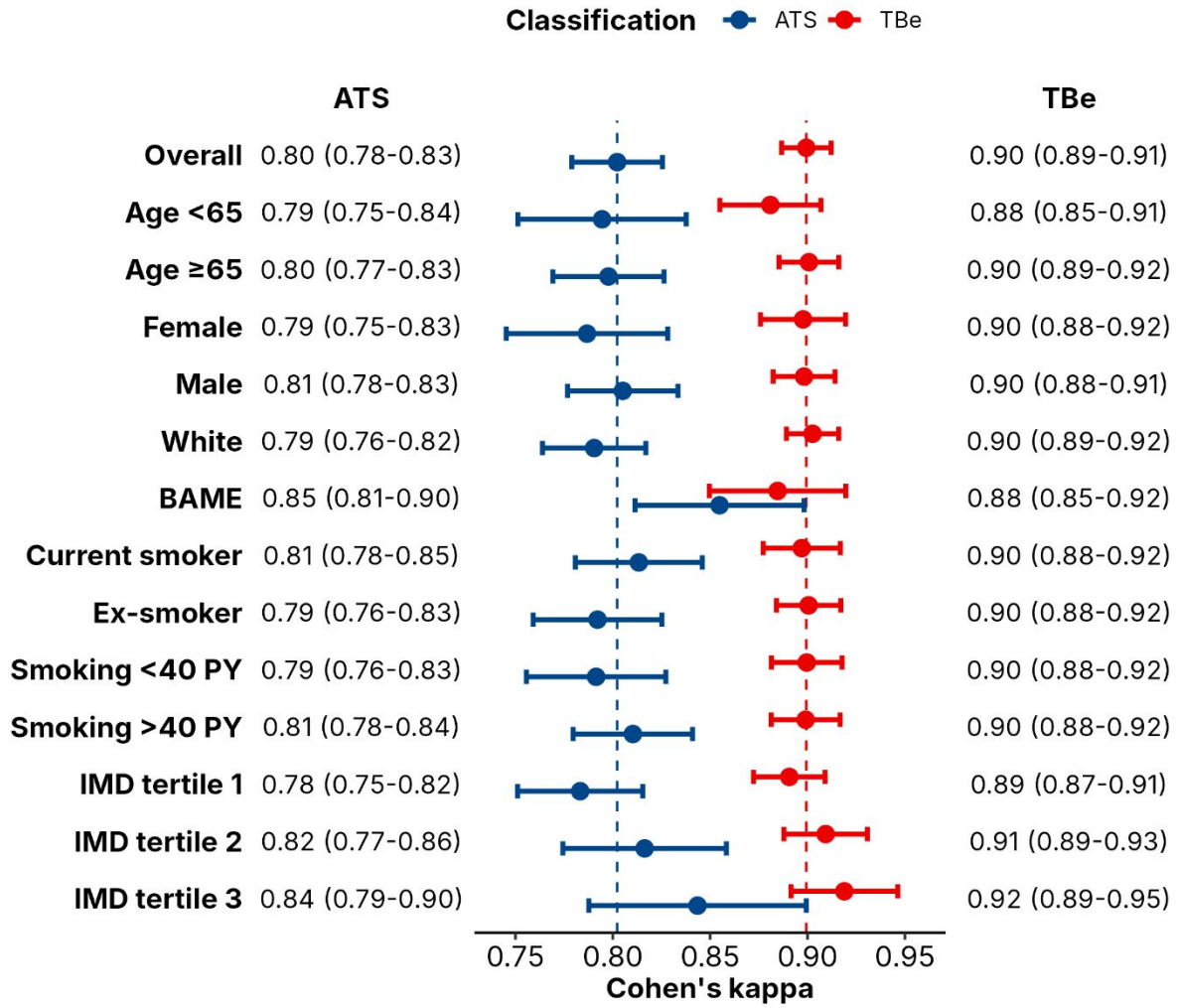


Figure 2. Comparison of interobserver agreement between methods including subgroup analyses in the SUMMIT cohort. Error bars represent weighted Cohen's kappa statistics and 95% confidence intervals. BAME, Black, Asian and Minority Ethnic; PY, pack years; IMD, Indices of Multiple Deprivation.

