UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE

Academic Unit of Clinical and Experimental Sciences

Volume 1 of 1

Characterisation of lung sounds for early identification and monitoring of fibrotic Interstitial Lung Disease

by

Giacomo Sgalla

Thesis for the degree of Doctor of Philosophy

February 2017

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE

Clinical and Experimental Sciences

Thesis for the degree of Doctor of Philosophy

CHARACTERISATION OF LUNG SOUNDS FOR EARLY IDENTIFICATION AND MONITORING OF FIBROTIC INTERSTITIAL LUNG DISEASE

Giacomo Sgalla

Chronic Interstitial Lung Diseases (ILD) are characterised by the interstitial involvement of the lungs, often resulting in the aberrant deposition of fibrotic tissue which causes progressive impairment of lung function. Earlier diagnosis of Idiopathic Pulmonary Fibrosis (IPF), the most frequent and severe among Idiopathic Interstitial Pneumonias, is warranted for a timely start of effective antifibrotic treatment. New tools for a more accurate prediction of disease progression and poor outcome are also required to improve the management of the individual patients.

Audible “Velcro-type” crackles on chest auscultation represent a typical finding in patients with fibrotic ILD. Although non-specific, their assessment performed by healthcare practitioners is invaluable to prompt the diagnostic work up and limit the diagnostic delay in these patients. Nevertheless, their value toward the early identification and the monitoring of fibrotic ILD has never been explored. Despite the modern advances in the field of electronic auscultation and computerised analysis of lung sounds offer a reliable, objective characterisation of normal and adventitious lung sounds, such methods have never been successfully translated into clinical practice.

This thesis provides a comprehensive characterisation of lung sounds recorded in patients with fibrotic ILD using a digital stethoscope, and tries to clarify whether this approach may valuably inform the diagnostic process and the management of these patients in the clinical practice.

A prospective case-control study investigated the association between the presence of “Velcro-type” crackles, evaluated by respiratory physicians, and radiologic signs of pulmonary fibrosis in three consecutive large cohorts of patients undergoing a High-Resolution Computerised Tomography (HRCT) scan of the chest. The presence of “Velcro-type” crackles predicted the

presence of ILD patterns on HRCT and was independently associated with distinct radiologic features suggestive of pulmonary fibrosis. Such evidence indicates that lung sounds have potential to improve early identification of ILD, while future work is needed to determine the real diagnostic accuracy of “Velcro-type” crackles, either assessed subjectively or via quantitative computerised analysis.

A longitudinal cohort study collected serial recordings of lung sounds from a cohort of 19 patients with IPF and 10 healthy controls over a 12-month observation period. Computer-aided lung sound analysis was used to extract almost 500 different acoustic features from each recording. A narrow set of acoustic features of lung sounds recorded via a digital stethoscope was found to be highly reproducible, distinctive of pulmonary fibrosis and responsive to disease progression assessed via validated clinical and functional parameters. As such, this study provided the first evidence on the longitudinal changes occurring in the acoustic features of lung sounds from IPF patients, and paves the way to further research aimed to determine the validity of lung sounds as a new prognostic marker in patients with progressive ILD.

Table of Contents

List of Tables

List of Figures

List of accompanying material - pendrive

List of the appendices included in the pendrive in separate folders and sub-folders:

- Appendix 1 – Contingency tables of radiologic assessment (case-control study)
- Appendix 2 – Pairwise comparisons of clinical parameters between study time points (cohort study)
- Appendix 3 – Acoustic features (cohort study)
    - Spreadsheets
    - Descriptives
    - Pairwise comparisons of repeatable acoustic features between study time points

DECLARATION OF AUTHORSHIP

I, Giacomo Sgalla declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Characterisation of lung sounds for early identification and monitoring of fibrotic Interstitial Lung Disease

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission

Signed: .....

Date: .....

Acknowledgements

I would like to express my gratitude to a number of people and organisations without whom this research and overall very challenging journey would not have been possible.

First, my thought goes to the dear departed Professor Borislav Dimitrov. I will always remember Borislav as one of the nicest persons I have met in Southampton. His enthusiasm, expertise and willingness to take part to this research were unique. His advice and support were crucial for this research, and his teaching the most inspirational I had over these years.

I would like to thank Francesca for her love and unconditional support throughout these years, and for being constantly on my side, even in the most unfavourable circumstances this journey put us through sometimes.

I would like to thank my parents for their love and patience, and all my friends for never making me feel alone during this time abroad.

My gratitude goes to my Supervisors, especially Professor Luca Richeldi, who's been guiding me for the last ten years now and gave me the opportunity of pursuing this experience, and Professor Donna Davies, for being a constant source of invaluable advice and so kind to take over my supervision during the last months. A special thank you goes to Dr. Anna Barney and Dr. Dragana Nikolic of the Institute of Sound and Vibration Research of the University of Southampton, for their wisdom, their availability and for the precious support which made this collaborative research possible.

Thank you to all the people who actively collaborated to this research providing their great expertise, the respiratory physicians Dr. Sophie Fletcher and Dr. Stefania Cerri, and the thoracic radiologists Dr. Simon Walsh, Professor David Hansell and Professor Nicola Sverzellati.

I would like to thank the Southampton Centre for Biomedical Research of the University Hospital of Southampton for supporting and funding this research, all the members of staff of the Respiratory Biomedical Research Unit and the ILD clinic for being so welcoming and kind with me. I would like to thank especially Louise Stanley, Research Nurse at the Respiratory Biomedical Research Unit, for her dedication, her expertise and for assisting me in all the phases of this project, from ethics to recruitment and data collection.

Main definitions and Abbreviations

AEC – Alveolar Epithelial Cells

AE-IPF - Acute Exacerbation of Idiopathic Pulmonary Fibrosis

ALAT - Latin-American Thoracic Society

ALI - Acute Lung Injury

ANOVA - Analysis Of Variance

ATS - American Thoracic Society

AV – Alveolar Volume

BAL - Bronchoalveolar Lavage

BMI - Body Mass Index

Cm – Centimetres

COPD - Chronic Obstructive Pulmonary Disease

CORSA - Computerised Respiratory Sound Analysis

CTC – Crackle Transmission Coefficient

CTD - Connective Tissue Diseases

DAD - Diffuse Alveolar Damage

Db – Decibel

DICOM – Digital Imaging and Communications in Medicine

DL_{CO} - Diffusion Lung Capacity for Carbon Monoxide

EMD – empirical Mode Decomposition

ERS - European Respiratory Society

IMFs – Intrinsic Mode Functions

F – Test statistic for analysis of variance (ANOVA)

FFT - Fast Fourier Transform

FDA – Food and Drug Administration agency

FEV_1 – Forced Expiratory Volume 1^{st} second

FVC – Forced Vital Capacity

GAP staging index – Gender Age Physiology staging index

GERD - Gastro-Esophageal Reflux Disease

GLIMMIX - Generalised Linear Mixed Modelling

HRCT – High Resolution Computerised Tomography

Hz – Hertz

K_{CO} - Rate of uptake of carbon monoxide

kPa – Kilo Pascal

Kg – Kilograms

k_w – Cohen's kappa statistic

ICC – Intraclass Coefficient

IDW – Initial Defection Width

IIP – Idiopathic Interstitial Pneumonias

ILD – Interstitial Lung Disease

IPF – Idiopathic Pulmonary Fibrosis

IPFnet - IPF Clinical Trials Network

ISVR – Institute for Sound and Vibration Research

L= Litres

LDW – Largest Deflection Width

JRS - Japanese Respiratory Society

Max – Maximum

MCID - Minimal Clinically Important Difference

MFCC - Mel Frequency Cepstral Coefficient

min - minute

Min – Minimum

mmHg – Torr (millimetre of Mercury)

mmol – Millimole

ms – Milliseconds

N - Number

NIH - National Institute of Health

NIHR – National Institute for Health Research

NPV = Negative Predictive Value

NREC - National Research Ethics Committee

OR - Odds Ratio

PCR - Polymerase Chain Reaction

PFTs - Pulmonary Function Tests

PPV – Positive Predictive Value

r – Pearson's correlation coefficient

RBRU – Respiratory Biomedical Research Unit

RCT – Randomised Clinical Trial

RDS - Research Design Service

RMS – Root Mean Square

RV – Residual Volume

R&D – Research and Development Department

SaO2 – Saturation of Oxygen

SCBR - Southampton Centre for Biomedical Research

SD – Standard Deviation

SE – Standard Error

SGRQ – St. George's Respiratory Questionnaire

SLB – Surgical Lung Biopsy

TBB -Transbronchial Lung Biopsy

TBLC - Transbronchial Lung Cryobiopsy

TEWA - Time-Expanded Waveform Analysis

TLC – total Lung Capacity

UCSD-SOBQ – University of California San Diego Shortness Of Breath Questionnaire

UIP - Usual Interstitial Pneumonia

VC – Vital Capacity

ZCR - Zero-crossing rate

$\chi 2$ – Chi-Squared test statistic

2CD – 2-Cycle Deflection width

6MWD - Six Minute Walk Distance

6MWT – Six Minute Walk Test

95% CI - 95% Confidence Intervals

Chapter 1: Introduction

This chapter will present the personal motivation of the author in conducting the research and the reasons behind the choice of the research topic, consisting in the study of the role of pathological lung sounds in fibrotic Interstitial Lung Diseases (ILD) for early identification and monitoring purposes. An overview of the structure of the thesis will be also presented.

1.1 Personal research commitment and career intentions

The author of this research obtained the title of Bachelor of Medicine at the University of Modena & Reggio Emilia, Italy in 2007. Since then, he developed a specific interest in ILD and IPF, driven by the challenges of the diagnostic process, the paucity of therapeutic options and the need for better management strategies for these patients. His graduation thesis focused on the role of fibrocytes, bone marrow-derived progenitor mesenchymal cells found to be highly represented in the blood of patients with IPF. During the appointment as specialty registrar in Respiratory Medicine at the University Hospital of Modena the author attended the Centre for Rare Lung Disease, led by Professor Luca Richeldi and specialised in the diagnosis and management of ILD. There he started getting in touch with translational research and participated to for several randomised clinical trials (RCTs) in IPF as sub-investigator. After obtaining the specialty degree in Respiratory Medicine in 2012, the author was appointed as research fellow at the Centre for Rare Lung Disease. In January 2013, he registered to the postgraduate course in Clinical and Experimental Medicine at the University of Modena and Reggio Emilia. This allowed him develop expertise in clinical research under the supervision of Professor Luca Richeldi. During this period, the author designed an exploratory, proof-of-concept research project investigating the association between lung sounds and imaging in fibrotic ILD. Soon after Professor Richeldi moved to the University of Southampton as Professor of Respiratory Medicine and Chair of Interstitial Lung Disease in the summer of 2013, the author started an appointment as clinical research fellow at the National Institute of Health Research (NIHR) Southampton Respiratory Biomedical Research Unit (RBRU) based at the University Hospital of Southampton. In order to develop his research further, he transferred the postgraduate course to the Department of Clinical and Experimental Sciences of the University of Southampton. Building on a partnership with academics at the Institute of Sound and Vibration Research (IVSR) of the University of Southampton, a centre with renowned expertise in computerised analysis of adventitious lung sounds, during the last three years of his postgraduate course the author designed and conducted

a second research project that employed computerised methods for analysis of lung sounds to characterise quantitatively lung sounds (namely "Velcro-type" crackles) in patients with IPF.

The author's ambition for the future is to further improve his skills in translational research in the field of ILD, with a focus on addressing emerging medical issues related to diagnosis and management of such disorders. The long-term aim is to generate a broad range of research projects to assume a role as a research leader. This purpose might be fulfilled through the following milestones:

- Implementing the research network between tertiary ILD centres in Europe and worldwide by promoting multi-centre international studies. This will allow gain the best evidence by recruiting independent cohorts from different social and cultural contexts.
- Pursuing collaborative research between experts in distinct academic fields, including engineering and bioinformatics, to combine methodological approaches and assess the validity of modern technologies for improving diagnosis and management in ILD. This will require strong partnerships with advocacy groups and members of the public, acknowledging their pivotal role in the design and implementation of successful research.
- Developing new, cost-effective methodologies and tools for disease stratification and management of patients with ILD, through the integration of medical knowledge for customising existing technologies and devices for the study of fibrotic ILD in the clinical practice.

1.2 Overview of the research

Interstitial Lung Disease (ILD) are a large and highly heterogeneous group of different conditions characterised by the interstitial involvement of the lungs, classified together because of similar clinical, radiological and pathological features. Many of these entities are classified as rare diseases, which are defined by the European Union as chronically debilitating disorders affecting less than 1 in 2000 people. However, all rare lung diseases combined account for about one third of the overall respiratory morbidity (Du Bois et al., 2009). Some of these conditions are known to be associated with specific exposures (environmental, occupational, or drug-related), while others are recognised as part of an auto-immune systemic disorder (such as connective tissue diseases, sarcoidosis or vasculitis). However, there is a group of ILD which doesn't identify a precise aetiology. Idiopathic Interstitial Pneumonias (IIP) comprise several ILD of unknown cause characterised by varying patterns of inflammation and fibrosis, which are sufficiently different from one another to be designated as separate disease entities (Travis et al., 2013). Among these, Idiopathic Pulmonary Fibrosis (IPF) is a chronic and progressive lung disorder characterised by the

aberrant deposition of extracellular matrix leading to extensive lung remodelling (King et al., 2011b). It accounts for about 20% of all cases of ILD and represents the most frequent and severe among IIP. The differential diagnosis of ILD often represents a medical challenge, as the rarer the disease, the higher the number of tests and the greater efforts by physicians are required to reach a definitive diagnosis. This is particularly true for IPF, whose social, healthcare and economic burden is far from irrelevant. It's estimated that in Europe approximately 40,000 new cases are being diagnosed each year, with more than 5,000 in the UK only (Navaratnam et al., 2011). Despite being a clinically heterogeneous disease with multiple variable courses, the prognosis of patients with IPF is poor, with a median survival time of 3 to 5 years from the time of diagnosis (Flaherty et al., 2002,King et al., 2001b,Nicholson et al., 2000) and a 5-year survival estimated around 20%, worse than those observed in several types of cancer (Vancheri et al., 2010). As a consequence, the total direct treatment cost for IPF is estimated to be around 25,000 $/person-year, more than the direct cost for breast cancer (Collard et al., 2015). Due to the subtle, invariably progressive course, the late onset of symptoms and the non-specificity of clinical and physiological signs, the diagnosis of IPF is often delayed, with a median duration of symptoms before diagnosis of more than 24 months (King et al., 2001a). The delay in the referral of patients to a tertiary care centre specialised in ILD has been found to correlate with poorer survival (Lamas et al., 2011), which makes diagnosing IPF at an earlier stage an urgent matter, especially now that the first safe and effective treatments, namely pirfenidone and nintedanib, have become available (King et al., 2014b,Richeldi et al., 2014).

This research articulated around two separate studies addressing a range of objectives. A case-control study with cross-sectional design investigated the relationships between audible "Velcro-type" crackles and signs of fibrosis at HRCT. The main hypothesis was that "Velcro-type" crackles generate from specific alterations occurring in the lung parenchyma underneath, as such they may inform the presence of specific features of pulmonary fibrosis and therefore support identification of fibrotic ILD. Data was prospectively collected at two Italian centres, the University Hospital of Modena (coordinating centre) and the University Hospital of Parma (participating centre). Lung sounds were digitally recorded before the participants underwent HRCT. Imaging data were reviewed by expert thoracic radiologists to assess the presence and the extension of fibrotic abnormalities in the lungs. The recordings were assessed by ILD physicians for the presence of "Velcro-type" crackles. The analysis of the accrued data was performed by the author at the University of Southampton.

The second project consisted in a prospective cohort study with 1-year follow up in patients with IPF and aimed to assess longitudinal changes of lung sounds using computerised analysis in order to pilot further investigation of the potential role of "Velcro-type" crackles as a prognostic marker

in progressive ILD. Since IPF has an invariably progressive course, the hypothesis was that the acoustic properties of lung sounds (namely "Velcro-type" crackles) change over time across different regions of the lung. The study was designed and conducted by the author at the Southampton Centre for Biomedical Research of the University Hospital of Southampton, building on a collaboration with Dr. Anna Barney and Dr. Dragana Nikolic, academics at the Institute of Sound and Vibration Research (IVSR) of the University of Southampton. A broad set of acoustic features were extracted from the recordings taken at different time points in a cohort of IPF patients and healthy controls. A set of reproducible features were assessed for longitudinal changes and correlated with clinical parameters of disease progression.

Overall, this research aimed to provide new evidence on the potential of a new, cost-effective approach, based on the electronic auscultation of the chest and computerised analysis of lung sounds, toward the early detection and the management of these disorders.

1.3 Thesis structure

First, a comprehensive review of the literature on IPF, the paradigm of progressive ILD, will be presented in Chapter 2. Several topics will be covered, ranging from diagnostic approach to epidemiology, natural history, management and therapeutic approach. Chapter 3 will describe the main concepts of computerised analysis of lung sounds with a focus on the analysis of "Velcro-type" crackles in ILD. Chapter 4 and 5 will describe respectively the methods, the results for the case-control and the longitudinal study, and will discuss their findings. Chapter 6 will draw the conclusions of the work done so far and present the plans for future studies.

Chapter 2: Idiopathic Pulmonary Fibrosis

2.1 Introduction

This chapter is a review of the most relevant scientific literature on IPF, which represents the paradigm of progressive ILD. The focus will be on the current diagnostic limitations, the unpredictability of its natural course and the unmet need of a cure, which make its diagnosis crucial when the clinical suspect of ILD is being raised.

2.2 Diagnostic approach

2.2.1 Clinical Presentation

IPF generally occurs after 60 years of age, and is more prevalent in males (Raghu et al., 2011). The initial clinical presentation is not specific, consisting of progressive dyspnoea on exertion combined with dry cough. At chest auscultation, bibasilar inspiratory "Velcro-type" crackles are constant and appear early in the disease (Cottin et al., 2012). Finger clubbing is present in 25-50% of cases, while weight loss and alteration of the general status are less common. Cyanosis and signs of right ventricular failure may occur in more advanced stages together with respiratory failure, ultimately causing death in these patients (King et al., 2011b).

The pulmonary function tests (PFTs) typically show a restrictive pattern at spirometry with reduced forced vital capacity (FVC) and there is an impairment of gas exchange reflected by the reduction of the diffusion lung capacity for carbon monoxide (DL_{CO}), which may represent the only functional abnormality in milder stages of the disease. However, it is important to note that in patients with concomitant emphysema the predicted FVC values may appear falsely normal, due to the concomitance of restrictive and obstructive defects.

Laboratory findings are usually non-specific and are mostly used to rule out alternative diagnosis especially connective tissue diseases (CTD), although some patients with IPF may present low-positive rheumatoid factor and antinuclear antibodies titres. On the other hand, some patients with suspected CTD-related ILD and low titres of autoantibodies do not fulfil the criteria for any specific autoimmune disease, making the differential diagnosis problematic (Vij et al., 2011). It is still debated whether these IPF patients with features of a subclinical autoimmune disorder have a different clinical course and response to treatments: this topic is the subject of active clinical research.

2.2.2 Diagnostic criteria

In 2011, the American Thoracic Society (ATS), the European Respiratory Society (ERS), the Japanese Respiratory Society (JRS) and the Latin-American Thoracic Society (ALAT) jointly published an evidence-based statement providing recommendation for the diagnosis and management of IPF (Raghu et al., 2011). This document provided an update of the diagnostic criteria, ten years after the previous statement (ATS et al., 2000), which now consists of the following:

1) exclusion of other known causes of ILD (e.g. domestic and occupational environmental exposures, connective tissue disease and drug toxicity)

2) presence of Usual Interstitial Pneumonia (UIP) pattern on high resolution computed tomography (HRCT) of the chest in individuals where surgical lung biopsy (SLB) is not indicated or available

3) specific combinations of HRCT and biopsy pattern in individuals undergoing SLB.

The new diagnostic criteria originated from the evidence that in an appropriate clinical setting, the presence of a classical (or definite) UIP pattern on the HRCT scan has a very high positive predictive value (between 90 and 100%) for a histological diagnosis of UIP (Sundaram et al., 2008), and has been therefore considered sufficient for a diagnosis of IPF to be made. A classical UIP pattern on HRCT is defined by the following features:

1) the presence of sub-pleural abnormalities, with predominant distribution at the lung bases

2) reticular abnormalities

3) honeycombing with or without traction bronchiectasis, along with the absence of features that are inconsistent with a UIP pattern (such as upper or middle lobe predominance, peri-bronchovascular predominance, ground-glass abnormalities more extensive than reticulations, profuse micro-nodules, discrete multiple cysts away from areas of honeycombing, diffuse mosaic attenuation/air trapping, segmental or lobar areas of consolidation)(Hansell et al., 2008b).

If honeycombing is absent, then the diagnosis of IPF is regarded as “possible”, and further diagnostic evaluation on surgical lung biopsy is required (Raghu et al., 2011). In these cases, biopsies should be taken at least in two different lobes and selecting sites based on the HRCT findings (Larsen et al., 2012).

For a definitive diagnosis of UIP on histology, the following features are required, based on current guidelines: 1) marked fibrosis/architectural distortion with or without honeycombing in a

predominantly sub-pleural/paraseptal distribution; 2) patchy involvement with alternating areas of normal and scarred parenchyma; 3) presence of active fibroblast foci adjacent to areas of fibrosis. Additionally, the lack of any features considered to be inconsistent with a UIP pattern is required. Based on these features, 4 levels of confidence for a diagnosis of histological UIP are recognized: definitive, probable, possible and not UIP (Table 1). As such, although the current guidelines provide guidance for the different combinations of radiological and histological patterns (Table 2), a multidisciplinary approach involving ILD specialists, radiologists and pathologists is recommended in the evaluation of suspected IPF. This approach, by considering the clinical context as well as the radiological and histological findings, has been shown to improve diagnostic accuracy and is now widely accepted to be the gold standard for the diagnosis of IPF.

The 2011 evidence-based guidelines suggest a stepwise diagnostic algorithm for IPF (Raghu et al., 2011). Firstly, Individuals with suspected IPF should be carefully evaluated for identifiable causes of ILD. In the absence of any identifiable cause for ILD, a chest HRCT demonstrating a UIP pattern is diagnostic of IPF. Otherwise, patients should be advised to undergo SLB; in this case, IPF could be diagnosed by the combination of HRCT and histological patterns and considering the clinical context during a multidisciplinary discussion.

Table 1 - Diagnostic criteria for IPF: HRTC and histological patterns (Raghu et al., 2011)

HIGH-RESOLUTION COMPUTED TOMOGRAPHY CRITERIA FOR UIP PATTERN			
UIP Pattern (All Four Features)	*Possible UIP Pattern (All Three Features)*	*Inconsistent with UIP Pattern (Any of the Seven Features)*	
• Subpleural, basal predominance • Reticular abnormality • Honeycombing with or without traction bronchiectasis • Absence of features listed as inconsistent with UIP pattern (see third column)	• Subpleural, basal predominance • Reticular abnormality • Absence of features listed as inconsistent with UIP pattern (see third column)	• Upper or mid-lung predominance • Peri-bronchovascular predominance • Extensive ground glass abnormality (extent > reticular abnormality) • Profuse micronodules (bilateral, predominantly upper lobes) • Discrete cysts (multiple, bilateral, away from areas of honeycombing) • Diffuse mosaic attenuation/air-trapping (bilateral, in three or more lobes) • Consolidation in bronchopulmonary segment(s)/lobe(s)	
HISTOPATHOLOGICAL CRITERIA FOR UIP PATTERN			
UIP Pattern (All Four Criteria)	*Probable UIP Pattern*	*Possible UIP Pattern (All Three Criteria)*	*Not UIP Pattern (Any of the Six Criteria)*
• Evidence of marked fibrosis/ architectural distortion, ± honeycombing in a predominantly subpleural/ paraseptal distribution • Presence of patchy involvement of lung parenchyma by fibrosis • Presence of fibroblast foci • Absence of features against a diagnosis of UIP suggesting an alternate diagnosis (see fourth column)	• Evidence of marked fibrosis /architectural distortion, ± honeycombing • Absence of either patchy involvement or fibroblastic foci, but not both • Absence of features against a diagnosis of UIP suggesting an alternate diagnosis (see fourth column) OR • Honeycomb changes only‡	• Patchy or diffuse involvement of lung parenchyma by fibrosis, with or without interstitial inflammation • Absence of other criteria for UIP (see UIP PATTERN column) • Absence of features against a diagnosis of UIP suggesting an alternate diagnosis column)	• Hyaline membranes* • Organizing pneumonia*† • Granulomas† • Marked interstitial inflammatory cell infiltrate away from honeycombing • Predominant airway centered changes • Other features suggestive of an alternate diagnosis

* Can be associated with acute exacerbation of idiopathic pulmonary fibrosis.† An isolated or occasional granuloma and/or a mild component of organizing pneumonia pattern may rarely be coexisting in lung biopsies with an otherwise UIP pattern. ‡ This scenario usually represents end-stage fibrotic lung disease where honeycombed segments have been sampled but where a UIP pattern might be present in other areas. Such areas are usually represented by overt honeycombing on HRCT and can be avoided by pre-operative targeting of biopsy sites away from these areas using HRCT.

Table 2 - Combination of HRTC and surgical lung biopsy patterns for the diagnosis of IPF (Raghu et al., 2011)

HRCT Pattern	**Surgical Lung Biopsy Pattern (When Performed)**	**Diagnosis of IPF?**
UIP	UIP Probable UIP Possible UIP Non-classifiable fibrosis‡	*Yes*
	Not UIP	No
Possible UIP	UIP Probable UIP	*Yes*
	Possible UIP Non-classifiable fibrosis	Probable *
	Not UIP	No
Inconsistent with UIP	UIP	Possible *
	Probable UIP Possible UIP Non-classifiable fibrosis Not UIP	No

‡ Non-classifiable fibrosis: Some biopsies may reveal a pattern of fibrosis that does not meet the above criteria for UIP pattern and the other idiopathic interstitial pneumonias. These biopsies may be termed ‘‘non-classifiable fibrosis.’’ Multidisciplinary discussion should include discussions of the potential for sampling error and a re-evaluation of adequacy of technique of HRCT. NOTE: In cases with an ‘‘inconsistent with UIP’’ HRCT pattern and a ‘‘UIP’’ surgical lung biopsy pattern, the possibility of a diagnosis of IPF still exists and clarification by MDD among interstitial lung disease experts is indicated.

2.2.3 Alternative investigations: the transbronchial lung cryobiopsy

In real-life clinical practice and based on current guidelines, approximately in half of patients with suspected IPF, a histology confirmation of UIP pattern would be required to make a confident diagnosis of IPF, using tissue samples obtained by surgical lung biopsy (SLB) (Raghu et al., 2011). However, such procedure carries appreciable risks, with a mortality estimated between 2 and 6% within 90 days (Park et al., 2007). As such, a possible role for transbronchial lung biopsies (TBB) has been advocated: however, current guidelines do not recommend their use in the evaluation of IPF in most individuals, mainly due to the current uncertainty about their diagnostic accuracy, the best location to biopsy and the number of biopsies to be taken.

In a retrospective study evaluating the diagnostic yield for TBB, sensitivity and negative predictive value were found to be low (30 and 50% respectively), due to a high proportion of inadequate and non-diagnostic samples (Tomassetti et al., 2012). Recently, a new type of endoscopic procedure called cryobiopsy, that can extract larger pieces of tissue using the freeze-thaw cycle, has demonstrated superior diagnostic performance as compared to the conventional forceps in ILD evaluation (Pajares et al., 2014). A recent prospective study in 69 patients with a HRCT pattern of fibrotic ILD has shown that transbronchial lung cryobiopsy (TBLC) is feasible and safe in these patients, and pathologists could identify criteria sufficient to define a specific pattern in 63 patients out of 68 (93%), including 47 UIP, suggesting good diagnostic yield (Casoni et al., 2014). However, given the considerable costs and burden on patients related to TBLC, further studies are needed to assess the usefulness of this technique as a potential alternative to SLB in the evaluation of ILD and especially IPF.

2.2.4 Current limitations

Despite the availability of diagnostic guidelines based on international consensus (Raghu et al., 2011,Travis et al., 2013), accurate differential diagnosis of chronic fibrotic ILD remains challenging. A HRCT-based diagnosis of UIP/IPF requires the presence of honeycombing, which in most cases is a sign of advanced fibrosis. As such, the main challenge when diagnosing IPF is represented by discerning those cases where ascertainment by histopathology is needed. Since the risks of surgical lung biopsy (SLB) may outweigh the benefits of establishing a secure diagnosis of IPF (in particular when the disease is more advanced), and many individuals may refuse to undergo a surgical procedure, only a minority of patients with suspected IPF reach diagnosis via SLB (du Bois, 2012). It's estimated that up to 10% of ILD cases remains unclassifiable, the main reason being missing histopathology (Ryerson et al., 2013). Although the proportion of individuals

with histologically-proven UIP and honeycombing changes at HRCT cannot be established with certainty, it has been reported ranging between 60% and 80% in older studies (Flaherty et al., 2003,Hunninghake et al., 2003), while data from recent trials suggest this percentage is likely to be an overestimate (Richeldi et al., 2011), making reasonable to say that, according to current guidelines, a confident HRCT diagnosis can be made in about only half of the cases. Patients with a possible (or "atypical") UIP pattern at HRCT don't fall in any diagnostic category as per current guidelines; the need of classifying these individuals with a "possible disease" was already underlined shortly after the publication of the 2011 statement (du Bois, 2012), and the urgency of making a secure diagnosis in these patients without recurring to surgery has become now compelling in order to allow them access to treatments for IPF. A heterogeneous pattern of fibrosis on CT in absence of honeycombing was reported to be sufficient to secure a diagnosis of UIP at histopathology (Gruden et al., 2015). Data from late phase randomised clinical trials in IPF also offered some new interesting insights. A study assessing the diagnostic predictive value of HRCT patterns in a cohort of patients with suspected IPF from the ARTEMIS trial (Raghu et al., 2014b) has shown a positive predictive value of possible UIP on HRCT towards histological confirmation of 94%, supporting the idea that surgical lung biopsy might not be necessary to reach a diagnosis of IPF if HRCT scans are being assessed by experts. The phase 2 TOMORROW study (Richeldi et al., 2011), evaluating the safety and efficacy of nintedanib in IPF patients, included patients having both "definite" and "probable" IPF, the latter being defined by a possible (or "atypical") UIP pattern at HRCT scan. These patients represented more than half the study cohort (62%). Still, the study succeeded in demonstrating the efficacy of nintedanib in reducing the rate of decline of FVC in these patients. More recently, data from a pre-defined subgroup analysis of the phase 3 INPULSIS trials of nintedanib, which adopted the same eligibility criteria as the phase 2 study, confirmed that the treatment effect of nintedanib in terms of FVC decline is not different between patients with honeycombing and/or biopsy confirmation of UIP and patients with a possible UIP and no biopsy. Interestingly, these two groups of patients showed identical rates of lung function decline in the placebo groups (Raghu et al., 2017).

The same definition and ascertainment of radiological UIP as reported in the current guidelines is not without flaws. Although honeycombing represents the hallmark of established fibrosis, the smaller cysts that are visible on pathological samples and are referred to as microscopic honeycombing cannot be seen at the HRCT of the chest for being beyond the limits of resolution (Arakawa et al., 2011). As such, the absence of radiological honeycombing doesn't translate automatically into the absence of its histological equivalent. Moreover, the inter-observer agreement in recognising honeycombing is far from ideal even among expert chest radiologists, with the presence of emphysema and peripheral traction bronchiolectasis being indicated as the

main confounding factors for its ascertainment (Watadani et al., 2013). A recent study conducted in a large group of thoracic radiologists with diverse levels of experience confirmed and further corroborated this finding, as the inter-observer agreement for the current CT criteria for a diagnosis of UIP was only moderate, with no significant differences between observer subgroups of different expertise (Walsh et al., 2016). Moreover, prognosis in histologically confirmed UIP cases has been demonstrated being influenced more by the presence and severity of traction bronchiectasis than by a definite UIP pattern at the HRCT (Sumikawa et al., 2008). The high predictive value of traction bronchiectasis towards poor survival has been confirmed by two recent studies exploring the HRCT prognostic determinants both in connective tissue disease related fibrotic lung disease and in chronic hypersensitivity pneumonitis (Walsh et al., 2012,Walsh et al., 2014). These studies also reported a good inter-observer agreement for the identification of traction bronchiectasis, as opposed to honeycombing. Traction bronchiectasis, currently regarded as an optional feature in the radiological criteria for UIP, may therefore have a major role in future diagnostic algorithms.

2.3 Epidemiology

2.3.1 Incidence and prevalence

IPF is known for being the most common among the idiopathic interstitial pneumonias (IIP). Data from existing registries suggest that IPF accounts for 17% to 37% of all ILD diagnoses (Thomeer et al., 2001,Tinelli et al., 2005) .

Nevertheless, its prevalence and incidence remain unclear, with a large variability across reports due to several reasons. Firstly, a precise knowledge of IPF epidemiology is hurdled by the lack of an uniform definition of IPF in the studies conducted so far, as the older ones included cases diagnosed prior the 2000 consensus statement on IPF (ATS et al., 2000). Secondly, different methodologies have been used for the cases ascertainment in the populations examined, ranging from the use of different diagnostic ICD codes, to death registries and surveys of clinicians with varying degree of specialty, or a mix of the previous. Finally, studies adopted different designs and the results are not easily comparable.

A comprehensive understanding of IPF epidemiology worldwide is also limited by the fact that most studies have been carried out in the US and in Europe. Few data have been reported from Asia (Japan, mostly), and nothing has been published about the prevalence and incidence of IPF in other countries/continents such as Africa, South America or Australia.

One of the first large studies in New Mexico on a population-based ILD registry between 1988 and 1990, using the ICD-9 code 516.3 for idiopathic fibrosing alveolitis, found a prevalence of 13.2 and 20.2 cases per 100,000/year for women and men respectively, and an incidence of 7.4 and 10.7 (Coultas et al., 1994). A subsequent study investigating data from a large healthcare claims database between 1996 and 2000 in 20 different US counties reported variable rates according to two different definitions of IPF adopted. Using a broad definition, including the cases identified by the ICD-9 code 516.3 and exclusion of associated conditions, a prevalence of 42.7 and an incidence of 16.3 cases per 100,000/year were reported, suggesting an increasing trend. However, the analysis performed using of a narrow definition, requiring the evidence of a UIP pattern either at a surgical lung biopsy or on a HRCT scan, showed rates closer to those reported previously almost 10 years before (prevalence and incidence 14.0 and 6.8 cases per 100,000/year)(Raghu et al., 2006). More recently, a study on a population-based registry in Minnesota between 1997 and 2005 has shown slightly higher rates of prevalence and incidence for both broad and narrow definitions based on the 2000 IPF diagnostic criteria. However, this study reported a decreasing incidence in the period observed (Fernandez Perez et al., 2010).

Another recent study collecting data from US Medicare beneficiaries aged 65 or older between 2001 and 2011 has shown much higher rates than previously described (incidence of 93.7 cases per 100,000/year, prevalence ranging from 202.2 cases per 100,000 in 2001 to 494.5 cases per 100,000 in 2011). In the observed period incidence has remained stable while prevalence appears on the rise (Raghu et al., 2014a).

In continental Europe, both the prevalence and incidence of IPF seem to be somewhat lower, although the high variability of reported rates and the differences of methodologies between studies make such comparison not reliable. Prevalence and incidence of IPF ranged from 1.25 and 0.22 per 100,000/year in a Belgian study collecting data from a ILD registry in 20 centres between 1992 and 1996 (von Plessen et al., 2003), to 23.4 and 4.3 per 100,000/year respectively in a population from a single hospital in Norway between 1984 and 1998 (Gribbin et al., 2006). In the UK, two recent studies have reported a higher incidence though. A first study (Navaratnam et al., 2011) reported an incidence of 4.6 per 100,000/year by retrieving cases from a longitudinal primary care database between 1991 and 2003. An even higher incidence of 7.4 per 100,000/year has been reported in a follow up study using data from same database (but slightly different methods for case inclusion) for the period 2000-2009 (Navaratnam et al., 2011). Importantly, both studies have showed an increase of incidence rates throughout the observed periods.

Overall, the recent data suggests an increasing prevalence and a stable or increasing incidence of IPF in western countries. When data have been stratified per sex and age, most studies have showed higher prevalence and incidence rates among males and with increasing age, especially after 75.

As noted above, there have been few epidemiologic studies in Asian communities. A nationwide survey conducted in Japan (Ohno et al., 2008) have estimated an annual prevalence for IPF of 2.95 per 100,000 based on the investigation of the medical benefit certificates for all IIPs in 2005. However, the results for IPF were calculated based on the proportion of ascertained IPF cases among IIPs, and the milder cases were excluded due to the study methodology. More recently, another Japanese study (Natsuizaka et al., 2014) has explored the certificates of medical benefits for IPF in the Hokkaido prefecture between 2003 and 2007, reporting a prevalence and incidence of 10.0 and 2.23 per 100,000/year respectively. Whilst male predominance and an increase of frequency with age are confirmed in these studies, prevalence and incidence rates appear lower than those described in western countries, suggesting a possible role of ethnic differences in IPF epidemiology between Asian and Western populations.

2.3.2 Mortality

IPF-related mortality data share some common limitations with prevalence and incidence evaluations. Being derived from national registries built on information provided by death certificates, IPF-related mortality is largely affected by IPF under-recognition and misdiagnosis, and is likely to be underestimated (Mannino et al., 1996).

Several studies have pointed out that, similarly to prevalence and incidence, IPF-related mortality is higher in males and increases with age (Johnston et al., 1990,Olson et al., 2007), and seems to increase over time. Olson and co-workers (Olson et al., 2007), looking at IPF-related deaths in the US from 1992 to 2003, have calculated an age and sex-adjusted mortality of 50.8 per 100,000/year, with an increment rate of 28.4% for men and 41.3% for women over the study period. In the UK mortality has also steadily increased from 1968 to 2008 (Navaratnam et al., 2011). The reasons behind the observed increase in IPF-related mortality are not fully clear, although it can be hypothesized that an improvement in recognition of the disease, more than a real increase in mortality, plays a major role.

In terms of causes of death, respiratory failure from IPF accounted for the 60% of IPF-related deaths in a US-based study (Olson et al., 2007), followed by cardiovascular diseases (8.5%) and lung cancer (2.9%). Such proportion is even higher than in previous reports (Mannino et al., 1996), confirming IPF as the major underlying cause of death in these patients.

Another US study investigated ethnical differences in IPF-related mortality, reporting that IPF-related deaths appears to occur more in Caucasians than in Blacks (Swigris et al., 2012b). Interestingly, a recent study (Natsuizaka et al., 2014) has reported a percentage of death from acute exacerbation of IPF of 40%, much higher than previously described (Fernandez Perez et al., 2010).

2.4 Natural history

2.4.1 Patterns of disease progression

Notwithstanding the advances in the understanding of the underlying pathogenic mechanisms and the discovery of two agents effective in reducing the decline of pulmonary function, IPF remains a progressive disease with unfavourable prognosis. Median survival from diagnosis is only 2 to 3 years (Raghu et al., 2011), reportedly varying from 27.4 months for patients with severe disease (FVC <55% predicted) to 55.6 months for patients with mild disease (FVC ≥70% predicted) (Nathan et al., 2011).

The rate of progression of IPF is highly variable, both between patients and between different periods in a single individual. Many patients have a slow, but progressive clinical course over a period of years, while in 10-15% of patients the course of the disease is much more rapid, leading to death from respiratory failure in few months. Finally, a minority of patients present relative stability over long periods, punctuated by episodes of rapid acute deterioration (acute exacerbation of IPF, AE-IPF), either fatal or leading to a step down in pulmonary function (Ley et al., 2011,Raghu et al., 2011). However, it is impossible at present to predict, once a consensus diagnosis is reached, how the disease will behave in the single individual. This heterogeneity represents a major challenge for physicians, since traditional means of risk stratification are dependent on clinical variables (history and physical examination, pulmonary function, exercise testing, radiological findings) that are poorly reflective of disease pathogenesis and have insufficient power to accurately predict clinical outcome.

The presence of several comorbidities indeed influences prognosis in IPF. Emphysema and secondary pulmonary arterial hypertension, commonly present in IPF patients, are both associated with poor survival (Corte et al., 2009,Cottin, 2013). Lung cancer is also frequently associated with IPF and has a significant impact on survival (Tomassetti et al., 2015). Gastro-oesophageal reflux disease (GERD) has been proven being an important factor both in the pathogenesis and progression of the disease (Lee et al., 2013). As such, comorbidities should be timely evaluated to identify those patients at higher risk, and a proper treatment should be delivered whenever possible.

As new and more targeted therapies are being developed, clarifying the heterogeneity of IPF becomes more than ever compelling and tools for stratification of patients are strongly needed to tailor personalised treatments to those patients who could benefit the most.

