

Lessons from ECLIPSE: a review of COPD biomarkers

Rosa Faner,¹ Ruth Tal-Singer,² John H Riley,³ Bartolomé Celli,⁴ Jørgen Vestbo,^{5,6} William MacNee,⁷ Per Bakke,⁸ Peter M A Calverley,⁹ Harvey Coxson,¹⁰ Courtney Crim,¹¹ Lisa D Edwards,¹¹ Nick Locantore,¹¹ David A Lomas,¹² Bruce E Miller,² Stephen I Rennard,¹³ Emiel F M Wouters,¹⁴ Julie C Yates,¹¹ Edwin K Silverman,⁴ Alvar Agusti,^{1,15} on behalf of the ECLIPSE Study Investigators

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/thoraxjnl-2013-204778>).

For numbered affiliations see end of article.

Correspondence to

Dr Alvar Agusti, Institut del Tòrax, Hospital Clinic, Villarroel 170, Escala 3, Planta 5, Barcelona 08036, Spain; alvar.agusti@clinic.ub.es

RF and RT-S contributed equally

Received 8 November 2013
Accepted 21 November 2013
Published Online First
5 December 2013

ABSTRACT

The Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE) study was a large 3-year observational controlled multicentre international study aimed at defining clinically relevant subtypes of chronic obstructive pulmonary disease (COPD) and identifying novel biomarkers and genetic factors. So far, the ECLIPSE study has produced more than 50 original publications and 75 communications to international meetings, many of which have significantly influenced our understanding of COPD. However, because there is not one paper reporting the biomarker results of the ECLIPSE study that may serve as a reference for practising clinicians, researchers and healthcare providers from academia, industry and government agencies interested in COPD, we decided to write a review summarising the main biomarker findings in ECLIPSE.

INTRODUCTION

The Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE) study was a large 3-year observational controlled multicentre international study (Clinicaltrials.gov identifier NCT00292552; GSK study code SCO104960)¹ aimed at defining clinically relevant subtypes of chronic obstructive pulmonary disease (COPD) and identifying novel biomarkers and genetic factors.¹ It was a joint venture between academic researchers and GlaxoSmithKline, which fully funded it. So far, the ECLIPSE study has produced more than 50 original publications and 75 communications to international meetings (<http://www.eclipse-copd.com>), many of which have significantly influenced our understanding of COPD. This paper reviews the main biomarker findings in ECLIPSE (tables 1 and 2) in order to serve as a unified reference for practising clinicians, researchers and healthcare providers from academia, industry and government agencies interested in the field. Details of the design of the ECLIPSE study have been published elsewhere¹ and will not be reviewed here.

CELLULAR BIOMARKERS

Sputum neutrophils

Many previous studies have shown that sputum neutrophils are increased in smokers and in patients with COPD.² In a subset of patients included in ECLIPSE, we examined its long-term variability at 1 year follow-up and its relationship with a number of clinically relevant characteristics of COPD.³ In

168 subjects who provided valid induced sputum samples at baseline and 1 year later, the mean change over 1 year in neutrophils was an increase of 3.5%; however, most of the change was in patients with a low proportion of neutrophils at the baseline visit.³ On the other hand, in 488 patients with COPD we found that the proportion of neutrophils in sputum at baseline increased with GOLD stage. There was a weak but statistically significant association between percentage sputum neutrophils and both forced expiratory volume in 1 s (FEV₁) percentage predicted and health status (St George's Respiratory Questionnaire). By contrast, there were no associations between neutrophils and exacerbation rates or emphysema. Associations between sputum neutrophils and systemic biomarkers were non-significant or similarly weak.³ In summary, these observations suggest that sputum neutrophilia: (1) can be quantified reliably in multicentre trials using a standardised methodology; (2) is a relatively stable biomarker in COPD; and, (3) does not appear to be a major surrogate of clinical or pathophysiological abnormalities in COPD which limits its application in a clinical setting.

Circulating white blood cells

High circulating white blood cell (WBC) counts were weakly associated with persistent systemic inflammation,⁴ frequent exacerbations⁵ and mortality⁶ in ECLIPSE. A recent study in the general population has reported similar observations in relation to exacerbations.⁷ Although elevations in blood WBC were stable over time (at least in a subgroup of patients),⁴ numerical differences between patients with COPD and controls are small and often within the range of normal laboratory values. Overall, these results support the value of a high circulating WBC count as a relevant biomarker in COPD. Of note, similar results were obtained when neutrophil counts were assessed instead of total WBC.⁴⁻⁶ The potential role of other cellular biomarkers, such as circulating eosinophils, is currently being analysed.

PROTEIN BIOMARKERS

We initially explored 34 protein biomarkers in peripheral blood selected on the basis of previously published work⁸ and/or potential association with biological mechanisms believed to be relevant in the pathogenesis of COPD, including chemoattractants, tissue destruction/repair/remodelling and



To cite: Faner R, Tal-Singer R, Riley JH, *et al.* *Thorax* 2014;**69**:666–672.