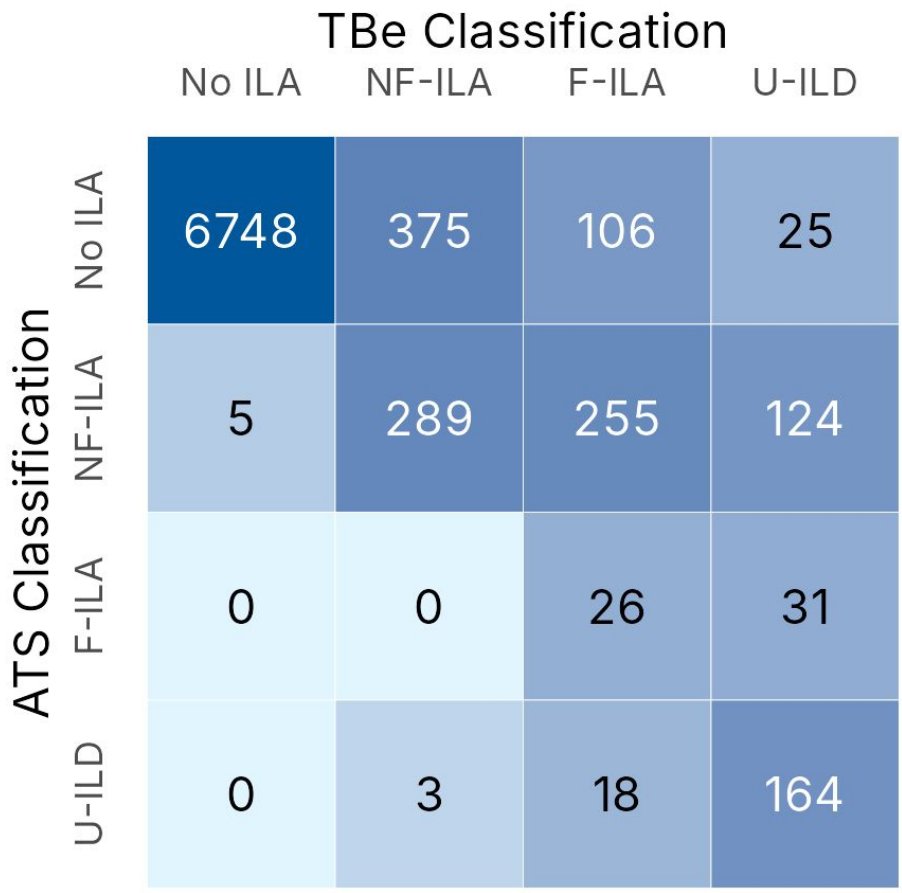


Figure 3. Cross tabulation of the cases classified by the new lobar traction bronchiolectasis (TBe) classification and the American Thoracic Society classification. Darker colours represent higher numbers of cases. NF-ILA: non-fibrotic ILA; F-ILA: fibrotic ILA; U-ILD: undiagnosed ILD.

References

1 New lung cancer screening roll out to detect cancer sooner - GOV.UK [Internet]. [cited 2025 Apr 14]. Available from: https://www.gov.uk/government/news/new-lung-cancer-screening-roll-out-to-detect-cancer-sooner.

2 Jin GY, Lynch D, Chawla A, *et al.* Interstitial lung abnormalities in a CT lung cancer screening population : Prevalence and progression rate. *Radiology* [Internet] Radiological Society of North America, Inc.; 2013 [cited 2025 Apr 14]; 268: 563–571. Available from: https://pubs.rsna.org/doi/10.1148/radiol.13120816.

3 Grant-Orser A, Min B, Elmrayed S, *et al.* Prevalence, Risk Factors, and Outcomes of Adult Interstitial Lung Abnormalities. *Am J Respir Crit Care Med* [Internet] American Thoracic Society; 2023 [cited 2025 Apr 29]; 208: 695–708. Available from: https://www.atsjournals.org/doi/pdf/10.1164/rccm.202302-0271OC?download=true.

4 Whittaker Brown SA, Padilla M, Mhango G, *et al.* Interstitial Lung Abnormalities and Lung Cancer Risk in the National Lung Screening Trial. *Chest* [Internet] Elsevier Inc; 2019 [cited 2025 Apr 29]; 156: 1195–1203. Available from: https://pubmed.ncbi.nlm.nih.gov/31404527/.

5 Sanders JL, Putman RK, Dupuis J, *et al.* The Association of Aging Biomarkers, Interstitial Lung Abnormalities, and Mortality. *Am J Respir Crit Care Med* [Internet] Am J Respir Crit Care Med; 2021 [cited 2025 Apr 14]; 203: 1149–1157. Available from: https://pubmed.ncbi.nlm.nih.gov/33080140/.