2.4.2 Acute exacerbation of IPF

2.4.2.1 Definition and diagnostic criteria

Every IPF patient may experience, at any point in the course of the disease and regardless of its severity, episodes of acute and severe respiratory deterioration. Some of these are idiopathic, as they do not recognise an immediately identifiable cause (such as infections, left heart failure, pulmonary embolism, etc.), and are referred to as acute exacerbation of IPF (AE-IPF). These are characterised by a subacute worsening of dyspnea, usually in the previous month (Collard et al., 2007), despite other symptoms (such as cough) are often present. HRCT shows a background of interstitial lung disease with superimposition of bilateral, diffuse ground glass opacities, with or without consolidation (Akira et al., 2008). The typical histological finding in AE-IPF is diffuse alveolar damage (DAD) superimposed on the underlying UIP pattern (Oda et al., 2014a), although surgical procedures are usually avoided being considered at high risk for mortality, since they don't usually modify the management of the exacerbations and carry a very high risk of complications. Acute exacerbations of IPF are more common in patients with advanced disease (Kondoh et al., 2010,Song et al., 2011,Sugiura et al., 2012), but rarely they can also represent the first manifestation of the disease (Daniels et al., 2008,Kim et al., 2006). These events represent an acceleration of the underlying fibrotic process, have no effective treatment and usually lead to a poor outcome (Huie et al., 2010,Song et al., 2011).

Diagnosis of AE-IPF is based on a combination of clinical, radiologic and laboratory data, as defined for the first time in a 2007 consensus statement published by IPF Clinical Trials Network (IPFnet) and sponsored by the National Institute of Health (NIH) (Kinder et al., 2007) which proposed the following criteria: 1) a previous diagnosis of IPF, 2) acute and unexplained onset or worsening of dyspnoea (within 30 days or less), 3) new bilateral ground glass or consolidative opacities at the HRCT of the chest superimposed on a background UIP pattern, and 4) the exclusion of alternative identifiable causes including congestive heart failure, pulmonary embolism or infection through endobronchial aspiration or bronchoalveolar lavage (BAL). However, procedures such as bronchoscopy and BAL are not feasible in the most critical patients with severe hypoxemia. As such, based on the availability of data AE-IPF was classified as "confirmed" (when all data was available) or "suspected" (if some of the data was not available). The 2007 criteria were applied in the following randomised clinical trials, where the large amount of data generated was of great help for reassessing such criteria and improve the understanding of the nature and relevance of such events. For example, data from the recent INPULSIS trials on nintedanib (Richeldi et al., 2014) suggest that these criteria are often challenging to satisfy in a very high percentage of patients, even in the context of a controlled clinical trial conducted in

specialised centres. In these studies, 7.6% of patients treated with placebo experienced at least one acute exacerbation of IPF over one year of observation, as reported by investigators; however, a careful central review of cases performed by an independent committee classified as "confirmed" or "suspected" AE-IPF only half of these events, while only 10% satisfied all the criteria for a "confirmed" AE-IPF. This classification issue carries some potential clinical repercussions, since the presence of "spurious" AE-IPF may mask a signal of efficacy for a specific treatment. In fact, while the effect of nintedanib on the time to first exacerbation wasn't found statistically significant when AE-IPF were adjudicated by the investigators, it became significant in the pre-specified analysis which considered only the "suspected" or "confirmed" AE-IPF as adjudicated by the reviewers of the central committee.

Apart from the diagnostic challenges and their implications, the definition of AE-IPF proposed in 2007 also presents a "conceptual" limitation. Firstly, the clinical picture doesn't seem to differ much between suspected AE-IPF and confirmed AE-IPF, which are both associated with high risk of short-term mortality. As such, suspected AE-IPF were proposed as an outcome measure in clinical trials, where the low rate of observed AE-IPF makes difficult to capture a potential reduction in the incidence of these events as determined by an investigational therapy (Collard et al., 2013). Recent data also reported no difference of outcome between hospitalised patients with idiopathic acute exacerbation of fibrotic ILD and those for whom aetiologies were identified (Huie et al., 2010). Such evidence has recently brought researchers active in this specific field to support the concept that the attention should shift to the pathobiologic features of AE-IPF, i.e. the acute lung injury resulting in DAD, instead of clinical aetiology (Ryerson et al., 2014). A working group of international experts has therefore provided an update on AE-IPF through the systematic review of the literature published after the 2007 IPFnet consensus statement, including a revision of the definition and the diagnostic criteria for AE-IPF (Collard et al., 2016). In such update, AE-IPF is more broadly defined as an "acute and clinically significant respiratory deterioration characterised by evidence of new widespread alveolar abnormality". The term "idiopathic" has been removed, being considered useless, while any acute respiratory event characterised by new ground glass or consolidative opacities not fully explained by cardiac failure or fluid overload is being included. The exclusion of such causes has been justified because cardiogenic pulmonary oedema has a distinct pathobiology and a more favourable prognosis as compared to other causes of acute respiratory worsening with bilateral radiologic involvement. Another change in the diagnostic criteria is the time interval required for the development of a worsening of dyspnea. The 30-day limit was considered arbitrary, and the criteria has been rephrased as "typically of the length inferior to a month", a more flexible expression which allows include those cases that despite falling outside the 30-day window can be still considered as acute exacerbations by clinicians. In

order to avoid heterogeneity in the definition of AE-IPF in clinical trials (as such complicating comparisons between studies), the working group suggested a more rigorous definition of AE-IPF as an endpoint restricted to one month or less of worsened dyspnea should be applied. The only qualifying aspect for the diagnosis of AE-IPF is the ascertainment that the radiologic findings represented by new bilateral ground glass opacities or consolidation are not fully explained by the presence of heart failure or fluid overload. This should not however diminish the clinical importance of identifying, when present, a triggering infection, since its treatment might be crucial to the general management of the patient.

In conclusion, the new definition of AE-IPF makes the diagnostic approach easier without the need to perform invasive procedures (such as bronchoscopy), which would be anyway difficult to perform in the clinical setting. Furthermore, the new definition has got the advantage of increasing the rate of AE-IPF events by reducing the need for invasive diagnostic tests and not excluding events caused by external factors, thus contributing to satisfy the clinical need – still unmet – of a treatment for AE-IPF, and making the outcome of AE-IPF (and not only its incidence) a more attractive efficacy endpoint for future clinical trials.

2.4.2.2 Epidemiology

AE-IPF represents the more frequent cause of death in patients with IPF (Natsuizaka et al., 2014), counting for about half of all death causes. Short term mortality is approximately 50% (Huie et al., 2010,Kishaba et al., 2014,Song et al., 2011) and usually reaches 90% for hospitalised patients in intensive care units (Al-Hameed et al., 2004). As such, these acute episodes differ substantially in terms of clinical features and prognosis when compared to acute exacerbations of other chronic respiratory disorders such as COPD and asthma.

The incidence of AE-IPF has been reported as varying between studies, with an annual incidence estimated being up to 20% (Ryerson et al., 2014). Different rates of exacerbations reported in the studies of the last decade likely mirror the differences in disease severity of the populations under study, since AE-IPF occurs predominantly in those patients with more severe functional impairment (Collard et al., 2013,Song et al., 2011). The definition of the endpoints used and the statistical methodology are other factors that can affect the incidence rates recorded (Ryerson et al., 2015). A recent meta-analysis of six clinical trials in patients with IPF revealed a weighted mean of 41 exacerbations per 1000 patients-year (Atkins et al., 2014). One of the major limitations of these estimates based on clinical trials is represented by the missed events due to missing or non-available clinical data. For example, in the STEP-IPF trial a post-hoc analysis of respiratory adverse events showed an incidence of only 40 exacerbations of IPF per 1000 patients-year when the diagnostic criteria were applied more rigorously (Collard et al., 2013).

Nevertheless, the incidence increased to 200 AE-IPF per 1000 patients-year when patients with both confirmed and suspected exacerbations were included. Cohort studies have generally reported higher incidence rates of AE-IPF as compared to randomised clinical trials (Johannson et al., 2014,Kim et al., 2006,Mura et al., 2012,Ohshimo et al., 2014,Schupp et al., 2015,Sugino et al., 2015). A US study based on clinical registries reported an annual incidence of 130 per 1000 patients-year (Fernandez Perez et al., 2010); a Korean cohort study reported incidence rates at one and three years of 14.2%, and 20.7%, respectively (Song et al., 2011); a Japanese study reported incidence rates at one, two and three years of 8.6%, 12.6%, and 23.9%, respectively (Kondoh et al., 2010). On the other hand, a limitation of cohort studies is the potential overestimation of events due to the wrong classification of respiratory deteriorations from known causes reported as AE-IPF, as demonstrated by the data from the INPULSIS trials previously described.

In conclusion, the incidence of AE-IPF is still unclear, varying from 5% to 15% per year in retrospective studies based on the placebo arm populations enrolled in clinical trials (Azuma et al., 2005,Martinez et al., 2014,Richeldi et al., 2014) and might increase with time, reaching 20.7% at 3 years from diagnosis (Song et al., 2011). Lower FVC and DL_{CO} have been shown to correlate with a higher risk of AE-IPF (Song et al., 2011), and the concomitant presence of emphysema or pulmonary hypertension are also associated with an increased risk (Judge et al., 2012).

2.4.2.3 Aetiology

Aetiology of AE-IPF remains unclear. The central question asked by the IPFnet committee in 2007 was whether AE-IPF represented an intrinsic acceleration of the underlying fibrotic condition or a response to external events (such as infections) which leads to an acute lung injury (ALI) and to the diffuse alveolar damage (DAD) at the histopathology. Indeed, AE-IPF and ALI as clinical entities share many features, such as the demand for high levels of oxygen and bilateral radiologic abnormalities (de Hemptinne et al., 2009,Ferguson et al., 2005). Cases accidentally diagnosed as ALI through clinical and radiological evaluation were found to represent an AE-IPF when lung tissue was examined histologically for the presence of an underlying fibrotic pattern consistent with UIP (Olson et al., 1990). However, ALI recognises several causes, including infections, aspiration of gastric content, drugs, massive transfusion and surgical procedures (Matthay et al., 2012), while AE-IPF is considered primarily idiopathic in nature.

Nevertheless, the role of potential triggers in driving occult injuries to the lung has been progressively recognised. It has been demonstrated that at least 9% of AE-IPF may be caused by clinically silent viral infections according to molecular or microarray findings in BAL (Wootton et al., 2011), and Polymerase Chain Reaction (PCR) multiplex and pan-viral matrixes have

demonstrated the presence of respiratory viruses RNA in a minority of patients with AE-IPF (Bando et al., 2001,Huie et al., 2010,Tomioka et al., 2007,Ushiki et al., 2014,Wootton et al., 2011). The assessment of post-mortem tissue samples from patients deceased following an AE-IPF has also shown the presence of occult respiratory infection in some (Oda et al., 2014a,Santos et al., 2013) even if not all (Konishi et al., 2009) the cohorts examined in the studies. Finally, epidemiological support towards an infective aetiology also comes from studies demonstrating that AE-IPF is significantly more frequent in winter and spring months (Collard et al., 2013,Simon-Blancal et al., 2012), and in patients in immunosuppressive treatment (Collard et al., 2013,Johannson et al., 2014,Petrosyan et al., 2015).

Indirect evidence linking gastric micro-aspiration to AE-IPF comes from a post-hoc analysis of the placebo arms of three clinical trials, showing that AE-IPF occurred only in subjects not in anti-acid treatment (Lee et al., 2013). This is presumably due to ant-acid therapy reducing the potential for a lung injury from micro-aspiration to occur. Such hypothesis is supported by a small case-control study demonstrating high levels of pepsin in the bronchoalveolar lavage of patients with AE-IPF (Lee et al., 2012), and by a small cohort study describing an increased prevalence of gastroesophageal reflux in patients with AE-IPF (Tcherakian et al., 2011).

Increased exposure to ozone and nitrogen dioxide has also been shown to be associated with an increased risk of AE-IPF (Johannson et al., 2014).

The pathobiology of AE-IPF can be correlated to intrinsic defects which make lungs of patients with IPF more susceptible to external insults as compared to non-IPF lungs. Indirect evidence in support of this concept comes from the comparison of data from thoracic surgery in patients with and without IPF. Retrospective studies have observed the development of acute respiratory deterioration after resection of lung cancer in patients with IPF, estimating an incidence varying from 7% and 32% (Choi et al., 2014,Mizuno et al., 2012,Suzuki et al., 2011,Watanabe et al., 2008). AE-IPF has been also reported in patients with IPF following surgical lung biopsy and bronchoscopy (Bando et al., 2009,Ghatol et al., 2012,Sakamoto et al., 2012,Sakamoto et al., 2011,Samejima et al., 2015). These data suggest that IPF lung is extremely susceptible to the stress caused by thoracic surgery (either due to the procedure itself or to the barotrauma, the volutrauma or to the hyperoxia induced by mechanical ventilation), supporting the hypothesis of an intrinsic pathobiologic contribution.

The identification of all the above mentioned triggers as likely causes of diffuse damage in lungs with inability to undergo a normal repair, together with the evidence of similar clinical characteristics and outcomes irrespectively of an identifiable aetiology have brought to the proposal of a new framework for AE-IPF and ultimately to the revision of its definition and

diagnostic criteria, and it is now accepted that acute worsening of IPF may recognise secondary causes, some of which are not identifiable, sometimes leading to an acceleration of the disease with the development of diffuse lung injury (Collard et al., 2016,Johannson et al., 2013).

2.4.2.4 Risk factors

As mentioned, the severity of the disease is the major risk factor for the development of AE-IPF, since it is more frequent in patients with advanced disease. A reduced FVC has demonstrated to represent the more consistent risk factor for AE-IPF (Collard et al., 2013,Johannson et al., 2014,Kishaba et al., 2014,Kondoh et al., 2010,Ohshimo et al., 2014,Schupp et al., 2015,Song et al., 2011). Other physiology parameters have been associated to an increased risk, including a reduced DLCO (Collard et al., 2013,Kishaba et al., 2014,Mura et al., 2012,Schupp et al., 2015,Song et al., 2011), low distance covered at the 6 minute walk test (Collard et al., 2013), the presence of pulmonary hypertension (Judge et al., 2012), low oxygenation (Collard et al., 2013,Kondoh et al., 2015), increased dyspnea (Collard et al., 2013,Kondoh et al., 2010) and a recent decline in FVC (Kondoh et al., 2015,Kondoh et al., 2010,Reichmann et al., 2015). Whether patients with more advanced IPF are more susceptible to AE-IPF or they are simply more prone to required unscheduled medical care as a result of AE-IPF (and therefore to its diagnosis) is not known. Further candidate risk factors for AE-IPF include young age (Schupp et al., 2015), ischaemic heart disease (Collard et al., 2013) and a high body mass index (Kondoh et al., 2010). Data on smoking habit and concomitant emphysema as risk factors are discordant (Mura et al., 2012,Ohshimo et al., 2014,Song et al., 2011,Sugino et al., 2015). A previous history of AE-IPF has been also associated to an increased risk (Johannson et al., 2014,Reichmann et al., 2015). Finally, increased serum levels of Krebs von Lungen-6 (KL-6) factor have been associated with higher risk for AE-IPF, also after adjustment for clinical characteristics including pulmonary vital capacity (Ohshimo et al., 2014).

2.4.2.5 Prognosis

Prognostic implications of an acute exacerbation of IPF are deeply relevant. Available data suggest that up to 46% of deaths from IPF are preceded by AE-IPF (Jeon et al., 2006,Kondoh et al., 2010,Natsuizaka et al., 2014), while the average survival of patients with IPF experiencing an AE is approximately of 3-4 months (Collard et al., 2013,Song et al., 2011). Respiratory failure from AE-IPF is associated with high intra-hospital mortality; in the majority of the reported cases it reaches the 50% (Akira et al., 2008,Al-Hameed et al., 2004,Horita et al., 2011,Kim et al., 2006,Kishaba et al., 2014,Song et al., 2011,Tomassetti et al., 2013,Tomioka et al., 2007,Usui et al., 2013). Limited data also suggest that patients with confirmed or suspected AE-IPF carry a similar prognosis. (Collard et al., 2013,Huie et al., 2010).

Several candidate prognostic factors have been identified in AE-IPF. These include reduced baseline FVC and DLCO (Kondoh et al., 2010,Simon-Blancal et al., 2012,Song et al., 2011), more extensive alterations at CT at the moment of presentation of AE-IPF (Akira et al., 2008,Fujimoto et al., 2012,Kim et al., 2006,Kishaba et al., 2014), a worse oxygenation (Abe et al., 2012,Song et al., 2011), and the percentages of neutrophils and lymphocytes at the bronchoalveolar lavage (Song et al., 2011). Different serum markers have been proposed for a prognostic role, including lactate-dehydrogenases, C-reactive protein, factor KL-6, circulating fibrocytes and anti-heat shock protein 70 antibodies (Akira et al., 2008,Fujimoto et al., 2012,Isshiki et al., 2015,Kahloon et al., 2013,Kishaba et al., 2014,Moeller et al., 2009,Simon-Blancal et al., 2012,Song et al., 2011,Tsushima et al., 2014). Recently, a staging system for AE-IPF which includes several of these factors has been proposed (Kishaba et al., 2014).

2.4.3 Predicting progression and poor outcomes

Many clinical measurements, imaging features and molecular biomarkers have been largely investigated and proposed as predictors of disease progression and poor outcome in IPF, both as individual parameters or incorporated in risk prediction models. Nevertheless, most of the available data comes so far from retrospective studies or post-hoc analyses of pooled data from the placebo groups of large clinical trials, which have indeed the advantages of providing high-quality data from well-characterised populations, but are biased by relatively short observation and the inclusion of patients who are not representative of the general IPF population for being healthier. None of the proposed parameters or models, in fact, has been validated in properly designed longitudinal studies.

2.4.3.1 Individual predictors

Among demographics, older age and male sex seem to correlate with a poorer prognosis in IPF patients (King et al., 2001a,Ley et al., 2011), while there is contrasting data about smoking status (Antoniou et al., 2008,King et al., 2001a). Respiratory symptoms, presented as basal dyspnea scores at validated questionnaires or their longitudinal changes, have also demonstrated to predict survival (Li et al., 2014,Manali et al., 2008,Nishiyama et al., 2010,Swigris et al., 2012a). An acute respiratory worsening requiring hospitalisation has been also associated to increased risk of subsequent death in many studies (Brown et al., 2015,Durheim et al., 2015,Song et al., 2011).

However, both severity and progression is IPF are best defined by means of physiology variables, in particular FVC. Decline in FVC is widely accepted as a reliable, easily measurable marker of disease progression and is linked to mortality in IPF (Collard et al., 2003,Jegal et al., 2005). Percentage predicted FVC has shown a minimal clinically important difference (MCID) of 2-6% (du

Bois et al., 2011a), and in clinical practice FVC has been consistently used by physicians as an indicator treatment initiation and as a criterion for referral for lung transplantation in the individual patient, and has been used as primary endpoint in most clinical trials so far. In particular, the categorical change in FVC carries significant prognostic information: 1-year mortality risk has been found to be up to 8 fold higher in IPF patients with significant decline (>10%) and 2-fold higher in those with a marginal (from 5% to 10%) decline in FVC at 6 months (du Bois et al., 2011a,du Bois et al., 2011b,Zappala et al., 2010). Moreover, it has been shown that using the relative >10% decline in predicted FVC (i.e., the change in percentage predicted FVC divided by the baseline value) rather than the absolute change may be more sensitive toward identification of a meaningful decline without affecting prognostic accuracy (Richeldi et al., 2012). A decline >15% in percentage predicted DLco has been also found associated with increased risk of mortality at 12 and at 6 months (Li et al., 2014,Zappala et al., 2010). Tolerance to exercise is another valid measurement of functional status with prognostic implications: a distance walked at the 6-minute walking test (6-minute walk distance, 6MWD) <250 meters at baseline and a 24-week change of >50 meters have been reported to be independent predictors of mortality (du Bois et al., 2014).

Radiologic features directly reflect the progression of the fibrotic process in the lungs, as such baseline or longitudinal HRCT indices of fibrosis represent an appealing alternative to physiology measures. The prognostic value of the presence and extent of different radiologic patterns on HRCT has been investigated by several studies. The extent of pulmonary fibrosis at baseline and its modifications on consecutive scans, evaluated semi-quantitatively using visual scoring systems, demonstrated to correlate well with parameters of functional impairment and predict poor survival in IPF patients (Best et al., 2008,Lynch et al., 2005). A recent study showed that a high fibrosis score on HRCT can predict a poor prognosis at 6 months even in patients with no significant functional deterioration, suggesting that the severity of radiologic findings may identify a subset of patients with a poor prognosis despite a stable FVC (Oda et al., 2014b). Furthermore, the presence and severity of individual abnormalities such as traction bronchiectasis has been found to predict mortality independently (Sumikawa et al., 2008).

Many candidate molecular biomarkers have been studied in IPF as indicators of disease severity, likelihood of disease progression and poor outcome (prognostic biomarkers). The term biomarker refers to a wide range of objectively quantifiable biological measurements that may act as surrogate for clinically meaningful variables, including proteins and cytokines in fluids such as blood or bronchoalveolar lavage, or genetic mutations and polymorphisms. An ideal biomarker should be easily acquired through non-invasive means, have high validity and reliability, and be available for serial monitoring. In order to be considered useful, the information it provides

should allow the provision of superior patient care beyond that of conventional practices. Despite the huge interest biomarkers gained over the last years in IPF, none has been implemented for use in the clinical practice due to the challenges offered by their validation and demonstration of clinical utility.

High baseline levels (>1000 U/ml) and serial increases in concentration of Krebs von den Lungen-6 (KL-6), a glycoprotein expressed on the surface of regenerating alveolar epithelial cells (AEC) type 2 cells and bronchiolar epithelial cells, have been found to be related to worse survival in patients with ILD (Satoh et al., 2006), but there are few and contrasting data as to it might improve the prediction of survival beyond known clinical parameters. Increased baseline levels of surfactant proteins A and D (SP-A and SP-D), lipoprotein complexes secreted into a liquid layer lining the epithelium and playing central roles in the host defence against pathogens, have also been demonstrated to be independent predictors of death or requirement for lung transplantation (Greene et al., 2002,Kinder et al., 2009), but when included in prediction models they didn't provide greater prognostic discrimination compared to routine clinical assessments (Song et al., 2013). Serial changes in levels of vascular endothelial growth factor (VEGF), a glycoprotein expressed in AEC promoting vascular permeability and regulating angiogenesis, seems to be correlated with decline of FVC and with poorer survival but the only available data come from a single-centre retrospective study (Ando et al., 2010). Cleaved cytokeratin 18 (CCK18), liberated during apoptosis of AEC, is significantly higher in IPF patients in comparison to normal controls and patients with other forms of ILD, even though its levels have not been associated so far with disease severity or outcome (Richards et al., 2012). Matrix metalloproteinases (MMPs) are a broad family of 23 different zinc-dependent proteinases involved in matrix turnover regulation, cell chemotaxis and mediator activation. MMP1 and MMP7 combined have shown potential to be used as a diagnostic marker in IPF, and MMP7 also demonstrated negative correlation with FVC, suggesting a role in predicting prognosis as well (Rosas et al., 2008). A study that prospectively evaluated the prognostic value of 95 different potential biomarkers demonstrated that high concentrations of MMP7, Intra-cellular adhesion molecule – 1 (ICAM-1), IL-8, vascular-cellular adhesion molecule – 1 (VCAM-1) and S100A12 were significantly associated with transplant-free survival (Richards et al., 2012). The validity of MMP7 as independent predictor of survival was further supported when inserted in a model including clinical parameters and MUC5B genotype (Peljto et al., 2013). Periostin, a protein secreted by bronchial epithelial cells in response to interleukin-13 (IL-13) and promoting extra cellular matrix deposition and mesenchymal cells proliferation, is elevated in patients with IPF and was demonstrated to correlate with disease progression and overall survival (Naik et al., 2012,Okamoto et al., 2011,Tajiri et al., 2015). YKL-40, a protein that regulates proliferation and survival of many cell types, seems to facilitate the

release of fibrotic and inflammatory mediators from alveolar macrophages and has a mitogenic effect on lung fibroblasts, and is significantly elevated in the lung tissue, BAL and serum of patients with IPF, and a small centre study demonstrated that both serum and BAL levels of YLK-40 are associated with poorer survival in IPF (Korthagen et al., 2011), suggesting a role as a prognostic marker. High baseline serum levels of CC chemokine ligand 18 (CCL18), a protein produced by alveolar macrophages that stimulates collagen production and differentiation in fibroblasts, were independently associated with physiological progression and, for levels > 150 ng/ml, with death, suggesting a potential as a prognostic marker. Moreover, serial measurements of CCL18 also correlated well with pulmonary function, suggesting that it could also possibly be used in the ongoing assessment of patients (Prasse et al., 2009). Lately, markers of adaptive immunity and inflammation have been investigated as they seem to provide useful information for patient stratification. IgG autoantibodies directed against heat-shock protein 70 (HSP70), which plays a protective role attenuating injury, inflammation and fibrosis, have been found in a proportion of patients with IPF and are associated with worse lung function and survival (Kahloon et al., 2013). Plasma concentration of circulating BLyS, a trophic factor for B-lymphocytes, are significantly greater in IPF patients than controls, and predicted transplant-free survival at 1 year (Xue et al., 2013). CXC chemokine 13 (CXCL13), implicated in the homing of B-lymphocytes to lymphoid aggregates and inflammatory foci, have been demonstrated to be elevated in IPF patients, and baseline levels and increases over time were associated with reduced survival (Vuga et al., 2014).

In 2009, the UK-based PROFILE (Prospective Observation of Fibrosis in the Lung Clinical Endpoints) study was launched with the aim to longitudinally evaluate disease behaviour in a large prospective cohort of well-characterised IPF patients to validate biological and clinical endpoints. Using cohorts from the PROFILE study, Jenkins and co-workers have recently focused their attention on collagen fragments generated by MMPs, and have demonstrated that the 3-months change in the serum levels of some of these fragments (called neoepitopes) is strongly predictive of mortality, suggesting a potential role for these biomarkers as early predictors of poor outcome (Jenkins et al., 2015). Circulating, bone marrow-derived mesenchymal progenitor cells called fibrocytes have been also investigated as potential biomarkers of progression in IPF. Fibrocytes can migrate into the lungs in response to different chemokines produced in the sites of injury. Once there, they differentiate into fibroblasts and myofibroblasts thus playing a role in tissue repair and fibrosis (Maharaj et al., 2013). Fibrocytes levels are higher in the peripheral blood of IPF patients as compared to healthy subjects, and the cut-off level of >5% (of total leucocytes) has been found to be associated with worse survival (Moeller et al., 2009). However,

the precise role of fibrocytes is not completely clear and there is still large uncertainty in literature about which cell surface markers to use to identify them (Maharaj et al., 2013).

The advances in modern "omics" technologies (genomics, epigenomics, transcriptomics, proteomics etc.) and the application of bioinformatics methodologies (enabling the automated analysis of large amounts of biological data) are indeed promising as to the identification of patients with more rapid progression. It has already been demonstrated that that different gene expression profiles correlate with different survival, irrespectively of clinical presentation at the time of diagnosis (Selman et al., 2007) It has also been reported that a molecular expression signature of 134 gene transcripts at the time of diagnosis has the potential to discern relatively stable IPF patients from those with a more rapid functional decline (Boon et al., 2009). The minor allele of the MUC5B polymorphism known to confer an increased risk of developing IPF (Seibold et al., 2011) has been found to be associated with better survival (Peljto et al., 2013). Similarly, IPF cases with the Toll-interacting protein (TOLLIP) major allele have been found to have a lower mortality risk as compared to minor allele carriers (Noth et al., 2013). These results pave the way to the stratification of IPF patients based on a genetically determined risk of progression, which might also enhance the design of future clinical trials.

Very recently it has been suggested that also the lung microbiome, i.e. the overall burden of bacteria in the lungs, may have a role in the pathogenesis and progression of IPF, namely specific members of the Staphylococcus and Streptococcus genera to be involved (Maoua et al., 2014). Molyneaux and co-workers have reported that a higher bacterial burden in BAL at the time of diagnosis allows identify patients with a more progressive disease (Molyneaux et al., 2014). The evidence of reduced mortality in patients with fibrotic ILD treated with co-trimoxazole supports the hypothesis of the influence of respiratory bacterial colonisation in the progression of the disease (Shulgina et al., 2013).

Tissue markers might be finally used for predicting prognosis - as well as response to specific treatments. Quantification by immunostaining of αvβ6 integrin, a key mediator for the activation of transforming growth factor β (TGF- β) and target of a novel monoclonal antibody, has been found associated with increased mortality (Saini et al., 2015), and Single-Photon Emission Computed Tomography (SPECT)/CT scanning techniques have been demonstrated to be capable of detecting and quantifying labelled αvβ6 integrin in the lungs (John et al., 2013).

2.4.3.2 Risk prediction scores and models

Several risk models, based on different combinations of demographic, clinical, physiology and radiologic parameters, have been proposed to stage IPF and predict survival. King and co-workers

firstly proposed a clinical-radiologic-physiologic (CRP) scoring system (King et al., 2001a) integrating 7 different parameters, but resulted too complicated for use in clinical practice. The composite physiologic index (CPI) developed by Wells and colleagues (Wells et al., 2003) was conceived on 3 pulmonary function measurements (DLco, the forced expiratory volume in 1 second - FEV_1 - and FVC) and was proven to correlate with fibrosis extent on HRCT scan of the chest. Mura and colleagues integrated the CPI in the Risk Stratification ScorE (ROSE) including two more variables, the Medical Research Council dyspnoea score and the 6MWD, and showed high specificity in predicting 3-years mortality (Mura et al., 2012).

Du Bois and co-workers identified 7 independent predictors of 1-year mortality using data from clinical trials and developed a simplified scoring system based on 4 of them: age, respiratory hospitalization, % predicted FVC and 24-week change in FVC (du Bois et al., 2011b).

Ley and co-workers (Ley et al., 2012) developed a simple, multidimensional prognostic staging system using data from 3 large and geographically distinct cohorts. 4 baseline variables were included in the model: gender (G), age (A) and two respiratory physiology variables (P) (FVC and DL_{CO}). The GAP index score identifies 3 stages of severity with 1-year mortality risk of 6%, 16% and 39%, respectively. Recently, the same group has proposed a longitudinal GAP model that incorporates the 24-weeks relative change in % predicted FVC and respiratory hospitalisations (Ley et al., 2015), with significant risk re-classification improvement. These same authors have also investigated the prognostic value of the scores of extent of fibrosis and emphysema on HRCT in the context of the GAP model (Ley et al., 2014). The fibrosis score has been proposed as potential replacement of DLco in a modified GAP model (the CT-GAP model) with comparable performance. This could be useful in clinical practice since DL_{CO} is difficult to obtain in the most severe IPF patients.

Table 3 - Proposed prognostic predictors in IPF

Category	Parameter	Evidence
Demographic/ Clinical	Age	Older age (King et al., 2001a)
	Sex	Male (Johnston I, 1990)
	Dyspnea and oxygen level	Baseline (Nishiyama et al., 2010) Changes at 6 months (Li et al., 2014)
	Comorbidities	Pulmonary hypertension (Corte et al., 2009) Pulmonary emphysema (Cottin, 2013) GERD (Lee et al., 2013) Lung cancer (Tomassetti et al., 2015)
Physiological	FVC	Baseline FVC <55% (Nathan et al., 2011) 6 months' decline > 10% or between 5-10% (Zappala et al., 2010)
	DLCO	Baseline (Nathan et al., 2011) 6 or 12 months' decline >15% (Li et al., 2014,Zappala et al., 2010)
	6MWD	Baseline < 250m (du Bois et al., 2014) 24-weeks' decline > 50m (du Bois et al., 2014)
Radiological	Fibrosis score	Baseline and changes at follow-up (Best et al., 2008,Lynch et al., 2005,Sumikawa et al., 2008)
	Traction bronchiectasis	Extent of traction bronchiectasis (Sumikawa et al., 2008)
Biomarkers	Serum and plasma biomarkers	Baseline levels of: SPA, SPD, KL-6, CCL18, YKL40, CXCL13, anti-HSP70 IgG, MMP7, MMP1, Periostin, etc.
	MMPs collagen fragments (neoepitopes)	Baseline levels and changes at 3 months (Jenkins et al., 2015)
	Fibrocytes	>5% (total leucocytes) (Moeller et al., 2009)
	αvβ6 integrin	Extent of immunostaining on lung IPF tissue (Saini et al., 2015)
Microbiome	Members of Staphylococcus and Streptococcus genera	Higher concentration in BAL at diagnosis (Maoua et al., 2014)
	Bacterial burden	Higher bacterial burden in BAL at diagnosis(Molyneaux et al., 2014)
Genetic	Muc5B promoter polymorphism rs35705950	Minor "risk" allele (improved survival) (Peljto et al., 2013)
	TOLLIP polymorphism rs5743890	Major "risk" allele (improved survival)(Noth et al., 2013)

2.5 Therapeutic approach

Historically, the pharmaceutical industry has been reluctant to invest in the development of drugs for rare diseases like IPF (Spagnolo et al., 2013). Nevertheless, following the publication of the first joint ATS/ERS consensus statement on the diagnosis and management of IPF more than fifteen years ago (ATS et al., 2000) the interest of stakeholders has progressively increased, leading to major advances in the understanding of IPF pathobiology thanks to the concerted efforts of dedicated academic institutions, patient organisations, healthcare authorities and pharmaceutical companies (Spagnolo et al., 2015a). The therapeutic approach to IPF management has indeed changed as the understanding of the pathogenesis of the disease has evolved over the last two decades. In fact, the initial thinking was in favour of a disease triggered by a persistent inflammatory process, resulting in the induction of fibrosis and scarring of the lungs. As such, several trials were performed evaluating the efficacy of drugs that primarily exert their functions by suppressing inflammatory or immune responses. Historically, patients with IPF have been (and most continue to be treated in many part of the world) with corticosteroids. A summary of the results available for efficacy of corticosteroids in IPF has been first published in 2003 as a Cochrane systematic review (Richeldi et al., 2003). At that time, no high-quality studies were identified and only non-randomised, retrospective, studies were available. Therefore, the conclusion was that there was a major lack of evidence supporting the use of corticosteroids in the treatment of IPF. An update of that same systematic review has been published in 2010 (Richeldi et al., 2010) and, surprisingly, did not identify any new additional randomised clinical trial on the use of steroids in IPF, thus confirming the persisting lack of evidence for their use in the management of this disease. This issue has been also reassessed in the evidence-based guidelines of 2011 (Raghu et al., 2011), in which a strong recommendation against the use of corticosteroid monotherapy in IPF has been made. This important recommendation relies on the availability of very low-quality evidence and places a high value on preventing treatment-related morbidity from long-term corticosteroid therapy. Low-quality evidence is also available for the use of non-steroid immunomodulatory drugs in IPF, such as colchicine, cyclosporine A, cyclophosphamide or azathioprine, either alone or in combination with corticosteroids (Spagnolo et al., 2010); as such, the 2011 guidelines placed strong recommendations against the use of immunomodulatory agents in the treatment of IPF patients. In 2005, a randomised clinical trial evaluated the effect of N-acetylcysteine (NAC) in patients already receiving combination therapy with prednisone and azathioprine (Demedts et al., 2005). Despite the evident limitations of the study, consisting in the substantial drop-out rate observed and the lack of a true placebo arm (i.e. not taking any potentially effective drug), the significant results in terms of change in vital capacity and diffusing capacity at 12 months supported the use of the so-called “triple therapy” in

the clinical practice for years; in 2011 no further evidence was available for this combination regimen, which therefore received a weak recommendation against use in the international guidelines (Raghu et al., 2011). Nevertheless, the increasing evidence that IPF was primarily fibrotic in nature progressively made a huge number of compounds being tested as new potential therapies for IPF, with more and more patients recruited into larger, well designed late phase randomised clinical trials. Despite majority of studies so far had disappointing results, the year 2014 has witnessed the approval of the first two effective anti-fibrotic agents, marking a turning point in the medical management of IPF patients and lighting up the hope for the development of a cure for this lethal disease (Hunninghake, 2014). However, so far no pharmacological therapy has been demonstrated to prolong survival in patients with IPF.

2.5.1 Current guidelines for treatment

Over the last five years, clinical management of IPF has been primarily based on the recommendations made by the 2011 ATS/ERS/ALAT/JRS joint statement (Raghu et al., 2011). In this document, recommendations for or against several pharmacological and non-pharmacological treatments have been formulated according to the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) evidence-based methodology (Schunemann et al., 2006), which rates the quality of available evidence and strength of recommendation following literature searches and assessment of the available information. According to the GRADE system, recommendations were made as either “strong” or “weak” based on the quality of available data and the evidence for efficacy. Back at the time, the panel of experts involved didn’t find enough evidence to recommend the use of any drug treatment for IPF, based on the negative or inconsistent results available. Lung transplantation was the only intervention receiving favourable recommendation given its proven survival benefit.

Since then though some key clinical trials have redefined the evidence for IPF therapeutic approach. Firstly, new important insights have been obtained from recent negative studies. The once widely accepted “triple therapy” (consisting in the combination of prednisone, azathioprine, and N-acetylcysteine), that received a weak recommendation against use in the 2011 statement, has ultimately proven to be associated to increased mortality as compared to placebo and to N-acetylcysteine alone in the three arms PANTHER-IPF study (Prednisone, Azathioprine, and N-acetylcysteine: A Study That Evaluates Response in IPF) (Raghu et al., 2012a). While the “triple therapy” arm was discontinued early due to safety concerns, the N-acetylcysteine arm completed the study but the monotherapy ultimately failed to demonstrate efficacy (Martinez et al., 2014). Anticoagulation, for years considered a potential therapy based on a small study that showed survival benefit in IPF hospitalised patients treated with warfarin (Kubo et al., 2005), has been

associated with increased mortality in the recent, early discontinued ACE-IPF (Anticoagulant Effectiveness in Idiopathic Pulmonary Fibrosis) trial (Noth et al., 2012). While dual endothelin receptor antagonists (ERAs) Bosentan and Macitentan have shown negative results in three phase 2 trials (King et al., 2008,King et al., 2011a,Raghu et al., 2013b), the ARTEMIS (a Randomised, Placebo-Controlled Study to Evaluate Safety and Effectiveness of Ambrisentan in IPF study) study on Ambrisentan, a selective ERA, has been stopped early for lack of benefit and a high likelihood of harm in the treated group (Raghu et al., 2013a). Most importantly, over the last few years two anti-fibrotic drugs, pirfenidone and nintedanib, have demonstrated to slow down the functional decline in IPF in the first successful phase 3 clinical trials in IPF (King et al., 2014b,Richeldi et al., 2014). Such evidence has led to the approval of pirfenidone firstly in Japan, then in Europe and ultimately in the US, while nintedanib has received approval by the US Food and Drug Administration (FDA) in late 2014 and has become available on the market across Europe in 2016.

Based on these and other recent findings, the 2011 clinical practice guidelines on treatment of IPF have been recently updated (Raghu et al., 2015a). In keeping with the 2011 document, the GRADE system has been adopted to express recommendations for or against the use of the treatments evaluated, with the only substitution of the term “weak” with “conditional”. While no treatment has been strongly recommended for use in IPF, some clarifying changes have been made. The combination of prednisone, immunosuppressive agents and N-acetylcysteine has received a strong recommendation against use, and so have other agents including warfarin. N-acetylcysteine in monotherapy has remained associated to a conditional recommendation against use. Most importantly, in the light of the recent positive trials pirfenidone, which received a weak recommendation against use in 2011, has been conditionally recommended in the update, and so has been nintedanib (which was not even addressed in the previous statement). Interestingly, nothing has changed as to the recommendation of anti-acid treatment in these patients (conditional recommendation for use), whose evidence is still relying though on retrospective data showing a functional or even survival benefit in patients treated with protonic pump inhibitors or histamine-2 receptor blockers (Lee et al., 2011) (Lee et al., 2013). In the next section, the main evidence supporting the therapies tested in phase 3 trials is being reported more in detail, while the evidence-based recommendations for the other treatments evaluated by the ATS/ERS/ALAT/JRS guidelines are summarised in Table 4.

Table 4 - Recommendations for pharmacological treatments of IPF - comparison between 2011 and 2015 ATS/ERS/ALAT/JRS guidelines

Evidence	**Drug (trial acronym)**	**Mechanism of action**	**2011 guidelines**	**2015 guidelines**
New	*Nintedanib (INPULSIS)*	*Tyrosine kinase inhibitor*	Not addressed	Conditional recommendation for use
	Ambrisentan (ARTEMIS)	*Selective endothelin A receptor antagonist*	Not addressed	Strong recommendation against use
	Imatinib	*Tyrosine kinase inhibitor*	Not addressed	Strong recommendation against use
	Sildenafil (STEP)	*Phosphodiesterase-5 inhibitor*	Not addressed	Conditional recommendation against use
Revised	*Pirfenidone (ASCEND)*	*Multiple mechanisms of action (unknown)*	Weak recommendation against use	Conditional recommendation for use
	Macitentan and Bosentan (BUILD-1 and -3)	*Dual endothelin receptor antagonists*	Strong recommendation against use	Conditional recommendation against use
	Warfarin (ACE)	*Anti-coagulant*	Weak recommendation against use	Strong recommendation against use
	Prednisone, azathioprine, and N-acetylcysteine (PANTHER)	*Immunosuppressant, anti-inflammatory, and antioxidant*	Weak recommendation against use	Strong recommendation against use
Unchanged	*Anti-Acid therapy*		Weak recommendation for use	Conditional recommendation for use
	N-acetylcysteine monotherapy (PANTHER)	*Anti-oxidant*	Weak recommendation against use	Conditional recommendation against use
	Interferon-γ-1β (INSPIRE)	*Anti-fibrotic, anti-proliferative, immunomodulatory*	Strong recommendation against use	Not addressed

2.5.2 Pirfenidone

Pirfenidone is an orally administered pyridine that has demonstrated combined anti-inflammatory, anti-oxidant and anti-fibrotic actions both in vitro and in experimental models of pulmonary fibrosis, consisting in regulation of the expression of transforming growth factor β (TGF-β) and inhibition of fibroblast and collagen synthesis. However, the precise mechanism of action remains unknown.