Table 1 Summary of main findings in different biomarker ECLIPSE studies

| Type of biomarker | Main findings |
|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Cellular biomarkers | |
| Sputum neutrophils | 3.5% variability at 1-year follow-up ³ |
| Sputum neutrophils | Weak/absent association with FEV ₁ %, SGRQ, emphysema, systemic inflammatory markers, exacerbation frequency or lung function decline ³ |
| Circulating WBC | Associated with persistent systemic inflammation, ⁴ frequent exacerbations ⁵ and mortality ⁶ |
| Blood protein biomarkers | |
| Fibrinogen | Significantly associated with symptoms, exercise capacity, exacerbation rate and BODE index. ^{6–10} Currently undergoing a regulatory qualification process ¹¹ |
| CC16 | Weakly associated with lung function decline, emphysema and depression ^{12–15} |
| SP-D | Weak association with COPD exacerbations ^{5, 16} ; sensitive to treatment with oral and inhaled corticosteroids ¹⁶ |
| CCL18 (PARC) | Increased risk of cardiovascular hospitalisation or mortality ¹⁷ |
| sRAGE | Lower circulating sRAGE levels are associated with emphysema severity, and genetic polymorphisms in the AGER locus are associated with circulating sRAGE levels ²² |
| Inflammome* | Patients with persistent systemic inflammation (16%) had higher mortality and exacerbation rate than patients without inflammation (30%). ⁴ Systemic inflammation was also associated with heart disease, hypertension and diabetes ¹⁸ |
| Adipokines | Leptin and adiponectin levels were (+) and (–) related to CRP, respectively; BMI and gender were the strongest determinants of both adipokines ²⁰ |
| Vitamin D | Low levels of vitamin D were related to emphysema, 6MWD, airways reactivity and CC-16 levels ²¹ |
| Sputum transcriptomics | |
| Airflow limitation | 277 genes associated with the severity of airflow limitation (GOLD stage) ⁴⁸ |
| Emphysema | 198 genes associated with the presence of emphysema ⁴⁸ |
| Blood biomarkers | SNPs affecting circulating CC16 protein levels were significantly associated with sputum mRNA expression of SCGB1A1, the CC16 coding gene on chromosome 11 ⁴⁷ |
| COPD susceptibility | An integrative genomics approach located potential functional variants in two genes located within a COPD GWAS locus on chromosome 15 (CHRNA5 and IREB2) and one locus in the HLA-C region on chromosome 6 ⁵¹ |
| Serum metabolomics | |
| Serum profile | Results indicate that ^{52–53} : (1) there is increased protein turnover in all COPD patients; (2) increased protein degradation in individuals with emphysema and cachexia |
| Exhaled breath condensate | |
| pH | Lower pH in COPD and smoking controls, but not related to FEV ₁ or sputum leucocyte counts and not responsive to steroid treatment ⁵⁴ |
| Adenosine/purines | Increased concentrations in COPD patients and adenosine levels correlated with FEV ₁ ⁵⁵ |

*Inflammome: combined panel of six inflammatory markers in serum (WBC, CRP, interleukin 6, interleukin 8, tumour necrosis factor α and fibrinogen).

BMI, body mass index; COPD, chronic obstructive pulmonary disease; CRP, C-reactive protein; FEV₁, forced expiratory volume in 1 s; FFMI, fat-free mass index; GWAS, genome-wide association studies; 6MWD, 6-minute walking distance; PARC, pulmonary and activation-regulated chemokine; SGRQ, St George's Respiratory Questionnaire; SNP, single nucleotide polymorphism; sRAGE, soluble receptor for advanced glycation end products; WBC, white blood cell.

Table 2 Summary of main findings in genetic biomarkers in ECLIPSE studies

| Gene studies | |
|---------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Smoking history | Suggestive associations identified for age at smoking initiation (chromosomes 2q21 and 6p21), lifetime mean number of cigarettes per day (CHRNA3/CHRNA5 and CYP2A6), current number of cigarettes smoked per day (CYP2A6) and smoking cessation (DBH) ³⁰ |
| COPD susceptibility | Several genomic regions associated with COPD susceptibility (FAM13A, HHIP, CHRNA3/CHRNA5/IREB2 and a region on chromosome 19). Others (ADAM19, FGF7 and SP-D) need replication in other populations |
| COPD subtypes | CHRNA3/5 significantly associated with pack-years, emphysema and airflow limitation ⁴² ; HHIP not associated with pack-years but related to FEV ₁ /FVC, lean body mass and exacerbation. ⁴² |
| Emphysema | Borderline genome-wide significant association with BICD1 ⁴³ |
| Cachexia | Suggestive association of BMI and FFMI with FTO gene. ⁴⁶ The latter also related to FEV ₁ ⁴⁶ |
| Blood biomarkers | Genome-wide significant associations identified only for CC16 (chromosome 11) and SP-D (SFTPD and SNPs on chromosomes 6 and 16) ⁴⁷ |

BMI, body mass index; COPD, chronic obstructive pulmonary disease; DBH, dopamine beta-hydroxylase; ECLIPSE, Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points; FEV₁, forced expiratory volume in 1 s; FFMI, fat-free mass index; FVC, forced vital capacity; HHIP, hedgehog interacting protein; SNP, single nucleotide polymorphism; SP-D, surfactant protein D.

other inflammatory markers. Prior to the assessment of samples in the full ECLIPSE cohort, each biomarker assay was validated in terms of its sensitivity, accuracy, precision and reproducibility because, except for plasma fibrinogen and C-reactive protein (CRP), there was a lack of validated clinical laboratory assays. Repeatability was assessed by calculating the proportion of values at 3 months that were within 25% of the baseline value, as this reflects the typical total error associated with 'research-grade' immunoassays. We did this in what we called 'the ECLIPSE biomarker cohort', which consisted of 201 former smokers with COPD and 74 controls (37 ex-smokers with normal lung function and 37 healthy non-smokers) selected as a representative sample of the full ECLIPSE cohort.⁹ Biomarkers fulfilling these validation criteria included plasma CRP and fibrinogen, serum interleukin (IL)-8 and the predominantly lung-derived biomarkers (so-called 'pneumoproteins') surfactant protein-D (SP-D), club cell protein (formerly Clara cell secretory protein-16 (CC16)) and CCL18 (also known as pulmonary and activation-regulated chemokine (PARC)).⁹