6 Ash SY, Harmouche R, Putman RK, *et al.* Clinical and Genetic Associations of Objectively Identified Interstitial Changes in Smokers. *Chest* [Internet] Chest; 2017 [cited 2025 Apr 14]; 152: 780–791. Available from: https://pubmed.ncbi.nlm.nih.gov/28506611/.

7 Putman RK, Hatabu H, Araki T, *et al.* Association Between Interstitial Lung Abnormalities and All-Cause Mortality. *JAMA* [Internet] American Medical Association; 2016 [cited 2025 Apr 14]; 315: 672–681. Available from: https://jamanetwork.com/journals/jama/fullarticle/2491637.

8 Hatabu H, Hunninghake GM, Richeldi L, *et al.* Interstitial lung abnormalities detected incidentally on CT: a Position Paper from the Fleischner Society. *Lancet Respir Med* [Internet] Lancet Publishing Group; 2020 [cited 2025 Apr 14]; 8: 726–737. Available from: https://www.thelancet.com/journals/lanres/article/PIIS2213-2600(20)30168-5/abstract.

9 Podolanczuk AJ, Hunninghake GM, Wilson KC, *et al.* AMERICAN THORACIC SOCIETY DOCUMENTS Approach to the Evaluation and Management of Interstitial Lung Abnormalities An Official American Thoracic Society Clinical Statement. 2025 [cited 2025 May 23]; . Available from: /doi/pdf/10.1164/rccm.202505-1054ST.

10 Hunninghake GM, Goldin JG, Kadoch MA, *et al.* Detection and Early Referral of Patients With Interstitial Lung Abnormalities: An Expert Survey Initiative. *Chest* [Internet] Elsevier Inc.; 2022 [cited 2025 Apr 14]; 161: 470–482. Available from: https://journal.chestnet.org/action/showFullText?pii=S0012369221012630.

11 Selvarajah B, Bhamani A, Azimbagihrad M, *et al.* Differentiating interstitial lung abnormalities and interstitial lung disease in lung cancer screening. *European Respiratory Journal* [Internet] European Respiratory Society; 2024 [cited 2025 Sep 1]; 64: PA3400. Available from: https://publications.ersnet.org/content/erj/64/suppl_68/PA3400.

12 Selvarajah B, Bhamani A, Azimbagirad M, *et al.* Differentiating Clinically Important Interstitial Lung Abnormalities in Lung Cancer Screening. 2024 [cited 2025 Sep 1]; . Available from: https://papers.ssrn.com/abstract=4874329.

13 Dickson JL, Hall H, Horst C, *et al.* Telephone risk-based eligibility assessment for low-dose CT lung cancer screening. *Thorax* [Internet] BMJ Publishing Group Ltd; 2022 [cited 2025 Sep 1]; 77: 1036–1040. Available from: https://thorax.bmj.com/content/77/10/1036.

14 Dickson JL, Hall H, Horst C, *et al.* Uptake of invitations to a lung health check offering low-dose CT lung cancer screening among an ethnically and socioeconomically diverse population at risk of lung cancer in the UK (SUMMIT): a prospective, longitudinal cohort study. *Lancet Public Health* [Internet] Elsevier Ltd; 2023 [cited 2025 Sep 1]; 8: e130–e140. Available from: https://www.thelancet.com/action/showFullText?pii=S2468266722002584.

15 Tammemägi MC, Katki HA, Hocking WG, *et al.* Selection Criteria for Lung-Cancer Screening. *New England Journal of Medicine* [Internet] Massachusetts Medical Society; 2013 [cited 2025 Apr 30]; 368: 728–736. Available from: https://www.nejm.org/doi/pdf/10.1056/NEJMoa1211776.

16 Brant R. Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics* JSTOR; 1990; 46: 1171.

17 Riley RD, Snell KIE, Ensor J, *et al.* Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med* [Internet] John Wiley & Sons, Ltd; 2019 [cited 2025 Apr 14]; 38: 1276–1296. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/sim.7992.

18 R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2025.

19 Hunninghake GM, Hatabu H, Okajima Y, *et al.* MUC5B Promoter Polymorphism and Interstitial Lung Abnormalities . *New England Journal of Medicine* [Internet] New England Journal of Medicine (NEJM/MMS); 2013 [cited 2025 Apr 14]; 368: 2192–2200. Available from: https://www.nejm.org/doi/full/10.1056/NEJMoa1216076.

20 Putman RK, Gudmundsson G, Araki T, *et al.* The MUC5B promoter polymorphism is associated with specific interstitial lung abnormality subtypes. *European Respiratory Journal* [Internet] European Respiratory Society; 2017 [cited 2025 Apr 14]; 50: 2025. Available from: https://publications.ersnet.org/content/erj/50/3/1700537.