Four placebo-controlled randomised trials (Azuma et al., 2005,Noble et al., 2011a,Taniguchi et al., 2010) have explored the therapeutic role of pirfenidone in IPF. Most recently, the CAPACITY trials (Noble et al., 2011a), were two almost identical 72-week phase 3 trials conducted to study the efficacy of pirfenidone in reducing decline in lung function of patients with IPF with mild-moderate impairment of lung function (FVC of 35–90%, DLCO of 50–90% and a 6-minute walking test distance of at least 150 m). An efficacy dose-response relationship was observed in one study, with an absolute difference in FVC of 4.4% in the high-dose group (2403 mg/day) as compared to placebo at 72 weeks. In the other study, no significant differences were noted between the pirfenidone and placebo groups as to the primary outcome of percentage predicted FVC change. Nonetheless, the analysis of pooled data from the two trials supported a pirfenidone treatment effect on change of percentage predicted FVC, progression-free survival and distance at the 6-minute walking test. Most commonly reported side effects were gastrointestinal (nausea, dyspepsia, vomiting and anorexia), skin disorders (rash and photosensitivity) and dizziness.

A systematic Cochrane review of these data showed that treatment with pirfenidone reduces the risk of disease progression (as defined by progression-free survival) by 30% (hazard ratio 0.70, 95% confidence interval 0.56–0.88) (Spagnolo et al., 2010). Pirfenidone has been firstly approved for clinical use in Japan, Europe, India and Canada following regulatory appraisal of the data investigating efficacy and tolerability. The approval by the US FDA, due to concerns including a perceived lack of efficacy as measured by change in FVC and lack of survival benefit, required the conduction a further placebo-controlled phase 3 trial, the ASCEND (Assessment of pirfenidone to Confirm Efficacy and Safety In Idiopathic Pulmonary Fibrosis) study. This trial ascertained the beneficial effect of pirfenidone on disease progression in 555 patients randomly assigned to receive either the study drug (2403 mg per day) or placebo for 52 weeks (King et al., 2014b). In the pirfenidone group there was a relative reduction of 47.9% in the proportion of patients who had disease progression (defined as decline of 10% or more in percentage predicted FVC, or death from any cause). Also, there was a relative increase of 132.5% in the proportion of patients with no functional decline in the pirfenidone group as compared with the placebo arm. Rates of death from any cause or from IPF didn't differ significantly in the two groups, although in a pre-specified

pooled data analysis incorporating results from the three phase 3 trials (CAPACITY 1, CAPACITY 2 and ASCEND) the between-groups difference favouring pirfenidone was significant for death from any cause and death from IPF. The use of pirfenidone was associated with adverse events of generally mild to moderate intensity, such as gastrointestinal symptoms (nausea, dyspepsia), raised liver function tests and photosensitivity. Discontinuation of study treatment due to such adverse events occurred in 14% of patients in the pirfenidone group, a slightly higher rate than that observed in the placebo group (10%).

Post-authorisation data on the use of pirfenidone in clinical practice are already available both from international collaborative studies and from real-life experiences in Japan and Europe, where pirfenidone has been approved first. Interim reports from RECAP, an open-label extension study enrolling patients who completed the CAPACITY program (Costabel et al., 2014) and from PASSPORT (Pirfenidone Post-Authorization Safety Registry) (Maher et al., 2014), collecting prospective data from patients treated with pirfenidone from ten different countries, confirmed a favourable safety profile and good tolerability for the anti-fibrotic drug. Findings from different European studies have also provided important information, despite most of them reflect the experience of single centres with small populations of patients. A German study conducted at a tertiary referral centre for ILD following up IPF patients on pirfenidone showed that the majority (62%) remained stable during the treatment, even though the reduction of the decline of pulmonary function as compared to the period before the treatment was not statistically significant (Oltmanns et al., 2014). Interestingly, reports from both Europe and Japan showed that patients with a clear progression of disease before starting pirfenidone therapy have a more favourable course under pirfenidone treatment (Loeh et al., 2015,Okuda et al., 2013). Overall, pirfenidone confirmed good tolerability in these studies, as discontinuation due to adverse events ranged between 13% and 18%, quite consistent with the results of the RCTs. The safety profile was also very similar, despite weight loss and fatigue occurred more frequently in these real-life settings (Bonella et al., 2013,Okuda et al., 2013,Oltmanns et al., 2014).

2.5.3 N-Acetylcysteine

N-Acetylcysteine (NAC) is proposed to act on the lung by increasing intracellular and extracellular levels of glutathione thus exerting an antioxidant effect which counters the high levels of oxidative stress (driven by reactive oxygen species and free radical exposure) which the alveolar epithelium is exposed to. In IPF there is evidence that these antioxidant mechanisms are impaired and that this in turn contributes to epithelial injury and apoptosis. Levels of the key endogenous antioxidant glutathione are four times lower in the BAL fluid of individuals with IPF as compared with healthy controls (Cantin et al., 1989). The 2005 IFIGENIA (Idiopathic Pulmonary Fibrosis International Group Exploring N-Acetylcysteine I) trial (Demedts et al., 2005) showed that the addition of N-acetylcysteine to what was the standard therapy back then, i.e. prednisone and azathioprine combined, significantly reduced the rate of deterioration of the primary endpoints of the study, i.e. vital capacity (VC) and DL_{CO}, as compared to prednisone and azathioprine alone. A between-groups relative difference of 24 % was observed for DL_{CO} and of 9 % for VC. These results made the IFIGENIA trial the first to report a positive outcome in IPF. However, there were several limitations in this study, including the lack of a placebo arm and a significant patient drop-out rate, with the consequent need for imputation of a large set of missing data.

To further investigate the efficacy of NAC in IPF, the American consortium IPF network (IPFnet) designed and carried out a three-arm placebo-controlled trial, the PANTHER-IPF (Prednisone, Azathioprine, and N-Acetylcysteine: a study That Evaluates Response in Idiopathic Pulmonary Fibrosis) trial, sponsored by the National Heart, Lung and Blood Institute. Patients with mild to moderate lung function impairment were randomised in a 1:1:1 ratio to prednisolone, azathioprine and NAC (“triple therapy”), NAC alone or placebo. The primary outcome was the longitudinal change in percentage predicted FVC over a 60-week period. Following a pre-specified efficacy and safety interim analysis the independent data and safety monitoring board recommended termination of the combination therapy group at a mean follow-up of 32 weeks, having identified that the combination therapy was associated with a statistically significant increase in all-cause mortality, all cause hospitalisations, and treatment-related severe adverse events (Raghu et al., 2012a). The evidence that the “triple therapy” was harmful, having been considered the standard of care in IPF for over a decade, should be regarded as ground-breaking and underlines that even if negative, a properly designed randomised clinical trial can importantly inform guidelines, future trial design, and ultimately clinical practice. The NAC and the placebo arms of the trial continued enrolment and the results of the two-groups study demonstrated no difference with respect to the change of percentage predicted FVC at 60 weeks and to the majority of the secondary outcome measures, although patients in the NAC group reported having better mental well-being (Martinez et al., 2014).

2.5.4 Nintedanib

Nintedanib is an inhibitor of different tyrosine kinase receptors including platelet derived growth factor (PDGF) receptors α and β, vascular endothelial growth factor (VEGF) receptors 1, 2 and 3, and fibroblast growth factor (FGF) receptors 1, 2 and 3 (Hilberg et al., 2008). Activation of these receptors has been implicated in lung fibrosis pathogenesis (Allen et al., 2002,Coward et al., 2010), and in the murine model of bleomycin-induced fibrosis nintedanib was shown to prevent the development of lung fibrosis (Chaudhary et al., 2007).

A 52-week phase 2 randomised, double-blind, placebo-controlled trial (TOMORROW - To Improve Pulmonary Fibrosis With BIBF 1120) (Richeldi et al., 2011) evaluated the safety and efficacy of nintedanib at four different oral doses (50 mg once a day, 50 mg, 100 mg, or 150 mg twice daily). The primary endpoint of the study was the annual rate of decline in FVC. Secondary endpoints included rate of acute exacerbations, quality of life and total lung capacity. Nintedanib at a dose of 150 mg given twice daily showed a trend toward reduction of the decline in lung function, with fewer acute exacerbations and a preserved quality of life as compared with the placebo arm. In the group receiving 150mg of nintedanib twice daily, FVC declined by 0.06 litres per year, as compared to 0.19 litres per year in the placebo group, a 68.4 % relative reduction in the rate of functional loss. In addition, patients treated with the higher dose of this drug experienced a lower incidence of acute exacerbations and a small reduction at the St. George's Respiratory Questionnaire (SGRQ) scores (as compared with an increase of the scores in the placebo group, suggesting a better or at least more preserved quality of life in treated patients). Gastrointestinal side effects (diarrhoea, nausea, and vomiting) and increase in the levels of liver enzymes were more frequent in the group receiving 150 mg twice daily than in the placebo group. In order to confirm these findings, two parallel, identical 52-week phase 3 placebo-controlled trials (INPULSIS 1 and 2) (Richeldi et al., 2014) evaluated the efficacy and safety of nintedanib 150 mg twice a day in patients with IPF, with the same primary endpoint of annual rate of decline in FVC (population enrolled: 1066 patients). The results of both studies confirmed the efficacy of nintedanib in reducing the rate of decline in FVC, prompting the approval of the drug for use in patients with mild-to-moderate IPF. As for INPULSIS 1, FVC declined by 114.7 ml per year in the nintedanib group, as compared with 239.9 ml per year in the placebo group (difference of 125.3 ml per year). In INPULSIS 2 FVC declined by 113.6 ml in the nintedanib group and by 207.3 ml in the placebo group (difference of 93.7 ml per year). However, differently from the previous phase 2 trial, no consistent nintedanib effect was observed on the time to the first acute exacerbation or on the change in the total SGRQ scores.

2.5.5 Treatment of AE-IPF

Clinical management of AE-IPF is one of the major challenges in respiratory medicine, since there are no therapies of proven efficacy. This leaves the clinician consider supportive care and interventions not supported by an appropriate base of evidence. Many patients with AE-IPF are still being administered systemic corticosteroids, but there is no clear scientific evidence in support of such approach. Guidelines on management of IPF provide a weak recommendation in favour of the use of systemic corticosteroids, clarifying that such recommendation is based exclusively on anecdotal reports of benefit and on the high mortality of exacerbations (Raghu et al., 2011). It is therefore essential that well-designed and robust clinical research is conducted over the next decade to provide an absolutely needed body of evidence. On the other hand, there is no doubt that the management of AE-IPF must include supportive therapy focused on the palliation of symptoms and the correction of hypoxemia with high oxygen flows. Nevertheless, there is debate upon the time which such supportive therapy should be extended for – particularly with regards to the use of mechanical ventilation. Guidelines provide a weak recommendation against the use of mechanical ventilation to treat respiratory failure in IPF, based on an intra-hospital mortality estimated to be up to 90% in this (Raghu et al., 2011). Guidelines specify that the decision of using mechanical ventilation in this context should be carefully considered and discussed together with the patient and its relatives upon the base of a very clear understanding of the objectives of care of the individual patient. More conservative approaches towards respiratory failure and lung injury (e.g. low tidal volume ventilation) might improve the intra-hospital survival (Rush et al., 2016); this remains an area of careful investigation.

There have been many studies published over the last decade that describe different potential treatments for AE-IPF. Unfortunately, these trials are small and uncontrolled. The only randomised clinical trial investigating an intervention for AE-IPF compared the procalcitonine-guided antibiotic treatment with standard, clinically-driven antibiotic treatment (Ding et al., 2013). In the procalcitonine-guided group antibiotics were administered if the levels of procalcitonine were above 0.25 ng/ml and the therapy was interrupted when procalcitonine became < 0.25 ng/ml. The procalcitonine group presented a reduced duration of the antibiotic treatment (on average 5.8 days shorter), but the duration of mechanical ventilation and mortality were similar between the two groups. Observational cohort studies compared each of the following treatments with a historical or parallel control arm: cyclosporine (Homma et al., 2005,Inase et al., 2003,Sakamoto et al., 2010), rituximab combined with plasmapheresis and intravenous immunoglobulins (Donahoe et al., 2015), orally administered tacrolimus (Horita et al., 2011), and intravenous trombomoduline (Isshiki et al., 2015,Kataoka et al., 2015,Tsushima et al.,

2014). Uncontrolled cohort studies have also been published reporting the outcomes of patients with AE-IPF treated with steroids (Akira et al., 1997,Al-Hameed et al., 2004,Suzuki et al., 2011,Tachikawa et al., 2012), steroids combined with - or followed by - other immunomodulatory therapies (Akira et al., 2008,Donahoe et al., 2015,Fujimoto et al., 2012,Tachikawa et al., 2012,Yokoyama et al., 2010), and with perfusion of polymyxin B-immobilized fiber column (Abe et al., 2012,Abe et al., 2011,Oishi et al., 2013,Seo et al., 2006,Tachibana et al., 2011). Unfortunately, it's not possible to determine which changes in the clinical status of these patients can be related to the intervention, to the natural history of their disease or to other non-measured factors. Potential treatments among those mentioned should be appropriately studied in prospective randomised controlled trials to understand whether they are potentially beneficial.

Given the absence of an effective treatment strategy for AE-IPF, prevention becomes particularly important in consideration of the extremely high mortality rate. Part of the prevention strategy is based on targeting of those specific triggers that can lead to the acute lung injury. In this context, the risk of infection can be decreased through control measures such as anti-influenza and anti-pneumococcal vaccination. Anti-acid therapy has been studied as preventive treatment in a large clinical cohort study, suggesting benefit (Lee et al., 2013). In the above-mentioned studies showing an association between the treatment and improved survival, efficacy remains uncertain due to possible confounding factors. For example, patients with a more advanced disease could be considered "too sick" for the treatment, or conversely only those who survive long enough to be considered for the experimental therapy could have been treated. Furthermore, historical control populations may have received a sub-optimal supportive care and more frequently immunosuppressive therapy. In both cases, such differences might lead to apparent benefit in favour of the intervention. Even environmental pollution control measures may have a role in reducing the risk of AE-IPF in those regions with poor air quality (Johannson et al., 2014). However, if AE-IPF is considered a component of the pathobiologic process in IPF, and considering that AE-IPF is more common in patients with more advanced disease, targeting the pathogenetic mechanisms of IPF with anti-fibrotic therapies might be crucial. The results of the clinical trials on nintedanib and pirfenidone suggest that these therapies, already approved for the treatment of IPF, might help prevent the development of AE-IPF. Data in favour of nintedanib are more solid, since AE-IPF was a key secondary endpoint in all three randomised controlled trials. The TOMORROW phase 2 study with 432 participants demonstrated in the nintedanib group a delay in the time to the first AE-IPF as reported by the investigators (Richeldi et al., 2011), a result not confirmed though by the phase 3 program (INPULSIS 1 e 2) with 1066 participants (Richeldi et al., 2014). More significant however in the phase 3 program was the reduction, in the nintedanib arm, of the incidence of AE-IPF centrally adjudicated as confirmed or suspected (5.7% for the

placebo arm versus 1.9% for the nintedanib group) (Richeldi et al., 2014). A phase 2 study on pirfenidone with 107 participants was discontinued prematurely due to a statistically significant (but numerically small) reduction of AE-IPF in subjects in treatment with pirfenidone, but a follow up study with 275 participants did not replicate these results (Azuma et al., 2005,Taniguchi et al., 2010). Conversely, the definitive trials on pirfenidone didn't report AE-IPF among the endpoints. (King et al., 2014b,Noble et al., 2011a). It has been suggested that pirfenidone may reduce the risk of post-surgery AE-IPF, but these data are only observational and thus at high risk for confounding factors (Iwata et al., 2015). Further data are needed to fully understand the impact of anti-fibrotic therapies on the risk of developing AE-IPF, and so it is yet to be clarified if such therapies should be continued or interrupted in patients experiencing an AE-IPF. Other potential therapies tested in randomised clinical trials didn't show any impact on the prevention of (NAC in monotherapy, bosentan, interferon gamma, sildenafil) or on the increase in the risk for (ambrisentan, imatinib, combination of prednisone, azathioprine and NAC, warfarin) AE-IPF (Bando et al., 2010,Daniels et al., 2010,King et al., 2009,King et al., 2011a,Kubo et al., 2005,Martinez et al., 2014,Noth et al., 2012,Raghu et al., 2012a,Raghu et al., 2013a).

2.5.6 Practical implications for therapeutic approach

The results of the ASCEND and of the two INPULSIS trials represent an outstanding breakthrough for management of IPF, as they opened to the possibility of offering two safe and effective treatment options to patients suffering from this dreadful disease. Nevertheless, there are some limitations to the current therapeutic approach, and several questions remain unanswered.

Firstly, neither pirfenidone nor nintedanib can be considered a cure for the disease, as most patients continue to experience disease progression over time despite treatment, and it is not certain whether these drugs may remain effective in the longer term. Furthermore, none of the studies has shown a statistically significant reduction in mortality in the treated groups: as such, it is not sure whether the reduction in functional decline can be translated into a survival benefit. Also, a consistent impact on symptoms and quality of life was not proven either. Finally, it's not known whether the beneficial effects of these drugs might be generalised to the whole IPF population, as the eligibility criteria of clinical trials allowed the inclusion of patients with specific characteristics.

Indeed, pirfenidone and nintedanib appear to have comparable efficacy in reducing the functional decline as well as similar tolerability, as the treatment-related adverse effects led to the discontinuation of treatment in the trials in about 15% of patients treated with pirfenidone and 19% of those treated with nintedanib (King et al., 2014b,Richeldi et al., 2014). Moreover, the side

effect profiles show some overlap regarding gastrointestinal effects (nausea and diarrhoea, mostly) and elevation of liver enzymes. Hence, it remains unclear which one of these molecules should be used first and in which patients, although the initial choice should be based on the careful consideration of the patient's individual clinical features, including age, performance status, stage of the disease and comorbidities, using a decision framework which takes into account personal preferences. Whatever the initial choice, a sequential approach should be sought considering the replacement of the first agent when side effects are not tolerable or there is clear evidence for treatment failure, e.g. by identifying rapid progression with a decline of FVC > 10% over 6 months of treatment. Indeed, the choice to switch from pirfenidone to nintedanib is currently a more frequent scenario to face in clinical practice, since pirfenidone was licensed earlier. A small retrospective study on 7 patients switching to nintedanib due to adverse effects or progression experienced under pirfenidone treatment has shown that nintedanib may be better tolerated despite the similar tolerability profiles (Milger et al., 2015). However, further data is needed to draw any conclusions, in particular nothing is known yet from patients starting treatment with nintedanib and eventually shifting to pirfenidone. Timing for treatment initiation remains also questionable. Theoretically, both drugs could be used to try to prevent progression in those subjects who have been diagnosed at earlier stages. However, should a patient remain stable once treatment has started, there is no marker to demonstrate the anti-fibrotic drug is to be accounted for such stability. A close clinical monitoring of asymptomatic patients with limited or no functional impairment might be also appropriate, although recent post-hoc analyses of pooled data from the big trials on the efficacy of both nintedanib and pirfenidone suggest not only that subgroups of patients with more preserved lung function at baseline have a similar rate of progression of the disease as compared to patients with more impaired lung function (as demonstrated by the observation of the placebo arms), but also that these subgroups receive the same benefit from both anti-fibrotic drugs (Albera et al., 2016,Kolb et al., 2016). This evidence is in support of an early commencement of anti-fibrotic therapy in patients with IPF regardless of their functional impairment at the moment of diagnosis, however a careful assessment should be always carried out on an individual basis to balance risks and potential benefits of an early start of treatment -also considering the high costs of both therapies - at least until there are markers that can predict individual disease course (Spagnolo et al., 2013). Conversely, whether nintedanib or pirfenidone might be safe and beneficial in IPF patients with more severe functional impairment is unknown, because of the exclusion of subjects with FVC <50% from most RCTs based on the assumptions that advanced disease is less likely to respond to treatment or more frequent and severe adverse events may occur. Since some regulatory agencies such as the FDA included this population when approving both pirfenidone and nintedanib, much is left again to the judgement of the clinicians.

When available and appropriate, enrolment in a clinical trial should be always considered as an option and discussed with the patient. This offers access to new, potentially beneficial therapies and gives the patients the opportunity to play an active role in their own management and to be followed by expert medical staffs in specialised centres. Of course, those trials excluding patients already on treatment with an approved drug should be reserved to those cases where pirfenidone and/or nintedanib have already failed or cannot be prescribed.

Combination therapy is an appealing approach, especially considering that most of the treatments currently used in other fields of respiratory medicine such as asthma, chronic obstructive pulmonary disease, pulmonary hypertension and lung cancer are based on combination of two or more molecules (Wuyts et al., 2014). Indeed, the association of drugs with proven efficacy or, alternatively, the addition of a promising agent to a background effective therapy are likely to represent the future of pharmacological therapy in IPF. However, it seems still a long way to go before several uncertainties are being unravelled. To date not much is known about the interactions between the two approved drugs when administered together, both in terms of tolerability and efficacy. A small Japanese study has recently evaluated for the first time the safety, tolerability and pharmacokinetics of the two drugs combined in a small cohort of 50 patients (Ogura et al., 2015). The combination was not associated to serious adverse events, but the cohort was too small to derive any other conclusion. Interestingly, it showed that while exposure to nintedanib decreased when added to pirfenidone, the latter was not affected. A larger, multicentre phase 2 open-label, multiple dosing trial to investigate the pharmacokinetics of nintedanib and pirfenidone when administered separately or in combination has been recently conducted in the UK (ClinicalTrials.gov Identifier: NCT02606877). Until further, solid data becomes available, it is advisable to avoid the concomitant use of the two drugs given the risk associated to the partly overlapping side effect profiles.

Despite the results of the PANTHER phase 3 trial were inconclusive as to the efficacy of NAC monotherapy (Martinez et al., 2014), its use in clinical practice is still largely diffuse, mostly due to its good tolerability profile and the low costs of treatment. Indeed, the rationale behind its use - to contrast the increased oxidative stress in the lungs of IPF patients - still stands. Nebulisation of NAC via aerosol has been studied as a supposedly more efficient way for delivering the compound directly to the lung parenchyma, as indicated by a few Japanese studies showing that NAC inhalation increases neutralisation of oxygen radicals in a mouse model of bleomycin-induced lung fibrosis and can improve the redox imbalance in IPF patients (Hagiwara et al., 2000,Yoko et al., 2012). Recently, a small case-control study in severe IPF patients showed that combination of inhaled NAC and pirfenidone might reduce functional decline and improve prognosis as compared to pirfenidone alone, suggesting the need for further investigation. A post-hoc analysis conducted

in patients from the PANTHER trial has studied the influence of polymorphisms in TOLLIP and MUC5B genes (already associated with IPF susceptibility and survival) on the effect of NAC antioxidant therapy, building on the rationale that such polymorphisms are linked to alteration in the lung immune response through oxidative signalling (Oldham et al., 2015). Despite the results were ambiguous (NAC treatment has been associated with improved survival only in patients with the TT genotype of a TOLLIP polymorphism, while those carrying the GG genotype have shown a worse prognosis) and it remains unclear whether a mechanistic ink exists between oxidative stress and TOLLIP-related signalling, this evidence suggests that NAC might be efficient in a subset of patients. On the other hand, data from recent studies have seriously hampered the possibility for a role of NAC in the therapeutic approach of IPF. Patients with concomitant corticosteroids and/or NAC treatment in single-centre studies on the use of pirfenidone in real-life settings have experienced more adverse events than patients on pirfenidone in monotherapy and a higher rate of disease progression (Bonella et al., 2013,Oltmanns et al., 2014). Most importantly, the first randomised phase 2 trial comparing the safety and tolerability of the combination of NAC and pirfenidone with pirfenidone in monotherapy (PANORAMA) has shown results with potential repercussions on the use of NAC in the clinical management of IPF. Despite no substantial differences have been reported in terms of general tolerability after 6 months of treatment, the study has registered an increase of photosensitivity events in the combination therapy arm. Also, this group has shown an increased rate of decline in FVC as compared to the group in treatment with pirfenidone alone (change in FVC 125.6 ml and 34.3 ml respectively, difference of 91,3 ml at 6 months) (Behr et al., 2016). Despite efficacy of the association represented an exploratory outcome, these data suggest not only that the addition of NAC is not beneficial in patients with IPF in anti-fibrotic treatment with pirfenidone, but also that such combination might be harmful. This evidence, together with the lack of efficacy for NAC monotherapy as reported by the PANTHER-IPF study, might represent the end for testing NAC as a treatment option in IPF.

Apart from the implications deriving from the last advances in the pharmacological development, it must be reminded that the management of IPF should be always also based on the best supportive care. Referral for lung transplantation should never be delayed as it still represents the only life-changing intervention for IPF patients, who seem to benefit from better long-term survival compared to patients undergoing transplant for other indications (George et al., 2011). Comorbid conditions such as gastro-esophageal reflux disease, pulmonary arterial hypertension and obstructive sleep apnoea are frequent and can significantly impact quality of life and survival in these patients (Fell, 2012). Furthermore, there is evidence suggesting that they might be directly involved in IPF pathogenetic mechanisms either contributing to or resulting from recurring alveolar epithelial injury (Farkas et al., 2011,Lederer et al., 2012,Lee et al., 2012). As

such, concomitant conditions should be always explored in the individual patient and properly treated if clinically indicated. Their presence should be also considered very carefully when choosing a drug in the individual IPF patient. Having these patients been often excluded from clinical trials so far, the tolerability of nintedanib and pirfenidone in the general population might be different from what we might expect. Finally, emotional and psychological disturbance seem to be under-recognised and underestimated in the IPF population. Both anxiety and depression appear to be very common in these patients (Akhtar et al., 2013,De Vries et al., 2001,Holland et al., 2014,Ryerson et al., 2012,Ryerson et al., 2011) and have been found to be associated with higher levels of breathlessness irrespectively of disease severity (Holland et al., 2014,Ryerson et al., 2012), potentially generating further disability and isolation as part of a dangerous vicious circle. As such, there is a compelling need for alternative interventions as part of a more comprehensive care of these patients. Non-pharmacological management strategies such as pulmonary rehabilitation have been advised (Ferreira et al., 2009), although convincing evidence of efficacy is currently lacking. Psychological and behavioural interventions that could have an impact on patient wellbeing, ranging from individual psychological counselling to participation to patient support groups and attendance to mindfulness programmes are still lacking substantial evidence to support their adoption in the clinical practice.

2.5.7 Challenges in the development of new treatments

2.5.7.1 Incomplete knowledge of pathogenesis

The understanding of IPF pathogenesis has significantly improved over the last two decades, thus changing the approach to drug development. The previous paradigm considered IPF as the development of an inflammation-based process, which justified – despite the lack of proper evidence - the use of corticosteroids and immunosuppressive agents as stated in the first 2000 international guidelines (ATS et al., 2000). Since then though there has been growing evidence that IPF is instead the result of an abnormal wound healing response where persistent alveolar epithelial cell micro-injury is the trigger for the activation and the self-sustenance of several pro-fibrotic cascades, ultimately leading to lung fibroblasts and myofibroblasts expansion with secretion of disproportionate amounts of extracellular matrix components (Coward et al., 2010,King et al., 2011b). Many molecules including mediators, grow factors, and cytokines are contributing to this aberrant response. Despite many have been proposed as potential therapeutic targets so far, it's still unsure which of the several pathways involved are the most important perpetrators of the fibrotic process. Most of the past RCTs assessed the efficacy of agents initially developed for use in other diseases, and tested them directly in phase 2 or even 3 trials with lack of strong, supportive pre-clinical data in the setting of IPF. Whilst some of these

drugs are still currently under investigation, others (like the endothelin receptor antagonists bosentan and ambrisentan, both approved for treatment of pulmonary arterial hypertension) have demonstrated no efficacy or even to be harmful in IPF patients (King et al., 2008,King et al., 2011a,Raghu et al., 2013b). Conversely, multiple novel therapies have distinct mechanisms based on more targeted anti-fibrotic actions. Given the wide range of cells and mediators involved and their redundant interactions, the development of compounds that can target different biologic processes seems preferable, a concept further supported by the failure of most single-target agents and concomitantly by the evidence of efficacy of pirfenidone and nintedanib, both carrying pleiotropic properties.

2.5.7.2 Lack of optimal models

Despite many agents proved to have anti-fibrotic properties when tested in experimentally induced lung injury in murine models, only two drugs have ultimately demonstrated efficacy in late phase RCTs on IPF patients. One of the reasons for this discrepancy is that current animal models, indeed useful to explore the potential efficacy of a drug in the pre-clinical phase, are far from optimal in IPF. The most extensively used murine model of bleomycin-induced fibrosis is not representative of the pathological features of IPF, and only partially recapitulates the mechanisms of progressive pulmonary fibrosis (Moeller et al., 2008). The initial chemical damage produced by the administration of bleomycin results in extensive epithelial cell apoptosis and necrosis, which trigger an inflammatory response mediated by neutrophils. The subsequent fibrogenic process evolves in a self-limiting and at least partially resolving alveolar damage (Moeller et al., 2008,Moore et al., 2013). Consequently, its usefulness for prompting the advancement of novel agents to human trials has been largely discussed.

In order to improve such model, the timing of administration of the candidate drug has been addressed: a “therapeutic” administration (i.e. once bleomycin has already caused the injury) seems ideally preferable to a “prophylactic” one, as it partially allows the decrease of the intense inflammatory response giving time for the fibrotic process to develop. However, the failure of human trials testing compounds that showed efficacy in the “therapeutic” bleomycin model undermines this hypothesis (Raghu et al., 2013a,Spagnolo et al., 2015b).

Advances in the current models of IPF have been made by the association of the bleomycin model with gene knockout/mutation (e.g. surfactant proteins A and C, telomerases, Periostin) or, conversely, gene augmentation strategies (TGF-β, IL-13), that can increase the susceptibility to fibrotic stimuli and help the understanding of gene-environment interactions and mechanistic pathways (Blackwell et al., 2014). The overexpression of TGF-β in alveolar epithelial cells can be achieved via viral infections or as a result of a doxycycline-regulated transgenic expression, but

further data about the subsequent alveolar damage and pro-fibrotic pathways are needed in order to validate the model.

The use of exogenous agents other than bleomycin also represent an intriguing alternative. The fibrotic response to inhaled silica or to asbestos in murine models better mimics the lung damage occurring in humans, but takes longer to develop than an intra-tracheal instillation model. However, the latter is limited by the development of an uneven and centrally predominant (rather than sub-pleural) pattern of fibrosis. Radiation-induced lung fibrosis in mice may also mimic an interstitial lung fibrosis as a result of direct lung injury, but the fibrotic response is likely to be systemic if other organs are not shielded, and the model remains associated with low mortality (Moore et al., 2013).

While young mice adults have been so far extensively used for experiments of induced lung injury, it has been demonstrated that the lungs of aged mice might show a more pronounced fibrotic response to bleomycin, as in Hermansky-Pudlak and some virus-driven models. Since IPF is a disease of aging, the use of elder animals may predict the clinical relevance of experimental drugs in a more reliable way. Main limitations lie though in the difficulties in the selection of appropriate aged animals and in the high breeding costs (Moore et al., 2013)

Apart from animal models, the development of complex *ex vivo* models of IPF, starting from tissue samples of living patients undergoing surgical lung biopsies and consisting in three dimensional cultures of fibroblasts and other cellular types should provide more clinically relevant readouts than simple immortalised fibroblasts cultures.

2.5.7.3 Diagnostic fallpits

As mentioned above, according to the 2011 evidence-based guidelines (Raghu et al., 2011) the diagnosis of IPF is based on the identification of a radiological and/or histological pattern consistent with UIP when all known causes of lung fibrosis have been excluded. When HRCT scan is non-diagnostic a surgical lung biopsy is required for confirmation of the diagnosis. However, in clinical practice such procedure is being performed in a minority of patients, due both to the associated risks – especially in elderly patients, functionally compromised and with comorbidities – and to the natural reluctance of many subjects to undergo surgery. While the need of making a non-invasive, definite diagnosis of IPF in these patients remains unmet, in some cases even the integration of clinical, radiological, and histological data may not be sufficient to reach a confident diagnosis of IPF. These groups of patients don't fall in any diagnostic category as per current guidelines and have been excluded from most of RCTs so far, thus limiting the generalisation of

the results and with the consequence that they might not gain access to specific therapeutic options when available.

2.5.7.4 Limitations of RCTs

With few exceptions, most RCTs conducted so far (even the successful ones like INPULSIS and ASCEND) have included highly selected patients with mild to moderate disease (as defined by functional parameters at baseline), without significant comorbidities and in a specified age range. On one hand these patients are not representative of the whole IPF population, hence it is not clear whether the results obtained may be applied to patients with a more severe functional impairment or presenting significant comorbidities. On the other hand, the patients satisfying these eligibility criteria are still very heterogeneous in terms of rate of progression of the disease. Efficacy endpoints in these trials were driven only by the minority who significantly declined over the limited period of the study, whereas approximately two thirds of them remained more stable. Since there are no tools to determine whether functional stability is due to a benign course of the disease or to a treatment-related effect, such lack of stratification limits the efficiency of trials by impeding the achievement of high event rates. Finally, the limited duration of RCTs (usually 52 or 72 weeks) makes impossible to ascertain whether the effectiveness of an anti-fibrotic therapy can hold for longer periods (Spagnolo et al., 2015c).

The choice of the most appropriate primary clinical endpoint when designing late phase RCTs remains highly debated. Despite being the primary endpoint in most trials, the use of longitudinal change in percentage predicted FVC carries some limitations. Firstly, in the individual patient it doesn't allow distinguish between stability resulting from a drug-related effect or as part of the natural history of the disease. Secondly, it's highly prone to missing data, leaving room for misinterpretation due to different methods for statistical imputation adopted in the analysis of the results (Thabut et al., 2015). Its validity has been also argued for not representing a good surrogate endpoint for mortality. In stark contrast, mortality has been proposed by some regulatory agencies as the only clinically meaningful endpoint for late phase RCTs in IPF (Raghu et al., 2012b). However, the feasibility of using mortality as primary endpoint is limited by the number of patients and the study duration required for the design of an adequately powered study (du Bois et al., 2012). A retrospective analysis of phase 3 RCTs data estimated that a hypothetical mortality trial, in order to detect a 25% reduction in all-cause mortality in the treatment group, should enrol more than 2500 patients over 3 years and a with a follow-up period of 5 years (King et al., 2014a). With the likely perspective of the inclusion of approved IPF therapies in future trials, it can be expected that the mortality rate would be even lower in a cohort of patients with mild to moderate disease. As such, the use of mortality as an endpoint

should be reserved to populations with end-stage disease or selected patients at high risk of disease progression or acute exacerbation in well-designed studies adopting cohort enrichment strategies (Nathan et al., 2014).

2.5.7.5 Strategies for optimisation of RCTs

Building on the advances in the understanding of the pathobiology and clinical behaviour of IPF achieved in the past decade, the availability of the first effective treatments and the lessons learned from the past unsuccessful trials, the hope is that the next few years will witness the rise of an increasing number of new effective therapies, with the goal of finding a cure for IPF. However, the conduction of effective RCTs in IPF is going to be a continuous challenge over the next years. As noted above, heterogeneity of IPF has not been addressed yet. Stratification of IPF patients based on disease progression and response to different treatments remains a challenge due to the complex network of different pathways interacting together and the unforeseeable influence of environmental exposures on the development of the disease. Furthermore, with two drugs approved and more possibly available in the next few years, the adoption of placebo arms in trials is becoming unacceptable, both for recruitment difficulties and for potential biases affecting patient selection. This will lead to the rise of comparative trials between a novel agent and an approved treatment (or a combination of treatments), whose conduction (either with design of superiority of non-inferiority) will be challenged by the requirement of larger sample size or longer duration due to smaller differences in treatment effect.

In order to overcome these limitations, greater efforts in the application of effective cohort enrichment strategies will be required. The validation of markers of likelihood of disease progression (prognostic biomarkers) will be critical, as the inclusion of patients at increased risk of progression would increase the event rates for different outcomes of interest thus allowing the reduction of the size and length required for detection of a drug's effect. A simple and logical strategy for prognostic cohort enrichment would be including patients with more severe physiological impairment at baseline or with evidence of previous functional decline, as they are supposedly more likely to progress over time. However, it has been discussed above (section 2.4.3) how baseline and longitudinal measurements, including physiology parameters, have failed so far in predicting subsequent progression in retrospective investigations on pooled data from past clinical trials. The establishment of large, longitudinal cohorts of patients and the adoption of a comprehensive approach for their characterisation, including demographic, clinical, and biological data, might allow define a specific "signature" for patients with IPF, providing valuable information on prognosis and expected response to available therapies (Maher, 2013).

On the other hand, the proper design of early phase, proof of concept studies will lead to the development of biomarkers that could inform the specific mechanism of action of the agents under study and represent short-term readouts of the expected biological effect in patients. This approach will enhance the efficiency of later development phases by identification and inclusion of those subjects who are more likely to respond to a specific compound (predictive cohort enrichment). Despite lacking strong mechanistic grounds, the afore-mentioned study on the correlation between TOLLIP and MUC5B polymorphisms and the response to NAC among IPF patients can be considered as the first pharmacogenomic study in patients with IPF, and has got relevant implications as to the importance of conducting prospective, enriched genotype-stratified trials to head towards a more personalised therapy. Although it is still a long shot, the definition of unique fingerprints for IPF individuals through the combined application of molecular techniques remains the most exciting goal in the field of IPF research, as it would dramatically impact disease stratification and resolve the (only apparent) paradox seeing combination and individualised treatment as mutually exclusive approaches.

Indeed, over the next few years the choice of primary endpoints in late phase trials will be based still on clinical measurements, bearing in mind that it should always adapt to the characteristics of patients included in the trial. Despite the mentioned debate on the clinical meaningfulness of FVC, the recent approval of two drugs based on studies adopting longitudinal change in % predicted FVC points out that this can be considered an acceptable endpoint for regulatory authorities. Yet, more efforts should be made to standardise the methods for handling missing data in order to make results between studies more comparable (Thabut et al., 2015). Indeed, the use of FVC as a stand-alone primary endpoint is likely to be an impractical choice in future placebo-free trials, and the incorporation of FVC in event-driven composite endpoints including other outcomes of interest (such as exacerbations, mortality and measures of symptoms burden and quality of life) is likely to be needed to reflect a wider range of pathophysiological consequences related to disease progression and maximise the chances of detection of a drug's effect by increasing the overall event rate. The development of composite end points for future trials should be specifically targeted to demonstrate either consequent bad outcome (e.g. including mortality, respiratory hospitalisation, major decline in FVC or the 6-minute walk distance, or the development of pulmonary hypertension) for more accurate short-term prediction, or chronic disease progression (e.g. associating FVC change with changes in DLco, symptoms and computed tomography indices), more relevant to long-term outcome.

As noted above, it is still uncertain whether the results of clinical trials conducted on patients with definite IPF also apply to those with "probable" or "possible" IPF. However, the example of the successful TOMORROW and INPULSIS, which adopted broader inclusion criteria, indicates that

these patients should be allowed entering future trials - at least as a subgroup - for testing pharmacological effects in a more representative population. Results from a pre-defined subgroup analysis of the INPULSIS trials have confirmed that the efficacy of nintedanib is not different between patients with confirmed UIP and patients with a possible UIP at HRCT and no biopsy, with identical rates of lung function decline in the respective placebo groups (Raghu et al., 2015b).