These validated biomarkers were then explored in the entire ECLIPSE population as a replication of the initial findings and to investigate their relationships with age, gender, smoking status, clinical characteristics and outcomes. The main results were: (1) plasma fibrinogen is the most robust (in terms of relative longitudinal stability) biomarker investigated so far. It is significantly associated with symptoms, exercise capacity,

exacerbation rate, the BODE index and mortality.^{6–10} Plasma fibrinogen is currently being considered a potential candidate for regulatory qualification as a prognostic biomarker¹¹; however, more research is needed to fully assess its utility as an outcome measure in clinical trials and as a personalised risk factor in clinical practice; (2) CC16¹² was weakly associated with lung function decline,¹³ emphysema¹⁴ and depression¹⁵; (3) SP-D showed some weak association with COPD exacerbations^{5–16} and appears to be sensitive to treatment with oral and inhaled corticosteroids¹⁶; and (4) in a joint analysis of the ECLIPSE and Lung Health Study cohorts,¹⁷ serum CCL18 (PARC) was associated with an increased risk of cardiovascular hospitalisation or mortality.

We used a combined panel of six inflammatory markers in serum (WBC, CRP, IL-6, IL-8, tumour necrosis factor α (TNF α) and fibrinogen) to define a systemic ‘inflammome’ and we observed that the systemic ‘inflammome’ of smokers with normal lung function (basically characterised by increased circulating levels of WBC, IL-8 and TNF α) was different from that of smokers with COPD (characterised by a further increase in WBC plus abnormal serum levels of CRP, IL-6 and fibrinogen).⁴ We also found that 30% of patients with COPD did not have evidence of systemic inflammation (neither at baseline nor after 1-year follow-up), whereas 16% of them had persistent systemic inflammation (defined by the presence of two or more abnormal inflammatory markers, both at baseline and after 1-year follow-up). Finally, and importantly, we observed that mortality in patients with persistent systemic inflammation was six times higher than in patients without inflammation and their rate of exacerbations was double.⁴

Other studies have directly investigated the relationship between multi-morbidity and systemic inflammation in COPD and found that systemic inflammation was associated with the presence of heart disease, hypertension and diabetes.¹⁸ Likewise, since obesity and cachexia are common and relevant clinical problems in COPD,¹⁹ we compared adipokine metabolism (and markers of systemic inflammation) in 136 patients with COPD and 113 controls from the ECLIPSE cohort matched for age, gender and body composition.²⁰ The main results showed that: (1) CRP, IL-6, fibrinogen and adiponectin serum levels were higher in patients with COPD; (2) CRP levels were positively related to leptin and inversely related to adiponectin; and (3) body mass index (BMI) and gender were the strongest determinants for both leptin and adiponectin levels.²⁰

Vitamin D has been related to lung function in several previous studies of COPD, and a relationship between low levels of vitamin D in blood and emphysema, 6 min walk distance, airways reactivity and blood CC-16 levels was confirmed in a subset of 498 patients in ECLIPSE.²¹

Finally, in a very recent combined analysis we assessed the relationships of soluble receptor for advanced glycation end products (sRAGE) and CT-defined emphysema and found that lower circulating sRAGE levels are associated with emphysema severity and genetic polymorphisms in the AGER locus (in the gene coding for RAGE) were associated with circulating sRAGE levels.²²

Overall, these studies have identified several systemic biomarkers that alone or in combination have potential relevance for the enrichment of clinical trials aimed at validating future therapeutic interventions.

GENETIC STUDIES

The large number of patients with COPD included in the ECLIPSE study allowed the performance of genetic studies. In

order to validate the findings from ECLIPSE or to reveal additional common variants that contribute to COPD susceptibility, analyses were often performed in collaboration with other large cohorts of patients with COPD and/or controls, including the GenKOLS cohort (Bergen, Norway),²³ the International COPD Genetics Network (ICGN),²⁴ the National Emphysema Treatment Trial (NETT),²⁵ the Normative Aging Study (NAS),²⁶ the Lung Health Study (LHS)²⁷ and the COPDGene cohort.²⁸ The number of patients and controls included in each of the following studies varied accordingly. Below we summarise the results of these genetic studies in relation to: (1) the smoking history; (2) the susceptibility to develop COPD among smokers; and/or (3) the occurrence of different COPD characteristics in those smokers who have already developed the disease. We refer to ‘genome-wide significant’ results as those associations with p values $<5 \times 10^{-8}$, in order to adjust for the multiple comparisons involved in testing a genome-wide single nucleotide polymorphism (SNP) panel.

Genes associated with smoking history

Tobacco smoking is the main risk factor for COPD.²⁹ To identify SNPs associated with smoking intensity and behaviour, genome-wide association studies (GWAS) were conducted in four independent cohorts encompassing 3441 ever-smoking patients with COPD (GOLD stage ≥ 2).³⁰ No genomic regions were identified that reached genome-wide significance for any of the smoking-related phenotypes in this population of smokers and ex-smokers with COPD. However, suggestive association results were found for several loci associated with age at smoking initiation (SNPs in an intergenic region on chromosome 2q21 and near the HLA region on chromosome 6p21), lifetime mean number of cigarettes per day (SNPs in *CHRNA3/CHRNA5* and cytochrome P450, family 2, subfamily A, polypeptide 6 (*CYP2A6*)), current number of cigarettes smoked per day (*CYP2A6*) and smoking cessation (SNP rs3025343 in dopamine β -hydroxylase (*DBH*) locus).³⁰ These results strongly support a genetic basis for smoking initiation, maintenance and cessation.