21 Regan EA, Hokanson JE, Murphy JR, *et al.* Genetic Epidemiology of COPD (COPDGene) Study Design. *COPD: Journal of Chronic Obstructive Pulmonary Disease* 2011; 7: 32–43.

22 Rose JA, Menon AA, Hino T, *et al.* Suspected Interstitial Lung Disease in COPDGene Study. *Am J Respir Crit Care Med* [Internet] NLM (Medline); 2023 [cited 2025 Apr 14]; 207: 60–68. Available from: www.copdgene.org.

23 Hata A, Hino T, Li Y, *et al.* Traction Bronchiectasis/Bronchiolectasis in Interstitial Lung Abnormality: Follow-up in the COPDGene Study. *Am J Respir Crit Care Med* [Internet] American Thoracic Society; 2023 [cited 2025 Sep 1]; 207: 1395–1398. Available from: /doi/pdf/10.1164/rccm.202211-2061LE?download=true.

Acknowledgements

This research was funded in whole or in part by the Wellcome Trust (227835/Z/23/Z). DC is supported by a Clinical Research Training Fellowship from the Alliance for Cancer Early Detection (ACEDAS-2023/100003). DY is supported by grants from the fellowship of Astellas Foundation for Research on Metabolic Disorders. JJ and SMJ were also supported by the NIHR UCLH Biomedical Research Centre. This project has also been made possible in part by grant number CZIF2024-009938 from the Chan Zuckerberg Initiative Foundation. The SUMMIT Study is funded by GRAIL LLC. through a research grant awarded to SMJ as Principal Investigator.

SUMMIT consortium: Aashna Samson[4], Adefiola Olabode[4], Alberto Villanueva[4], Ali Mohammed[13], Alice Cotton[4], Andrew Crossingham[4], Andrew Perugia[22], Anjeli Ketkar[4], Anna Sikorska[4], Anthony Edey[4], Antonette Andrews[4], April Neville[4], Burcu Ozaltin[20], Camilla Mitchell[4], Catherine Nestor[4], Celia Snell[4], Charlie Sayer[4], Charlotte Cash[7], Chimtom Nwaosu[4], Christine Hosein[4], Claire Levermore[4], Columbus Ife[4], Derya Ovayolu[4], Dominique Arancon[4], Domminique Arancon[4], Eleanor Hellier[4], Elena Stefan[4], Elodie Murali[4], Esther Arthur-Darkwa[4], Ethaar El-Emir[4], Fabia Ruggiero[4], Fahiza Chowdhury[4], Fahmida Hoque[4], Fanta Bojang[4], Farhida Begum[4], Geoff Bellingan[4], Georgia Bullock[4], Graham Robinson[4], Gulistan Keskinbicak[4], Gulten Geneci[4], Harriet Sorrell[4], Hasti Robbie[4], Helen Kiconco[4], Helen Victoria Bowyer[4], Henna Makwana[4], Hina Pervez[4], Houda Benmebarek[4], James Rusius[22], Jane Rowlands[4], Janine Zylstra[4], Jeannie Eng[4], Jhanara Begum[4], Jon Teague[2], Joseph Jacob[20], Judy Airebamen[4], Julian McKee[4], Karen Parry-Billings[4], Karen Sennett[23], Kate Davies[1], Kate Gowers[1], Kaylene Phua[4], Kim Sorley[4], Kitty Chan[2], Kylie Gyertson[4], Laila Ahmed[4], Laura Farrelly[2], Laura Green[4], Lucas Clarke[4],

Lynsey Gallagher[4], Magali Taylor[4], Malavika Suresh[2], Mamta Ruparel[1], Marie Calligaris[4], Marilyn Ijeomah-Orji[4], Mark Clark[4], Maureen Browne[4], Mehran Azimbagirad[20], Miguel Ferreira[4], Milica Rajkov[4], Moksud Miah[4], Monika Sommerrey[4], Navinah Nundlall[4], Nazim Miah[4], Neil Songsong[4], Nicholas Beech[4], Nick Screaton[19], Nikita Sharma[4], Olufunmilola Ajayi[4], Oluwadara Akerewusi[4], Patricia Castello[4], Paul Hickey[4], Paul Robinson[4], Qin Neville[4], Rabiya Patel[4], Rachael Sarpong[2], Ricardo McEwen[4], Rosie Dapaah Tweneboah[4], Rowan Faiers[4], Sahana Uthayakumar[4], Sam Hare[4], Samanjit Hare[7], Sara Lock[8], Sebri Hasaballa[4], Shaira Hassan[4], Sheetal Karavadra[4], Shrinkhala Dawadi[2], Shummi Begum[4], Simranjit Mehta[4], Sofia Nnorom[4], Stephen Ellis[4], Stuart Marnham[4], Sumiya Mahey[4], Suraiya Sharmin[4], Syeda Anam Hussain[4], Sylvia Patricia Enes[4], Tania Anastasiadis[21], Tania Anastasiadis[21], Tanita Limani[4], Tanya Patrick[1], Terry O'Shaughnessy[13], Thea Buchan[4], Theepan Visakan[4], Thomas Callender[1], Tracy Odigie[4], Trisha Enes[4], Tunku Aziz[13], Urja Patel[4], Vicky Bowyer[4], Ya-Fei Li[4], Yee Chin Lee[4], Zaheer Mangera[12], Zahra Awan[4], Zahra Hanif[4], Zelena Aziz[4].