2.5.8 New molecular targets in IPF

The scenario of early and late phase trials in IPF is continuously evolving, and multiple novel therapies currently under evaluation have more precise anti-fibrotic mechanisms than in the past. The exaggerate deposition of cross-linked collagen by activated fibroblasts and myofibroblasts is still considered the mainstay of the fibrotic process in IPF and has been addressed by most of the potential therapies developed so far, especially through direct or indirect targeting of transforming growth factor β1 (TGF-β1), a cytokine with pleiotropic pro-fibrotic actions central to wound healing processes. However, fibroblast activation must be placed in the context of the lung microenvironment and the complex relationships between epithelial injury, inflammatory cells and the extracellular matrix components (Friedman et al., 2013,Lota et al., 2013,O'Riordan et al., 2015,Steele et al., 2013). Indeed, the focus of new therapies should be on the selective inhibition of aberrant fibrotic mechanisms of repair without disrupting healthy connective tissue or interfering with normal wound healing. The role of oxidative stress is also well recognised in IPF, especially as driver of the initial epithelial cells injury and early apoptosis. However, N-acetylcysteine failed to demonstrate benefit in IPF patients when used in monotherapy (Martinez et al., 2014), and the efficacy of other anti-oxidant drugs like bardoxolone methyl, that has been shown to decrease oxidative stress in animal models (Kulkarni et al., 2013), has not been established yet in humans. Although inflammation and immune activation are no longer recognised as main actors in the pathogenesis of IPF, and despite the failure of several anti-inflammatory or immunomodulatory agents (King et al., 2009,Raghu et al., 2012a), they could still play a role as collateral mechanisms of disease. There is evidence now supporting the rationale for using pro-inflammatory cytokines and mediators (like lysophosphatidic acid - LPA), Th2-type immunity-related pathways and cells like pro-fibrotic M2 macrophages as targets of new compounds (O'Riordan et al., 2015). The stiff, fibrotic matrix itself, by self-sustaining the production of collagen through biomechanical stress and further activation of TGF-β1 represents a new potential target (Klingberg et al., 2014). However, if it’s true that the mechanisms of action of most of these new agents is well known, it’s still very hard to ascertain which of the targeted pathways represent primary drivers of the fibrotic process and which should be considered more

like epiphenomena. Since many of these molecules, unlike nintedanib or pirfenidone (both with pleiotropic actions), address single molecular targets, it gets more difficult to get proof of efficacy in the highly heterogeneous populations of IPF patients recruited in the trials. A proper disease stratification, operated by the validation of biomarkers of treatment efficacy would allow isolate those subjects with greater likelihood of response to therapy, thus increasing the effectiveness of clinical trials and paving the way to a personalised treatment approach.

The available evidence for the most promising drugs currently tested in early and late phase RCTs in IPF is summarised in Table 5 and Table 6 while their main mechanistic pathways are being shown in Figure 1.

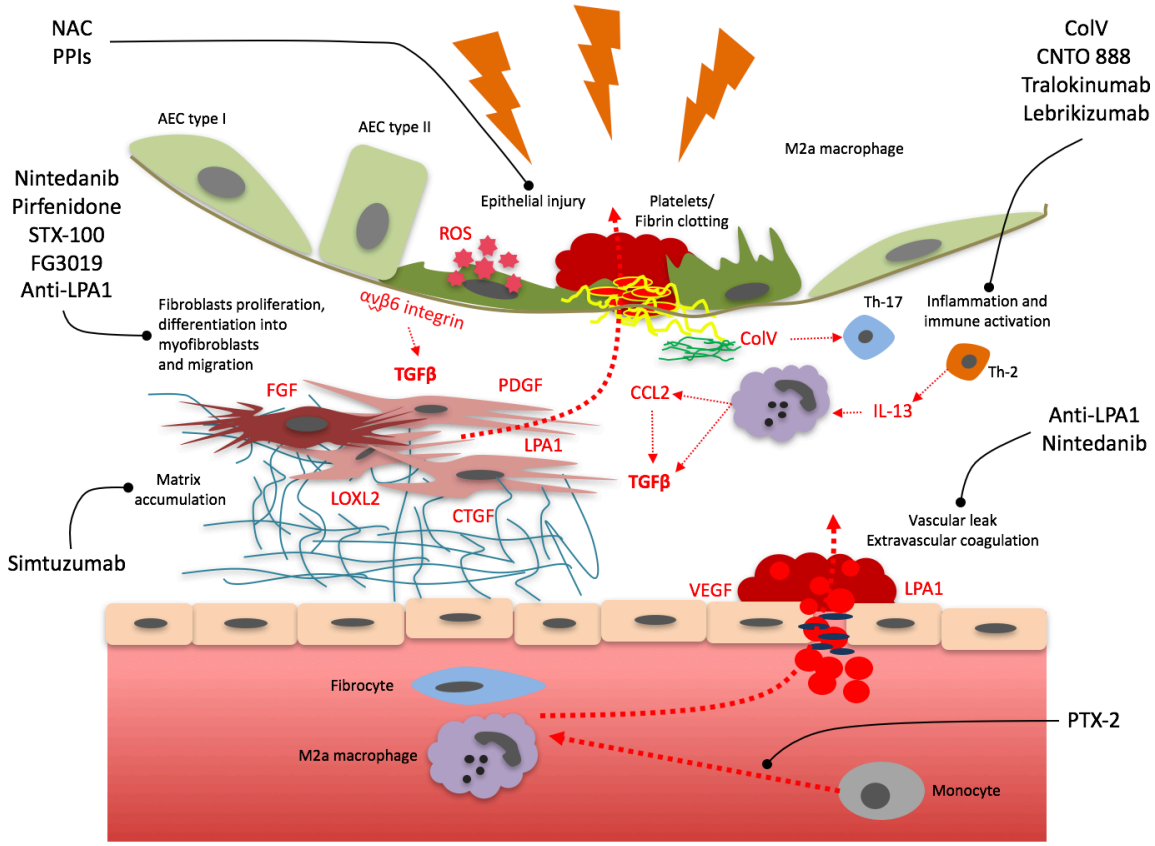


Figure 1 - Novel molecular targets in IPF and main related pathways. NAC=N-Acetylcysteine; PPIs=proton pump inhibitors; ColV=collagen type V; PTX-2=pentraxin 2; LPA1= lymphophosphatidic acid receptor type 1; ROS=reactive oxygen species; TGFβ =transforming growth factor β; FGF=fibroblasts growth factor; PDGF=platelet derived growth factor; CTGF=connective tissue growth factor; LOXL2=lysil oxidase-like type 2; CCL2=chemokine ligand 2; IL13=interleukine 13; VEGF=vascular endothelial growth factor.

Table 5 – Drugs tested in phase 2 and 3 trials in IPF

Drug	Mechanism of action	Study name - Sponsor	Identifier	Primary endpoints/duration	Planned enrolment	Outcome	Phase
Pirfenidone	Multiple	ASCEND - Intermune	NCT01366209	Δ%FVC at 52 weeks	500	Efficacy of Pirfenidone confirmed	3
STX-100	Anti-αvβ6 Integrin	Biogen/Stromedix	NCT01371305	Safety over 24 weeks	32	Ongoing	2
FG-3019	CTGF inhibitor	Fibrogen	NCT01890265	ΔFVC at week 48	90	Ongoing	2
BIBF1120 (Nintedanib)	Tyrosin kinase inhibitor	INPULSIS 1 and 2 - Boehringer Ingelheim	NCT01335464 and NCT01335477	ΔFVC at week 52	515 and 551	Efficacy and tolerability of Nintedanib confirmed. Reduction of rate of decline of FVC in the treated arm as compared to placebo.	3
NAC	Antioxidant	PANTHER-IPF - IPFnet	NCT00650091	ΔFVC at 60 weeks	390	Increased mortality in the triple therapy arm, stopped. No difference between NAC and Placebo arm with respect to the change of FVC.	2\3
BMS-986020	LPA1 receptor antagonist	Bristol-Myers Squibb	NCT01766817	ΔFVC at 24 weeks	300	Ongoing	2
GS-6624	Anti LOXL-2 Ab	Gilead	NCT01362231	Progression free survival at week 182	500	Ongoing	2
Lebrikizumab	Anti IL13 Ab	Roche	NCT01872689	Progression free survival up to 2.5 years	250	Ongoing	2
Tralokinumab	Anti IL13 Ab	Medimmune	NCT01629667	ΔFVC at week 72	186	Ongoing	2
QAX576	Anti IL13 Ab	Novartis	NCT00532233	IL13 serum levels	60	Terminated; awaiting results	2
CNTO888	Anti-CCL2 Ab	Centocor	NCT00786201	Safety and lung function	126	Terminated; awaiting results	2

Table 6 - Drugs tested in phase 1 trials in IPF

Drug	Mechanism of action	Study name Sponsor	Identifier	Primary endpoints/duration	Planned enrolment	Outcome
Allogeneic Human Cells (hMSC)	re-epithelization	AETHER - Joshua M Hare	NCT02013700	Safety after 1 month from infusion	25	Ongoing
Mesenchimal stem cells	re-epithelization	The Prince Charles Hospital	NCT01385644	Safety after 6 month from infusion	8	Ongoing
IW001	Col(V) - immune tolerance	ImmuneWorks	NCT01199887	Safety over 2 years	30	Terminated; awaiting results
Interferon gamma (nebulised)	Antifibrotic	New York University School of Medicine	NCT00563212	Safety over 1 year	12	Ongoing
PRM-151	Recombinant form of human pentraxin 2 (PTX2) - antifibrotic	Promedior	NCT01254409	Safety	21	Terminated; awaiting results
GSK2126458	inhibitor of PI3Kα and mTOR	GlaxoSmithKline	NCT01725139	PK/PD and safety over 8 days	24	Ongoing

2.6 Summary

IPF is the most frequent and severe among Idiopathic Interstitial Pneumonias (IIPs). Earlier diagnosis of IPF is crucial because of its relevant prognostic and management implications. Despite the standardisation of the diagnostic approach, it still represents a medical challenge and always requires a multidisciplinary approach involving professional figures of different expertise. The heterogeneity and unpredictability of IPF seriously hamper the clinical management of these patients, since current tools are defective in ensuring accurate prediction of disease behaviour in the single individuals. The emergence of the first effective treatments for limiting the functional decline in IPF makes the need of new strategies to improve diagnosis and monitoring more compelling than ever. Better disease stratification would also enhance the efficacy of randomised clinical trials, thus helping the development of novel therapeutic agents

Chapter 3: Computerised analysis of lung sounds

3.1 Introduction

This research involved the use of electronic auscultation for recording lung sounds and computerised methods for their quantitative analysis. This chapter will describe the advantages offered by these techniques in comparison with standard auscultation. Some basic concepts on computerised analysis of respiratory sound signals will be described. A review of the literature on computerised analysis of crackles will be presented.

3.2 Electronic auscultation

Since the invention of the stethoscope by René Théophile Hyacinthe Laënnec two centuries ago, chest auscultation represents the most important clinical assessment in patients with respiratory disorders. It provides immediate, reliable and low-cost information to clinicians, such as: the timing, the pitch and the duration of the breathing cycle; the presence of adventitious lung sounds and their location over different regions of the chest. However, its role in both clinical and research settings has been historically limited by its subjectivity and the lack of a uniform nomenclature for lung sounds, with significant variations among both healthcare professionals and different languages. To address the limitations of standard auscultation, computerised methods to record, analyse and classify respiratory sounds have been developed over the last thirty years, allowing the characterisation of normal and adventitious lung sounds. Recording lung sounds is a simple assessment, non-invasive, it usually requires the same level of patient collaboration as standard auscultation. There are immediate advantages to electronic auscultation: firstly, it allows to keep permanent records of the signals for future reference; secondly, while standard auscultation is generally limited by poor signal transmission due to the interference of external noise and the attenuation of higher frequency sounds (above 120 Hz, to which human ear is more sensitive) (Abella et al., 1992), electronic auscultation offers signal amplification and ambient noise reduction, and is independent on human ear sensitivity to different frequencies; thirdly, the methods used for quantitative analysis can be used to produce graphical representations with the potential to help physicians in the management of patients suffering from respiratory conditions (Sovijärvi et al., 2000).

Despite computerised analysis of respiratory sounds defined the properties of normal and adventitious lung sounds, for many years it has been hampered by the lack of standardisation of the methods for recording, the processing of signals and the reporting of theresults, which made it difficult to compare studies. For this reason the Computerised Respiratory Sound Analysis (CORSA) guidelines, promoted by ERS and published in 2000, standardised respiratory sounds analysis for research and clinical practice by collecting the evidence on definitions, environmental conditions, acquisition, pre-processing, digitalisation and analysis of lung sounds (Sovijärvi et al., 2000).

3.3 Origin and nomenclature of lung sounds

Although the precise origin of respiratory sounds is still not completely clear and likely involves multiple mechanisms, they are certainly generated by the turbulent airflow in the bronchial tree during ventilation. Since in smaller airways the airflow is believed to be laminar, thus silent, respiratory sounds are likely induced by a turbulence of the air at the lobar or segmental bronchi level. An ERS task force was recently established to provide updated recommendations for the standardisation of lung sounds nomenclature (Figure 2) (Pasterkamp et al., 2016) and suggested that the term “vesicular sound”, derived from Laënnec’s observations on breath sounds heard close to the chest wall (“a distinct murmur corresponding to the flow of air into and out of air cells”) and used in many languages to indicate normal sounds heard on the chest of healthy subjects, is not recommendable anymore. Instead, it has been proposed to replace it with “normal” or “basic” lung sounds: the airflow is assumed to be zero at the alveolar level (Pasterkamp et al., 2016). Because of this dependency from the airflow, respiratory sounds differ according to the location where they are recorded and vary with the breathing cycle; the expiratory phase resulting much less noisy than the inspiratory. For the same reason, lung sounds are also characterised by a large inter-individual variability, being the size of airways largely influenced by factors such as age, gender, body weight and height (Pasterkamp et al., 1997). The acoustic energy produced during normal breathing is transmitted through the airway walls, the lung parenchyma and the chest wall, which alter the characteristics of sounds by acting as filters. These structures together act as a low-pass filter, meaning that they attenuate the highest frequencies of lung sounds, with a cut off that is set around 300 Hz. Due to all these factors, the resulting acoustic system is non-periodic and non-stationary, with respiratory sounds being highly non-stochastic signals (i.e. the result of a process made of random variables); which makes them not suitable for conventional signal processing methods. Overall, respiratory sounds are characterised by a wide frequency range, usually between 50 Hz and 2500 Hz, with the main frequency band of normal lung sounds recorded over the chest going up to 200-250 Hz. However,

sounds recorded over the trachea, where there is almost no filter, frequencies components can reach values up to 4000 Hz (Reichert et al., 2008). Among adventitious lung sounds, crackles (the English term superseding the original Laënnec's "dry" *râles*) have been defined as brief (lasting less than 20 milliseconds), discontinuous, explosive and transient in character, occurring frequently in different cardiorespiratory diseases (Piirila et al., 1995); they can be referred to as "fine" or "coarse" according to the perceived high or low frequency/pitch (ATS, 1977). Their frequency content is wide, ranging from 100 to 2000 Hz or higher (Sovijärvi et al., 2000). Various hypotheses have been proposed over the years to explain the precise origin of crackles, which are thought to be generated during inspiration by the sudden opening of abnormally closed airways (Forgacs, 1967). In fact, they are assumed to be the result of the acoustic energy generated by a change in elastic stress after a sudden opening or closing of the distal airways, according to the stress-relaxation quadrupoles hypothesis developed by Fredberg in 1983 (Fredberg JJ, 1983). It has been demonstrated that the frequency, the duration and the timing of crackles within the breathing cycle are influenced by their origin. Smaller airways produce fine, late inspiratory crackles with short duration; larger airways produce coarse, longer crackles audible in the early inspiratory or in the expiratory phase (Kompis et al., 2001). As described in detail in the next section, fundamental frequency, duration, number and distribution per breath are all basic characteristics of crackles that wouldn't be available by the means of standard auscultation and that can be measured using computerised methods.

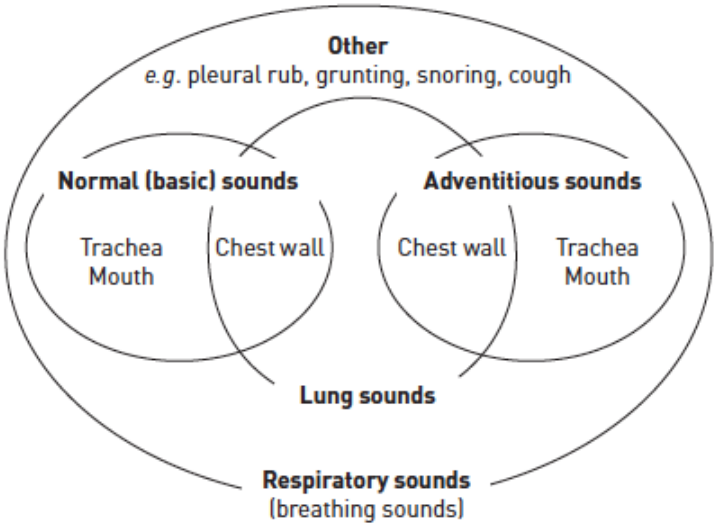


Figure 2 - Suggested classification of sounds, proposed by (Sovijärvi et al., 2000) and modified by (Pasterkamp et al., 2016).

3.4 Quantitative analysis of crackles

3.4.1 Basic concepts

In general, there are two main types of analysis of an acoustic signal, one carried out in the time domain and one carried out in the frequency domain. Consequently, there are two main types of features that can be extracted from lung sounds analysis.

The most immediate representation of sound signals is represented by the display of their amplitude (perceived as the loudness of the signal and consisting in the degree of displacement of the air molecules) as a function of time, which allows relation between the respiratory cycles and the identification of characteristic waveforms. Whilst waveforms of normal lung sounds present an irregular signal shape without repetitive patterns, adventitious sounds such as wheezes and stridors have a periodic waveform, sinusoidal or more complex. Conversely, crackles have a very distinctive appearance, with a short initial deflection in amplitude followed by other deflections with greater amplitude (Figure 3). Such detailed information can be retrieved through a widely used technique called time-expanded waveform analysis (TEWA), which consists in digitally zooming in the waveform (Forgacs, 1967,Sovijärvi et al., 2000) . This technique allowed quantitative classification of crackles in “fine” (high-pitched) or “coarse” (low-pitched) based on different time waveform intervals (Figure 3): initial deflection width (IDW), two cycle duration (2CD, the duration of the first two cycles of the crackle) and the largest deflection width (LDW, the width of the of the largest deflection) (Holford, 1981,Murphy RL Jr, 1977,Sovijärvi et al., 2000) . According to CORSA, 2CD of fine crackles is <10 ms, whilst that of coarse crackles is >10 ms (Sovijärvi et al., 2000). This analysis also allowed the validation of different methods to automatically detect and count crackles (Murphy et al., 1989,Vannuccini et al., 1998), whose number and distribution per breath has been associated with the severity of the disease in patients with ILD (Piirila et al., 1995) and pneumonia (Murphy RL, 2004). Moreover, it helped distinguish different types of crackles generated by different portions of the breathing cycle, thus proving their mechanistic origin: the highest deflection (or peak) of the crackle waveform can be used to define the polarity of crackles, and it has been demonstrated that the majority of crackles have negative polarity during the inspiration phase and positive polarity during the expiration (Vyshedskiy et al., 2009). This is consistent with the hypothesis of the stress-relaxation quadrupole mentioned above (Fredberg JJ, 1983), according to which crackles have different polarities that are dependent on the deflection of the tissue: positive polarity means that tissue is being displaced towards the microphone (occurring during expiration), while negative polarity means that the tissue is displaced away from it (inspiration). This represents sufficient proof that

crackles are being produced by a sudden opening and closing of the airways due to rapid changes in the airflow rate.

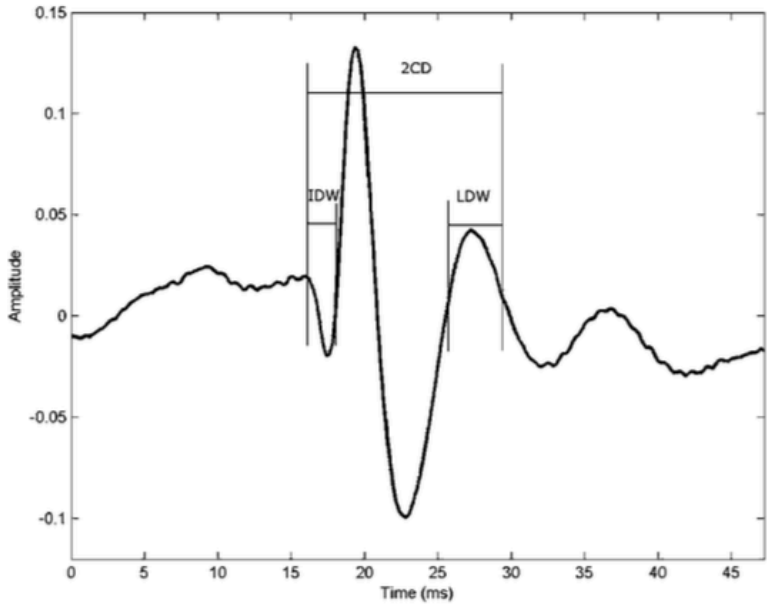


Figure 3 – Single crackle's waveform and parameters in TEWA (time-expanded waveform analysis). IDW: initial deflection width, 2CD: 2-cycle deflection, and LDW: largest deflection width.

Most importantly, the characterisation of the morphological features of crackles in the time-domain allowed to distinguish those generated in different processes such as pneumonia, chronic heart failure (CHF), pulmonary fibrosis, Chronic Obstructive Pulmonary Disease (COPD) and bronchiectasis (Piirila et al., 1991,Ponte et al., 2013).

Nevertheless, analysis carried out in the frequency domain provides lots of further information which is not readily available in the time domain, and it allows easier comparisons between signals and, as such, the findings reported by studies. Assuming that a continuous-time signal *s(t)* can be sampled to obtain a discrete-time signal *s(nT)* – or time series, where *n* is an integer variable and *T* is the sampling interval - time can be quantised and signals can be represented mathematically as sequences of numbers and used for frequency analysis. The Fast Fourier Transform (FFT) is a mathematical operation which decomposes a physical time-function signal into several discrete frequencies over a continuous range. The distribution plot of the amplitude (or power) versus the frequency components of a time series takes the name of power spectrum Figure 4. Consequently, the statistical average of a certain signal as analysed in terms of its frequency content is called spectrum of that signal. The term power spectral density, more commonly used, applies to signals existing over an infinite time interval, as such it represents an

estimation of the distribution of the frequency components composing a given signal if that signal existed over an infinite time interval (Figure 4).

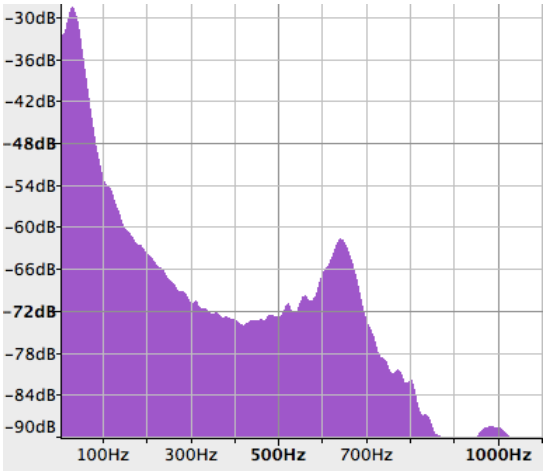


Figure 4 – Power spectrum of an acoustic signal. X axis = frequency (Hertz, Hz), Y axis = amplitude (decibel, dB).

The Fourier Transform and its derivatives are the most used spectral analysis algorithms, signal processing methods for the extraction of acoustic features that can be used for the characterisation and classification of different types of adventitious sounds. More and more advanced spectral analysis techniques have been increasingly used over the years, and algorithms involving machine learning methods (such as neural networks) have been extensively employed for the automatic classification of lung sounds based on the features selected from frequency decomposition and associated statistical parameters (Kandaswamy A, 2004). A recent meta-analysis of eight studies showed that while quality data on computerised lung sound analysis is relatively limited, analysis of existing information suggests that this technique can provide a relatively high sensitivity and specificity for detecting abnormal lung sounds such as crackles and wheezes. Inclusion of a broader spectrum of detection parameters from both the time and the frequencies domains might further improve the sensitivity and specificity of the computerised lung sound analysis algorithm in detecting respiratory disorders (Gurung et al., 2011). The spectral analysis of frequencies with inclusion of crackle features such as counts, duration, and the timing within the breathing cycle was found to improve the diagnostic accuracy in patients with pneumonia (Murphy RL, 2004). Altogether, this evidence suggests that these methods have the potential to be validated in the future for diagnostic purposes of different pulmonary conditions. A recent systematic review of twelve studies evaluating different types of adventitious sounds before and after a therapeutic intervention showed that the specific variables or characteristics have the potential to be used as outcome measures too (Marques et al., 2013). With regard to the crackles, the LDW showed high effect sizes in one study (Piirilä, 1992), while the crackles number (counts) and 2CD presented conflicting information. Notwithstanding the relatively limited

evidence available, this data shows that there are specific features that can be extracted from lung sounds to monitor the underlying disease, supporting the rationale for further exploration of lung sounds as clinical endpoint in respiratory diseases. This might be especially true for fibrotic ILD, where crackles are the expression of a progressive fibrotic process and are not subject to daily variations due, for example, to movement of secretions (such as in bronchiectasis).

3.4.2 Current limitations and future directions

Despite the advances in the technologies underpinning computerised analysis of lung sounds, the application of these methods in the clinical setting has been limited by several factors. Firstly, most studies so far have been carried out in highly controlled settings, adopting complex recording systems which are not easily applicable in a real-life clinical setting. Secondly, the populations of patients recruited were not always well characterised. Thirdly, the outputs of the automatic analysis have been often adjudicated against standard chest auscultation or historical diagnosis of ILD, thus lacking a proper reference standard as the direct evidence of pulmonary fibrosis as shown at HRCT. Finally, most studies on lung sounds in ILD so far applied different methods for analysis and automatic classification and focused on few acoustic features, sometimes chosen arbitrarily.

A few studies, however, support the potential clinical utility of lung sounds analysis. The intra-subject reliability of crackles recorded using an electronic stethoscope in a clinical setting from a population of patients with bronchiectasis and cystic fibrosis has been demonstrated (Marques et al., 2009), providing ground for the application of computerised sound analysis using a simple, point of care tool for the signal acquisition process. The inter-rater reliability (or agreement) between physicians in identifying crackles' time-related features such as IDW, LDW and 2CD using a simple graphical interface showing the crackle waveform has also been demonstrated (Hoevers et al., 1990). Recent evidence further supports the feasibility of combining the use of a simple tool for recording (a digital stethoscope) with analytical methods providing readouts that can be quickly and easily interpreted by users (physicians) such as graphical representations, thus delivering ready, useful clinical information (Bhattacharyya et al., 2015,Ohshimo et al., 2016).

3.5 Summary

This chapter described the advantages of electronic auscultation, providing physicians with readily available data both for real-time and future reference, offering the chance for further quantitative analysis of lung sounds. A few concepts at the basis of computerised analysis of lung sounds were also introduced. Over the past decades, these methods defined the fundamental acoustic characteristics of normal and adventitious breath sounds, thus clarifying their pathobiology. The characterisation of pathologic lung sounds in the time and frequency domains allowed to develop methods for the automatic quantification and classification, which have potential to assist the diagnosis and management of respiratory disorders. Despite the recent standardisation of techniques for the quantitative analysis of lung sound signals, translating these systems into clinical practice has proven to be elusive so far due to several factors such as the complexity of the systems used for recording, the poor interpretability of the outputs and, most importantly, the lack of a robust research design aimed to deliver clinical applicability.

Chapter 4: “Velcro-type” crackles and HRCT imaging: a correlation study

4.1 Introduction

This chapter will describe a study aimed to characterise ““Velcro-type”-type” crackles in fibrotic ILD using HRCT findings. First, the rationale of the study will be presented in section **Error! Reference source not found.**. Section 4.3.1 will describe the methods of the research including the aims, the procedures performed to address the proposed objectives, the main outcomes and the methods used for data analysis. The results will be presented in section 4.4 and discussed in section 4.5. Finally, in section 4.6 conclusions will be drawn and plans for future research in this field will be presented.

4.2 Rationale

4.2.1 The need for an earlier diagnosis of fibrotic ILD and IPF

Determining a diagnosis is a fundamental step to inform patients about their prognosis and to direct the physician in making decisions about the correct management. In a time when no effective treatments were available for IPF, there was already evidence that the delay in the referral of patients to a tertiary care centre specialised in ILD was associated with poor survival (Lamas et al., 2011), suggesting the importance of setting an earlier diagnosis of IPF. Prompt referral provides patients with access to centres with expertise in diagnosis and management, including efficient monitoring and non-pharmacological support. The approval of the first therapies capable of modifying the natural course of IPF – namely pirfenidone and nintedanib – revolutionised the landscape of IPF management, since physicians can now offer two safe and effective treatment options to these patients. However, some patients remain stable for several months and it might be argued that in a stable patient with mild disease, treatment should be postponed until functional decline begins, thus suggesting that an earlier diagnosis of IPF does not translate automatically into a therapeutic intervention. Nevertheless, there are no tools at present to predict the rate of the disease progression in each patient or to know when they will experience an acute exacerbation. Moreover, deleterious subclinical changes might take place in patients without signs of significant functional deterioration. Recently, some post-hoc analyses of pooled data from the large randomised trials that led to the approval of pirfenidone and nintedanib demonstrated that those patients with preserved lung function at baseline had similar

rate of functional decline and received the same benefit from anti-fibrotic treatment as compared to patients with a more advanced disease (Albera et al., 2016,Costabel et al., 2016,Kolb et al., 2016). Such evidence significantly reinforced the importance of an earlier diagnosis of IPF for prompting the start of a treatment regardless of the presence of functional impairment.

When considering strategies to determine earlier diagnosis of a disease, the difference between diagnosing and screening for a disease should be recognised, and the opportunity for these processes properly evaluated in the first place. The diagnostic process in the daily practice starts with a patient presenting with a particular symptom or sign that makes the physician suspicious of the patient’s having a particular disorder (target disease) (Sackett et al., 1985). The term “screening” in clinical epidemiology refers to the identification, among asymptomatic, apparently healthy individuals, of those who are sufficiently at risk of a specific disorder to justify a subsequent diagnostic test or procedure. In order to justify the efforts, the costs and eventually the risks related to the conduction of a screening program, the target disease should be common enough, it should be accompanied by significant morbidity if not treated, and an effective therapy should exist to alter its natural history. Given the prognostic implications and the availability of effective therapies, might a screening program for IPF considered to be feasible, and if so, useful? IPF is certainly burdened by high morbidity and mortality, with a socio-economic impact far from irrelevant (Lee et al., 2014). Nevertheless, the disease is relatively rare in the general population, and the high costs and radiation-related risks of the current key diagnostic tool – namely the High Resolution CT scan of the chest - rule out, for now, a dedicated screening program. In the light of the evidence of significant findings of fibrotic lung abnormalities in lung cancer screening programs conducted using low-dose CT scan of the chest (Jin et al., 2013), schemes similar to or as a by-product of lung cancer screening have been recently proposed (Cordier et al., 2013,Cottin et al., 2014), taking advantage of shared risk factors such as age and smoking. Indeed, some concerns regarding a screening program for IPF might be raised due to the potential detection of cases with mild, non-specific alterations at HRCT. Terms like “interstitial lung abnormalities” (ILA) or “subclinical ILD” have been used to describe the whole of the patterns of interstitial lung alterations that can be found on HRCT in absence of a significant clinical presentation. Although a few patients presenting with ILA can progress to a UIP pattern over some years, most subject with ILA will not develop IPF, either going towards resolution, keeping on with stable or minimally progressive ILA, or evolving to other forms of ILD/IIP (Putman et al., 2014). Over the last years, several genome-wide associations studies have identified several common and rare genetic variants in more than a dozen loci on different chromosomes that appear to contribute to the risk of developing IPF. The most consistent and reproducible genetic finding in IPF is represented by a common single nucleotide polymorphism in the promoter of the gene encoding for mucin 5B

(MUC5B) located on the chromosome 11, which has been found to be associated with a 6 to 8-fold increased risk for both sporadic and familiar forms of IPF (Seibold et al., 2011). However, this variant is estimated to be present in 20% of the general population, and has been reported as a risk factor also for asymptomatic ILA (Hunninghake et al., 2013), confirming a large genetic effect for this variant. As such, despite ILA and IPF seem to share some common genetic grounds, it is still impossible to say with certainty what is the biological effect of the single genetic variants isolated so far, and most importantly what is the influence of environmental factors on such variants in the development of a progressive fibrosis. Large, longitudinal studies are needed to determine which combination of genetic variants and clinical, physiology and imaging features predict an increased risk of progression to IPF, thus defining a true “early IPF” phenotype (Putman et al., 2014). For the time being, the benefit of treating subjects with non-specific ILA is highly questionable, and likely overweighed by the risks related to the side effects.

Having excluded the opportunity of running a screening program for IPF in the general population, the prompt recognition of signs and symptoms suggestive of ILD operated by healthcare practitioners represent the most advisable approach for limiting the diagnostic delay. In recent years the importance of “low-cost” approaches to the diagnosis of disease as opposed to the over use of high-tech, expensive tests or invasive tests has been emphasised (Martinez et al., 2005). Frequently, data which is routinely available in all patients may be the most useful promoting diagnostic confidence. Recent reports actually suggest that demographic or clinical data may provide significant contribution toward prediction of histopathologic diagnosis of UIP regardless of specific HRCT appearances. A retrospective study demonstrated that an age greater than 70 in patients with some extent of reticular abnormalities at HRCT has the 95% of positive predictive value of UIP on surgical lung biopsy, and it becomes 100% for people older than 75 (Fell et al., 2010). A cross-sectional analysis in patients with suspected fibrotic ILD and absence of honeycombing on CT showed that the extent of reticular opacities can confidently predict a diagnosis of IPF when age is taken into account, with a probability of confirmed IPF above 80% in subjects over age 60 years with reticular opacities involving one-third of total lung volume (Salisbury et al., 2016). Very recently, it has been demonstrated that additional clinical and radiographic information, such as age > 60years and extent bronchiectasis, can significantly increase the predictive value of non-definitive imaging findings (possible UIP) toward confirmation of histopathological UIP across different population prevalences (Brownell et al., 2017). Building on such body of evidence, different groups of experts have recently advocated the need to overcome the limitations of strict diagnostic criteria by proposing more flexible diagnostic approaches that include diagnostic confidence based on clinical judgment (Martinez et al., 2017,Ryerson et al., 2017). Despite an objective guidance on how to formally incorporate

different source of information in the evaluation of suspected ILD is still lacking, in the near future the diagnosis of fibrotic ILD and IPF will likely rely upon a more comprehensive approach including clinical data and, possibly, specific molecular signatures. As such, it is expected that the validation of reliable, readily available clinical findings in properly designed diagnostic research will help pursue an earlier, less invasive and more confident diagnosis of these disorders.

4.2.2 “Velcro-type” crackles: an early detection tool of fibrotic ILD?

“Velcro-type” crackles, named after their similarity to the sound generated by Velcro strips separating, have been defined as brief, discontinuous lung sounds, explosive and transient in character, and are universally considered representative of established lung fibrosis. Their careful assessment may represent a practical, reliable and cost-effective way for prompting a proper diagnostic work up (Cottin et al., 2012), and international consensus guidelines recommend that IPF be considered in all patients with bibasilar inspiratory crackles (Raghu et al., 2011). Nevertheless, their real value in assisting diagnosis and clinical management of ILD has never been properly explored, likely due to factors such as the subjectivity of chest auscultation, the lack of a uniform nomenclature for lung sounds (Pasterkamp et al., 2016) and the poor signal transmission of standard stethoscopes (Abella et al., 1992). Several studies involving electronic recording and computerised analysis of lung sounds attempted to determine the diagnostic accuracy of automatic classification systems toward fibrotic ILD. None of these however was robustly designed diagnostic research. These studies often enrolled small, convenience sample populations made of selected patients with historical diagnosis of ILD and healthy controls, and in most cases the tools used for signal acquisition and analysis were not applicable in everyday clinical practice. As such, the outputs of performance of the systems were of little, if any, clinical utility. Most importantly, there has never been yet a systematic assessment of the relationships between “Velcro-type” crackles and radiologic patterns and features on HRCT - the key diagnostic test for fibrotic ILD - and whether they are predictive of fibrotic abnormalities occurring in the lungs has been unknown. If such linkages exist, the assessment of lung sounds in routine clinical practice may provide the first step toward early detection of pulmonary fibrosis and more specifically IPF. As such, this study sought to investigate the relationships between audible “Velcro-type” crackles and distinct radiologic findings representative of pulmonary fibrosis on HRCT. In this study, lung sounds were digitally recorded via electronic stethoscope. As compared to standard auscultation, electronic recording offers the opportunity of storing normal and pathological lung sounds: acoustic data is not part of the electronic medical record yet, but might prove to be very useful to assist physicians in everyday clinical practice. Acoustic data might represent an important source of information available for immediate or future reference: lung

sounds recording could be even performed by properly trained healthcare professionals or by the patients’ carers and transferred to the physicians for a real-time or deferred evaluation. This study used this approach to assess the reproducibility of the subjective assessment of “Velcro-type” crackles, which has never been explored so far.

Overall, this study might provide ground evidence for further research aimed to quantify the added value of the subjective or computer-aided assessment of “Velcro-type” crackles toward accurate detection of fibrotic ILD.

4.3 Methods

4.3.1 Research objectives

4.3.1.1 Primary objective

To investigate the relationships between the presence of audible “Velcro-type” crackles, electronically recorded via digital stethoscope, and features and patterns of fibrotic ILD findings on High Resolution Computed Tomography (HRCT) of the chest.

4.3.1.2 Secondary objectives

- To assess the inter-rater reliability of the assessment of “Velcro-type” crackles on electronic auscultation between respiratory physicians.
- To assess the accuracy of the subjective assessment of “Velcro-type” crackles on electronic auscultation toward detection of signs of pulmonary fibrosis on HRCT, used as standard reference.

4.3.2 Research design

A prospective, case-control study was conducted to determine the relationships between the presence of audible “Velcro-type” crackles and different radiologic features and patterns on HRCT scan of the chest.

4.3.3 Research procedures

4.3.3.1 Ethics and governance

The original protocol was submitted to the local Ethics Committee of the University Hospital of Modena, Italy. Full approval was obtained prior to recruitment. The University Hospital of Parma

was subsequently involved as participant centre for the recruitment of a further, independent cohort of subjects.

4.3.3.2 Setting, timeline and selection of population

Study participants were recruited prospectively between January 2013 and February 2015 at the University Hospitals of Modena and Parma, Italy. Patients were eligible if they were referred to perform a HRCT scan of the chest, either for diagnostic or monitoring purposes, and they were aged 18 years or over with capacity of providing valid consent. The choice of such inclusion criteria was made to enrich the study population with patients with ILD, since the HRCT scan of the chest represents the key diagnostic test when ILD is suspected, and a higher prevalence of fibrotic ILD as compared to the general population was therefore expected. This was indeed useful to explore the association of audible “Velcro-type” crackles with HRCT findings of pulmonary fibrosis. On the other hand, chest HRCT is also routinely performed for the diagnosis and the follow up of other parenchymal lung diseases, which allowed the inclusion of subjects with a broad spectrum of conditions (such as pneumonias, bronchiectasis, or lung nodules in follow-up) together with HRCT-negative controls. The potential presence in the recordings of pathological sounds that may act as confounders (such as “wet” crackles) was actually considered essential to determine the relationships between the specific assessment of “Velcro-type” crackles toward the confirmation of pulmonary fibrosis on HRCT, ensuring wider applicability and generalisability of the findings.

4.3.3.3 Recruitment strategies

Recruitment took place in the Radiology Units of both centres involved in the project. Weekly sessions dedicated to chest HRCT scans were established in both hospitals to speed up the recruitment process. The staff of the CT scan unit, comprised of thoracic radiologists, radiology technicians and nurses, liaised with the author-researcher to allocate the proper time to allow the collection of data without affecting the efficiency of the clinical service. An information sheet was provided in advance to the study candidates. For those who were interested in taking part to the study, a discussion took place in the CT scan patient preparation room before the test. The researcher personally recruited subjects and collected the data until October 2013. Dr. Sofia Taddei, a respiratory physician working at the Department of Respiratory Medicine of the University Hospital of Modena, completed the recruitment and data collection for the study.

4.3.3.4 Population sample and eligibility criteria

As described above, participants were recruited prospectively from a population referred to undergo a chest HRCT either by general practitioner or respiratory consultant. No sample size and

power calculations were performed for this study. The annual incidence rate of ILD diagnoses at the two centres was kept into consideration, however it was not feasible to anticipate the frequencies of radiologic signs of pulmonary fibrosis in the population under study. Moreover, while power calculations do exist for test research (i.e., studies aiming to estimate the diagnostic value of a single test or to compare the properties of two single tests (Hanley et al., 1983)) this study was primarily aimed to determine the relationships between lung sounds and radiologic findings, as such those requirements were not applied. The size of the cohorts enrolled was therefore determined by the time allocated to recruitment. A first cohort of 113 subjects was recruited at the University Hospital of Modena between January and June 2013. A second, consecutive cohort of 114 subjects was recruited between July 2013 and October 2014. A third independent cohort consisting of 46 further subjects was recruited at the Radiology Unit of the University Hospital of Parma between April 2014 and February 2015. 5, 14 and 3 subjects for the first, second and third cohort respectively were initially excluded because of missing imaging or acoustic data at the time of data extraction. This was due to technical issues such as irretrievable HRCT scans on local servers, metallic marks accidentally removed before HRCT scan (see next section) or faulty recordings.