Genes associated with COPD susceptibility

Because not all smokers develop COPD²⁹ and the disease clusters in families,³¹ it has long been proposed that there are specific genetic abnormalities that increase the susceptibility of some smokers to develop the disease. Several analyses explore this hypothesis in ECLIPSE.

An initial GWAS meta-analysis (including ECLIPSE) in 2940 patients with COPD and 1380 current or former smokers with normal lung function identified a new susceptibility locus at 4q22.1 in *FAM13A* and replicated this association in one case-control group ($n=1006$) and two family-based cohorts ($n=3808$) (rs7671167).³² Two previously reported genome-wide significant COPD GWAS regions near hedgehog interacting protein (*HHIP*) and *CHRNA3/CHNRA5/IREB2* also showed evidence for association. Because a larger sample size in GWAS may identify additional loci associated with COPD susceptibility, we extended our GWAS to 3499 cases and 1922 control subjects from four different cohorts (ECLIPSE, NAS and NETT, GenKOLS and COPDGene) pooled together.³³ The results identified a new genome-wide significant locus on chromosome 19q13 (rs7937). Genotyping this SNP and another nearby SNP in linkage disequilibrium (rs2604894) in 2859 subjects from the ICGN demonstrated supportive evidence of their association with COPD, pre-bronchodilator FEV₁ and severe COPD (GOLD stages 3 and 4). This region includes *RAB4B*, *EGLN2*,

MIA and CYP2A6, and has previously been identified in association with cigarette smoking behaviour.³⁴ Finally, previous GWAS meta-analyses of lung function in general population samples^{35–36} identified genome-wide significant evidence for association of multiple novel loci with two key spirometric variables describing airflow limitation in COPD (FEV₁ and FEV₁/forced vital capacity (FVC)). To investigate if a subset of these markers could also be associated with COPD susceptibility, 32 SNPs in or near 17 genes in 11 previously identified GWAS spirometric genomic regions were tested for association with COPD status in four COPD case-control study samples (NETT/NAS, GenKOLS, ECLIPSE and the first 1000 subjects in COPDGene; the total sample thus consisted of 3456 cases and 1906 controls).³⁷ Three loci harboured SNPs with suggestive evidence for an association with COPD susceptibility at a 5% false discovery rate: the 4q24 locus including FLJ20184/INTS12/GSTCD/NPNT, the 6p21 locus including AGER and PPT2 and the 5q33 locus including ADAM19.

Because SP-D is an immunomodulatory pneumoprotein essential to host defence and because we had identified it as a potentially relevant biomarker in COPD^{5–16} (see above), we hypothesised that polymorphisms in SP-D could influence the susceptibility to COPD. Indeed, we found that four SP-D SNPs (rs2245121, rs911887, rs6413520 and rs721917) showed suggestive associations with susceptibility to COPD (but not at genome-wide levels of significance), and that multiple SP-D SNPs were strongly associated with serum SP-D levels.³⁸

Homozygosity haplotype analysis is a very efficient and effective methodology for identifying potential disease-linked regions.³⁹ Using this approach, we identified 2318 regions of conserved homozygosity haplotype, of which 576 were significantly ($p < 0.05$) over-represented in patients with COPD.⁴⁰ After applying the weights constructed from these regions in the abovementioned GWAS of COPD,³² we identified two SNPs (rs12591300 and rs4480740) in a novel gene (fibroblast growth factor-7 (FGF7)) with suggestive evidence for an association with COPD susceptibility.⁴⁰

Finally, we used 'mediation analysis' in 3424 COPD cases and 1872 controls to estimate the direct (ie, independent from smoking) and indirect (ie, mediated by smoking) effects of three loci previously associated with COPD development.⁴¹ The results showed that the AGPHD1/CHRNA3, IREB2, FAM13A and HHIP loci had direct effects on COPD development, although the association of the AGPHD1/CHRNA3 locus is significantly mediated by cumulative exposure to tobacco smoke.⁴¹

Overall, these studies have contributed to the identification of several genomic regions that are associated with COPD at genome-wide significance, including FAM13A, HHIP, CHRNA3/CHRNA5/IREB2 and a region on chromosome 19. Several other genes and gene regions including ADAM19, FGF7 and SP-D showed evidence for an association with the development of COPD in smokers but will need to be replicated in additional populations.

Genes associated with different COPD subtypes

COPD is a complex and heterogeneous disease. The potential genetic contributions to this heterogeneity are unknown. To investigate it, we explored the association of a number of genetic loci with different COPD-related characteristics.

We tested the association of several key SNPs within the COPD GWAS regions identified above, such as the cholinergic nicotinic acetylcholine receptor (CHRNA3/5) and FAM13A genes and variants near HHIP, with several clinically relevant characteristics of COPD in the ECLIPSE cohort and then

validated the results in the ICGN cohort.⁴² We found that CHRNA3/5 was significantly associated with cumulative smoking exposure (pack-years), emphysema and airflow limitation in both populations.⁴² By contrast, HHIP was not associated with pack-years but it was related to the FEV₁/FVC ratio in both populations and with lean body mass and COPD exacerbations in ECLIPSE.⁴²

To explore the genetic basis of emphysema (as determined by chest CT), we used GWAS in 2380 patients with COPD from ECLIPSE, the GenKOLS cohort and NETT, and identified a borderline genome-wide significant association of BICD1 SNPs with the presence of emphysema as assessed by radiologist scores.⁴³ Since variants in BICD1 are associated with telomere length,⁴³ this observation suggests accelerated ageing as a potential mechanism involved in the development of emphysema.⁴⁴

The occurrence of cachexia in some patients with COPD is associated with increased mortality and increased emphysema.⁴⁵ To identify genetic susceptibility loci potentially related to cachexia in COPD (assessed by BMI or fat-free mass index (FFMI)), we performed a GWAS in patients with COPD pooled from the ECLIPSE study (n=1734), the GenKOLS cohort (n=851) and the NETT study (n=365).⁴⁶ We identified a suggestive association between an SNP (rs8050136) located in the first intron of the fat mass and obesity-associated (FTO) gene and both BMI and FFMI; this observation was replicated in 502 patients with COPD from the COPDGene cohort.⁴⁶ Interestingly, we also found a significant relationship between FEV₁ and FTO genotype.⁴⁶ All in all, these observations suggest a role for the FTO locus in the determination of body composition in COPD.