Affiliations: 1- Lungs For Living Research Centre, UCL Respiratory, University College London, London; 2- CRUK & UCL Cancer Trials Centre, University College London, London; 3- Centre for Cancer Screening, Prevention, Detection and Early Diagnosis, Wolfson Institute of Population Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London; 4- University College London Hospitals NHS Foundation Trust, London; 5- Royal Brompton and Harefield NHS Foundation Trust, London; 6- National Heart and Lung Institute, Imperial College, London; 7- Royal Free London NHS Foundation Trust, London; 8- Whittington Health NHS Trust, London; 9- Barking, Havering and Redbridge University Hospitals NHS Trust, Essex; 10- Homerton University Hospital Foundation

Trust, London; 11- The Princess Alexandra Hospital NHS Trust, Essex; 12- North Middlesex University Hospital NHS Trust, London; 13- Barts Health NHS Trust, London; 14- North Bristol NHS Trust, Bristol; 15- Royal United Hospitals Bath NHS Foundation Trust, Bath; 16- Surrey and Sussex Healthcare NHS Trust, Surrey; 17- King's College Hospital NHS Foundation Trust, London; 18- University Hospitals Sussex NHS Foundation Trust, Sussex; 19- Royal Papworth Hospital NHS Foundation Trust, Cambridge; 20-Satsuma Lab, Hawkes Institute, London; 21- Tower Hamlets Clinical Commissioning Group, London; 22- Noclor Research Support, London; 23- Islington Clinical Commissioning Group, London.

Declarations and Conflicts of Interest:

JJ declares consultancy fees from Boehringer Ingelheim, F. Hoffmann-La Roche, GlaxoSmithKline, NHSX; fees from advisory Boards for Boehringer Ingelheim, F. Hoffmann-La Roche; lecture fees from Boehringer Ingelheim, F. Hoffmann-La Roche, Takeda; grant funding from GlaxoSmithKline, Wellcome Trust, Microsoft Research, Gilead Sciences, Chan Zuckerberg Initiative and UK patent application numbers: 2113765.8 and GB2211487.0.

SMJ is supported by CRUK programme grant (EDDCPGM\100002), and MRC Programme grant (MR/W025051/1). SMJ receives support from the CRUK Lung Cancer Centre of Excellence (C11496/ A30025) and the CRUK City of London Centre, the Rosetrees Trust, the Roy Castle Lung Cancer foundation, the Longfonds BREATH Consortia, MRC UKRMP2 Consortia, the Garfield Weston Trust and University College London Hospitals Charitable Foundation. SMJ's work is supported by a Stand Up To Cancer-LUNGevity- American Lung Association Lung Cancer Interception Dream Team Translational Research Grant and Johnson and Johnson

(grant number: SU2C-AACR-DT23-17 to S.M. Dubinett and A.E. Spira). Stand Up To Cancer is a division of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the Scientific Partner of SU2C.This work was partly undertaken at UCLH/UCL who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centre's funding scheme. SMJ has received fees for advisory board membership in the last three years from Bard1 Lifescience. He has received grant income from GRAIL Inc. He is an unpaid member of a GRAIL advisory board. He has received lecture fees for academic meetings from Cheisi and Astra Zeneca. His wife works for Astra Zeneca.

AN has received a consultation fee from MSD and honoraria speaking engagements from Mevis Medical Solutions. AN is an Honorary Executive Committee member for the British Society of Thoracic Imaging, Co-chair of the British Thoracic Society Pulmonary Nodule Guideline Development Group, Lung Taskforce member for the British Lung Foundation and clinical lead for the NHS England Targeted Lung Health Checks Program.