4.3.3.5 Research protocol

Written informed consent was obtained from all participants before proceeding to collect any data for the study. The study, having a cross-sectional design, consisted of a single visit during which all the data was collected.

Demographics and other data including past medical history, smoking history and family history for ILD were collected. When available, the clinical indication for performing the HRCT scan, as reported on the referral slips or letters from the general practitioner or the respiratory consultants, was also collected.

The patients were invited to assume the sitting position on a plinth in the patient preparation room of the Radiology Unit, to expose the chest and place both hands on the knees. Lung sounds were consecutively recorded using an electronic stethoscope (Littmann 3200, 3M, USA) over six different anatomical sites identified on the posterior chest. The recording sites were selected based on the indications of the guidelines for Computerised Respiratory Sounds Analysis – CORSA (Sovijärvi et al., 2000), and taking into account the lower lobes-predominant distribution of many ILD. As such, for the left and right side of the chest two recordings were performed at the lung bases (at two and five cm from the paravertebral line, at seven cm below the scapular angle) and one recordings was taken at mid chest (at two cm from the paravertebral line, at the fourth or fifth intercostal space) (Figure 5).

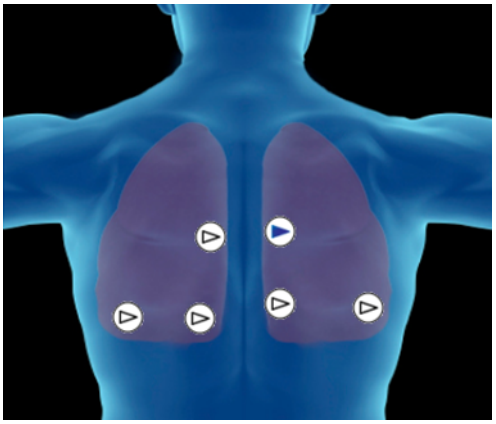


Figure 5 - Recording sites in the case-control study. For each side of the chest, 2 recordings were performed at the lung bases at 7 cm below the scapular angle, at 2 and 5 cm from the paravertebral line respectively; another recording was taken at mid chest in correspondence of the fourth or fifth intercostal space, at 2 cm from the paravertebral line.

Sounds were recorded approximately for ten seconds at each site, a time sufficient to catch approximately two or three full breathing cycles. Patients were asked to breathe deeply through the mouth during the recordings to increase flow levels and therefore increase intensity of lung sounds. Care was taken to ensure the stethoscope was kept as still as possible during the recordings; while a more secure fixing for the stethoscope might have been ideal to minimise background noise, manual placement of the stethoscope was preferable to develop a measurement protocol sufficiently robust to be used within routine clinical practice. After each recording, a small, radio-opaque metallic mark (bio-compatible electrocardiography electrode) was applied to the skin to keep track of the recording sites after the HRCT scan was performed (Figure 6).

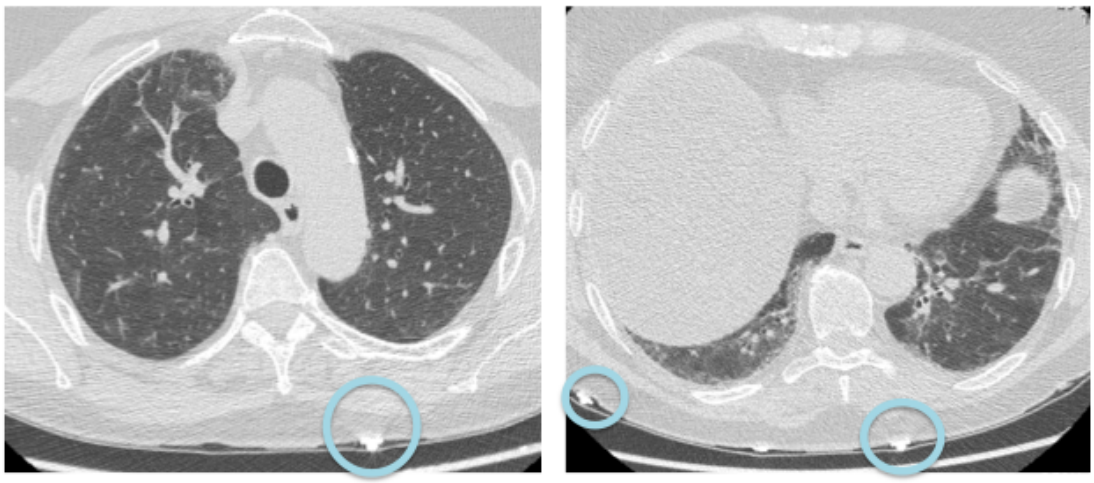


Figure 6 – High Resolution Computed Tomography (HRCT) sections showing metallic marks (electrodes) applied to the posterior chest of the patient.

Single HRCT sections corresponding to the 6 marked sites of recording were extracted from the full volume HRCTs of patient enrolled in the study. Some images however couldn’t be retrieved because the metallic mark accidentally fell out of the HRCT viewing area, making the recording site not recognisable. The HRCT images were cropped as to show left and right chest and then uploaded in random order on a database. A review was performed by two thoracic radiologists with high experience in ILD (one working at the Royal Brompton and Harefield Hospital in London, UK, the other working at the University Hospital of Parma, Italy), blind to the clinical characteristics of the study participants. The observers independently assessed the images for the presence of pulmonary fibrosis. Then, individual radiologic patterns (reticulation, bronchiectasis, honeycombing, ground glass and emphysema, as defined in the Fleischner society glossary of thoracic imaging (Hansell et al., 2008a)) were semi-quantitatively scored using a visual scoring method) combining the severity and extent of the alterations, similarly to systems used in previous studies (Kazerooni et al., 1997,Oda et al., 2014a). Qualitative visual scores of fibrosis were combined, adjudicating an image as fibrotic in case of disagreement between the two observers. The semi-quantitative scores for the individual radiologic patterns were averaged between the two observers.

For each cohort, cases were identified based on the evidence of pulmonary fibrosis on one of the HRCT sections. A group of age- and sex- matched controls was then formed from the remaining subjects that had no signs of fibrosis on the HRCT sections. The imaging and acoustic data from the case and control groups of the three cohorts formed the final data set used in the analysis. Since the assessment of the whole data set of recordings would have proved to be significantly time consuming for the physicians, the case-control design was considered a more suitable and convenient option as it would allow study the cases in detail together with a sample of those subjects who turned out to be free from fibrotic abnormalities.

The initial study population consisted of 253 consecutively enrolled in the three cohorts (108, 102 and 43 subjects for the first, second and third cohort respectively). Following radiologic review, three groups of 28, 31 and 17 fibrotic subjects (cases) and corresponding groups of 27, 28 and 17 non-fibrotic matched subjects (controls) were formed. A final data set of 805 images and corresponding sound files was built for subsequent analysis of correlation between presence of “Velcro-type” crackles and radiologic patterns of pulmonary fibrosis. A flowchart showing details of the selection process of imaging and acoustic data is shown in Figure 9.

The full volume HRCT scans of case and control groups were also downloaded from the local servers of the Radiology Units in the DICOM (Digital Imaging and COmmunications in Medicine) format. A third thoracic radiologist (working at King’s College Hospital in London, UK), blind to the

clinical information, assessed the HRCTs for the presence of fibrotic ILD. The fibrotic scans were further classified into UIP, possible UIP or inconsistent with UIP pattern. This assessment was used in a secondary analysis to explore the association between the presence of bilateral “Velcro-type” crackles with the radiologic categories of pulmonary fibrosis at the full HRCT.

Two respiratory physicians with several years of experience in ILD, one working at the University Hospital of Southampton, UK, and one working at the University Hospital of Modena, Italy, performed a qualitative assessment of “Velcro-type” crackles by listening to the recordings of each data set. The sound files were randomised prior to assessment to ensure that recordings from the same subject were not consecutive.

4.3.4 Data anonymisation, management and storage

The data collected in this study was anonymised prior to any review or analysis using unique codes, and securely stored on a University password-protected laptop. Paper clinical research forms were used for collection of demographics and clinical data. Since this documentation included patients’ personal information and details, they were stored in locked filing cabinets together with the consent forms. After collection, clinical data was temporarily transferred onto a digital spreadsheet (Microsoft Excel) on a University laptop for data management before analysis.

4.3.5 Outcome measures

4.3.5.1 Measurement of lung sounds

Figure 7 - 3M Littmann 3200 electronic stethoscope.

The recordings were taken using an electronic stethoscope (Littmann 3200, 3M, USA, Figure 7) **Error! Reference source not found.**. The chest piece contains a microphone and a small hard drive where up to 12 different recordings can be stored. The stethoscope can be paired via Bluetooth technology to a dedicated software (Littmann StethAssist, 3M, USA). Recordings can be either transferred in real time to the software platform or stored in the stethoscope’s drive and transferred to the software at a second time. The StethAssist software allows create single

encounters with as many recordings as desired (as shown in Figure 8**Error! Reference source not found.**), save them in a specific format (.zsa), and export them to other format. For this study, the .wav format (sampled at 4 kHz with a resolution of 16-bit) was used for analysis.

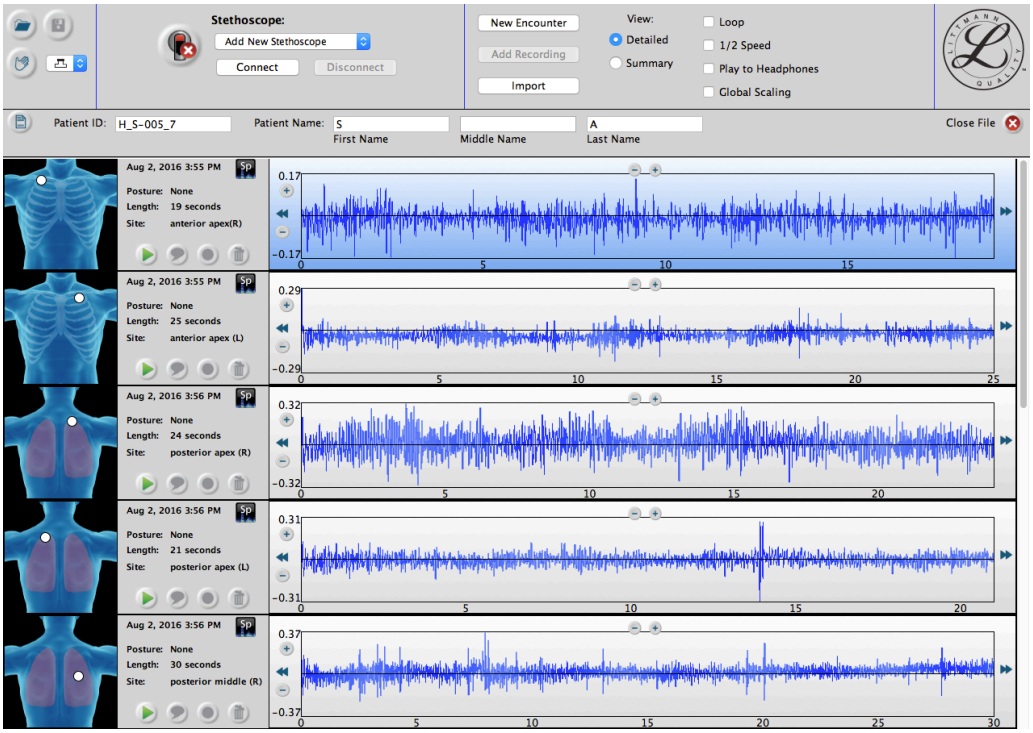


Figure 8 - StethAssist software platform. Visual displays of multiple recordings for a single encounter (patient).

Two respiratory physicians with several years of experience in ILD listened to the sound data sets. The three data sets were provided consecutively with an interval of few months between them, to minimise the chances of errors due to fatigue and to avoid interfere with the clinical duties of the physicians. The physicians, blind to the imaging data and other clinical information, were provided with a customised sounds assessment chart (Appendix C) and were invited to rate each sound file for the absence or presence of “Velcro-type” crackles. The files were listened via Personal Computer. The same media player (Audacity, a free open source software available for download online) and the same over-ear headphones (Sennheiser HD201 closed dynamic stereo headphones) were provided to both physicians for this assessment.

4.3.5.2 Demographic data

Demographics including gender, age, and race/ethnicity were recorded as reported by the patients. Information on past medical history, smoking habits, familiarity for ILD and other conditions, history of occupational and environmental exposures, current medications were also provided directly by the patients.

4.3.5.3 Measurement of pulmonary fibrosis

The evidence and extent of pulmonary fibrosis and individual abnormalities characteristic of the fibrotic process in the lung parenchyma were assessed via HRCT of the chest, which represents the elective test for diagnosis of diffuse parenchymal lung diseases and an important tool for monitoring the progression of the disease in these patients. HRCT is routinely performed using a conventional CT scanner, however imaging parameters are chosen to maximise spatial resolution. A narrow slice width is used (usually 1-2 mm), a high spatial resolution image reconstruction algorithm is used and field of viewing is minimised, to minimise the size of each pixel. Conventionally, as the aim of HRCT is to assess a diffuse lung disease, the test is performed by taking thin sections 10–40 mm apart (spaced axial technique). The result is a few images that should be representative of the lungs in general, but that cover only approximately one tenth of the lungs. However, a volumetric HRCT technique, where thin sections are acquired continuously in a single breath hold (full inspiration), is preferred as it provides greater diagnostic information by examining the entire lung, and permits the use of multi-planar reconstruction techniques. The HRCT scans performed in this research were acquired through volumetric techniques.

Since HRCT was performed as part of the routine clinical assessment of the patients in two different centres, it was not feasible to use a single HRCT protocol for the study. As such, minor technical differences in terms of image acquisition were present between the scans, depending on the centre and the adjustments made by the radiologists and radiology technicians present on the day of the test. The HRCT sections extracted from the full volume scans in correspondence of the marked sites of recording were independently reviewed by two thoracic radiologists with expertise in ILD imaging. They scored each section for the absence or presence of pulmonary fibrosis. Then, they also semi-quantitatively scored the images for individual radiologic abnormalities – namely reticulation, ground glass, traction bronchiectasis, honeycombing and emphysema, as defined in the Fleischner society glossary of thoracic imaging (Hansell et al., 2008b) and according to a scoring system (shown in Table 7**Error! Reference source not found.**) which combines the severity and extent of the alterations in a similar fashion to semi-quantitative visual scoring systems adopted in previous studies (Kazerooni et al., 1997,Oda et al., 2014a). For reticulation, honeycombing, ground glass and emphysema the scores went from 0 to 4 and were defined by the proportion of bronchopulmonary segments involved, as such a greater number of segments involved corresponded to a higher score. Traction bronchiectasis was assigned with a categorical severity score that took into account the average degree of airway dilatation within areas of fibrosis as well as the extend of dilatation throughout the lobe and was given an overall score ranging from 0 to 3 (none=0, mild=1, moderate=2, severe=3).

The full volume HRCTs of the groups defined for the 3 cohorts were reviewed by a third ILD radiologist, who independently assessed the scans as to the presence of fibrotic ILD and classified the radiologic pattern as UIP, possible UIP or inconsistent with UIP according to validated criteria (Travis et al., 2013).

Table 7 - Scoring system used for individual radiologic abnormalities in the case-control study.

Feature	Score				
	0	1	2	3	4
Traction Bronchiectasis	No	Mild	Moderate	Severe	-
Reticulation Honeycombing Ground glass Emphysema	No	≤ 25%	> 25% ≤ 50%	> 50% < 75%	≥ 75%
Fibrosis	No	Yes	-	-	-

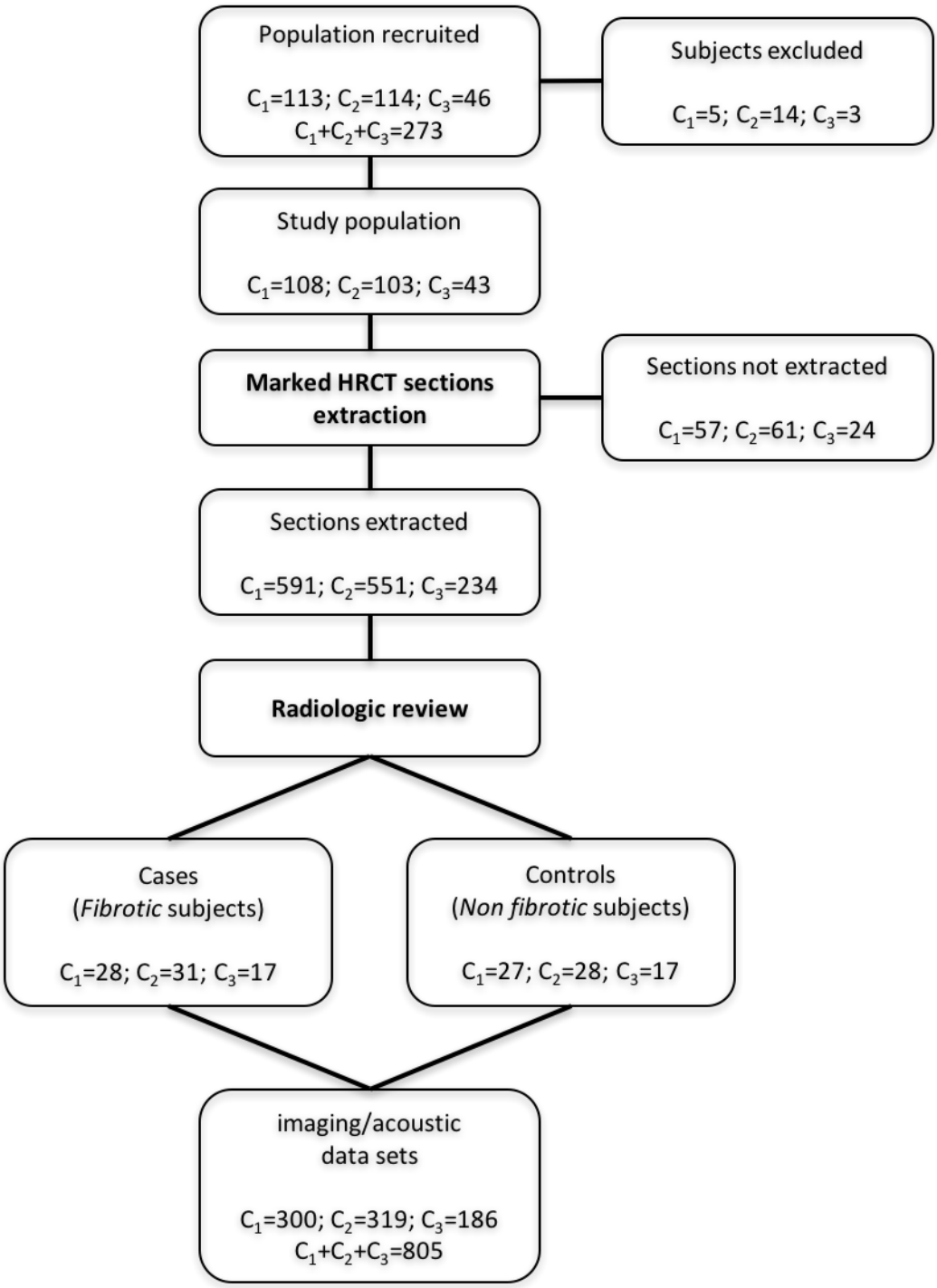


Figure 9 – Flowchart of definition of study groups (cases and controls) and data sets. C_1=first cohort (University Hospital of Modena). C_2 = second cohort (University Hospital of Modena). C_3=third cohort (University Hospital of Parma).

4.3.6 Data analysis

4.3.6.1 Overview of analysis and types of data

The analysis of the accrued data was performed in Southampton by the author. Statistical support was provided throughout the analysis process by Dr. Borislav Dimitrov, Professor of Medical Statistics at the University of Southampton and co-supervisor of the author.

The analysis focused on the correlation between nominal or ordinal variables, consisting in the outputs of the qualitative assessment of “Velcro-type” crackles in the recordings and the scores of radiologic patterns of fibrosis at HRCT.

The data was presented using descriptive statistics for continuous variables such as mean with standard deviations and minimum and maximum values, while categorical variables are described using frequencies and percentages as appropriate. The distribution of ordinal or continuous data collected in the studies was assessed for normality using the Shapiro-Wilk test for variables with a high number of observations (above 200) or the Kolmogorov-Smirnov tests for variables with a lower number of observations. For all analyses performed, statistical significance level was set at $p < 0.05$. When not indicated otherwise, the data collected was entered in SPSS version 24 for analysis.

4.3.6.2 Analysis of demographic and medical history data

Information including age, gender, race/ethnicity, smoking status/history, clinical indication for HRCT and familiarity for ILD were all entered in SPSS (version 24) and descriptive statistics were used to characterise the study population. Where appropriate, categorical data were contrasted using Chi-squared test, while continuous variables were compared using independent samples t-test.

4.3.6.3 Analysis of imaging data

The presence of pulmonary fibrosis in the data set of single HRCT sections was assessed visually using a binary score. The scores were combined between the two radiologists: in case of disagreement, the image was classified as fibrotic. The semi-quantitative scores expressing the extension/severity of individual radiologic patterns were averaged between the two observers. The final scores (combined or averaged) were described using frequencies to explore their distribution in the data set. The scores were used in the primary analysis for correlation with the presence of audible “Velcro-type” crackles.

Weighted Cohen’s kappa statistic (k_w) was used to evaluate the inter-observer agreement for categorical variable. Weighting the k coefficient allows the degree of disagreement to be quantified by assigning greater emphasis to large differences between scores. The level of agreement was categorised as follows: poor ($0 < k_w < 0.19$), fair ($0.20 < k_w < 0.39$), moderate ($0.40 < k_w < 0.59$), good ($0.60 < k_w < 0.79$), and optimal ($k_w > 0.81$)(Brennan et al., 1992).

In the secondary analysis, the frequencies of the diagnostic pattern identified at the full volumetric HRCTs (non-ILD, UIP, possible UIP and inconsistent with UIP) of the patients from the selected study groups were presented and then correlated with the presence of bilateral “Velcro-type” crackles in the same patients.

4.3.6.4 Analysis of sound data

The binary scores provided by the 2 raters indicating the absence or the presence of “Velcro-type” crackles in the recordings were first presented using descriptive statistics and the inter-rater agreement for categorical variables was expressed using Cohen’s kappa statistic (k), and categorised as described above for the radiologists.

For each physician and each data set, the overall performance (sensitivity, specificity, negative and positive predictive values, and overall accuracy) of the assessment of “Velcro-type” crackles towards the identification of signs and patterns of pulmonary fibrosis (used as reference standard) on single HRCT sections and full volume HRCT scans was calculated.

For the analysis of correlation with different radiologic patterns of pulmonary fibrosis, the scores given by the physicians were combined by adjudicating “Velcro-type” crackles as present in case of disagreement.

The assessments of single recordings were also used to identify two groups of patients based on the presence of bilateral “Velcro-type” crackles (i.e. “Velcro-type” crackles scored as present in at least one right and one left site in the same patient). The derived scores were used in the secondary analysis of the study, consisting in the correlation of bilateral “Velcro-type” crackles with the diagnostic ILD pattern at the full volume HRCT scan.

4.3.6.5 Correlation between “Velcro-type” crackles and HRCT features and patterns

In the primary analysis of the case-control study the presence of “Velcro-type” crackles was associated with the evidence of fibrosis and individual radiologic patterns in the lung parenchyma below the recording sites at single HRCT sections. As such, the scores for left or right HRCT sections were used as reference according to the side of the chest the sound was recorded from.

First, cross-tabulation was used to explore differences between expected and observed proportions. These were tested for statistical significance using Chi-square test. Spearman’s rank correlation coefficient was then used in a univariate correlation analysis between the categorical (ordinal) variables.

Simple and multiple logistic regression modelling was used to determine the relationships between the extension/severity of individual radiologic patterns and “Velcro-type” crackles at the corresponding sites. Odds ratios (OR) and 95% confidence intervals were calculated for the outcome variables. Further models were built to assess whether the presence of crackles could predict the presence and severity of individual radiologic patterns. Because of the non-parametric nature of the outcome variables (the semi-quantitative scores of the radiologic patterns) generalised linear mixed modelling (GLIMMIX) was used.

Following identification of patients presenting bilateral “Velcro-type” crackles on chest auscultation, the relationships with the presence of fibrotic ILD on full volume HRCT scans and specific patterns (UIP, possible UIP, and inconsistent with UIP) were explored. Observed proportions were tested for statistical significance using Chi-square test. Logistic regression was used to further examine the relationship of the diagnostic patterns with the presence of bilateral crackles.

4.4 Measurements and Results

4.4.1 Characteristics of study population

The characteristics of the study groups defined according to the presence of fibrosis on single HRCT sections are reported in Table 8. The data collected included demographics, smoking history, family history for ILD and the clinical indication for undergoing a HRCT scan of the chest.

The fibrotic groups from the three cohorts were not significantly different in terms of age (mean age 69.7 ± 9.2, 70.1 ± 7.2, and 69.2 ± 10.3 respectively, p=0.931). Males were predominant in the two consecutive cohorts enrolled in Modena (61% and 67.7%), whilst females were slightly predominant in the Parma cohort (58.8%). As described in the Methodology chapter, within each cohort controls were matched with the fibrotic patients based on age and sex.

In Modena’s cohorts, fibrotic patients had more frequently a positive smoking history irrespectively to the radiologic evidence of pulmonary fibrosis. In contrast, in the independent Parma cohort non-smokers were predominant (64.7%).

In each cohort, most fibrotic patients underwent HRCT for known or suspected ILD (such as IPF, Connective Tissue Disease-related ILD, Pneumoconiosis), while most of non-fibrotic subjects were sent to HRCT for other reasons, including general symptoms such as dyspnea and cough, bronchiectasis, COPD, follow up of lung nodules, and haemoptysis. On the other hand, there was still a minority of subjects who underwent HRCT for suspicion of ILD but didn’t show signs of pulmonary fibrosis at the radiologic assessment. Conversely, a proportion of patients who didn’t perform HRCT for suspect of ILD were found to have radiologic signs of pulmonary fibrosis. In a few circumstances, it was not possible to retrieve the clinical indication for HRCT due to missing or unreliable information.

Approximately between 20% and 25% of fibrotic patients in each cohort had a positive family history for autoimmune disease, while a positive family history for ILD was much less common.

Table 8 - Characteristics of the case and control groups from the three cohorts enrolled in the study. Data are expressed as counts (%) or mean ± standard deviation (SD).

	First cohort (Modena)		**Second cohort (Modena)**		**Third cohort (Parma)**	
	Fibrotic (n=28)	Non-fibrotic (n=27)	Fibrotic (n=31)	Non-fibrotic (n=28)	Fibrotic (n=17)	Non-fibrotic (n=17)
Age, years	69.7 (± 9.2)	69.0 (± 10.6)	70.1 (± 7.2)	68.8 (± 7.3)	69.2 (± 10.3)	66.8 (± 8.8)
Sex						
Male	17(61%)	17(63%)	21(67.7%)	20(71.4%)	7 (41.2%)	6 (35.3%)
Female	11(39%)	10 (37%)	10 (32.3%)	8(28.6%)	10 (58.8%)	11 (64.7%)
Smoking history						
Current/Former	16 (57%)	16 (59%)	20 (64.5%)	23 (82.1%)	6 (35.3%)	12 (70.6%)
Never smoker	12 (43%)	11 (41%)	11 (35.5%)	5 (17.9%)	11 (64.7%)	5 (29.4%)
Indication for HRCT						
ILD	19(67.8%)	1(3.7%)	23(74.2%)	8 (28.6%)	8 (47.1%)	6 (35.3%)
Other	8(28.6%)	9(33.3%)	3 (9.7%)	17(60.7%)	7 (41.2%)	11 (64.7%)
Unknown	1(3.6%)	17(63%)	5 (16.1%)	3(10.7%)	2 (11.8%)	0 (%)
Family History						
Pulmonary Fibrosis	1 (3.6%)	0 (0%)	2 (6.5%)	0(0%)	0 (0%)	0 (%)
Autoimmune disease	6 (21.4%)	1 (3.7%)	6 (19.4%)	1(3.6%)	4 (23.5%)	4 (23.5%)

Given the demographic similarities between the three cohorts, the characteristics of cases and controls as single groups are also reported below in Table 9. The fibrotic patients from the three cohorts had a mean age of 69.78 years and were predominantly males (59.2%). They were more frequently ex- or current smokers (55.3%) and the more common indication for HRCT was the clinical suspect of ILD (64.5%). Approximately 1/5 of these subjects (21.1%) had familiarity for autoimmune disease.

Table 9 - Characteristics of cases and controls in the study (joined data from the three cohort groups). Data are expressed as counts (%) or mean ± standard deviation (SD).

	Fibrotic (n=76)	**Non-fibrotic (n=72)**
Age, years	69.78(± 8.6)	68.36(± 8.94)
Sex		
Male	45(59.2%)	43(59.7%)
Female	31 (40.8%)	29(40.3%)
Smoking history		
Current/Former	42 (55.3%)	51(70.8%)
Never smoker	34 (44.7%)	21(29.2%)
Indication for HRCT		
ILD	49(64.5%)	16(22.2%)
Other	13(17.1%)	36(50.0%)
Unknown	14(18.4%)	20(27.8%)
Family History		
Pulmonary Fibrosis	3(3.9%)	0(0.0%)
Autoimmune disease	16(21.1%)	6(8.3%)

4.4.2 Radiologic assessment of HRCTs and inter-observer agreement

Overall, 805 single HRCT sections were independently assessed by the two observers. The scores of pulmonary fibrosis and individual radiologic patterns were combined or averaged as described in the Methodology chapter. The frequencies of the scores are reported in Tables 10-12, while histograms in Figure 10 show their distribution in the imaging data set.

As mentioned in the research protocol, the evidence of pulmonary fibrosis in one of the HRCT sections from each patient allowed the identification of cases (“fibrotic” subjects) in the study. As such, despite the matching with “non-fibrotic” controls (who therefore had only non-fibrotic images), the imaging data set was still unbalanced towards a higher number of non-fibrotic images. Consequently, the semi-quantitative scores for individual radiologic features were not normally distributed, with lower scores more represented in the data set. The non-normal distribution was confirmed by the Shapiro-Wilk test, which was significant for all features evaluated (p=0.000). As such, non-parametric tests were chosen when these variables were used as outcomes in the analysis.

The semi-quantitative assessment is also reported for the two data sets of fibrotic and non-fibrotic images separately in Table 12. Even in the fibrotic data set, none of the individual patterns was normally distributed in terms of severity (p=0.000 at the Shapiro Wilk test), with milder grades of each alteration being more represented in the dataset. This was particularly evident for honeycombing - present in only 15.8% of the fibrotic images – but also for ground glass opacities and emphysema. Conversely, the different grades of reticulation were more fairly distributed, with only 2.3% images being without evidence of interstitial thickening. On the other hand, although low grades of individual radiologic abnormalities could be found in images evaluated as non-fibrotic, honeycombing was never present.

Table 10 – Qualitative assessment of pulmonary fibrosis (scores combined) in the imaging dataset. Data are expressed as counts and percentages (%).

Score	
0 (No fibrosis)	544 (67.6%)
1 (Fibrosis)	261 (32.4%)
Total	805(100%)

Table 11 – Semi-quantitative assessment of individual radiologic features (scores averaged) in the full imaging data set (n=805). Data are expressed as counts and percentages (%).

Score	Ground glass	Reticulation	Honeycombing	Emphysema	Traction bronchiectasis
0	648(80.5%)	427(53%)	761(94.5%)	690(85.7%)	593(73.7%)
0.5	72(8.9%)	95(11.8%)	11(1.4%)	49(6.1%)	85(10.6%)
1	34(4.2%)	108(13.4%)	15(1.9%)	30(3.7%)	60(7.5%)
1.5	19(2.4%)	52(6.5%)	3(0.4%)	19(2.4%)	44(5.5%)
2	16(2%)	68(8.4%)	7(0.9%)	9(1.1%)	16(2%)
2.5	10(1.2%)	18(2.2%)	3(0.4%)	3(0.4%)	6(0.7%)
3	5(0.6%)	25(3.1%)	1(0.1%)	4(0.5%)	1(0.1%)
3.5	1(0.1%)	5(0.6%)	2(0.2%)	1(0.1%)	-
4	0(0%)	7(0.9%)	2(0.2%)	0(0%)	-

Table 12 – Semi-quantitative assessment of individual radiologic features (scores averaged) in the fibrotic and non-fibrotic imaging datasets. Data are expressed as counts and percentages (%).

	Score	Ground glass	Reticulation	Honeycombing	Emphysema	Traction bronchiectasis
Fibrotic (n=261)	**0**	166(63.6%)	6(2.3%)	217(83.1%)	229(87.7%)	54(20.7%)
	0.5	42(16.1%)	18(6.9%)	11(4.2%)	13(5.0%)	80(30.7%)
	1	21(8.0%)	78(29.9%)	15(5.7%)	11(4.2%)	60(23%)
	1.5	12(4.6%)	43(16.5%)	3(1.1%)	3(1.1%)	44(16.9%)
	2	9(3.4%)	63(24.1%)	7(2.7%)	2(0.8%)	16(6.1%)
	2.5	7(2.7%)	18(6.9%)	3(1.1%)	2(0.8%)	6(2.3%)
	3	4(1.5%)	24(9.2%)	1(0.4%)	1(0.4%)	1(0.4%
	3.5	0(0%)	4(1.5%)	2(0.8%)	0(0%)	-
	4	0(0%)	7(2.7%)	2(0.8%)	0(0%)	-
Non-fibrotic (n=544)	**0**	482(88.6%)	421(77.4%)	0(0%)	461(84.7%)	539(99.1%)
	0.5	30(5.5%)	77(14.2%)	0(0%)	36(6.6%)	5(%)
	1	13(2.4%)	30(5.5%)	0(0%)	19(3.5%)	0(0%)
	1.5	7(1.3%)	9(1.7%)	0(0%)	16(2.9%)	0(0%)
	2	7(1.3%)	5(0.9%)	0(0%)	7(1.3%)	0(0%)
	2.5	3(0.6%)	0(0%)	0(0%)	1(0.2%)	0(0%)
	3	1(0.2%)	1(0.2%)	0(0%)	3(0.6%)	0(0%)
	3.5	1(0.2%)	1(0.2%)	0(0%)	1(0.2%)	-
	4	0(0%)	0(0%)	0(0%)	0(%)	-

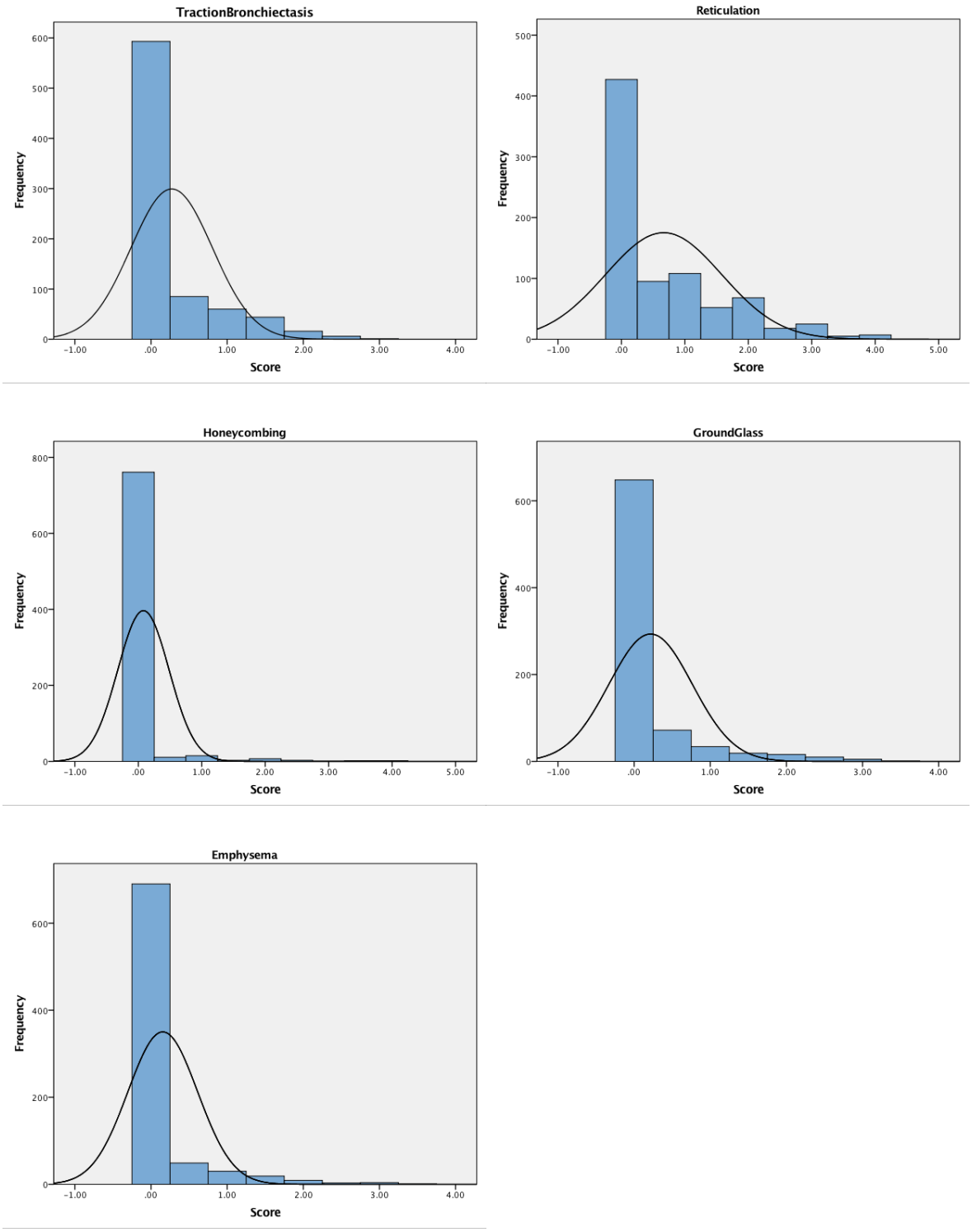


Figure 10 – Histograms showing the distribution of averaged scores of the individual radiologic features evaluated in the study. The black line shows normal distribution.

The inter-observer agreement was calculated for the different interstitial patterns using weighted Cohen's Kappa (k_w)for ordinal variables. The level of agreement between the thoracic radiologists was good for the qualitative evaluation of fibrosis (k_w=0.69, 95% CI 0.65-0.73). Among the individual radiologic features, the level of agreement was good for honeycombing (k_w=0.71, 95% CI 0.63-0.79) and reticular opacities (k_w 0.65, 95% CI 0.62-0.68), while it was moderate for traction bronchiectasis (K_w=0.51, 95% CI 0.47-0.55), fair for emphysema (K_w=0.38, 95% CI 0.31-0.44) and poor for ground glass opacities (K_w=0.28, 95% CI 0.22-0.34).The contingency tables built to calculate the level of agreement between the two radiologists are fully reported in Appendix 1 in the accompanying material of this thesis.

With regard to the full volume HRCT scans, assessed by a third independent observer, 66 over 148 subjects (44.6%) had a pattern consistent with fibrotic ILD on HRCT. The characteristics of the study groups defined according to the evidence of a fibrotic pattern on full volume HRCT scans are shown in Table 13. As for specific ILD diagnostic patterns, possible UIP was the most represented (47%). Overall, the characteristics were consistent to those of the groups originally defined according to the evidence of fibrosis on single HRCT sections. The clinical suspicion of ILD was strongly associated with the evidence of ILD on HRCT, with 81.3% of the patients with UIP having such indication.

Table 13 - Characteristics of study groups defined according to the evidence of fibrotic ILD on full volume HRCT scan. Data are expressed as counts (%) or mean ± standard deviation (SD). FILD = Fibrotic Interstitial Lung Disease.

	FILD (n=66)	**Non FILD (n=82)**
Age, years	71 (± 8.2)	67.55 (± 8.96)
Sex		
Male	43 (65.2%)	45 (54.9%)
Female	23 (34.8%)	37 (45.1%)
Smoking history		
Current/Former	37 (56.1%)	56 (68.3%)
Never smoker	29 (43.9%)	26 (31.7%)
Indication for HRCT		
ILD	46 (69.7%)	19(23.2%)
Other	14 (21.2%)	43 (52.4%)
Unknown	6 (9.1%)	20 (24.4%)
Family History		
Pulmonary Fibrosis	3 (4.5%)	1 (1.2%)
Autoimmune disease	13 (19.7%)	10 (12.2%)
HRCT pattern		
UIP	15 (22.7%)	
Possible UIP	31 (47%)	-
Inconsistent with UIP	20 (30.3%)	

4.4.3 Subjective assessment of "Velcro-type" crackles and inter-rater reliability

The qualitative assessment of "Velcro-type" crackles on electronic auscultation of single recordings is presented in Table 14. The presence of "Velcro-type" crackles was reported in 18.1% and 30.4% of the recordings by the first and second physician respectively. The recordings containing "Velcro-type" crackles in the opinion of either one of the two physicians were approximately 1/3 of the data set (32.7%). The level of agreement between the two raters was moderate (Cohen's kappa = 0.553, Table 15).