Finally, to investigate potential genetic determinants of the circulating levels of the protein biomarkers discussed above, we performed GWAS for two pneumoproteins (CC16 and SP-D) and five inflammatory markers (CRP, fibrinogen, IL-6, IL-8 and TNF α) in 1951 COPD subjects from ECLIPSE.⁴⁷ Genome-wide significant associations were identified only for the bloodstream levels of CC16 and SP-D. For CC16, two discrete genetic loci were identified; one was near the CC16 coding gene (SCGB1A1) on chromosome 11 while the other was located more than 20 Mb away on the same chromosome. Multiple SNPs near the coding gene (SFTPD) were associated with SP-D levels at genome-wide significance. In addition, SNPs on chromosomes 6 and 16 also demonstrated genome-wide significant associations with SP-D serum levels.⁴⁷

Overall, these studies have identified a number of novel genes associated with different COPD-related characteristics including: (1) airflow limitation (eg, CHRNA3/5, IREB2, HHIP, FTO); (2) emphysema (eg, CHRNA3/5, BICD1); (3) exacerbation frequency (eg, HHIP); (4) BMI (eg, HHIP, FTO); and (5) serum levels of two COPD-related pneumoproteins, CC16 and SP-D. Many of these associations were suggestive rather than genome-wide significant and will require replication in additional studies.

SPUTUM TRANSCRIPTOMIC STUDIES

ECLIPSE performed the largest sputum transcriptomic analysis ever reported in COPD. This has produced novel data (discussed below) but also highlights the feasibility of this type of analysis in large multicentre trials.

COPD characteristics

To identify candidate genes associated with the degree of airflow limitation and the extent of emphysema, sputum gene expression profiling was assessed in 148 former smokers with COPD (GOLD stages 2–4) from ECLIPSE and the findings were

replicated in a separate population of 176 patients using real-time PCR.⁴⁸ The results identified significant changes in 277 genes associated with the severity of airflow limitation (GOLD stage) and 198 genes with emphysema. Twelve candidate genes were selected from the microarray data set (based on a twofold change in expression between GOLD stage 2 vs GOLD stages 3 and 4) and 11 of them were validated by PCR in the replication cohort.⁴⁸ To illustrate the potential of these findings, one selected gene (IL-18R) was further analysed using immunohistochemistry in lung tissue, which demonstrated increased expression of IL-18R in COPD airway macrophages.⁴⁸ These results therefore have potential functional implications, given the role of IL-18 in neutrophil and macrophage activation as well as T cell development.^{49 50}

We also explored the relationship between sputum gene transcriptomics and several SNPs affecting circulating levels of several protein biomarkers (see above) including CC16, SP-D, CRP, fibrinogen, IL-6, IL-8 and TNF α . We found that several SNPs affecting circulating CC16 protein levels were significantly associated with sputum mRNA expression of SCGB1A1, the CC16 coding gene on chromosome 11.⁴⁷ This supports a coordinated regulation of CC16 expression, both systemically and in the lungs.⁴⁷

COPD susceptibility

To identify potential functional effects of known COPD susceptibility genes and to find novel disease gene candidates, we used an integrative genomics approach that combined analysis of GWAS (from ECLIPSE, Bergen, NETT and NAS) and gene expression data from relevant tissue samples (sputum) from 131 patients with COPD.⁵¹ This strategy located potential functional variants in two genes located within a COPD GWAS locus on chromosome 15 (CHRNA5 and IREB2) and has provided suggestive evidence for a novel COPD susceptibility locus in the HLA-C region on chromosome 6.⁵¹

Taken together, the above reviewed results on sputum RNA biomarkers show that: (1) sputum transcriptomics studies may be feasible in multicentre trials; (2) sputum gene expression profiling identified 277 genes associated with airflow limitation and 198 with emphysema⁴⁸; (3) several SNPs affecting circulating CC16 levels were associated with sputum mRNA expression of the CC16 gene, suggesting coordinated CC16 regulation systemically and in the lungs⁴⁷; and (4) an integrative genomics approach can identify potential COPD susceptibility loci.⁵¹

SERUM METABOLOMIC BIOMARKERS

Several ECLIPSE studies investigated the serum metabolomic profile of patients with COPD. We used proton nuclear magnetic resonance (¹H NMR) to compare the metabolomic serum profile of 1678 patients with COPD and 566 healthy smokers.⁵² The results showed: (1) decreased lipoproteins, N, N-dimethylglycine and increased glutamine, phenylalanine, 3-methylhistidine and ketone bodies in patients with COPD, with decreased branched chain amino acids (BCAAs) observed in patients with GOLD stage 4; (2) BCAAs and their degradation products (3-methylhistidine, ketone bodies and triglycerides) correlated negatively with body weight and positively with systemic inflammation; and (3) patients with emphysema displayed decreased serum creatine, glycine and N, N-dimethylglycine.⁵² Liquid chromatography with tandem mass spectrometry (LC-MS/MS) confirmed most of the ¹H NMR findings.⁵²

In a follow-up study⁵³ we used quantitative LC-MS/MS to measure 34 amino acids and dipeptides in different subgroups

of patients with COPD classified according to: (1) the severity of airflow limitation (GOLD stage 4 (n=30) vs controls (n=30)); (2) presence (n=38) or absence (n=21) of emphysema; and (3) cachexia (n=30) vs normal BMI (n=30). Targeted LC-MS/MS distinguished groups in all three categories.⁵³ In particular, glutamine, aspartate and arginine were significantly increased in patients with GOLD stage 4, emphysema or cachexia whereas amino adipate was decreased.⁵³

These results indicate that: (1) there is increased protein turnover in all patients with COPD, with increased protein degradation in individuals with emphysema and cachexia; and (2) while there are some promising metabolomic signals detected in the serum of certain subtypes of patients with COPD, replication in other cohorts is required.