DA is Director and a shareholder of Queen Square Analytics.

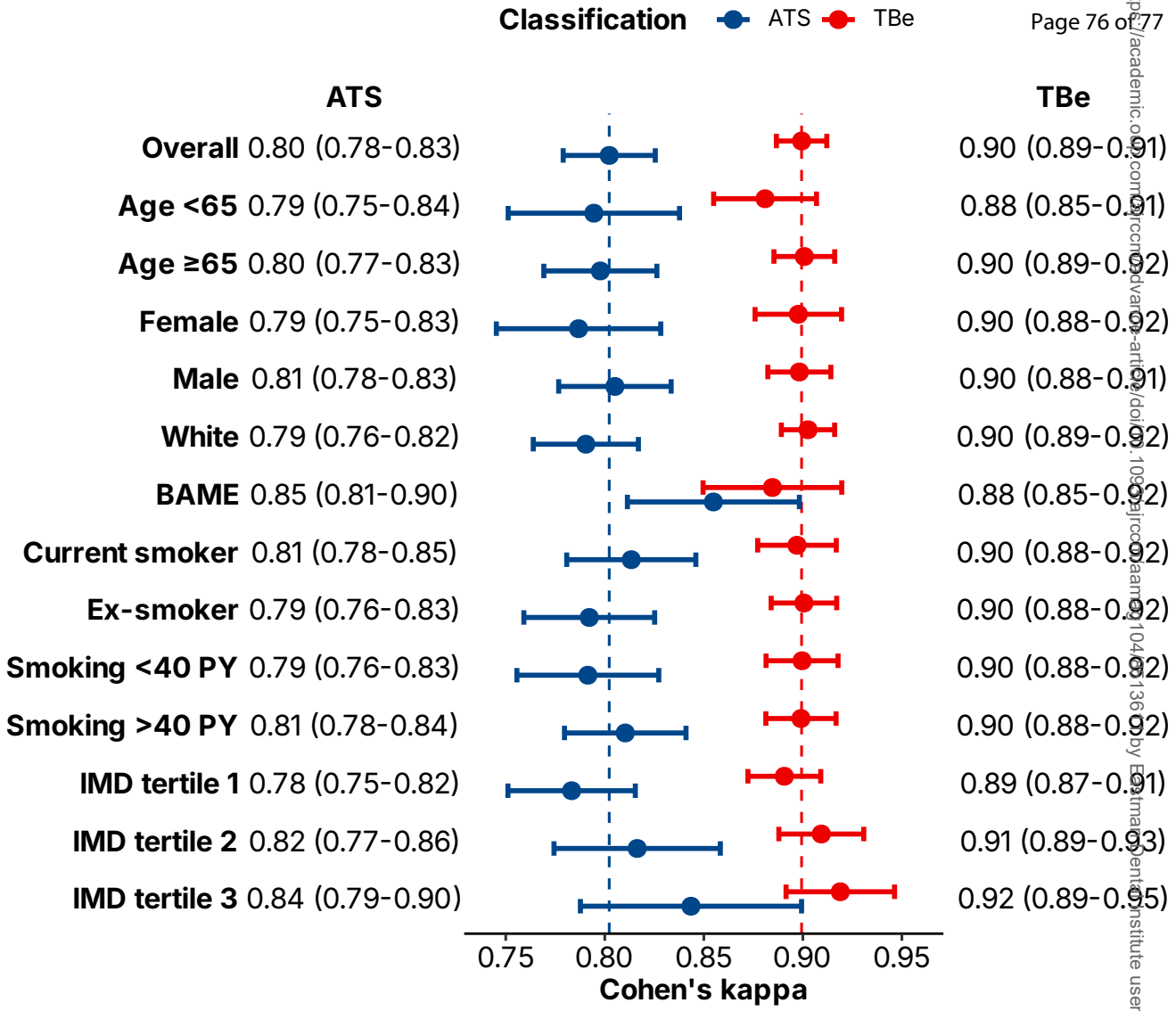
Competing interest declaration: "All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that (1) no authors have support from a company for the submitted work; (2) SMJ, JJ, AN have specified relationships with one or more of: AstraZeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, F. Hoffmann-La Roche, Gilead Sciences, Sanofi, Sanofi-Regeneron, Microsoft Research, Johnson and Johnson, Bard1 Lifescience , GRAIL, Mevis Medical