Table 14 – Assessment of "Velcro-type" crackles in single recordings as reported by the two physicians separately and combined. Data are expressed as counts and percentages (%).

"Velcro-type" crackles	Physician 1	Physician 2	Combined
Absent	659 (81.9%)	560 (69.6%)	542 (67.3%)
Present	146 (18.1%)	245 (30.4%)	263 (32.7%)

Table 15 – Contingency table for inter-rater agreement of assessment of "Velcro-type" crackles on electronic auscultation of single recordings. Data are expressed as counts.

		Physician 2		
	"Velcro-type" crackles	**No**	**Yes**	**Total**
Physician 1	**No**	542	117	659
	Yes	18	128	146
	Total	560	245	805

Patients with bilateral "Velcro-type" crackles were identified based on the assessment of "Velcro-type" crackles in the single recordings (Table 16). The agreement between the two physicians for the assessment of bilateral "Velcro-type" crackles was good (Cohen's kappa= 0.69, 95% CI 0.57-0.82) (

Table 17).

Table 16 – Assessment of bilateral “Velcro-type” crackles in individual patients as reported by the two physicians. Data expressed as counts and percentages (%).

Bilateral “Velcro-type” crackles	Physician 1	Physician 2	Combined
No	112(75.7%)	96(64.9%)	95(64.2%)
Yes	36(24.3%)	52(35.1%)	53(35.8%)

Table 17 - Contingency table for calculation of inter-rater agreement for the assessment of bilateral “Velcro-type” crackles on electronic auscultation. Data expressed as counts.

		Physician 2		**Total**
	Bilateral “Velcro-type” crackles	No	Yes	
Physician 1	No	95	17	112
	Yes	1	35	36
	Total	96	52	148

4.4.4 Correlation of "Velcro-type" crackles with HRCT

4.4.4.1 Relationships between "Velcro-type" crackles and radiologic features on single HRCT sections

The association between the radiologic abnormalities and the presence of "Velcro-type" crackles in the recordings obtained from corresponding sites was first explored via crosstabulation. As shown in Table 18 and displayed in the bar charts in Figure 11 and Figure 12, the observed proportions suggest a linkage exists between "Velcro-type" crackles and the extent of all individual radiologic features evaluated, except for emphysema.

- Table 18 – Association between extent of different radiologic patterns and "Velcro-type" crackles on corresponding sites. Data expressed as counts and percentages (%) within imaging scores, presented with Pearson's Chi-Squared test statistics ($\chi 2$) and relative p values.

Pattern	Absence of "Velcro-type" crackles (n=542)	Presence of "Velcro-type" crackles (n=263)	$\chi 2$	p value
Fibrosis			132.61	0.000
Yes	104(39.8%)	157(60.2%)		
No	438(80.5%)	106(19.5%)		
Traction bronchiectasis			136.33	0.000
0	467(78.8%)	126(21.2%)		
0.5	30 (35.3%)	55 (64.7%)		
1	23 (38.3%)	37 (61.7%)		
1.5	17 (38.6%)	27(61.4%)		
2	4(25.0%)	12(75.0%)		
2.5	1(16.7%)	5(83.3%)		
3	0(0.0%)	1(100%)		
Reticulation			135.1	0.000
0	349(81.7%)	78(18.3%)		
0.5	71(74.7%)	24(25.3%)		
1	57(52.8%)	51(47.2%)		
1.5	26(50.0%)	26(50.0%)		
2	21(30.9%)	47(69.1%)		
2.5	7(38.9%)	11(61.1%)		
3	9(36.0%)	16(64.0%)		
3.5	2(40.0%)	3(60.0%)		
4	0(0.0%)	7(100%)		

Honeycombing			33.87	0.000
0	528(69.4%)	233(30.6%)		
0.5	2(18.2%)	9(81.8%)		
1	8(53.3%)	7(46.7%)		
1.5	0(0.0%)	3(100%)		
2	2(28.6%)	5(71.4%)		
2.5	1(33.3%)	2(66.7%)		
3	0(0.0%)	1(100%)		
3.5	1(50.0%)	1(50.0%)		
4	0(0.0%)	2(100%)		
Ground glass opacities			39.00	0.000
0	462(71.3%)	186(28.7%)		
0.5	42(58.3%)	30(41.7%)		
1	22(64.7%)	12(35.3%)		
1.5	7(36.8%)	12(63.2%)		
2	4(25.0%)	12(75.0%)		
2.5	4(40.0%)	6(60.0%)		
3	1(20.0%)	4(80.0%)		
3.5	0(0.0%)	1(100%)		
Emphysema			4.08	0.770
0	457(66.2%)	233(33.8%)		
0.5	35(71.4%)	14(28.6%)		
1	22(73.3%)	8(26.7%)		
1.5	14(73.7%)	5(26.3%)		
2	8(88.9%)	1(11.1%)		
2.5	2(66.7%)	1(33.3%)		
3	3(75.0%)	1(25.0%)		
3.5	1(100%)	0(0.0%)		

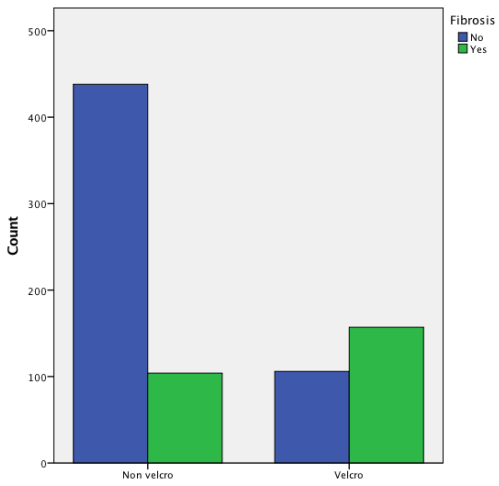


Figure 11 – Frequencies of HRCT images with or without pulmonary fibrosis stratified by presence or absence “Velcro-type”-crackles at the corresponding recordings.

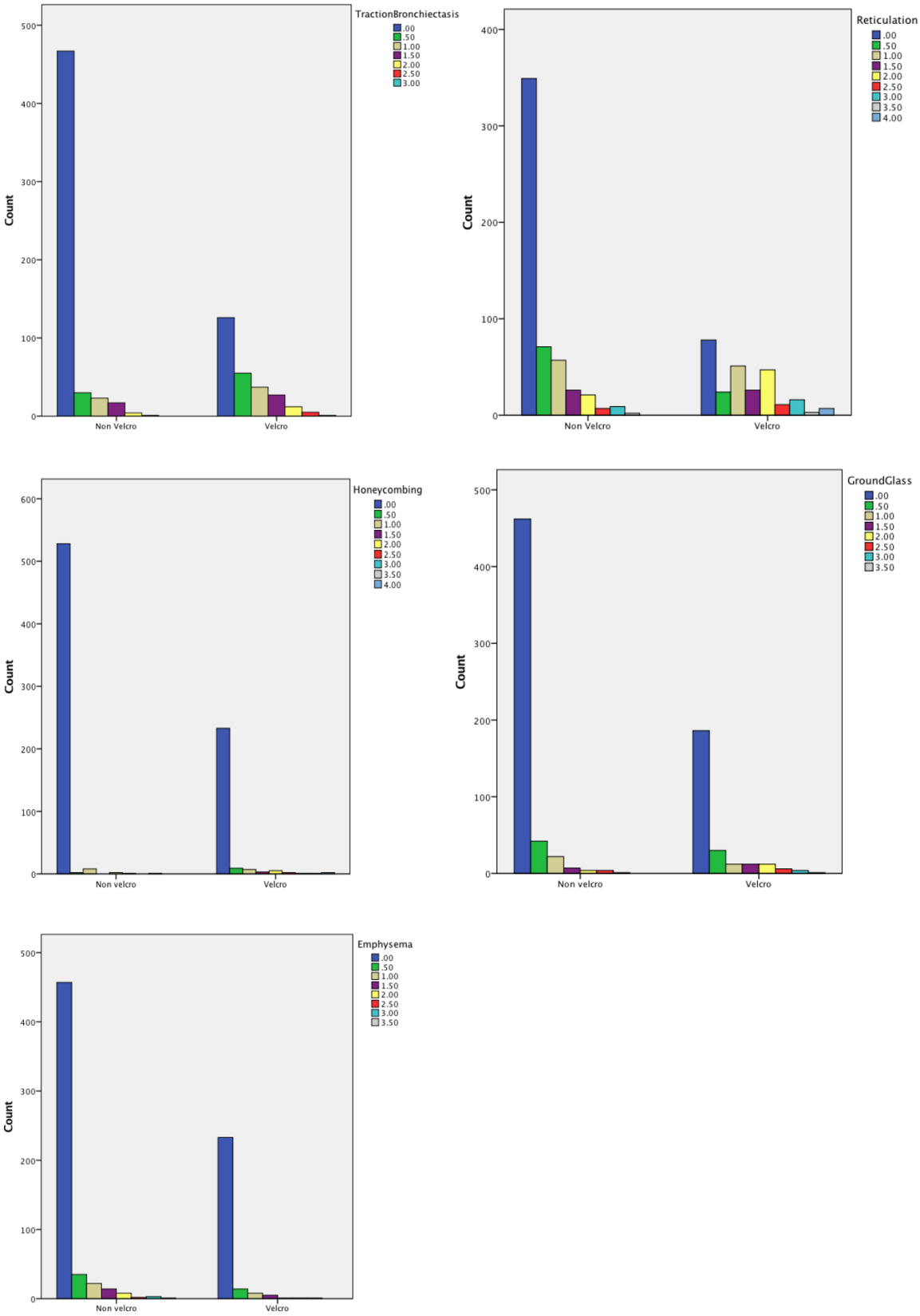


Figure 12 – Frequencies of HRCT images with different grades of severity of individual radiologic features stratified by presence or absence of “Velcro-type”-crackles at the corresponding recording.

On univariate correlation analysis, the strongest relationships were found between crackles and the presence of fibrosis (Spearman’s rank correlation coefficient = 0.406) and between crackles and the severity of traction bronchiectasis and reticulation (Spearman’s coefficient = 0.405 and 0.385, respectively), as shown in Table 19. A slightly negative, although not significant correlation was found between emphysema and crackles.

Table 19 – Univariate correlation between radiologic patterns and “Velcro-type” crackles, expressed as Spearman’s rank correlation coefficient (r).

	r	P value
Fibrosis	0.406	0.000
Traction bronchiectasis	0.405	0.000
Reticulation	0.385	0.000
Honeycombing	0.182	0.000
Ground glass opacities	0.183	0.000
Emphysema	-0.055	0.171

A simple logistic regression model was then used to estimate the relationships between the radiologic patterns and the presence of “Velcro-type” crackles, used as dependent variable (Table 20). With the exception of emphysema, which didn’t show significant correlation, all the remaining individual radiologic patterns were positively and significantly associated with crackles. The evidence of pulmonary fibrosis was the most strongly associated (odds radio 6.24, 95% CI 4.5;8.66); among the individual patterns, traction bronchiectasis had the strongest association with the presence of “Velcro-type” crackles (odds ratio 4.37, 95% CI 3.17-6.02). Since not all subjects had the same number of images in the data set, the analysis was adjusted by the number of images per patient.

Table 20 – Simple logistic regression of individual radiologic features at HRCT images toward presence of “Velcro-type” crackles on corresponding recording sites. Data presented as odds ratios with p values and 95% confidence intervals (CI).

Feature	Odds ratio	P value	95% CI	
			Lower bound	Upper bound
Fibrosis	6.24	0.000	4.5	8.66
Ground glass opacities	2.13	0.000	1.61	2.81
Reticulation	2.57	0.000	2.14	3.09
Traction bronchiectasis	4.37	0.000	3.17	6.02
Honeycombing	2.39	0.000	1.52	3.76
Emphysema	0.72	0.077	0.5	1.04

A multiple logistic regression analysis was then performed including the features that showed to be significantly correlated with crackles on univariable analysis. Apart from the overall evidence of fibrosis, both the extent of reticulation, ground glass and honeycombing were independently associated with “Velcro-type” crackles (fibrosis: odds ratio 1.86, 95% CI 1.05-3.33; reticulation: odds ratio 1.72, 95% CI 1.3-2.27; ground glass: odds ratio 1.66, 95% CI 0.84-1.98; honeycombing: odds ratio 1.66, 95% CI 1.08-2.53). Traction bronchiectasis, which showed the strongest correlation with crackles in the single factor model, was no longer significantly associated with crackles when entered in the multivariable model.

Table 21 – Multiple logistic regression of individual radiologic features at HRCT images towards presence of "Velcro-type" crackles on corresponding recording sites. Data presented as odds ratios with p values and 95% confidence intervals (CI).

Feature	Odds ratio	P value	95% CI	
			Lower bound	Upper bound
Fibrosis	1.86	0.034	1.05	3.3
Ground glass opacities	1.66	0.001	1.23	2.23
Reticulation	1.72	0.000	1.3	2.27
Traction bronchiectasis	1.29	0.249	0.84	1.98
Honeycombing	1.66	0.020	1.08	2.53

Generalised linear mixed models (GLIMMIX) were built using individual radiologic features as dependent variables to confirm whether the auscultation of "Velcro-type" crackles could predict the presence and severity of different abnormalities on HRCT. Just as in the former analysis, the number of images per patient was entered in each model as covariate for adjustment. The presence of crackles significantly predicted the severity of the radiologic patterns with the sole exception of emphysema, as demonstrated by the significant difference between estimated mean scores for reticulation, traction bronchiectasis, ground glass and honeycombing in the absence or presence of crackles. The estimated mean scores are presented in Table 22 to Table 26, and plotted in the graphs shown in Figure 13 to Figure 17.

Table 22 – Estimated mean scores of reticulation based on "Velcro-type" crackles' assessment. Data presented as means with standard error and 95% confidence intervals (CI).

"Velcro-type" crackles	Mean	Std. error	95% CI	
			Lower bound	Upper bound
Absent	0.407	0.303	-0.187	1.002
Present	1.171	0.307	0.570	1.073

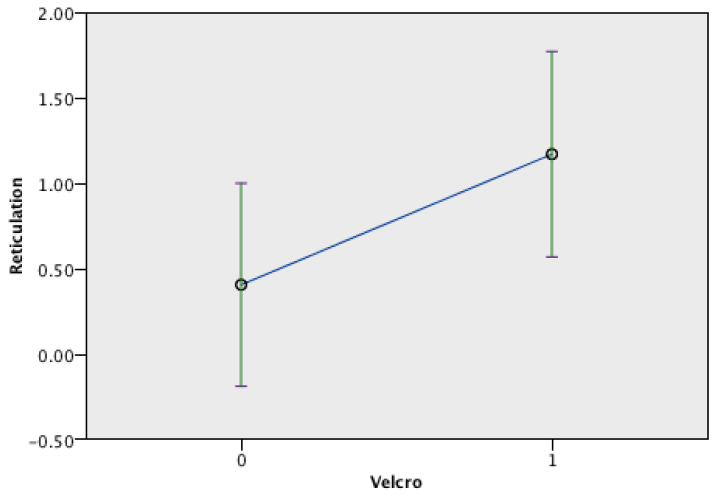


Figure 13 - Estimated mean scores of reticulation as predicted by assessment of “Velcro-type” crackles.

Table 23 - Estimated mean scores of traction bronchiectasis based on “Velcro-type” crackles’ assessment. Data presented as means with standard error and 95% confidence intervals (CI).

“Velcro-type” crackles	Mean	Std. error	95% CI	
			Lower bound	Upper bound
Absent	0.263	0.168	-0.066	0.592
Present	0.678	0.172	0.341	1.015

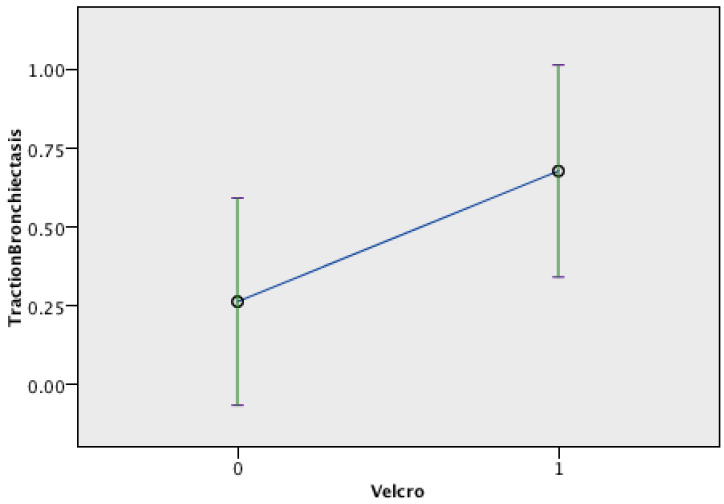


Figure 14 - Estimated mean scores of traction bronchiectasis as predicted by assessment of “Velcro-type” crackles.

Table 24 - Estimated mean scores of honeycombing based on “Velcro-type” crackles’ assessment. Data presented as means with standard error and 95% confidence intervals (CI).

“Velcro-type” crackles	Mean	Std. error	95% CI	
			Lower bound	Upper bound
Absent	0.157	0.229	-0.292	0.606
Present	0.297	0.234	-0.162	0.756

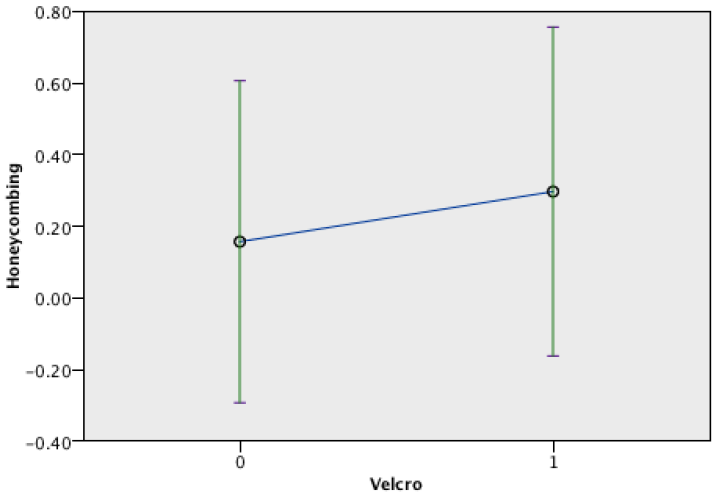


Figure 15 - Estimated mean scores of honeycombing as predicted by assessment of “Velcro-type” crackles.

Table 25 - Estimated mean scores of ground glass opacities based on “Velcro-type” crackles’ assessment. Data presented as means with standard error and 95% confidence intervals (CI).

“Velcro-type” crackles	Mean	Std. error	95% CI	
			Lower bound	Upper bound
Absent	0.370	0.076	0.220	0.520
Present	0.607	0.079	0.453	0.761

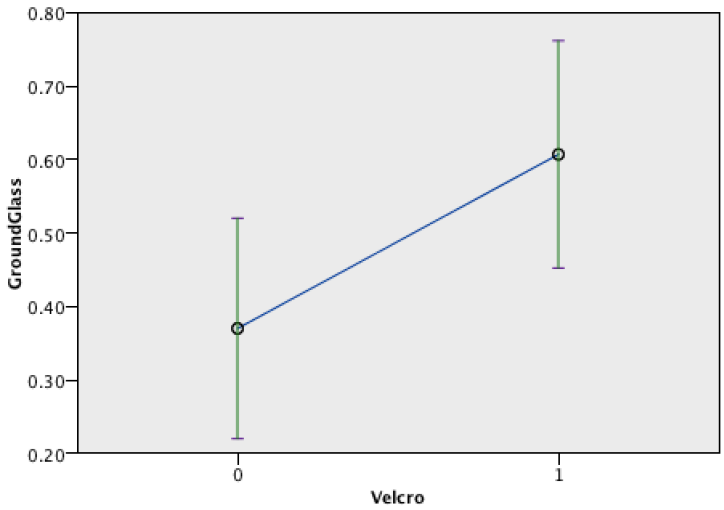


Figure 16 - Estimated mean scores of ground glass opacities as predicted by assessment of “Velcro-type” crackles.

Table 26 - Estimated mean scores of emphysema based on “Velcro-type” crackles’ assessment. Data presented as means with standard error and 95% confidence intervals (CI).

“Velcro-type” crackles	Mean	Std. error	95% CI	
			Lower bound	Upper bound
Absent	0.412	0.495	-0.059	1.383
Present	0.352	0.511	-0.651	1.356

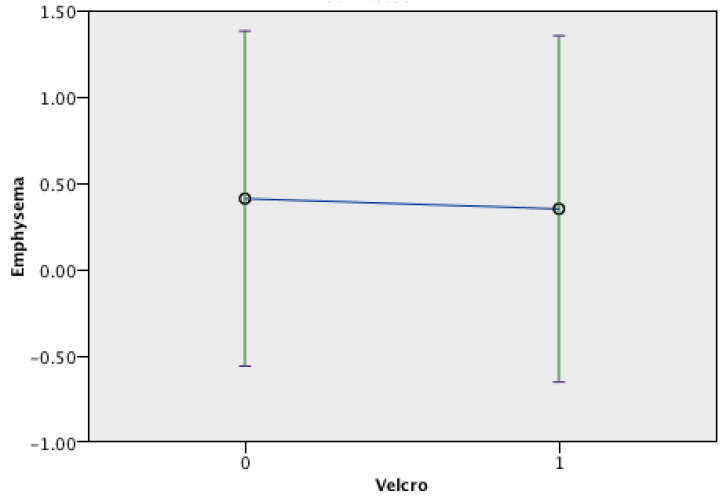


Figure 17 - Estimated means of scores of emphysema as predicted by assessment of “Velcro-type” crackles.

4.4.4.2 Relationships between “Velcro-type” crackles and radiologic patterns on full volumetric HRCTs scans

The diagnostic patterns evaluated on full volumetric HRCT scans were associated with the presence of either unilateral or bilateral “Velcro-type” crackles at the combined assessment of physicians (Table 27). Among patients with fibrotic ILD, the majority (65.2%) of patients had bilateral “Velcro-type” crackles, while only 16.7% had audible crackles on a single side of the chest. Notably, almost one fifth of the subjects (18.2%) with fibrotic ILD did not present “Velcro-type” crackles at all. Among patients without radiologic evidence of fibrotic ILD, 48.8% had at least some crackles heard on auscultation. On univariate analysis, the presence of ILD on HRCT significantly correlated with the presence of bilateral, but not unilateral crackles (Table 28).

As for specific diagnostic patterns, bilateral crackles showed the strongest correlation with the possible UIP pattern (r = 0.308). Conversely, the presence of a normal lung parenchyma or another non-fibrotic pattern was negatively correlated with the presence of bilateral crackles (r = -0.442).

Table 27 - Association between bilateral “Velcro-type” crackles and different diagnostic patterns on full HRCT scans. Data expressed as counts (% within patients with or without bilateral “Velcro-type” crackles), presented with Pearson’s Chi-Squared test statistics ($\chi 2$) and relative p values.

“Velcro-type” crackles	No fibrotic ILD	Fibrotic ILD	$\chi 2$	P value
No	42 (51.2%)	12 (18.2%)		
Unilateral	30 (36.6%)	11 (16.7%)	44.81	0.000
Bilateral	10 (12.2%)	43 (65.2%)		
Total	82 (100%)	66 (100%)		

Table 28 – Univariate correlation analysis between unilateral or bilateral "Velcro-type" crackles and patterns on full volume HRCT scans. Univariate correlations are expressed as Spearman's correlation coefficient (r).

	Unilateral "Velcro-type" crackles		Billateral "Velcro-type" crackles	
HRCT pattern	**r**	**P value**	**r**	**P value**
Fibrotic ILD	-0.127	0.123	0.549	0.000
UIP	-0.073	0.378	0.263	0.001
Possible UIP	-0.087	0.295	0.308	0.000
Inconsistent with UIP	0.124	0.134	-0.442	0.000

Emphysema had a trend towards a negative association with the presence of "Velcro-type" crackles (OR 0.72, CI 95% 0.5-1.04, p=0.077), and the absence "Velcro-type" crackles confirmed to predict higher emphysema scores (Table 26). To further investigate whether emphysema influences the transmission of lung sounds on chest auscultation in patients with coexisting fibrosis the mean scores for emphysematous alterations on single HRCT sections were compared between subgroups of ILD subjects with either bilateral, unilateral or absence of crackles on chest auscultation. The scores of emphysema were overall low in this population (mean score 0.11, 0.38 SD), and higher scores were found in patients with unilateral crackles (mean 0.3, SD 0.67) as compared to patients with bilateral or without crackles (mean 0.06, SD 0.23 and 0.11, SD 0.38, respectively), suggesting that in ILD subjects no clear link existed between the extent of emphysema and the absence of "Velcro-type" crackles on auscultation.

Table 29 - Scores of emphysema across subgroups of fibrotic ILD subjects with different acoustic findings on chest auscultation. Data are expressed as means with standard deviation (SD) and 95% confidence intervals (95% CI).

"Velcro-type" crackles	N	Mean	SD	95% CI	
				Lower Bound	Upper Bound
No	54	0.11	0.38	0.0140	0.21
Unilateral	41	0.30	0.67	0.1330	0.47
Bilateral	53	0.06	0.23	0.0322	0.09
	148	0.11	0.38	0.0728	0.15

Logistic regression was then used to further explore the relationships between the diagnostic patterns on full volume HRCT and bilateral “Velcro-type” crackles (Table 30). The presence of bilateral “Velcro-type” crackles confirmed to strongly predict fibrotic ILD pattern (OR 13.46, 95% CI 5.85-30.96, p<0.001). As for specific patterns, both definite and possible UIP were independently associated with the presence of bilateral “Velcro-type” crackles on multivariable regression analysis, with the UIP pattern showing the strongest association (UIP: OR 10.0, 95% CI 2.9-34.48, p<0.001; possible UIP: OR 6.61, 95% CI 2.76-15.84, p<0.001).

Table 30 – Regression analysis of presence of bilateral “Velcro-type” crackles for pattern on full volume HRCT scan.

HRCT pattern	Bilateral “Velcro-type” crackles	
	OR (CI 95%)	p
Fibrotic ILD	13.46 (5.71-29.182)	<0.001
Definite UIP	10.00 (2.9-34.98)	<0.001
Possible UIP	6.61 (2.76-15.84)	<0.001

4.4.5 Diagnostic accuracy of “Velcro-type” crackles assessment

The parameters of accuracy of the assessment of “Velcro-type” crackles toward the evidence of pulmonary fibrosis on single HRCT sections is reported in Table 31. Since the physicians listened to the three data sets of recordings at different times and the recordings were obtained in two different clinical settings, the performance was initially calculated for each data set separately.

Overall, specificity and negative predictive values were higher than sensitivity and positive predictive values, indicating that the absence of fibrosis was discriminated more efficiently than its presence. While the accuracy was similar in the first two data sets, the results from the third data set show lower sensitivity (reaching a bottom of 26.3%) and positive predictive value. The parameters of accuracy had substantial discrepancies between the two physicians. The first physician reached lower sensitivity and negative predictive value (meaning more false negatives) but higher specificity and positive predictive value (meaning less false positives) than the second physician.

Table 31 - Accuracy of the subjective assessment of “Velcro-type” crackles at electronic auscultation toward the detection of pulmonary fibrosis on single HRCT images. ACC=accuracy; SEN=sensitivity; SPEC=specificity; PPV=positive predictive value; NPV=negative predictive value. Data expresses as percentages with 95% confidence intervals.

Dataset		SEN	SPEC	PPV	NPV
1 (n=300)	Physician 1	40.4(31.2;50.2)	85.9(79.9;90.3)	62(49.6;73)	71.6(65.2;77.2)
	Physician 2	56(46.1;65.4)	77(70.2;82.6)	58.1(48.1;67.5)	75.4(68.6;81.1)
2 (n=319)	Physician 1	39.5(30.6;49.1)	94.1(89.8;96.8)	79.8(65.8;88.2)	73.7(67.8;78.8)
	Physician 2	64.9(55.3;73.5)	85.9(80.1;90.2)	71.8(61;80)	81.5(75.5;86.3)
3 (n=186)	Physician 1	26.3(14;43.4)	94.6(89.3;97.5)	55.6(31.3;77.6)	83.3(76.6;88.5)
	Physician 2	36.8(22.3;54)	84.5(77.4;89.7)	37.8(22.9;55.2)	83.9(76.8;89.2)
Average		38	87	60.9	78.2

The parameters of accuracy of bilateral “Velcro-type” crackles toward evidence of fibrotic ILD on full volume HRCT scans based are presented in Table 32. The presence of bilateral “Velcro-type” crackles predicted a pattern consistent with fibrotic ILD in the vast majority of patients (84.2% on average). Nevertheless, sensitivity was still low, as almost half of the patients with fibrotic ILD (55.3% on average) did not have bilateral crackles on auscultation.

Table 32 - Accuracy of bilateral “Velcro-type” crackles on chest auscultation toward detection of fibrotic ILD on full HRCT volumetric scans (n=148). ACC=accuracy; SEN=sensitivity; SPEC=specificity; PPV=positive predictive value; NPV=negative predictive value. Data expresses as percentages with 95% confidence intervals.

	SEN	SPEC	PPV	NPV	LR+
Physician 1	45.5 (33.1;58.2)	93.9 (86.3;98)	85.7(71.1;93.6)	68.1(62;72.9)	7.45(3.1-18.1)
Physician 2	65.1(52.4;76.5)	89.0(80.2;94.9)	82.7(69.6;88.5)	75(71.6;90.1)	5.94(3.1-11.3)
Average	55.3	91.5	84.2	71.6	6.7

4.5 Discussion

4.5.1 Demographic data

Overall, the collected demographics and clinical data showed that the three final cohorts were quite homogeneous, which justified unifying the data sets in the analysis.

The average age of the three fibrotic groups was consistent with the clinical picture of most fibrotic ILD, which are known to be diseases of ageing. The predominance of males was also consistent with the epidemiology of most ILD. The clinical suspect for ILD represented the primary indication for performing HRCT in all fibrotic groups, which confirms HRCT as an elective test for diffuse parenchymal lung disease. Not surprisingly, a positive history of smoking was present in most patients from both fibrotic and non-fibrotic groups, since cigarette smoke represents a common risk factor for respiratory disorders.

On independent full HRCT scan review, 66 subjects (44.6%) had radiological findings consistent with an ILD, with 16 (24.2%) of these consistent with UIP. As such, 10 subjects had evidence of isolated pulmonary fibrosis at single HRCT sections but did not have radiological findings consistent with an ILD on full HRCT review. The demographic characteristics of the ILD population were found to be very similar to the fibrotic groups originally defined according to the evidence of pulmonary fibrosis at single HRCT sections, both in terms of age and sex, smoking history and clinical indication for HRCT, thus supporting the robustness of the original study design. The evidence that the clinical suspect for ILD was strongly associated with the actual presence of ILD (and particularly UIP) at HRCT further confirms the value of clinical judgement (including demographics, symptoms, and the evidence of physical signs such as “Velcro-type” crackles) towards diagnosis of ILD.

4.5.2 Radiologic data and inter-observer agreement

Approximately two thirds (67.6%) of the HRCT sections forming the final data set were scored as non-fibrotic as a result of the combined assessment of radiologists. Such large proportion of non-fibrotic images was due to the case selection approach followed in this study. Most ILD present a peculiar or predominant distribution: for example, the fibrotic process in IPF usually starts in the basal posterior regions of the lungs and progress upwards, while other disorders such as Chronic Hypersensitivity Pneumonitis tend to involve the upper regions first. Only the most severe, advanced cases would present fibrosis involving the entirety of the lung parenchyma. As such, many patients classified as “fibrotic” had also some non-fibrotic images extracted from the HRCT scan, which was responsible for this unbalanced data set.

The prevalence of milder grades of radiologic abnormalities in the data set of fibrotic images might be due to the better awareness of ILD and IPF among physicians, who are nowadays more prone to suspect these disorders when a combination of symptoms and signs are present – which allows reach diagnosis at earlier stages. More and more patients are being evaluated at a stage where architectural distortion of the lung parenchyma is still very limited, or has not even occurred yet, explaining the low number of fibrotic images containing some degree of honeycombing. On one hand this is positive, however over the last years the number of patients diagnosed with “atypical” UIP at HRCT (i.e. with no honeycombing) has significantly increased: these patients cannot receive specific anti-fibrotic treatment though, which represents an unsolved clinical issue that should to be faced by the provision of new criteria for diagnosing IPF, as recently suggested by experts (Martinez et al., 2017).

Emphysema is known to affect the transmission of normal and pathological lung sounds, making their assessment more challenging. The absence of a significant quote of emphysematous alterations in the imaging data set evaluated in this study is reassuring as to the lack of potential interference with the assessment of “Velcro-type” crackles by the physicians.

In this study, the agreement between the two expert thoracic radiologists was found to be good with regard to the qualitative assessment of pulmonary fibrosis. Since the radiologic diagnosis of UIP can secure – together with an appropriate clinical picture - a diagnosis of IPF, high levels of inter-observer agreement for UIP are crucial to deliver clinical usefulness. Unfortunately, inter-observer agreement for the radiologic diagnosis of UIP (defined according the most recent ATS/ ERS/JRS/ALAT guidelines) has been found to be at best moderate across different level of expertise (Walsh et al., 2016). However, the purpose of this study was not to prove the diagnostic yield of “Velcro-type” crackles towards radiologic UIP, but to match lung sounds with the evidence of pulmonary fibrosis at single HRCT sections. Such assessment was therefore easier as compared to the evaluation of the distribution of different patterns at the full HRCT scan - which is instead required for the definition of UIP. It must be also pointed out that both radiologists had a high level of experience (>10 and >20 years respectively), which might also concur to explain this level of agreement. As for individual patterns like honeycombing and traction bronchiectasis, the levels of agreement reached by the two radiologists in this study reflected the values showed in a recent study (Walsh et al., 2016), where Cohen’s kappa values for the presence of honeycombing ranged from 0.56 to 0.65 (good), while for bronchiectasis they ranged from 0.32 to 0.45 (moderate) across groups of radiologists with different levels of expertise. This confirms that inter-observer agreement for honeycombing - whose presence is still required for diagnosis of UIP at HRCT - is far from optimal, as also reported in previous studies (Watadani et al., 2013).

4.5.3 Inter-rater reliability

The electronic recording of lung sounds allowed an accurate and unbiased assessment of “Velcro-type” crackles performed by two physicians with several years of expertise in ILD, who were blindly provided with the recordings to the clinical information of the patients. Importantly, this approach offered the opportunity to explore for the first time the reproducibility of the assessment of digitally recorded “Velcro-type” crackles between expert raters. Both physicians had comparable levels of expertise in diagnosis and management of ILD (>10 years); one had already used a digital stethoscope in clinical practice before, while the other physician was naive to the tool. The level of agreement for the assessment of “Velcro-type” crackles was only moderate, which can be explained by several factors. Firstly, the ability of discriminating sounds may not uniquely depend on the level of experience of physicians. They may have different ways of interpreting the transmission of lung sounds, which can be also influenced by one’s individual hearing capabilities. Secondly, the artefacts and the external noise audible in the recordings may represent a potential obstacle to the correct interpretation of the signal. Lung sounds were recorded in a clinical setting (the preparatory room of a radiology department), a non-controlled environment where external noise (e.g. doors opening, voices in the background) could affect the overall quality of the recordings. Despite efforts were made to keep the chest piece as still as possible, the friction produced by the manual placement of the stethoscope on the skin or produced by slight movements of the chest wall during deep breaths is another potential source of disturbance. Finally, despite instructions were provided to the patients, in some cases breathing might have been not deep enough to generate an adequate airflow and therefore ensure proper transmission of sounds across the thoracic wall. It is therefore possible that some disagreement was due to the different ability of physicians in discriminating noise from real lung sounds. The previous experience with digital stethoscopes of one of the physicians might have also played a role in the discrepancies.

On the other hand, it should be pointed out that the level of agreement for the identification of bilateral “Velcro-type” crackles in the individual patients indicated higher reproducibility of electronic auscultation for such assessment, which is indeed a more clinically relevant finding.

In conclusion, two expert respiratory physicians had from moderate to good agreement in evaluating “Velcro-type” crackles on electronic auscultation, which supports the applicability of this technique in clinical practice. Another measure of reproducibility of a measurement is represented by the intra-subject reliability, i.e. when the same recordings are blindly assessed by a physician at different times. Intra-subject reproducibility was not explored in this study as it might have implied further, significant amount work for the physicians. This study also lacked to

evaluate the reproducibility of “Velcro-type” crackles assessed via standard chest auscultation, which remains unknown. The comparison between the two approaches would help determine the real value of electronic auscultation. Indeed, such investigation would not be without biases as standard auscultation could not be totally blinded. Finally, only two physicians were involved in this study, and it is not known about whether and how the concordance would change for a larger number of physicians with different levels of expertise. In particular, valuable information would be obtained by having general practitioners listen to the recordings, since primary care is the ideal setting where to assess the diagnostic value of “Velcro-type” crackles.

4.5.4 Correlation between “Velcro-type” crackles and radiologic features and patterns of pulmonary fibrosis on HRCT

Quantitative analysis of lung sounds has been extensively used to distinguish ILD patients from healthy subjects or from subjects with other respiratory conditions, mostly via automatic classification methods relying on supervised machine learning techniques. A study by Filiestra and co-workers demonstrated that crackles recorded from IPF patients can be distinguished from those in patients with pneumonia with a sensitivity of 82%, a specificity of 88% and an accuracy of 86%, and from those in patients suffering from chronic heart failure patients with a sensitivity of 77%, a specificity of 85% and an accuracy of 82%, using supervised classifiers such as neural networks and vector machines (Flietstra et al., 2011). However, recording of lung sounds was performed using a complex device (a multichannel lung sounds analyser), and the population consisted of a convenience sample instead of being consecutively enrolled. A recent study using a novel two-dimensional grey-scale imaging system called vibration response imaging to classify lung sounds also suggested that crackles present in IPF patients have distinct features as compared to those found in other pulmonary or cardiac disorders (Liu et al., 2014). Nevertheless, these methods have never been applied in appropriate clinical settings, and the studies were not properly designed for diagnostic validation of the examined tools. As such, their clinical usefulness has remained elusive so far. A more clinically-oriented approach was pursued in a small recent study that combined the use of a customised digital stethoscope and computerised analysis of lung sounds to build acoustic signatures of lung sounds in ILD. The use of three statistical properties of signals (skewness, lacunarity and entropy) was validated via classification methods to distinguish 20 ILD patients from 8 healthy controls. The system showed optimal performance, and the features were used to visually transform the sound signals into graphical displays that were easily discriminated by physicians. The choice of the acoustic features used for discrimination has been also characterised by a substantial lack of standardisation and has been often either arbitrary or based on empirical methods. The use of time domain-related features

(such as IDW, 2CD and LDW), crackles counts and their distribution per breath, and a series of frequency domain-related features (such as fundamental frequency, peak frequency, zero-cross rate) varied from study to study. Finally, the information used to label the recordings in the data sets used to train the system (which is the basis of supervised machine learning) usually consisted in the historical diagnosis of ILD, whilst automatic classification systems have never relied on the characterisation of lung sounds via HRCT, which is the key diagnostic test for fibrotic ILD.

A recent study correlated the subjective assessment of bilateral Velcro-type” crackles with radiologic and clinical characteristics in a cohort of patients with clinical suspicion of ILD, and demonstrated a significant association with the presence of UIP at HRCT (Sellarés et al., 2016). This study did not make use of computerised methods for analysis, however it represents the first attempt to characterise “Velcro-type” crackles upon radiologic diagnosis. However, the design of the study had several limitations, for example the assessment of lung sounds was performed as part of the routine examination of the patients, as such it was unblinded. Furthermore, the direct relationship between the presence of “Velcro-type” crackles and the severity of specific fibrotic abnormalities in different areas of pulmonary parenchyma was not assessed.

A proper characterisation of audible “Velcro-type” crackles through HRCT imaging may inform the presence and extent of specific fibrotic abnormalities and patterns, therefore supporting their further investigation as an early identification and management tool in ILD. Such approach may also further clarify the mechanisms behind the generation of crackles and provide guidance for a more focused quantitative analysis of crackles. Generally, crackles are assumed to result from the acoustic energy produced by a change in elastic stress after a sudden opening or closing of distal airways – this according to the stress-relaxation quadrupoles hypothesis developed by Fredberg in 1983 (Forgacs, 1967,Fredberg JJ, 1983). Computer-aided lung sounds analysis (CALSA) techniques demonstrated that crackles in ILD have distinctive features as compared to crackles being generated in other disorders such as chronic heart failure and pneumonia, possibly reflecting the fact that the pathobiological mechanisms involved may be partly different (Vyshedskiy et al., 2005). The use of a multichannel analyser, capable of recording and analysing sounds simultaneously from several sites on the patient’s chest, allowed demonstrate that crackles occurring within few milliseconds over different regions are likely to represent the same event, triggered by the crackle with the highest amplitude (*mother* crackle) and transmitted in the form of smaller crackles (*daughter* crackles) over the chest. This phenomenon (that was given the name of “family of crackles”) enabled the definition of a crackle transmission coefficient (CTC), that represents the degree of transmission of sound from the sound source over different areas and was found to be inversely correlated with the mean frequency of the sound signal. Crackles in IPF patients were found to have higher frequency and to be transmitted over smaller areas than in patients with

chronic heart failure and pneumonia. The lower crackle transmission coefficient in IPF may be due to the higher frequency itself, or to the higher fluid concentration present in the airways of patients with other disorders.