EXHALED BREATH CONDENSATE

Several ECLIPSE studies explored exhaled breath condensate (EBC) as a potential source of valid COPD biomarkers. Using conventional methodology, we were not able reliably to measure protein biomarkers in EBC. However, we found that the pH of the EBC was consistently lower in COPD (but also in smoking controls) than in non-smokers, although it was not related to FEV₁ or sputum leucocyte counts, and that it was unresponsive to oral steroid treatment.⁵⁴ Using mass spectrometry, we found that the relative concentrations of adenosine and AMP were elevated in patients with COPD, and the former correlated with FEV₁.⁵⁵ Finally, EBC contains a complex mixture of volatile organic compounds (VOCs), some of which could potentially represent biomarkers for lung diseases.⁵⁶ We developed a special sampling methodology for collecting concentrated samples of exhaled air from participants with impaired respiratory function. By using two-stage thermal desorption gas chromatography-differential mobility spectrometry (GC-DMS) analysis, we then showed that it is possible to discriminate between healthy smokers and patients with COPD.⁵⁶ Mass spectrometry is then required to identify and quantify any discriminative biomarker. In summary, so far these studies do not support EBC as a useful source of biomarkers in COPD, although novel methodologies may help to explore potential new ones if care is taken to assess reproducibility with time and to validate the results in different cohorts.

OTHER LESSONS FROM THE ECLIPSE BIOMARKER EXPEDITION

The results discussed above clearly identify novel cellular, proteomic, genetic, transcriptomic and metabolomic biomarkers in COPD. Needless to say, many of them require validation in appropriately designed studies before routine use in clinical practice and as drug development tools. However, the 'ECLIPSE biomarker expedition' has also provided other equally important lessons for conducting biomarker analyses in large multicentre clinical studies. We believe that the following are worth considering:

1. Standardised methodology and strict quality control are critical for the assessment of biomarkers in a large multicentre trial.
2. Replication of data in different cohorts is imperative, highlighting the need for open collaborations. Plasma fibrinogen, serum CCL18 and sRAGE, as well as GWAS data on FAM13A, are specific examples of the power of a partnership approach in the data reviewed above.
3. ECLIPSE identified a panel of biomarkers associated with risks of poor clinical outcome. For newly emerging biomarkers, it is imperative to develop robust and cost-effective

diagnostic tests that can be qualified and that will allow comparison between studies as a replacement to research grade laboratory assays.

4. Transcriptomics and metabolomics pave the future by identifying pathways of interest, but data require replication and, very importantly, the use of cluster and network analysis for comprehensive systems biomedicine.⁵⁷ There are still several pending manuscripts of the ECLIPSE saga, but two of them address precisely this field of knowledge.^{58 59}

Author affiliations

¹Fundació Investigació Sanitària Illes Balears (FISIB), Ciber Enfermedades Respiratorias (CIBERES), Barcelona, Catalunya, Spain

²GlaxoSmithKline Research and Development, King of Prussia, Pennsylvania, USA

³GlaxoSmithKline Research and Development, Stevenage, UK

⁴Channing Division of Network Medicine and Pulmonary and Critical Care Division, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA

⁵Department of Respiratory Medicine, Odense University Hospital, and Clinical Institute, University of Southern Denmark, Odense, Denmark

⁶Respiratory and Allergy Research Group, Manchester Academic Health Sciences Centre, University Hospital South Manchester NHS Foundation Trust, Manchester, UK

⁷University of Edinburgh, MRC Centre for Inflammation Research, The Queen's Medical Research Institute, Edinburgh, UK

⁸Department of Thoracic Medicine, Institute of Clinical Science, University of Bergen, Haukeland University Hospital, Bergen, Norway

⁹Division of Infection and Immunity Clinical Sciences Centre, University Hospital Aintree, Liverpool, UK

¹⁰Department of Radiology, Vancouver General Hospital, University of British Columbia, Vancouver, British Columbia, Canada

¹¹GlaxoSmithKline Research and Development, Research Triangle Park, North Carolina, USA

¹²Division of Medicine, University College London, London, UK

¹³Department of Pulmonary and Critical Care Medicine, University of Nebraska Medical Center, Omaha, Nebraska, USA

¹⁴Department of Respiratory Medicine, Maastricht University Medical Center, Maastricht, The Netherlands

¹⁵Thorax Institute, Hospital Clinic, IDIBAPS, Univ. Barcelona, Barcelona, Spain

Acknowledgements The authors thank all the subjects, investigators and study site staff who participated in ECLIPSE and our many collaborators on the manuscripts cited herein.

Collaborators See online supplementary appendix for full list of ECLIPSE study investigators.

Contributors RF, RT-S, JV and AA planned the review, wrote the first draft, collated comments from the co-authors and are responsible for the overall content as guarantors. All the other co-authors read the draft, made comments and suggestions and approved the final manuscript. All authors (except RF) participate in the Steering Committee of ECLIPSE.