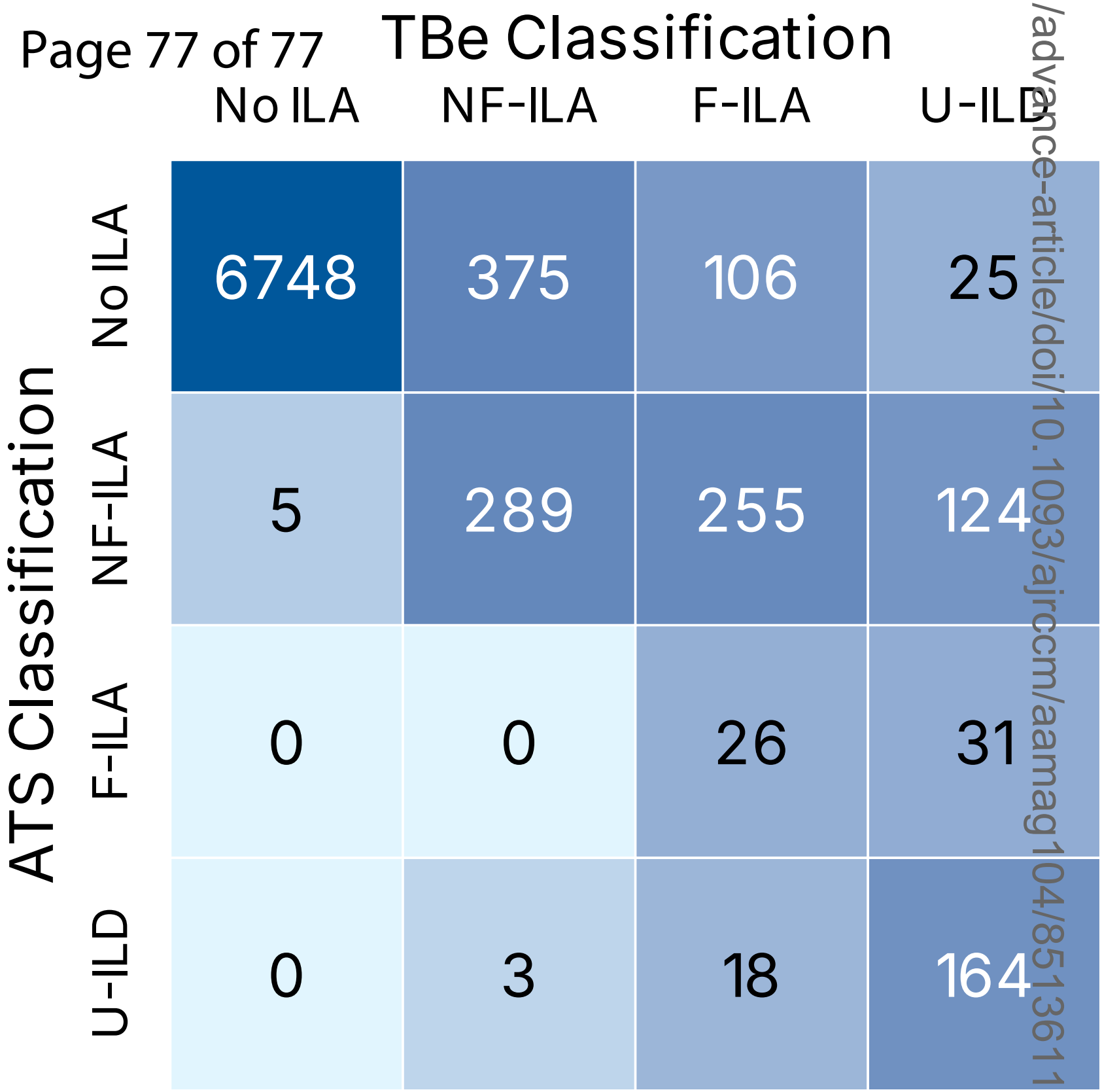
Solutions, MSD, Queen Square Analytics, Takeda that might have an interest in the submitted work in the previous 3 years; (3) SMJ's spouse works for AstraZeneca, but no other coauthors spouses, partners, or children have financial relationships that may be relevant to the submitted work; and (4) All authors have no non-financial interests that may be relevant to the submitted work."