In this study, when single HRCT sections were matched with the corresponding acoustic findings on auscultation of the chest, the evidence of pulmonary fibrosis was found to be significantly associated with “Velcro-type” crackles on electronic auscultation. This evidence served as a basis for the further correlation analysis with individual radiologic patterns of pulmonary fibrosis on HRCT. The most important finding was that the presence and severity of different patterns, representing both advanced (such as honeycombing) and milder interstitial abnormalities (such as reticulation and ground glass opacities) were found to be independently associated with the presence of “Velcro-type” crackles.

The finding that the extent of reticular opacities is independently and strongly associated with “Velcro-type” crackles suggests that any degree of abnormal deposition of fibrotic tissue might cause the collapse of distal airways, according to the original theory of crackles’ generation (Fredberg JJ, 1983). Honeycombing is a term used to indicate the appearance of destroyed lung parenchyma and late stage fibrosis, presenting as piled, thick-walled cysts with a predominant distribution in the subpleural regions of the lungs. Assuming that airflow is maintained in these regions, it is likely that these dilated, distorted airspaces collapse during expiration and reopen during inspiration, causing the generation of the crackles. Ground glass opacities are areas of increased radiologic attenuation with preservation of bronchial and vascular margins, which may arise from partial filling of airspaces, increased capillary volume or interstitial intralobular thickening due to fluids or fibrosis or a combination of these, the common factor being the partial displacement of air that results in a hazy increased opacity of lung, with preservation of bronchial and vascular margins in the lobule (Hansell et al., 2008b). The positive, independent association with “Velcro-type” crackles in this study strengthen the idea that even early stage interstitial involvement can be identified on chest auscultation. Traction bronchiectasis is a term identifying irregular bronchial dilatation caused by surrounding retractile pulmonary fibrosis (Hansell et al., 2008b). In this study, this pattern had the strongest correlation with crackles at the univariate analysis, but such association lost statistical significance when entered in the multifactorial model. Traction bronchiectasis may be considered as an advanced radiologic sign of pulmonary fibrosis and often accompany other signs such as honeycombing - actually, they can be seen as cysts or micro cysts in the lung periphery, making the distinction from honeycombing difficult - but the results of this study suggest that they are not responsible for the production of “Velcro-type” crackles, confirming that this acoustic finding is likely to result from something happening in the most distal airways. This evidence also reflects the different pathobiology of traction bronchiectasis as

compared to other types of bronchiectasis, e.g. those resulting from chronic infection or congenital bronchial abnormality, which are usually characterised by accumulation of inflammatory secretions causing production of coarse, "wet" crackles due to their mobilisation. Emphysematous alterations were the only radiologic feature showing a trend toward inverse correlation with crackles. Emphysema is characterised by destruction of alveolar walls resulting in a weaker transmission of lung sounds, either normal or adventitious. As such, it's not surprising that emphysema was not associated with audible "Velcro-type" crackles in this study. Conversely, coexisting emphysematous alterations may mask "Velcro-type" crackles, decreasing the sensitivity of such assessment. The potential interference of emphysema on "Velcro-type" crackles assessment has been addressed and is discussed below in section 4.5.5

The presence of bilateral "Velcro-type" crackles on chest auscultation represents a typical acoustic finding of patients with fibrotic ILD. While this assessment is endorsed by the current guidelines to prompt further investigation for ILD, their association with fibrotic patterns on HRCT has never been quantified. The results of the patient-based analysis showed that bilateral crackles correlate with the presence of fibrotic ILD on full volume HRCT scan, and most strongly predict the presence of a UIP pattern on HRCT. Conversely, unilateral "Velcro-type" crackles did not predict fibrotic ILD. The findings of this prospective study emphasise that the assessment of bilateral "Velcro-type" crackles should trigger further investigation via chest HRCT in patients presenting with chronic respiratory symptoms. In particular, the demonstration that bilateral "Velcro-type" crackles were associated with both possible and UIP patterns on HRCT, and most strongly with the former, might have relevance as to the incorporation of lung sounds in the diagnostic algorithm of IPF.

4.5.5 Accuracy of "Velcro-type" crackles assessment toward evidence of pulmonary fibrosis on HRCT

The ultimate goal of this analysis was to provide pilot, useful information as to the significance and limitations of the assessment of "Velcro-type" crackles on electronic auscultation operated by healthcare professionals. As such, this approach should not be regarded as an attempt to determine the real added diagnostic value of this assessment toward fibrotic ILD, an aim this research was clearly not designed for.

Generally, when the test results are dichotomous like in this case (absence or presence of "Velcro-type" crackles), four basic parameters are usually calculated in test research to estimate diagnostic accuracy:

1. Positive predictive value, or the probability of the presence of the disease in those with a positive test result

2. Negative predictive value, or the probability of absence of the disease in those with a negative test result
3. Sensitivity, or the probability of a positive test given presence of the disease (true positive rate)
4. Specificity, or probability of a negative test in those without the disease (the true negative rate)

Sensitivity and specificity are attributes of a specific test, as such they have the advantage of being independent of disease prevalence in a certain population (pre-test probability), but fail in reflecting decision making in clinical practice, based on the probability of a certain outcome given the results of a test (post-test probability). In stark contrast, positive and negative predictive values can be considered more relevant clinically, but are strictly dependent on the prevalence of the clinical outcome in the population examined. Diagnostic probabilities can also be expressed in terms of odds; the likelihood ratio (LR) of a positive test is the probability of a positive test in the subjects with the clinical outcome (sensitivity) divided by the probability of a negative test in the subjects who do not have the clinical outcome (1 – specificity or false positive rate), while the likelihood ratio of a negative test is the probability of a negative test in the diseased (1-sensitivity or false negative rate) divided by the probability of a negative test in the non-diseased (specificity). The higher the likelihood ratio for a positive test and the lower the likelihood ratio for a negative test, the more accurate the test will be. The advantage of using likelihood ratios is that they are more informative than the four parameters described above and overcome their respective limitations. In fact, using the Bayes’ theorem, the likelihood ratio can be used to easily estimate the post-test odds (or probability) of a clinical outcome:

$$Post\_test\ odds = Pre\_test\ odds \times LR$$

Nevertheless, the real clinical value of a test can be measured in the change it can produce between pre-test and post-test probability. As such, LR should not be used without information about the prevalence of the outcome; while a high LR can usually produce great changes between pre-test and post-test probability, greatest changes are achievable for moderate prevalences. For example, when a disease is extremely common or extremely rare in a population, it is unlikely that a test, no matter how accurate, will significantly modify the probability of a diagnosis, and its clinical usefulness will be lost. A nomogram can be used to perform such calculations easily (Fagan, 1975).

In the case of a tool proposed for screening or early identification of fibrotic ILD, primary importance should be placed on not missing any patient with pulmonary fibrosis at subsequent HRCT scan. In other terms, false positives (affecting the positive predictive value of a test) might be acceptable to a certain extent, although sending more patients to subsequent testing would

have repercussions on costs and radiation-related risks and should be always taken into account. On the other hand, false negatives (affecting the negative predictive value) would heavily affect the clinical value of a test meant for early identification of a disease. As such, an excellent negative predictive value is crucial together with high sensitivity and likelihood ratio for a positive test, to ensure good performance of the test across different populations of patients. If these requirements are met, in case of a negative result the absence of the disease can be confidently ruled out.

In this study, the assessment of “Velcro-type” crackles toward identification of pulmonary fibrosis produced high, although not excellent, negative predictive values, indicating that the absence of “Velcro-type” crackles over a specific region of the chest may exclude with good confidence the presence of fibrotic alterations on HRCT. Moreover, it might be expected that these values would be higher in a general population where prevalence of pulmonary fibrosis is low. Nevertheless, sensitivity was low, and even in the combined assessment (which should be more sensitive as compared to the isolated assessment of the physicians) the absence of “Velcro-type” crackles was reported in 39.8% of fibrotic images, as shown in Table 18. As compared to the assessment of single recordings, the presence of “Velcro-type” crackles bilaterally on chest auscultation demonstrated higher predictive values toward confirmation of fibrotic ILD on subsequent HRCT scan, reflecting the results of the regression analysis. Importantly, this assessment also showed higher sensitivity, indicating that the number of false negatives was considerably lower in such patient-based approach. Still, a non-negligible number (18.2%, almost a fifth) of patients with fibrotic ILD were “missed” by auscultation of “Velcro-type” crackles on chest auscultation.

Several factors might contribute to explain the overall poor sensitivity of “Velcro-type” crackles assessment in this study. The potential interference of emphysema on sound transmission was addressed in this study, however emphysematous abnormalities were not extensive overall in the imaging data set, and the possibility that these alterations could mask “Velcro-type” crackles was confidently excluded by comparing the emphysema scores between subgroups of fibrotic ILD patients with different acoustic findings. Body composition and weight can be also hypothesised to affect sound transmission in two ways. One is the potential interference operated by the increased thickness of the chest wall’s soft tissue. The other is represented by the reduction of pulmonary compliance which may limit the achievement of adequate pulmonary volumes and flow rates during inspiration. Nevertheless, whether and to what extent weight alters sound transmission is unknown, and body mass was not measured in this study either. The high proportion of false negatives could have alternative explanations that are related to the limitations of the methods used for electronic recording in this study. A post-hoc, unblinded assessment performed by the author identified that breath sounds had low intensity in several

recordings, likely due to superficial breathing. Low volumes and flow rates are known to affect sound transmission and might actually be responsible for a proportion of false negatives in this study, where breathing was not flow-standardised during lung sound recording. As for false positives, some type of background noise (e.g. skin friction produced by breathing movements) might have mimicked the presence of fine “Velcro-type” crackles, although these events did not seem to occur frequently in the data. It cannot be excluded that some images did not actually show evidence of fibrosis, but crackles were still audible for being generated in a close region of the lung parenchyma (e.g., above or below the site of recording). In both approaches (single images and patient-based), the overall accuracy of the assessment of “Velcro-type” crackles was very similar between the two physicians , however one physician produced less false negatives - thus reaching higher sensitivity and negative predictive values, while the other produced less false positives, reaching higher specificity and positive predictive values. It may be hypothesised that the latter had a more “cautious” approach toward the assessment of “Velcro-type” crackles. Interestingly, this physician was the one with some previous experience in the use of the digital stethoscope, and it is possible that this factor played a role in producing the discrepancies: although being a speculative hypothesis, the judgment of this physician might have been driven by artefacts in the recordings mimicking crackles. Notably, while the performance of either physician in the first data set was replicated in the second data set, a significant decrease in sensitivity was recorded in the third data set. The first two data sets of sounds were recorded from two consecutive cohorts enrolled in the same radiology unit, while the third data set was recorded from an independent cohort enrolled in a different centre. A possible explanation is that the radiology unit of the second centre was characterised by higher level of background noise, which might have made the assessment of the “Velcro-type” crackles in the recordings more challenging.

In conclusion, in this study the subjective assessment of “Velcro-type” crackles on electronic auscultation demonstrated good, but not optimal accuracy toward the identification of signs of pulmonary fibrosis in the lung parenchyma directly below the sites of auscultation. The overall low sensitivity and high specificity values seem to point out that the assessment of “Velcro-type” crackles might be a better “rule-in” than a “rule-out” test. The high positive predictive value of the assessment of bilateral “Velcro-type” crackles toward the presence of ILD on full volume HRCT scan seem to confirm these findings and indicate that bilateral “Velcro-type” crackles might be more useful for increasing diagnostic confidence than for screening purposes. For the time being though, this data should be considered purely informative and do not rule out the value of “Velcro-type” crackles in the first evaluation of a patient with chronic respiratory symptoms. First, this study was not designed to investigate the diagnostic yield of “Velcro-type” crackles toward

fibrotic ILD. Secondly, these findings depend on several factors, ranging from the limited number of physicians involved to the non-standardised quality of the recordings, that may have influenced the physicians’ assessment. As such, a deeper understanding of the relative contribution of these factors and the real yield of “Velcro-type” crackles assessment on electronic auscultation performed with a digital stethoscope should be sought through the comparison with standard auscultation and involving a larger number of clinicians with different levels of expertise.

4.5.6 Lung sounds recording: ideal quality requirements and practical advices

The post-hoc unblinded assessment of the acoustic data set underpinned the limitations of this study with regard to lung sounds recording. In order to implement electronic auscultation in clinical practice and facilitate further research the signal acquisition process should be improved and standardised, and some training in the use of digital stethoscope is advised for those physicians approaching the tool for the first time. Evidence-based recommendations for environmental and subject conditions and breathing manoeuvres for short- and long-term respiratory sound recordings have been made in the guidelines for computerised respiratory sounds analysis (Sovijärvi et al., 2000). Indeed, optimal techniques for monitoring respiration are still uncertain, and many of the proposed measures were intended to optimise lung sounds recording in controlled environments for research purposes; however, a few suggestions can be made based on such recommendations to enhance the quality of the recordings in the clinical practice without affecting applicability.

Generally, deep breathing with slow vital capacity manoeuvres is considered ideal for recording lung sounds, as it increases the sensitivity of the appearance of adventitious lung sounds. This has particular relevance in ILD where crackles are strongly related to lung volume (Kaisla et al., 1991), and can sometimes appear at the end of deep inspiration (al Jarad et al., 1993). The flow rate has also an influence on respiratory sounds amplitude and the frequency spectrum of the sound signal. Since the route of breathing has an effect on overall intensity of breath sounds, mouth breathing is preferred over nose breathing since higher levels of flow can be obtained. In phonopneumography, pulmonary sounds are recorded and analysed using a pneumotachograph signal that defines phase of respiration and airflow velocity, analysis of the signal intensity, and correlation with factors influencing sound intensity (Banaszak et al., 1973). The basic recommended standard is to record sounds during tidal breathing with a volume of 1.0 L or 15–20% of predicted vital capacity, and with an expiratory and inspiratory peak flow of 1.0–1.5 L/s or 10–15% of the predicted maximum peak expiratory flow for 7–10 respiratory cycles (Sovijärvi et al., 2000). When recording lung sounds, the respiratory effort should be therefore monitored using displays of flow-volume loops or via flow-targeting systems. A significant amount of

background noise can be produced either by environmental conditions or non-respiratory sounds such as chest motion, respiratory muscle sound or skin friction. The best way to eliminate ambient noise would be the use of a soundproof room or, alternatively, a body plethysmograph. While such ideal conditions represent a requirement to ensure quality of the quantitative study of lung sounds and can be reproduced in physiology testing laboratories, the mentioned devices and techniques are not easily applicable in most clinical settings. Moreover, it might prove to be challenging to achieve proper tidal volumes in ILD patients with severely reduced pulmonary compliance, and these manoeuvres may also exacerbate cough thus making recording of good quality sound signals very difficult. As such, a good compromise to improve standardisation of electronic auscultation in the clinical practice without affecting applicability would be to provide some guidance to physicians approaching the digital stethoscope. Recommendations should be made to instruct patients to breath deeply through the mouth with slow vital capacity manoeuvres, as long as they are comfortable in order to avoid include cough in the recordings. Clinicians should be also recommended to keep the stethoscope’s diaphragm as still as possible to reduce skin friction and therefore the generation of possible artefacts in the recordings. As long as it is feasible, the importance of choosing a room that is free from transient noises and having devices and computers as silent as possible should be stressed. These simple expedients might be effective in providing good quality recordings to be included in the patients’ electronic medical records, thus maximising the yield and usefulness of electronic auscultation in the clinical practice.

4.6 Conclusions and future work

4.6.1 Main conclusions

This study demonstrated for the first time that “Velcro-type” crackles, subjectively assessed on electronic auscultation by expert physicians, correlate with the presence and extent of distinct radiologic abnormalities and patterns of pulmonary fibrosis on HRCT of the chest. Both signs of advanced and less severe interstitial involvement were independently associated with “Velcro-type” crackles, suggesting that they may be an early acoustic finding in patients with fibrotic ILD. Extending this finding, the presence of “Velcro-type” crackles was highly predictive of fibrotic ILD and specifically UIP on full volume HRCT, supporting the international consensus guideline recommendation to consider IPF in patients presenting with bibasilar crackles. This study also demonstrated the good reproducibility of the assessment of bilateral “Velcro-type” crackles on electronic auscultation, and provided insights into the potential limitations of such approach and how to address them.

These results were obtained through the recruitment of a large population of patients with a relatively high prevalence of ILD in a secondary care setting. Sounds were recorded using a simple, point-of-care tool easily applicable in clinical practice. An unbiased review of the imaging and acoustic data sets was achieved through an independent, blinded assessment performed by thoracic radiologists and chest physicians, respectively.

In conclusion, this study provides grounds for further investigation of "Velcro-type" crackles as a novel tool for assisting the diagnostic process of fibrotic ILD and their early identification. Furthermore, this research supports the implementation of lung sounds recording as an integrative part of the electronic medical record in the clinical practice.

4.6.2 Summary of study limitations

While this research established for the first time the direct relationships existing between lung sounds and distinct radiologic findings on HRCT, it did not demonstrate the real diagnostic value of "Velcro-type" crackles assessment toward confirmation of fibrotic ILD, which limits the clinical usefulness of the findings. This is mainly due to the characteristics of the study design itself: the cross-sectional approach followed for data analysis is ideal for diagnostic research, where the prevalence of the outcome disease is estimated, and not its future occurrence. The inclusion of patients referred to HRCT scan of the chest for any clinical indication was however not the best fit to determine the diagnostic accuracy of an assessment. The ideal population for a diagnostic study in this field should include all consecutive patients presenting with signs and symptoms suggestive of fibrotic ILD, such as chronic exertion dyspnea and cough; these were not systematically investigated in this study, where the collection of clinical data was limited due to the little time available. The referral to chest HRCT – a key diagnostic and monitoring test for lung parenchyma diseases – may be therefore considered only a surrogate for the identification of a population at risk to develop ILD. Moreover, the recruitment of the study population occurred in the radiology unit of two centres specialised in ILD; in such secondary (or tertiary) care settings the prevalence of ILD is indeed higher than in the population referring to primary care services, which would be instead the first-line choice to estimate the real value of a proposed tool for early detection of ILD. On the other hand, the relatively high prevalence of ILD cases ensured by the inclusion criteria and the study setting was very convenient to fulfil the main objective of the study, as it provided statistical power to the correlation analyses.

As for the methods used for lung sound recording, the main limitations consisted in the lack of simultaneous measurement of tidal volumes and airflow rates, which would have allowed standardisation of breathing efforts and, therefore, lung sounds' intensity. Moreover, lung sounds

were recorded in a clinical setting uncontrolled for the presence of background noise. These factors might have influenced the good transmission of lung sounds and therefore affected the overall quality of recording.

In the primary analysis of the data, a “two-dimensional” approach was followed, with lung sounds being paired to the radiologic abnormalities present in the lung parenchyma below the sites of auscultation. Such approach allowed establish a direct association between sounds and imaging features, but could not take into account the real, three-dimensional transmission of sounds, which can be perceived as a limitation. Nevertheless, defining what is the ideal correspondence between alterations in the different areas of the lung parenchyma and sounds heard over the chest is indeed challenging and has been largely unknown. The design used in this study appeared to be the most reasonable, building on the evidence that the transmission of “Velcro-type” crackles across different areas of the chest is more limited than in other conditions such as chronic heart failure or pneumonia (Vyshedskiy et al., 2005). As discussed, potential effect modifiers of lung sounds transmission such as measurements of body composition were not recorded, and it remains therefore unknown whether and how these may have impacted the strength of the associations found between sounds and fibrotic alterations.

4.6.3 Diagnosing fibrotic ILD using lung sounds: future steps

The studies conducted involving methods of computerised analysis and testing the accuracy of automatic classification systems toward diagnostic confirmation of ILD had several limitations in terms of concept and design, which hampered the applicability of these systems in clinical practice. They often enrolled small, convenience sample populations (e.g., selected ILD patients and healthy controls); the reference standard was often represented by historical diagnosis of ILD and not by the actual evidence of fibrotic patterns on HRCT; the outputs of the analysis were often represented by isolated indices of performance of different systems, with the lack of a multivariable approach taking into account other clinical variables; finally, in most cases the tools used for signal acquisition and analysis were far from being easily applicable in everyday clinical practice.

Building on the robust design and the findings of the study described in this chapter, a practical proposal for further, sequential steps into diagnostic research aimed to demonstrate the validity of lung sounds as a new diagnostic tool in fibrotic ILD, is presented below.

Diagnostic research should be always performed in close adherence to daily clinical practice to ensure wide applicability of the findings. Thus, the typical features of a diagnostic process should be taken into account in the design of a study, with important consequences for the choice of the

study population, the diagnostic determinants and the data analysis. Generally, in diagnostic research the ability of a single test or combination of tests to distinguish between diseased and non-diseased should be studied in subjects who are representatives of the clinically relevant domain of patients suspected of having that disease – as such, with symptoms or signs that would bring the physician to consider diagnostic testing. Consequently, patients in whom the presence of the disease has already been established or in whom the probability of the disease is considered high enough to initiate adequate therapeutic actions fall outside the domain, similar to when the probability is deemed sufficiently low to exclude the diagnosis. Consequently, the ideal study population for evaluating the accuracy of “Velcro-type” crackles toward diagnostic confirmation of fibrotic ILD should be restricted to consecutive patients with signs or symptoms such as chronic dyspnea and cough (identified using a short, simple questionnaire) possibly with exposure to relevant risk factors (e.g., cigarette smoking) and with older age (e.g., > 60 years). On the other hand, exclusion criteria should be few to ensure wide applicability of the findings. Diagnostic accuracy of tests can vary across care settings due to differences in the distribution of severity of the disease; hence, primary care should be ideally preferred over secondary or tertiary care for this investigation, since the findings would fit the setting where the diagnostic work up of these patients actually starts. Any potentially relevant diagnostic determinant as well as the “true” presence or absence of the target disease should be measured in all patients following a cross-sectional approach. Although the results of the proposed diagnostic test should be ideally obtained at the same time with the outcome of the diagnostic reference standard, sometimes it can take several days or weeks before the definitive diagnosis becomes known. This is particularly true in the case of ILD, where it might take some time before a HRCT is performed and a final diagnosis is made. The use of a digital stethoscope to record lung sounds is advisable. First, it would allow the independent, blinded subjective assessment performed by healthcare professionals with different levels of expertise – in particular, the assessment operated by different general practitioners would optimise the applicability of the study findings. Second, this approach would allow quantitative analysis could be performed using computerised methods. Standard auscultation should be also performed in addition to determine the reproducibility and the validity of this assessment as compared to the assessment of digital recording. Ideally, the lung sounds should be directly recorded by general practitioners: this way, the collected data would resemble the quality of information obtainable in daily clinical practice, and the feasibility of digital recording in this setting could be evaluated by defining the training needs for the implementation of such technique. Indeed, such design would require the commitment of general practitioners, the allocation of significant time for the training and likely higher study costs. The outpatient clinic of a respiratory medicine department might be a more practical, convenient setting to conduct such investigation. Moreover, the subsequent diagnostic information would be

easier to retrieve, especially if patients perform HRCT scan at the same centre. On the other hand, as mentioned above the risk of bias would be higher due to the higher prevalence of ILD in the study population base.

The goal of determining a diagnosis is to estimate the probability of disease given the results of a certain diagnostic test. Such probability should be high (or low) enough to confirm (or exclude) with adequate confidence a certain diagnosis. To determine the validity of a new diagnostic tool, the outputs of the test under scrutiny must be compared with the results of the reference standard test, which provides, at the time of study initiation, the best estimate of the presence or absence of a disease. The reference standard to use in this case should be the HRCT scan of the chest, assessed independently by thoracic experts blinded to the results of the assessment of lung sounds or other clinical information to avoid biases.

The first analysis to perform would be aimed to assess whether the subjective assessment of “Velcro-type” crackles can confidently predict, as a single test, the presence or absence of fibrotic ILD on subsequent HRCT scan. As discussed above, in the case of a test with dichotomous output (absence or presence of “Velcro-type” crackles), the diagnostic value can be determined by assessing various parameters of accuracy. For an early detection tool, an optimal negative predictive value (ideally >90%) would be required to ensure that when “Velcro-type” crackles are not detected it would be very unlikely to find fibrotic alterations on subsequent HRCT. Since predictive values are influenced by the prevalence of pulmonary fibrosis in the study population, high sensitivity is also desirable to confirm the validity of such assessment across different prevalences. However, while the use of sensitivity, specificity and predictive values individually may be misleading as to the value of a diagnostic test, a likelihood ratio for a positive test >10 usually guarantees large changes from pre-test to post-test probabilities of disease presence (Jaeschke et al., 1994). Indeed, the performance might vary significantly across different levels of expertise and the level of agreement between such categories (e.g., general practitioners and respiratory physicians with different levels of experience in ILD) should be sought to provide further information as to the generalisability of the findings and the value of such approach across different settings of care.

The use of electronic recording of lung sounds in this research would allow objective, quantitative measurements, and the diagnostic yield of supervised classification systems in a clinical setting could be assessed. Such approach would overcome the issue of subjectivity, and the performance might prove to be equal or superior than physicians. As discussed in Chapter 2, several studies attempted to define the diagnostic accuracy of algorithms for automatic classification developed using machine learning techniques. Such techniques are usually supervised, meaning that they a

pre-specified portion of the data set is used to train the system in discriminating between subjects with and without the disease. In the field of lung sounds analysis this is achieved by providing to the system recordings labelled with the corresponding outcome provided by the reference standard test. The system is then provided with the remaining, unlabelled portion of the data set and the overall performance is calculated based on its ability to correctly discriminate between ILD and non-ILD patients. When the test results are on a continuous scale or the test provides a probability of having the disease, these may be used as cut-off points and dichotomised into positive or negative on a receiver-operator characteristics (ROC) curve, which will provide indices of accuracy for the different threshold values and, most importantly, an estimation of the overall discriminatory power of the test under scrutiny, expressed as the area under the curve. Whilst the reference standard used in most studies so far usually consisted in the historical diagnosis of ILD or IPF, the radiologic review of patterns and features at HRCT would be much more robust, unbiased, and accurate. Notably, the finding that “Velcro-type” crackles are independently associated with individual interstitial abnormalities provides further insights into the opportunity of using the presence of milder fibrotic alterations (instead of patterns of established fibrosis such as UIP) to label the ground truth, which might eventually improve the accuracy of automated systems toward detection of early ILD.

Since every diagnostic process is multivariable in nature, the clinically more relevant research aim would actually be to assess whether the detection of “Velcro-type” crackles on electronic auscultation appreciably adds to the diagnostic information that is readily available in clinical care such as demographics, symptoms and other signs. As such, along with the definition of the test properties described above, multivariable logistic regression would be the ideal strategy for data analysis. This approach would enable determine which diagnostic determinants contribute to the estimation of the probability of the disease presence, to quantify their relative accuracy or weight in such estimate, and, most importantly, to develop and eventually validate a diagnostic model or rule to estimate the probability of disease based on the combination of test results in individual patients in clinical practice. In order to determine the added value of a new tool or test, the diagnostic performance (i.e., the ability to discriminate between patients with or without ILD) of a model including all diagnostic information available to the physicians including the assessment of “Velcro-type” crackles should be compared to the performance of a model without the lung sounds assessment, either calculating the area under the relative ROC curves or via metrics such as the c-index, which numerically expresses the discriminatory power of models. Despite no straightforward methods exist to estimate the required patient number for such a study, several authors have stipulated that in multivariable prediction research for each diagnostic determinant studied at least 10 subjects positive for the outcome variable are needed to allow statistical

modelling (Harrell et al., 1996), otherwise the analysis tends to overestimate the accuracy of the diagnostic strategy or model. A final but crucial step to take before a model can be used in practice with confidence consists in the external validation of the model, performed by application of the model in new patients. This could be performed either in population of patients enrolled in other centres (ideally, in other countries in the context of a multi-centre international study), or across different care settings: a practical approach could be to develop a diagnostic model in a secondary care setting and to perform validation by applying the model in patients suspected for ILD accessing primary care services, to see whether its performance is retained. The suggested multivariable process gains even more relevance if computerised methods are involved: the algorithms used for sound classification – which are based on the same basic criteria of statistical modelling - might be designed to incorporate the demographic and clinical information as well, providing the opportunity to build a comprehensive, self-learning tool for diagnosis of ILD.

Chapter 5: Longitudinal assessment of lung sounds in IPF

5.1 Introduction

This chapter will describe a pilot study aimed to characterise the longitudinal changes occurring in the acoustic features of lung sounds recorded digitally from a cohort of patients with IPF. First, the study rationale will be presented in section 5.2. Section 5.3 will describe the methods of the research including the aims, the procedures performed to address the proposed objectives, the main outcomes and the methods used for data analysis. The main findings will be presented in section 5.4 and discussed in section 5.5.

5.2 Rationale

5.2.1 The need for an accurate prediction of disease progression in IPF

After setting a confident diagnosis, prognostication is the next challenge every physician has to face. An accurate prediction of risk of disease progression and death in progressive ILD such as IPF has enormous clinical relevance. Firstly, it enables patients to anticipate the future and make informed plans. Secondly, it forms the point of departure for all subsequent aspects of patient management, ranging from optimisation of timing for treatment and supportive care to referral to lung transplantation. Thirdly, as discussed in section 2.5.7.5, the validation of valuable predictive tools would allow better disease stratification, thus enhancing the efficiency of randomised clinical trials via enrichment of study cohorts with populations at higher risk of bad outcomes, or more likely to respond to specific experimental therapies.

As described in the introduction of this thesis, several baseline measures of disease severity have been extensively used to predict survival in IPF, including clinical features, pulmonary function measures, computed tomography findings and composite scoring indices. Dynamic surrogates of disease progression, such as longitudinal changes in physiological indices (e.g. decline in FVC and DLco over 6–12 months) have also been validated and seem to be superior in predicting mortality (du Bois et al., 2011a,Jegal et al., 2005,Richeldi et al., 2012). Nevertheless, prognostication in IPF is still very challenging due to the unravelled heterogeneity of the disease and the significant intra-patient variability of the disease natural course, which is known to be non-linear and unrelated to patients' previous disease behaviour (Ley et al., 2011) – in fact, none of the tools proposed so far

has been successfully translated into clinical practice for prediction of disease progression and mortality in the individual patient.

Physiology parameters such as FVC have the advantages of being simple to measure, highly reproducible, and represent the most validated tools to assess disease severity and monitor disease progression in both clinical practice and randomised trials, where they have been extensively used as important endpoints of efficacy. On the other hand, they are hampered by significant intra-individual variability and may be not sensitive enough to detect subclinical progression of fibrosis. Furthermore, these parameters are prone to the influence of conditions such as coexisting emphysema and pulmonary hypertension. With regard to emphysema in particular, the increased compliance due to emphysematous alterations compensates for the fibrosis-related stiffness and tend to normalise lung volumes, while the effect on diffusion is decreased by both conditions. Several retrospective studies demonstrated that patients with coexisting fibrosis and emphysema have higher need for long-term oxygen therapy and occurrence of pulmonary hypertension, and suggested that the decline in FVC might be reduced by this compensatory effect (Jankowich et al., 2010,Jegal et al., 2005,Kurashima et al., 2010). The effect of emphysema on the decline of FVC over time has been quantified in a recent study: the longitudinal change in FVC is significantly decreased with an emphysema extent of ≥ 15 % on HRCT (Cottin et al., 2017), implying that in these patients the functional decline may not be a reliable measure of disease progression, and would not be suitable as a trial endpoint if the study population includes subjects with ≥ 15 % emphysema on HRCT.

The use of radiologic indices for risk prediction in IPF is indeed appealing, as appreciable changes on serial HRCT scans directly reflect the progression of the fibrotic process (Best et al., 2008,Lynch et al., 2005,Oda et al., 2014a). So far however most studies conducted to determine the prognostic value of such indices are retrospective, with short-term follow ups, and the visual assessment of radiologic abnormalities was often based on techniques with large gaps in imaging (i.e. non-contiguous scanning), definitely not as accurate as it would be with full volume HRCT scans. Another hurdle consists in the limited applicability of such approach, which is time-consuming and requires the commitment of expert thoracic radiologists; moreover, the visual adjudication of HRCT findings is affected by substantial intra- and inter-observer variability even among expert observers, and it can be argued that even the radiologists with the highest level of expertise in ILD might be able to identify subtle, short-term changes in the fibrotic alterations. In order to obviate these limitations, computerised systems allowing objective, quantitative measurement of the extent of fibrosis and individual radiologic abnormalities on HRCT have been recently developed. Indeed, such systems might help the radiologist in the precise and quick identification of areas presenting subtle changes over short periods of time, which would be very

useful especially considering the high number of images provided by HRCT scan performed with spiral techniques. Measures of CT attenuation, obtained from CT density histograms (such as mean lung attenuation, skewness and kurtosis) demonstrated to change over time, and were validated against other measurements of disease progression such as severity of dyspnea and physiologic impairment (Best et al., 2003,Hartley et al., 1994). However, CT histogram-based measures highly depend upon the level of inspiration achieved during the scan, and provide only global measures of fibrotic alteration without measuring the type and extent of individual features. Most importantly, the histogram computer analysis did not add significant prognostic information to the subjective visual assessment of the extent of fibrosis (Best et al., 2008), limiting its clinical utility. Recently, advanced and more promising techniques of computer-aided texture-based imaging analysis, able to discriminate and quantify individual parenchymal abnormalities, have been developed (Yoon et al., 2013,Zavaletta et al., 2007). A retrospective study using the software CALIPER (Computer-Aided Lung Informatics for Pathology Evaluation and Rating) demonstrated that on consecutive HRCT scans short-term changes in the volume of reticular opacities, other than in absolute and percentage volume of fibrosis, were predictive of survival in IPF patients, providing substantial ground for the validation of quantitative assessment of specific radiologic abnormalities for prognostication in IPF (Maldonado et al., 2014b). The validity of such technique was proven in a recent study demonstrating that CALIPER-based estimates of extent of fibrosis in IPF have stronger correlations with physiology measurements than semi-quantitative visual CT scoring. Interestingly, a novel CALIPER-based parameter, the pulmonary vessel volume (PVV), showed strong correlation with the functional parameters, and was proposed as as new HRCT feature for the assessment and monitoring of IPF patients (Jacob et al., 2016). A further, retrospective study from the same group of researchers compared quantification by CALIPER, visual CT scoring and physiology measurements against survival in 283 IPF patients, and demonstrated that CALIPER-based parameters are more accurate toward prediction of mortality than traditional visual CT scores (Jacob et al., 2017). Importantly, two prediction models, including two CALIPER-derived independent predictors (PVV and honeycombing) with or without the Composite Physiology Index (Wells et al., 2003), demonstrated better calibration and similar prognostic accuracy as compared to the GAP index staging system. The very latest developments in quantitative imaging include the study of lung mechanics by dynamic CT, the measurement of gas transfer by magnetic resonance spectroscopy and measurement of tissue biomarkers by positron emission tomography, which represent some new promising approaches for monitoring disease progression and allowing disease stratification (O'Riordan et al., 2015).

Despite several risk prediction algorithms including clinical, physiologic and radiologic parameters have also been developed and validated through fine, robust statistical modelling (du Bois et al.,

2011b,Ley et al., 2014,Ley et al., 2012,Wells et al., 2003), their utility in individual patients has remained elusive so far. In most cases, the validation of such models has been sought using retrospective data from large randomised trials in IPF enrolling patients with mild to moderate disease, who are not fully representative of the general IPF population. Whilst changes in several variables may be reasonably reliable on average, in individual patients may be largely affected by the intra-patient variability of the disease course and by comorbidities, which are under-represented in clinical trials. On the other hand, in clinical practice prognostic evaluation might take into account parameters that in a mild to moderate IPF trial population are not found useful for prediction purposes. Finally, whilst mortality prediction has been thoroughly investigated in IPF, much less is known about prediction of future disease progression expressed as decline in FVC or other meaningful outcomes of interest, such as occurrence of acute exacerbations or significant worsening of quality of life, that might guide interventions such as supportive and palliative care. Such approach would be arguably more helpful for individual patients' counselling and short-term management, and for powering trials using endpoints of disease progression. A retrospective analysis has demonstrated that prior trends in FVC are poor predictors of subsequent decline in pulmonary function (Schmidt et al., 2014). This evidence has been further expanded by a recent study using basal and longitudinal measurements of several clinical variables from pooled data of placebo groups of large clinical trials to develop and test prediction models for disease progression in IPF, which ultimately failed in predicting functional worsening at 6 months. (Ley et al., 2016).

In conclusion, further, properly designed prognostic research is warranted in IPF to develop and validate novel prediction tools, which should be reliable, simple to measure, and valuable both in clinical practice for predicting disease progression and in interventional research as endpoints of response to treatment.

5.2.2 "Velcro-type" crackles: a valuable prognostic tool of fibrotic ILD?

In clinical practice, lung sounds represent a reliable piece of information for physicians when monitoring respiratory disorders. Chest auscultation is routinely used for the short-term follow up of patients with pathological conditions such as pneumonia or congestive heart failure, as it allows appreciate acoustic changes in a simple, non-invasive and cost-effective way. In the lon-term follow up of progressive ILD such as IPF, the time interval between consecutive clinical appointments may vary from a few weeks to several months, roughly based on the severity of the disease and the previous disease behaviour. Indeed, it is impossible for physicians to recall the properties of sounds heard on auscultation at the previous visit, and therefore to appreciate any change occurring within this timeframe, such as the spreading of "Velcro-type" crackles across

different lung regions. Consequently, the value of standard chest auscultation for monitoring of disease in ILD has been historically ruled out.

Nowadays, digital stethoscopes offer the opportunity to record breathing sounds and include them into the electronic medical record of patients. While collecting lung sounds for future reference is of immediate, clear clinical utility, methods of computerised analysis of lung sounds can provide further, objective information as to the serial changes occurring in the acoustic features of lung sounds, similarly to the systems for automated quantification of radiologic abnormalities on CT scan. Notably, as compared to CT techniques, lung sounds recording is easier to perform, totally non-invasive, and cheaper. In the first study described in this research thesis, "Velcro-type" crackles have been found to predict the presence of distinct fibrotic abnormalities in the lung parenchyma underneath. Building on such evidence, it can be hypothesised that the quantification of the changes in acoustic features of "Velcro-type" crackles may represent a valid surrogate marker for quantitative changes of fibrosis extent on HRCT. So far, the evidence on the role of lung sounds as a clinical outcome in respiratory disorders is limited to one single study which demonstrated substantial changes before and after intervention (respiratory physiotherapy aimed to clearance of secretions) in time domain-related features of crackles in patients with bronchiectasis and cystic fibrosis (Marques et al., 2013). The longitudinal quantitative evaluation of acoustic features in patients with progressive ILD has never been performed so far. Such assessment, together with the correlation with validated markers of disease progression would provide pilot evidence on the potential role of lung sounds for informing disease management in these patients.

5.3 Methods

5.3.1 Research objectives

5.3.1.1 Primary objective

To assess the longitudinal changes occurring in acoustic features of lung sounds serially recorded in a cohort of IPF patients and explore their potential as a marker of disease progression through correlation with validated clinical and functional parameters.

5.3.1.2 Secondary objectives

1. To assess the reproducibility of the measurement of acoustic features of lung sounds in IPF recorded using a digital stethoscope in a clinical setting.
2. To compare acoustic features of lung sounds longitudinally measured in a cohort of IPF patients with those measured in healthy controls.

5.3.2 Research procedures

5.3.2.1 Research design

This is a pilot, single-centre cohort study with a longitudinal design, with lung sounds measured at different time points over 12 months of clinical observation together with other clinical, physiology and radiologic parameters.

5.3.2.2 Ethics and governance

The study was conducted subject to National Research Ethics Committee (NREC) approval obtained on 22nd January 2015, including provisions of Site Specific Assessment (SSA), and local Research and Development (R&D) approval obtained on 13th March 2015 (see Appendix A for documentation regarding ethics). The project was conducted in accordance with the recommendations for physicians involved in research on human participants adopted by the 18th World Medical Assembly, Helsinki 1964 as revised and recognised by governing laws and EU Directives; and the principles of Good Clinical Practice and in accordance with all applicable regulatory requirements including but not limited to the Research Governance Framework and the Medicines for Human Use (Clinical Trial) Regulations 2004, as amended in 2006 and any subsequent amendments. The project was granted funding from the NIHR Respiratory Biomedical Research Unit of Southampton.

5.3.2.3 Setting, timeline and selection of population

Recruitment of subjects and data collection took place at the WellCome Trust facility of the Southampton Centre for Biomedical Research (SCBR) at the Southampton General Hospital, UK. Patients with a diagnosis of IPF according to the most recent international guidelines (Raghu et al., 2011) were enrolled in the study between March and September 2015 and were followed up for 12 months, with last visit of last patient occurring on 5th July 2016. IPF is invariably progressive despite treatment, with a compelling need for accurate prognostication: hence, the choice of including patients with an established diagnosis of IPF. The time of clinical observation was decided upon the average duration of most large randomised trials in IPF, where the placebo and treated arms showed a significant functional decline over 52 weeks (King et al., 2014b,Noble et al., 2011b,Richeldi et al., 2011,Richeldi et al., 2014).