Funding The ECLIPSE study was funded by GlaxoSmithKline. Disclosures for each author are found in supplementary material.

Competing interests None.

Ethics approval Local Ethics Boards.

Provenance and peer review Not commissioned; internally peer reviewed.

REFERENCES

- Vestbo J, Anderson W, Coxson HO, *et al.* Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE). *Eur Respir J* 2008;31:869–73.
- Stanescu D, Sanna A, Veriter C, *et al.* Airways obstruction, chronic expectoration, and rapid decline of FEV₁ in smokers are associated with increased levels of sputum neutrophils. *Thorax* 1996;51:267–71.
- Singh D, Edwards L, Tal-Singer R, *et al.* Sputum neutrophils as a biomarker in COPD: findings from the ECLIPSE study. *Respir Res* 2010;11:77.
- Agusti A, Edwards LD, Rennard SI, *et al.* Persistent systemic inflammation is associated with poor clinical outcomes in COPD: a novel phenotype. *PLoS ONE* 2012;7:e37483.
- Hurst JR, Vestbo J, Anzueto A, *et al.* Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med* 2010;363:1128–38.
- Celli BR, Locantore N, Yates J, *et al.* Inflammatory biomarkers improve clinical prediction of mortality in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012;185:1065–72.
- Thomsen M, Ingebrigtsen TS, Marott JL, *et al.* Inflammatory biomarkers and exacerbations in chronic obstructive pulmonary disease. *JAMA* 2013;309:2353–61.
- Pinto-Plata V, Toso J, Lee K, *et al.* Profiling serum biomarkers in patients with COPD: associations with clinical parameters. *Thorax* 2007;62:595–601.
- Dickens JA, Miller BE, Edwards LD, *et al.* COPD association and repeatability of blood biomarkers in the ECLIPSE cohort. *Respir Res* 2011;12:146.
- Duvoix A, Dickens J, Haq I, *et al.* Blood fibrinogen as a biomarker of chronic obstructive pulmonary disease. *Thorax* 2013;68:670–6.
- Casaburi R, Celli B, Crapo J, *et al.* The COPD Biomarker Qualification Consortium (CBQC). *COPD* 2013;10:367–77.
- Lomas DA, Silverman EK, Edwards LD, *et al.* Evaluation of serum CC-16 as a biomarker for COPD in the ECLIPSE cohort. *Thorax* 2008;63:1058–63.
- Vestbo J, Edwards LD, Scanlon PD, *et al.* Changes in forced expiratory volume in 1 second over time in COPD. *N Engl J Med* 2011;365:1184–92.
- Coxson HO, Dirksen A, Edwards L, *et al.* Change in lung density measured using computed tomography in chronic obstructive pulmonary disease. *Lancet Respir (In press)* 2013.
- Hanania NA, Mullerova H, Locantore NW, *et al.* Determinants of depression in the ECLIPSE chronic obstructive pulmonary disease cohort. *Am J Respir Crit Care Med* 2011;183:604–11.
- Lomas DA, Silverman EK, Edwards LD, *et al.* Serum surfactant protein D is steroid sensitive and associated with exacerbations of COPD. *Eur Respir J* 2009;34:95–102.
- Sin DD, Miller BE, Duvoix A, *et al.* Serum PARC/CCL-18 concentrations and health outcomes in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2011;183:1187–92.
- Miller J, Edwards LD, Agusti A, *et al.* Comorbidity, systemic inflammation and outcomes in the ECLIPSE cohort. *Respir Med* 2013;107:1376–84.
- Rutten EP, Breyer MK, Spruit MA, *et al.* Abdominal fat mass contributes to the systemic inflammation in chronic obstructive pulmonary disease. *Clin Nutr* 2010;29:756–60.
- Breyer MK, Rutten EP, Locantore NW, *et al.* Dysregulated adipokine metabolism in chronic obstructive pulmonary disease. *Eur J Clin Invest* 2012;42:983–91.
- Berg I, Hanson C, Sayles H, *et al.* Vitamin D, vitamin D binding protein, lung function and structure in COPD. *Respir Med* 2013;107:1578–88.
- Cheng DT, Kim DK, Cockayne DA, *et al.* Systemic sRAGE is a biomarker of emphysema and associated with AGER genetic variants in COPD patients. *Am J Respir Crit Care Med* 2013;188:948–57.
- Bakke PS, Zhu G, Gulsvik A, *et al.* Candidate genes for COPD in two large data sets. *Eur Respir J* 2011;37:255–63.
- Zhu G, Warren L, Aponte J, *et al.* The SERPINE2 gene is associated with chronic obstructive pulmonary disease in two large populations. *Am J Respir Crit Care Med* 2007;176:167–73.
- National Emphysema Treatment Trial Research Group. A randomized trial comparing lung-volume-reduction surgery with medical therapy for severe emphysema. *N Engl J Med* 2003;348:2059–73.
- Litonjua AA, Lazarus R, Sparrow D, *et al.* Lung function in type 2 diabetes: the Normative Aging Study. *Respir Med* 2005;99:1583–90.
- Anthonisen NR, Connett JE, Kiley JP, *et al.* Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV₁. The Lung Health Study. *JAMA* 1994;272:1497–505.
- Regan EA, Hokanson JE, Murphy JR, *et al.* Genetic epidemiology of COPD (COPDgene) study design. *COPD: J Chron Obstruct Pulmon Dis* 2010;7:32–43.
- Vestbo J, Hurd SS, Agusti AG, *et al.* Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease: GOLD Executive Summary. *Am J Respir Crit Care Med* 2013;187:347–65.
- Siedlinski M, Cho MH, Bakke P, *et al.* Genome-wide association study of smoking behaviours in patients with COPD. *Thorax* 2011;66:894–902.
- DeMeo DL, Carey VJ, Chapman HA, *et al.* Familial aggregation of FEF₂₅₋₇₅ and FEF₂₅₋₇₅/FVC in families with severe, early onset COPD. *Thorax* 2004;59:396–400.
- Cho MH, Boutaoui N, Klanderman BJ, *et al.* Variants in FAM13A are associated with chronic obstructive pulmonary disease. *Nat Genet* 2010;42:200–2.
- Cho MH, Castaldi PJ, Wan ES, *et al.* A genome-wide association study of COPD identifies a susceptibility locus on chromosome 19q13. *Hum Mol Genet* 2012;21:947–57.
- Thorgeirsson TE, Gudbjartsson DF, Surakka I, *et al.* Sequence variants at CHRN3-CHRNA6 and CYP2A6 affect smoking behavior. *Nat Genet* 2010;42:448–53.
- Hancock DB, Eijgelsheim M, Wilk JB, *et al.* Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet* 2010;42:45–52.
- Repapi E, Sayers I, Wain LV, *et al.* Genome-wide association study identifies five loci associated with lung function. *Nat Genet* 2010;42:36–44.
- Castaldi PJ, Cho MH, Litonjua AA, *et al.* The association of genome-wide significant spirometric loci with chronic obstructive pulmonary disease susceptibility. *Am J Respir Cell Mol Biol* 2011;45:1147–53.
- Foreman MG, Kong X, DeMeo DL, *et al.* Polymorphisms in surfactant protein-D are associated with chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2011;44:316–22.