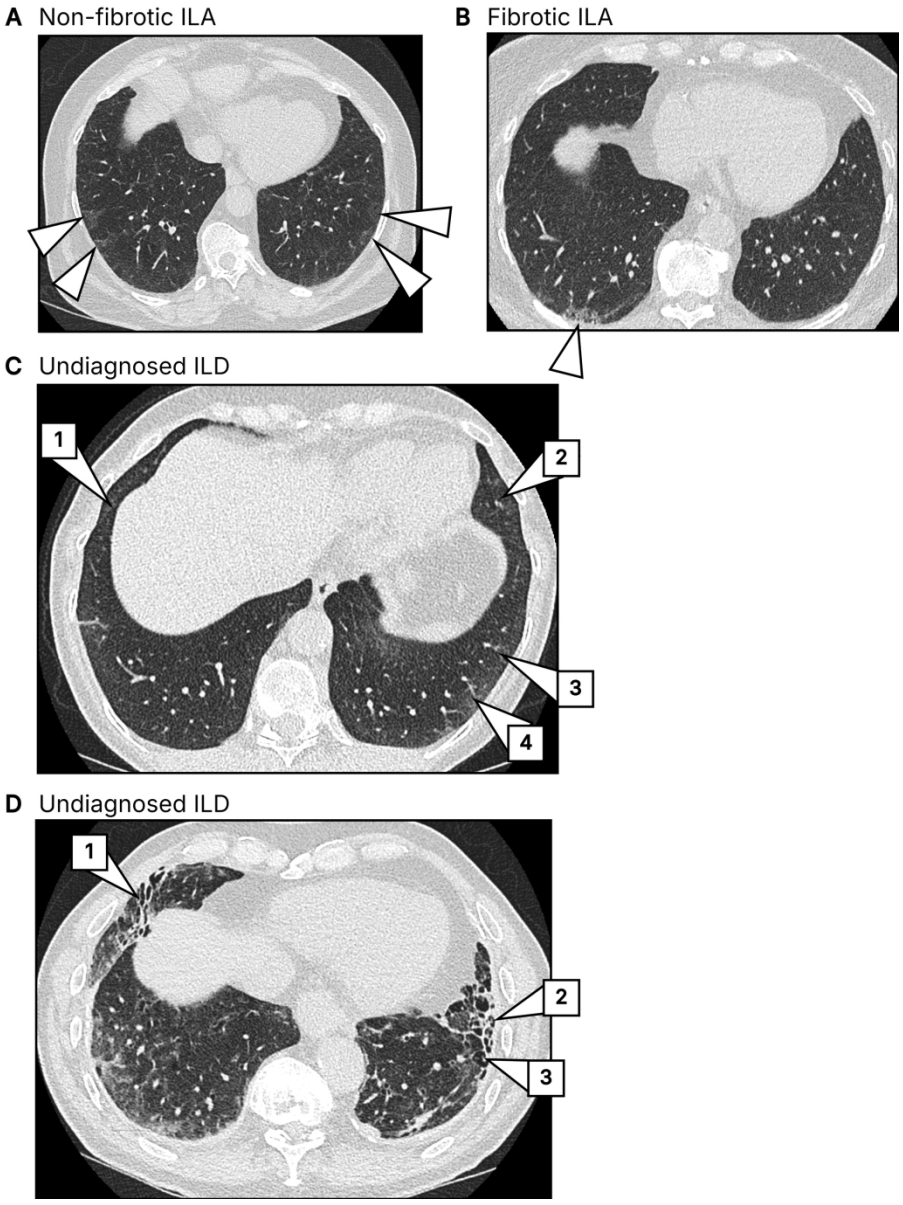
Role of study sponsors: The study sponsors did not have a role in design or in the collection, analysis, and interpretation of data.

Other declarations: The investigators were independent from the funders; JJ had full access to the data and can take responsibility for the integrity of the data and the accuracy of the data analysis; JJ (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Data sharing: No additional data available.







Axial CT images of cases classified by TBe score as A) non-fibrotic ILA (NF-ILA), B) Fibrotic ILA (F-ILA), and C and D) undiagnosed ILD (U-ILD). In panel A, arrows point to bilateral foci of ground glass change. In panel B, the enlarged image shows a focus of traction bronchiolectasis (arrowhead) occurring on the background of dense lung parenchyma. Panel C, shows a case where the American Thoracic Society (ATS) criteria showed discrepant reads whereby one radiologist designated this as no ILA and a second radiologist designated this non-fibrotic ILA. The consensus ATS read was non-fibrotic ILA. The TBe classification was UILD based on bronchiole dilatation within dense lung parenchyma in the right middle (arrow 1), lingula (arrow 2) and left lower lobes (arrow 3 and 4). Panel D shows a case concordantly classified as UILD by the TBe score and ILD by the ATS criteria with traction bronchiectasis and bronchiolectasis in the right middle (arrow 1), lingula (arrow 2) and left lower lobes (arrow 3) occurring on the background of ground glass opacities and reticulation.

192x254mm (300 x 300 DPI)