- 39 Jiang H, Orr A, Guernsey DL, *et al.* Application of homozygosity haplotype analysis to genetic mapping with high-density SNP genotype data. *PLoS ONE* 2009;4:e5280.
- 40 Brehm JM, Hagiwara K, Tesfaigzi Y, *et al.* Identification of FGF7 as a novel susceptibility locus for chronic obstructive pulmonary disease. *Thorax* 2011;66:1085–90.
- 41 Siedlinski M, Tingley D, Lipman PJ, *et al.* Dissecting direct and indirect genetic effects on chronic obstructive pulmonary disease (COPD) susceptibility. *Hum Genet* 2013;132:431–41.
- 42 Pillai SG, Kong X, Edwards LD, *et al.* Loci identified by genome-wide association studies influence different disease-related phenotypes in COPD. *Am J Respir Crit Care Med* 2010;182:1498–505.
- 43 Kong X, Cho MH, Anderson W, *et al.* Genome-wide association study identifies BICD1 as a susceptibility gene for emphysema. *Am J Respir Crit Care Med* 2011;183:43–9.
- 44 Faner R, Rojas M, MacNee W, *et al.* Abnormal lung aging in chronic obstructive pulmonary disease and idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2012;186:306–13.
- 45 Schols AM, Broekhuizen R, Weling-Scheepers CA, *et al.* Body composition and mortality in chronic obstructive pulmonary disease. *Am J Clin Nutr* 2005;82:53–9.
- 46 Wan ES, Cho MH, Boutaoui N, *et al.* Genome-wide association analysis of body mass in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2011;45:304–10.
- 47 Kim DK, Cho MH, Hersh CP, *et al.* Genome-wide association analysis of blood biomarkers in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012;186:1238–47.
- 48 Singh D, Fox SM, Tal-Singer R, *et al.* Induced sputum genes associated with spirometric and radiological disease severity in COPD ex-smokers. *Thorax* 2011;66:489–95.
- 49 Imaoka H, Hoshino T, Takei S, *et al.* IL-18 production and pulmonary function in chronic obstructive pulmonary disease. *Eur Respir J* 2008;31:287–97.
- 50 Kang MJ, Choi JM, Kim BH, *et al.* IL-18 induces emphysema and airway and vascular remodeling via IFN γ , IL-17A, and IL-13. *Am J Respir Crit Care Med* 2012;185:1205–17.
- 51 Qiu W, Cho MH, Riley JH, *et al.* Genetics of sputum gene expression in chronic obstructive pulmonary disease. *PLoS ONE* 2011;6:e24395.
- 52 Ubhi BK, Riley JH, Shaw PA, *et al.* Metabolic profiling detects biomarkers of protein degradation in COPD patients. *Eur Respir J* 2012;40:345–55.
- 53 Ubhi BK, Cheng KK, Dong J, *et al.* Targeted metabolomics identifies perturbations in amino acid metabolism that sub-classify patients with COPD. *Mol Biosyst* 2012;8:3125–33.
- 54 MacNee W, Rennard SI, Hunt JF, *et al.* Evaluation of exhaled breath condensate pH as a biomarker for COPD. *Respir Med* 2011;105:1037–45.
- 55 Esther CR Jr, Lazaar AL, Bordonali E, *et al.* Elevated airway purines in COPD. *Chest* 2011;140:954–60.
- 56 Basanta M, Jarvis RM, Xu Y, *et al.* Non-invasive metabolomic analysis of breath using differential mobility spectrometry in patients with chronic obstructive pulmonary disease and healthy smokers. *Analyst* 2010;135:315–20.
- 57 Barabasi AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet* 2011;12:56–68.
- 58 Menche J, Sharma A, Cho MH, *et al.* Development of a novel superVISED Shuffling Approach (VISA) to identify complex disease subtypes using gene expression data and application to chronic obstructive pulmonary disease. *Mol Syst Biol* (In press 2013).
- 59 Rennard SI, Vestbo J, Silverman EK, *et al.* Determination of COPD characteristics via unsupervised clustering on the ECLIPSE cohort. *Eur Respir J* 2011;38(Suppl 55):762s.