5.3.2.4 Recruitment strategies

The recruitment strategies were initially discussed with a recruitment officer at the NIHR Respiratory Biomedical Research Unit of Southampton, a person with experience in connecting with general practitioners and patients to spread awareness about clinical trials and research studies. After further consultation with members of the Research Design Service (RDS) South Central, patient and public involvement strategies were used to assess the sustainability of the study. Two patients with IPF referred to the ILD clinical service at the Southampton General Hospital and who previously agreed to be contacted for research activities were contacted via phone call by the author/researcher to discuss the overall acceptability and feasibility of the study design. Basic information about the study was sent in advance and some days were left for them to read and make their own considerations. Since IPF is a rare disease, it is not unusual that patients with IPF travel from a distance to get to a tertiary specialised centre: as such, the two patients were selected for living outside Southampton. One had been diagnosed with IPF about one year prior to the contact, while the other was more recently diagnosed (one month). Understandably, the latter seemed to be less informed about the nature and the course of the disease. Both patients expressed willingness to participate to the study, despite being aware they would not derive any direct benefit. They both gave positive feedbacks regarding the number and the duration of the visits, and the overall sustainability of the assessments. One patient expressed concern about the possibility not to live long enough to complete the study.

Outpatients with IPF referred to the ILD Clinic at Southampton General Hospital were invited to participate to the study. The author/researcher liaised with the respiratory consultants running the ILD service who kindly agreed to inform patients about the research. When consent was given, patients were contacted by the research staff (the researcher or a RBRU research nurse)

and were provided with an information sheet and about a week was left for the candidates to read it carefully. Then, they were invited to a screening visit. During this visit, the researcher discussed with the patients the design of the study, the assessments, the risks and the benefits, and answered their questions. At the end of the discussion, the patients willing to take part signed an informed consent. In order to speed up the recruitment process, informative posters and leaflets providing contacts for the study (phone numbers and e-mail addresses of the researcher and research nurses) were used.

5.3.2.5 Population sample and eligibility criteria

This study was designed as a pilot, single centre study, as such no sample size calculations were made. Before the beginning of the study, the expected recruitment rate was discussed with the respiratory consultants of the ILD clinic of the Southampton General Hospital. The time allocated for recruitment (approximately five months) was defined according to such expectations and taking into account the time required for study completion (two years). The inclusion criteria consisted of a diagnosis of IPF made according to the 2011 IPF international guidelines (Raghu et al., 2011) and age between 40 and 90 years, as IPF is unlikely to occur before 40 and an older age than 90 is usually accompanied by significant comorbidities or physical impairment that could affect participation. The exclusion criteria were any of the following: a diagnosis of clinically significant pulmonary disease other than IPF (e.g. asthma, COPD, bronchiectasis, history of lung cancer with history of radiation or of lung resection), history of New York Heart Association (NYHA) class III or IV heart failure, history of uncontrolled pulmonary hypertension (mean pulmonary arterial pressure > 25 mmHg), and a history of frequent chest infections (more than two episodes within six months prior screening) as it was felt that all these could affect the transmission of lung sounds related to the fibrotic disease. Patients with a Residual Volume (RV)/Total Lung Capacity (TLC) ratio > 100 measured at baseline via plethysmography were excluded since emphysema might affect the transmission of lung sounds. Also, in order to exclude a significant pulmonary obstructive disease, the FVC/FEV_1 ratio had to be >70 at baseline. 19 patients were screened and were all enrolled into the study – as such, there were no screen failures.

10 healthy volunteers were also enrolled in the study. These were recruited by the research staff via phone calls made to healthy subjects who previously agreed to be contacted for research, or through study advertisement via posters distributed in the hospital. The inclusion of this group was useful to discriminate between longitudinal physiological modifications in the acoustic features of lung sounds and the changes due to the progression of the fibrotic disease in IPF patients. Moreover, these acoustic data represent a valuable database of healthy respiratory

sounds for future research. These subjects underwent the same visits schedule as IPF patients. The eligibility criteria were age between 40 and 90 years and the absence of ILD or other clinically significant pulmonary and heart conditions (e.g. asthma, COPD, bronchiectasis, history of lung cancer with history of radiation or of lung resection, pulmonary hypertension, congestive heart failure, recurrent chest infections) that could affect the transmission of physiological respiratory sounds. Apart from recording lung sounds from these healthy volunteers, spirometry and a Diffusion Capacity for Carbon Monoxide (DL_{CO}) test were also performed at each visit to check that pulmonary function and gas exchange were normal.

5.3.2.6 Research protocol

The candidates (19 IPF patients and 10 healthy controls) contacted through the ILD clinic and willing to take part to the study were invited to attend a screening/baseline visit. Written informed consent (see Appendix A.3) was obtained before proceeding to any data collection. The subjects enrolled in the study were followed up for approximately 12 months. Each participant attended visits occurring approximately every two months for a total of 7 visits. Such timing allowed maximise the number of observations without affecting the feasibility of the study and the burden for patients (explored prior to the study beginning via interviews of patients' representatives). All the study visits were performed in a consulting room in the WellCome Trust research facility of the Southampton Centre for Biomedical Research. A graphical flowchart of study timeline is shown below in Figure 18, while a detailed schedule of the assessments is reported in Table 33.

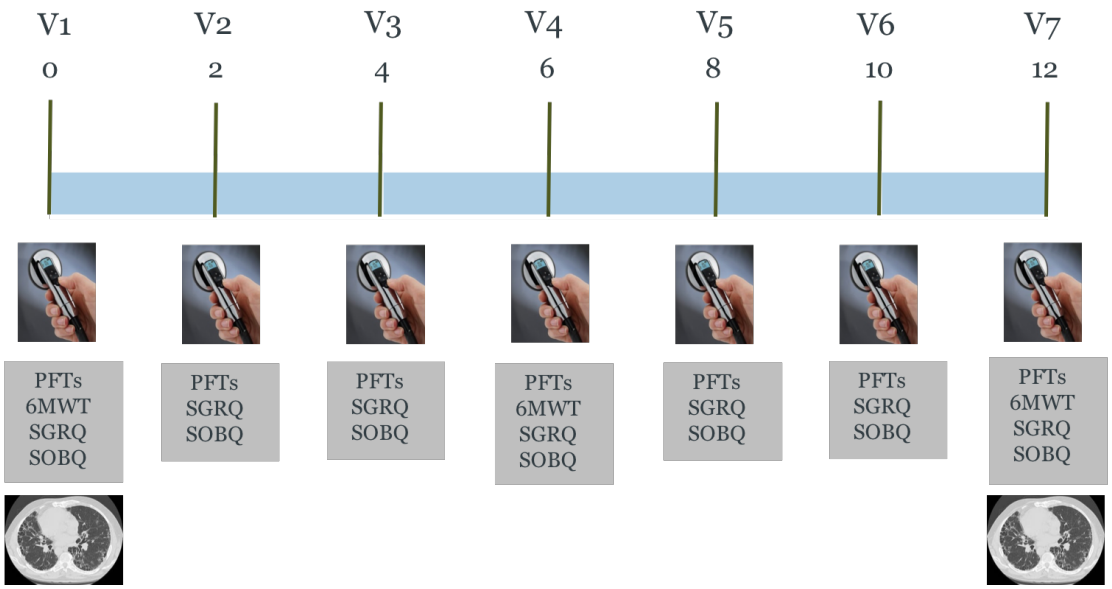


Figure 18 - Flowchart of the study. V1, V2 etc. indicate visit number; numbers indicate months of clinical observation.

At screening, the eligibility was determined through the collection of demographics, past medical history, current medications, smoking history. Once the eligibility was confirmed, the vital signs were taken. The arterial blood pressure was measured together with the body temperature, the respiratory rate and the oxygen saturation via a pulse oximeter placed on the index finger of the participant.

In IPF patients, breathlessness was measured using the University of California San Diego Shortness of Breath Questionnaire (UCSD-SOB) (Eakin et al., 1998), and respiratory symptoms-related quality of life was measured via the Saint George's Respiratory Questionnaire (SGRQ) (Jones et al., 1991). Both questionnaires were read and completed by the participants themselves. Lung sounds were recorded consecutively using a digital stethoscope (Littmann 3200, 3M, USA) held manually over different locations identified on the patients' chest using anatomic criteria. 10 sites were selected according to the guidelines for computerised respiratory sounds analysis (CORSA) (Sovijärvi et al., 2000) (Figure 19). 6 recordings were taken on the back of the chest - 2 at the apexes (right and left at 2 cm from the paravertebral line, in one of the first intercostal spaces), 2 at mid chest (right and left at 2 cm from the paravertebral line, in the fourth or fifth intercostal space) and 2 at the lung bases (right and left, at 5 cm from the paravertebral line and 7 cm below the scapular angle); 2 recordings were taken on anterior chest, right and left in the second intercostal space, mid-clavicular line; 2 recordings were finally taken on the side of the chest, right and left in the fourth or fifth intercostal space, in correspondence of the mid-axillary line. Sounds were recorded approximately for 15 seconds, a time considered sufficient to capture at least 3 breathing cycles. Since there is no indication as to the optimal recording time for computerised analysis, this choice was empirical and based upon discussion with the acoustic engineers of the Institute of Sound and Vibration Research of the University of Southampton. Patients were asked to breathe deeply through the mouth to maximise the intensity of the respiratory sounds, although the breathing effort was not airflow-targeted. Since IPF patients may get tired from deep ventilation, time was allocated to allow patients recover between the recordings. Care was taken to ensure the stethoscope was kept as still as possible during the recordings to minimise background noise. In 4 IPF patients, the set of 10 recordings was repeated 3 times during the same visit, leaving approximately 5-10 minutes between the measurements. This was performed in to assess the repeatability of a broad set of acoustic features of lung sounds recorded using a digital stethoscope.

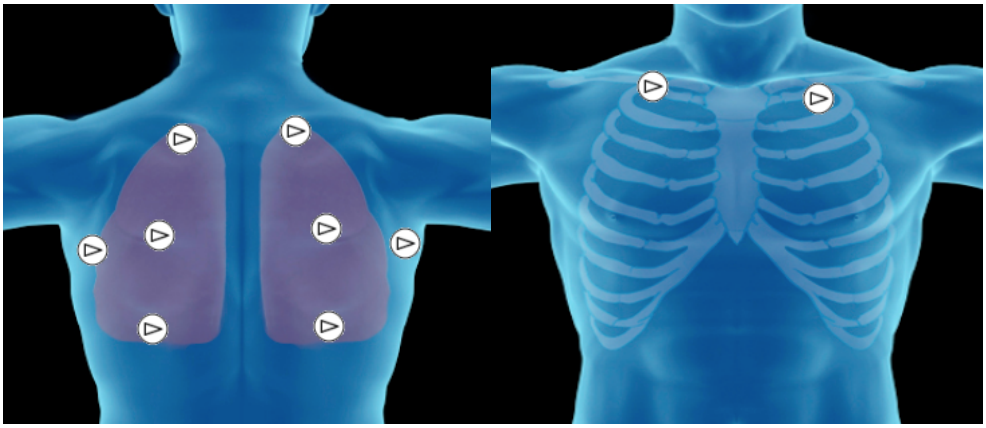


Figure 19 - Recording sites used in this study. For each side of the chest, 2 recordings were taken at the apexes (2 cm from the paravertebral line, in one of the first intercostal spaces); 2 were taken at mid chest (2 cm from the paravertebral line, in the fourth or fifth intercostal space); 2 at the lung bases (5 cm from the paravertebral line and 7 cm below the scapular angle); 2 on anterior chest, in the second intercostal space, mid-clavicular line; 2 on the side of the chest, in the fourth or fifth intercostal space, in correspondence of the mid-axillary line.

Pulmonary function was measured using spirometry, performed 3 times at each visit according to international recommendations (Miller et al., 2005). Gas exchange was measured using the Diffusion Capacity for Carbon Monoxide (DL_{CO}) test, also measured 3 times as recommended by international guidelines (MacIntyre et al., 2005). Exercise tolerance was measured through the 6-minute walk test (6MWT), performed according to guidelines (ATS, 2002).

Venous peripheral blood samples were taken and processed by the researcher for plasma and serum extraction. The samples collected during the study are currently being stored at the Clinical and Experimental Sciences Department facilities of the University of Southampton and will be used for future research in IPF.

Finally, IPF patients underwent a chest HRCT scan at the Radiology Unit at Southampton General Hospital. When HRCT scan was performed within a year prior to enrolment, the patient did not undergo a new test and the last HRCT scan was retrieved from the local servers/archives behind the patient's consent. The radiologic data collected will be used in future research to determine the correlation between changes in acoustic properties of lung sounds in IPF and CT indices of fibrosis, obtained either via visual assessment performed by expert thoracic radiologists or via automatic quantification of ILD features (Maldonado et al., 2014a).

The measurements of pulmonary function, lung sounds, breathlessness and quality of life were performed at each visit. The blood samples were collected and 6MWT was performed at 3 time points during the study – at baseline and at 6 and 12 months. A follow up chest HRCT scan was

performed in all IPF patients after 12 months (end of study). The timing of the visits and assessments was kept tighter than in the usual clinical practice to increase the statistical power of the analyses. Due to the lack of previous studies on the quantitative longitudinal assessment of lung sounds in ILD, this was an empirical choice, discussed with the statistician and confirmed after discussion of the study sustainability with 2 patients' representatives.

The healthy individuals enrolled in the study underwent the same visits schedule, but only performed measurement of lung sounds and pulmonary function. The full schedule of assessments is reported in Table 33.

Table 33 - Schedule of visits and assessments in the cohort study. * = only IPF patients; ** = if not performed within 12 months. SGRQ = Saint George Respiratory Questionnaire; UCSD-SOBQ = University of California San Diego Shortness of Breath Questionnaire; 6MWT = 6-minute walk test; HRCT = High Resolution Computerised Tomography.

Visit number	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Month	**0**	**2**	**4**	**6**	**8**	**10**	**12**
Obtain informed consent	X						
Demographics	X						
Smoking history	X						
Relevant medical history	X						
Check inclusion/Exclusion criteria	X						
Review of current medications	X	X	X	X	X	X	X
SGRQ*	X	X	X	X	X	X	X
UCSD-SOBQ*	X	X	X	X	X	X	X
Lung sounds recording	X	X	X	X	X	X	X
Spirometry	X	X	X	X	X	X	X
DLCO	X	X	X	X	X	X	X
6MWT*	X			X			X
Chest HRCT*	X**						X
Blood samples*	X			X			X

5.3.3 Data anonymisation, management and storage

The data collected in this study was anonymised prior to any review or analysis using unique codes, and securely stored on a University password-protected laptop. Paper clinical research forms were used for collection of demographics and clinical data. Since this documentation included patients' personal information and details, they were stored in locked filing cabinets together with the consent forms. After collection, the clinical data was temporarily transferred onto a digital spreadsheet (Microsoft Excel) on a University laptop for data management before analysis.

5.3.4 Outcome measures

5.3.4.1 Quantitative measurement of lung sounds

Lung sounds were recorded using the same electronic stethoscope (Littmann 3200, USA) described in section 4.3.5.1. The quantitative analysis of lung sounds involved methods for sound signal processing and feature extraction consisting in both standard and customised algorithms written in Matlab (version R2015a). The author/researcher was trained by Dr. Anna Barney and Dr. Dragana Nikolic, the acoustic engineering academics at the Institute for Sound and Vibration Research (ISVR) of the University of Southampton, to run the algorithms and extract the acoustic features that were used in the cohort study. A set of different techniques were automatically applied by the algorithms to pre-process the recordings, decompose the sound signals into different sound components and generate a broad set of acoustic features. Processing stages and the analyses performed for feature selection are described in detail in section 4.3.6.

5.3.4.2 Demographics, anthropometric data and vital signs

Demographics including gender, age, and race/ethnicity were recorded as reported by the patients. Information on past medical history, smoking habits, familiarity for ILD and other conditions, history of occupational and environmental exposures, current medications were also provided directly by the patients. Height and weight were determined using a digital scale previously calibrated, for use in the calculation of reference values for pulmonary function tests. Body Mass Index (BMI), an accepted index to assess adiposity in adults, was calculated as weight/height^2. Vital signs were taken as part of a general health check at each visit, while the patient was at rest in a sitting position. These included measurements of respiratory rate through observation, temperature via a digital thermometer, systolic and diastolic blood pressure and heart rate through a digital sphygmomanometer. Oxygen saturation was also recorded non-

invasively through a pulse oximeter probe attached to the finger of the patient at rest and after the 6MWT for assessing desaturation induced by exercise.

5.3.4.3 Measurement of lung function and gas exchange

Pulmonary function tests (PFTs) represent the most used and validated tool to assess IPF patients in terms of severity and progression of the disease. As such, spirometry and DL_{CO} test were performed in this research according to international recommended standards (MacIntyre et al., 2005,Miller et al., 2005) every 2 months.

As for spirometry, the main parameter used in this research was Forced Vital Capacity (FVC), as it represents the most reliable measurement of the pulmonary capacity in these patients. The Forced Expiratory Volume in the first second (FEV_1) was also measured to exclude an obstructive airways disease through the calculation of the FEV_1/FVC ratio. Percentages of predicted FEV1 and FVC values were calculated using prediction equations for lung function provided by Falaschetti et al. (Falaschetti et al., 2004). Percentages of predicted FVC values between 100 and 80 represent a mild impairment of lung function; between 80 and 50 a moderate impairment; between 50 and 30 a severe impairment; below 30 a very severe impairment. Absolute or relative changes in percentages of predicted FVC were used as appropriate in the analyses, with relative changes calculated as the following:

$$\frac{\Delta\ \%\ pred\ FVC}{baseline\ \%\ pred\ FVC}$$

Where $\Delta\ \%\ pred\ FVC$ is the difference between the percentage of predicted FVC measured at last visit and the percentage of predicted FVC measured at baseline. As described in section 2.4.3.1, a decline of ≥10% in absolute or relative FVC at 6 months is associated with a worse prognosis.

Total Lung Capacity (TLC) and Residual Volume (RV) were also measured at baseline via plethysmography and used to calculate the emphysema index (RV/TLC). When >100, such index suggests the presence of a significant quote of emphysema and represented an exclusion criteria for this study.

DL_{CO}, which measures the integrity of the blood-gas barrier in the lungs, is calculated by multiplying the rate of uptake of carbon monoxide (K_{CO}) and the alveolar volume (AV – measured by gas dilution during the DL_{co} measurement) and absolute values are expressed as mmol/Kpa/min. When DL_{CO} is below 60% of the predicted value, gas exchange is impaired.

5.3.4.4 Measurement of breathlessness and quality of life

As discussed in sections 2.2.1 and 2.4.3, chronic, progressive dyspnea is the most common symptom of patients with chronic ILD/IPF and often causes an impairment in the patients' daily activities with obvious repercussions in their quality of life. Both breathlessness and poor quality of life represent validated clinical indicators of progression of disease and poor outcome, as such the measurement of these outcomes was considered in this research.

Breathlessness was measured using two different tools. The University of California San Diego Shortness of Breath Questionnaire (UCSD-SOBQ) is a 24 items questionnaire for the assessment of the level of dyspnea experienced at rest and while performing a series of daily life activities, which has been validated to assess change in dyspnea over time in patients with IPF (Swigris et al., 2012a). For each item the patient is asked to indicate the level of dyspnea related to that specific activity on a scale from 0 (not at all) to 5 (unable to do the activity because of breathlessness). As such, the highest the score the highest is the level of dyspnea experienced. A longitudinal increase of 5 units in the score has been indicated as the minimal clinically important difference for this questionnaire, although this is not a specific threshold for IPF patients (Kupferberg et al., 2005). Since not all the questions were answered at all time points by participants and there is no indication in the literature for imputing missing data, for this study a "relative" score was calculated by dividing each score by the highest possible score based upon the number of questions answered. The Borg scale, a subjective, self-reported measure of breathlessness initially conceived as a measure of perception of exertion during exercise (Borg, 1982), was used for a quick assessment of dyspnea and overall fatigue at rest and after performing a 6MWT as recommended by ATS guidelines (ATS, 2002).

Quality of life was measured using the St. George's Respiratory Questionnaire (SGRQ), a standardised 50 items questionnaire for measuring impaired health and perceived well-being initially designed and validated for use in patients with chronic obstructive pulmonary disease (Jones et al., 1991). It is structured in two parts: part 1 (questions 1-8) addresses the frequency of respiratory symptoms, while part 2 (sections 9-16) addresses the patient's current state. The final scores are obtained using an automatic calculator (built onto a Microsoft Excel spreadsheet) and are structured into a symptoms score, an activity score (measuring disturbances of daily physical activity), an impact score (measuring disturbances of psycho-social function) and an overall score incorporating the 3 previous ones. The overall score was used in this research as it offers a comprehensive assessment of health status related to respiratory symptoms. In terms of minimal clinically important difference (MCID), the first estimates derived for SGRQ in IPF were greater than the widely accepted 4-points change in patients with obstructive diseases, and ranged from

6 to 13 units (Swigris et al., 2010). Despite an IPF-specific version, the SGRQ-I, has been recently developed (Yorke et al., 2010) by removing 16 items from the questionnaire, in this study the original version of SGRQ was used for being readily available online and for its widely proven reliability in respiratory disorders.

The tools used for measuring dyspnea and quality of life are reported in full in Appendix B.

5.3.4.5 Measurement of exercise tolerance

Exercise tolerance is routinely measured in IPF patients as it can reflect the functional impairment resulting from the progression of the disease. In this study, it was recorded every 6 months through the 6-minute Walk Test (6MWT) (ATS, 2002), performed in accordance with the RBRU standard operating procedures. A flat, straight, corridor in the WellCome Trust facility, long enough to ensure a walking course of thirty meters in length, was used. The length of the corridor was marked every three meters. Pulse oximetry was performed before and after the test to determine oxygen saturation. The distance walked (6-minute walking distance, 6MWD) was calculated counting the laps and the additional distance covered. As mentioned in section 2.4.3.1, a decline of >50 m in 6MWD after 6 months has been found to be of prognostic significance in patients with IPF (du Bois et al., 2011c). As mentioned above, baseline dyspnea and overall fatigue were also determined before and after the test using the Borg scale (Borg, 1982).

5.3.5 Data analysis

5.3.5.1 Overview of analysis and types of data

The analysis of the accrued data was performed in Southampton by the author. Statistical support was provided throughout the analysis process by Dr. Borislav Dimitrov, Professor of Medical Statistics at the University of Southampton and co-supervisor of the author. Dr. Anna Barney and Dr. Dragana Nikolic provided guidance and assistance in the computerised analysis of lung sounds.

In the cohort study, the primary analysis focused on continuous variables represented by acoustic features of lung sounds recorded from IPF patients and controls. Intra-subject reliability was assessed first and a comparative analysis of acoustic features between IPF patients and controls was performed. Then, acoustic features were correlated with other continuous variables consisting in clinical parameters reflecting disease progression in the IPF group.

5.3.5.2 Analysis of demographic, anthropometric and medical history data

Information including age, gender, race/ethnicity, BMI (calculated from height(cm)/weight(kg)^2) smoking status/history, years from IPF diagnosis, comorbidities and familiarity for ILD, IPF

medications, concomitant medications were all entered in SPSS (version 24) and descriptive statistics were used to characterise the study populations at baseline. Where appropriate, categorical data were contrasted using Chi-squared test, while continuous variables were compared using independent samples t-test.

5.3.5.3 Analysis of lung function, gas exchange, oxygen saturation, breathlessness, quality of life and exercise tolerance

In the cohort study, the relevant parameters measured via PFTs, DL_{CO} test, pulse oximetry, questionnaires and scales, and 6-minute walk test were also entered into SPSS (version 24).

Firstly, they were presented using descriptive statistics such as means, standard deviations, minimum and maximum values for baseline and other time points during the observation period, for both the IPF and the healthy volunteers groups where applicable.

In order to determine whether such parameters significantly changed overtime in the study population, the longitudinal changes of such parameters were explored using ANOVA for repeated measures. No imputation was made for missing data in this study, as such the observations that were available for all patients were used. Mauchly's test was used to assess sphericity for the sample. Estimated means of the parameters were calculated for each time point after adjusting for demographic and anthropometric parameters such as age, sex and BMI of the study subjects, all entered as covariates in the model. Pairwise comparisons between single study time points were made. Differences of estimated means, standard errors and 95% confidence intervals were reported. The Bonferroni method for multiple testing was used for correcting the values of the 95% confidence Intervals.

As for the exercise tolerance, the mean differences of the measurements taken before and after the 6-minute walk test – dyspnea and fatigue scores at BORG scale, and oxygen saturation – were reported together with the 95% confidence intervals and tested for significance using paired samples t-test.

5.3.5.4 Computerised analysis of lung sounds

In the cohort study, lung sound data was analysed via sound signals processing algorithms written in Matlab (version 2015a). MIR Toolbox, an open source software available online for extraction of acoustic features from sound signals (Lartillot et al., 2007), and a customised algorithm written for this project by Dr. Anna Barney and Dr. Dragana Nikolic were both used in this research. Each recording, represented by a file in .wav format, was processed via Matlab through the sequential application of such algorithms. This approach enabled the automatic extraction of almost 500

different variables corresponding to distinct acoustic features. A block scheme summarising the methods for lung sound analysis used in the cohort study is shown below in Figure 20.

Figure 20 - Block scheme of proposed analysis of lung sounds.

This section will firstly provide some detailed information about the fundamental steps of sound processing and feature extraction, performed automatically by applying the algorithms mentioned above. Then, the methods used for selection of the relevant features will be described. Finally, the analysis of correlation between the longitudinal changes in the set of selected features and the modifications in the other outcome measures will be presented.

5.3.5.4.1 Pre-processing and decomposing of lung sounds

Each lung sound signal was initially pre-processed via Matlab through a high-pass filtering with a cut-off set at 75 Hz (i.e. a filter which let only frequencies higher than 75 Hz pass). Then, it was decomposed using Empirical Mode Decomposition (EMD) into a number of intrinsic mode functions (IMFs) (Chen et al., 2014). The Empirical Mode Decomposition (EMD) technique is employed to assist in identifying “Velcro-type” crackles (if they are present) by separating the crackle sound from the background respiratory sound (Charleston-Villalobos et al., 2007,Chen et al., 2014). This technique corresponds to an automatic and adaptive filter method that empirically identifies and extracts the intrinsic oscillatory modes of the non-stationary and non-linear sound signals. The extracted modes, so-called Intrinsic Mode Functions (IMFs), represent physical properties of the signal across different frequencies.

As reported in the literature, due to differences in stationarity and frequency content between the crackles and the respiratory sounds, combining information from different IMFs can help to extract the fine crackle sound component (and its features) from the overall breath sound signal. This is illustrated in Figure 21**Error! Reference source not found.** where the crackle-related information in the inspiration sound is spread over the low-numbered IMFs, which have a higher frequency content. The crackle component shown in plot (c) represents the sum of the first three IMFs while the breath sound component in plot (d) corresponds to the sum of other IMFs and the residual signal. The use of the first 3 or 4 IMFs to represent the crackle component is an empirical choice (Charleston-Villalobos et al., 2007,Chen et al., 2014). Since the number of IMFs extracted for each sound file is variable, only the first 10 IMFs from any file were chosen to be used in further analysis. In this study, EMD was applied to each sound file individually to extract the IMFs, and these identified separate acoustic features/variables.

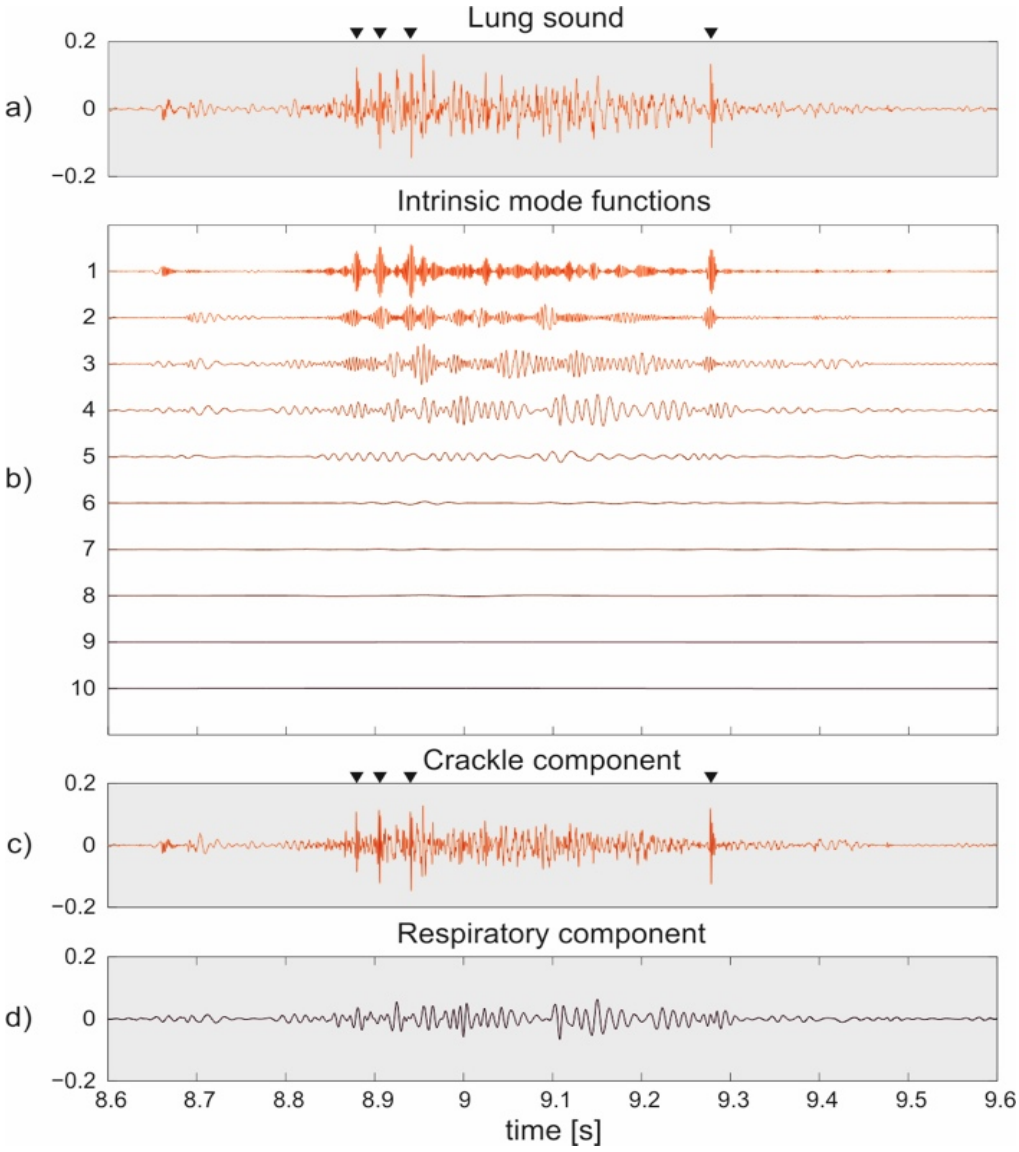


Figure 21 – Visualisation of Electronic Mode Decomposition (EMD) analysis applied to the inspiration phase of the lung sound recorded from an IPF patient: (b) the first 10 extracted Intrinsic Mode Functions (IMFs); (c) the crackle component obtained as a sum of the first 3 IMFs and (d) the respiratory component.

5.3.5.4.2 Feature generation

For each sound file analysed, 481 features were generated from the filtered lung sound signal as a whole and from the extracted IMFs and signal components using the MIR Toolbox as well as the custom-written code in Matlab (Lartillot et al., 2007). These features were calculated for the original acoustic signal as well as for different IMFs and for the signal components (crackle and respiratory) at different frequency bands, which explains the high total number of features generated. The features are derived either from the statistical properties of the signal and its components (such as skewness and kurtosis), from the energy content, and from the frequency

domain (such as cepstral features). The features' domains are summarised in Table 34**Error! Reference source not found.**, while a full list of the features extracted and their description is reported in Appendix D. The glossary reported after the appendices of this thesis describes more in detail a few technical concepts behind some relevant features used in the analysis.

The continuous data representing the acoustic features was automatically extracted onto an Excel spreadsheet, and then exported into SPSS for further analysis. At the first run of the algorithms, while 170 "global" features (i.e. those derived from the whole original sound signal) were generated for the entire data set, approximately 20% of the recordings (246 files for the IPF subjects, 146 files for the healthy controls) were not processed via EMD, thus generating a consistent amount of missing data. In order to address this issue, these files were edited by the author into shorter bits of recording using Audacity software, making sure that the recordings included at least one full breathing cycle. Depending on the length of the original recording, between 3 and 5 shorter recordings were edited from each file. The algorithms were run again to analyse the edited recordings and EMD performed successfully in most cases. For each recording, one of the edited bits that were successfully processed was chosen at random (using the RAND function in Excel), and the features extracted were used for the analyses.

Table 34 - List of main categories (vectors) of acoustic features used in the study. Hz = Hertz; IMF = Intrinsic Mode Function; EW = energy weight; RMS = Root Mean Square.

	Feature vector	Description	N	
1	Energy of each IMF individually	energy of each IMF vs the original signal	2N	
		ratio of energy of each IMF vs the original signal		
	Total energy of all IMFs	energy of the original signal	1	
2	Energy of crackle and respiratory components	energy ratios of the crackle and respiratory components vs the original signal	12	
		energy ratio of the crackle and respiratory components		
3	Average power of each IMF in the 150-450 Hz range	average power of each IMF in the 150-450 Hz frequency band	N	
4	Average power of the crackle and respiratory components	average power of the crackle and respiratory component in the 150,450 Hz frequency band	6	
		ratio of average power of the crackle and respiratory component in the 150,450 Hz frequency band		
5	Energy weights (EW) of the IMF components	energy weights EW in specific frequency bands	9N	
6	Energy weights (EW) of the crackle and respiratory component	energy weights EW in specific frequency bands	36	
		TOTAL (group I)	**12N+55**	
1	General features of the original signal	RMS, Low Energy, Zero-cross, Roll-off85, Roll-off95, Brightness1000Hz, Brightness1500Hz, Centroid, Spread, Skewness, Kurtosis, Flatness, Entropy, Regularity, MFCC, meanFRAME (mean, std, med), medianFRAME (mean, std, med)	34	
2	General features of the signal in specific frequency bands	RMS, Low Energy, Zero-cross, Roll-off85, Roll-off95, Brightness1000Hz, Brightness1500Hz, Centroid, Spread, Skewness, Kurtosis, Flatness, Entropy, Regularity, MFCC, meanFRAME (mean, std, med), medianFRAME (mean, std, med)	4*34	
3	General features of the crackle and respiratory components	RMS, Low Energy, Zero-cross, Roll-off85, Roll-off95, Brightness1000Hz, Brightness1500Hz, Centroid, Spread, Skewness, Kurtosis, Flatness, Entropy, Regularity, MFCC, meanFRAME (mean, std, med), medianFRAME (mean, std, med)	4*34	
		TOTAL (group II)	**306**	
		TOTAL (where N is the number of IMFs extracted per file)	**12N+361**	**481**

5.3.5.5 Screening of discriminating acoustic features of IPF

An analysis was performed to identify, among the 481 features generated, those that might represent the most clinically relevant in IPF.

The first step consisted in selecting those features that show good reliability when sounds are being recorded using an electronic stethoscope in IPF patients. An intra-subject (test-retest) reliability analysis was performed. The following step was to compare such repeatable features between IPF patients and controls, to determine which were distinctive of a fibrotic disease.

The reliability analysis was performed using 3 repeated measurements made at the 10 recording sites (used as independent variables) in 4 patients selected at random from the IPF cohort. For each patient, the recordings were obtained during the same visit, leaving approximately 5-10 minutes between the recording sessions.

The Intraclass Correlation Coefficient (ICC), which includes an analysis of variance, was used to assess intra-subject reliability of acoustic features. An ICC equalling zero means no reliability, while an ICC equalling one indicates perfect reliability. It is generally accepted that values above 0.75 represent excellent reliability, values between 0.4 and 0.74 moderate to good reliability and values below 0.4 represent poor reliability (Fleiss, 1986). For each feature, the Intraclass correlation coefficient (ICC) was calculated for all available observations with 3 measurements (n=37). Since it cannot be interpreted clinically (because it gives no indication of the magnitude of disagreement between measurements), a value > 0.5 was arbitrarily chosen to indicate acceptable reliability in this study.

A multivariate analysis of covariance (ANCOVA) model was used to determine which, among the repeatable features, entered as dependent variables, could discriminate between the IPF and controls subjects. Study group was the independent (fixed) factor in the model, while recording sites and time points (corresponding to the study visits) were entered as random factors (covariates) for adjustment. The model was finally also adjusted for demographics and anthropometric parameters (age, sex and BMI) of the subjects. Estimated means of the features were calculated as well as their differences between groups.

5.3.5.6 Longitudinal assessment of acoustic features and correlation with parameters of disease progression

The acoustic features found to be repeatable in IPF were assessed for longitudinal changes over 12 months of clinical observation. The variables measured at the different study time points were assessed using ANOVA for repeated measures. Just as for the clinical parameters, no imputation was made for missing data. Estimated means were calculated for each time point after adjusting for recording site and other demographic and anthropometric parameters such as age, sex and BMI of the study subjects, all entered as covariates in the model.

For each variable, pairwise comparisons between the single study time points were made. Differences of estimated means, standard errors and 95% confidence Intervals were reported. The Bonferroni method for multiple testing was used for correction of the values of the 95% confidence Intervals.

The features showing a significant change over the study observation period were correlated with clinical parameters, also collected longitudinally. For a specific study time point, every patient in the study had multiple measurements for a single acoustic feature, one for each recording site. As such, the values calculated for each feature from the different recordings were individually correlated with the value of a clinical parameter collected at the same study time point. The relationship between individual acoustic features and other clinical and physiology variables were examined calculation Pearson' correlation coefficients and via univariate linear regression. Multivariable linear regression analysis was also undertaken to determine the relationships between the reproducible acoustic features and other parameters of disease progression.

5.4 Results

5.4.1 Baseline characteristics of study population

The population enrolled in the cohort study consisted of subjects with an established diagnosis of IPF and healthy volunteers. The baseline demographics characteristics and other clinical information of the study population are reported in Table 35, while the baseline physiology measurements are reported in Table 36.

In the IPF group, 14 of the 19 patients recruited completed the study. Among the 5 patients who dropped out, 3 patients died during the observation period due to progressive pulmonary fibrosis, while 2 patients withdrew because of poor health conditions which made impossible for them to attend further clinical visits. No visits were skipped from the patients remaining in the study.

Among the healthy volunteers, 7 subjects out of 10 completed the study. 3 subjects withdrew from the study at some point for personal or otherwise non-specified reasons.

Mean age was higher in the IPF group than in the controls (70.8 and 55.2 years, respectively). In the IPF group there was a large male predominance (84.2%), whilst in the control group males were only 20% (2 subjects out of 10 participants). Body Mass Index (BMI) was higher in the control group (average value of 30.9), indicating a trend towards overweight. In both groups majority of patients were lifelong non-smokers (63.2% of subjects in the IPF group and 60% in the control group). Approximately 2/3 of patients with IPF were in anti-fibrotic treatment at baseline, with pirfenidone or nintedanib in similar proportions (6 and 7 patients, respectively). Among non-IPF related treatments, 3 patients were in treatment with steroids, although at low dosages (<10 mg prednisone daily) to improve fatigue. Majority of patients with IPF (63.2%) were taking proton pump inhibitors. Accordingly, gastro-oesophageal reflux disease was the most frequent comorbidity, while cardiovascular diseases (mostly arterial hypertension and atrial fibrillation) were the second most common. Only 1 patient was on supplemental oxygen at baseline.

Table 35 - Baseline demographics of IPF and control groups. Data are counts (%) or mean ± standard deviation (SD). BMI= Body Mass Index; GERD = Gastro-Esophageal Reflux Disease

	IPF group (n=19)	Healthy group (n=10)
Mean age, years	70.8(±6.53)	55.2(±9.99)
Sex		
Men	16(84.2%)	2(20%)
Women	3(15.8%)	8(80%)
Ethnicity		
White Caucasian	19(100%)	10(100%)
BMI	26.5(±4.16)	30.9(±5.45)
Smoking history		
Current	0(0%)	1(10%)
Former	7(36.8%)	3(30%)
Never smoker	12(63.2%)	6(60%)
Anti-fibrotic treatments		
Pirfenidone	6(31.6%)	0(0%)
Nintedanib	7(36.8%)	0(0%)
Investigational drug/placebo	1(5.3%)	0(0%)
Other treatments		
Prednisone	3(15.8%)	0(0%)
Azathioprine	1(5.3%)	0(0%)
NAC	4(21.1%)	0(0%)
PPI	12(63.2%)	0(0%)
Supplemental Oxygen		
Yes	1(5.3%)	0(0%)
No	18(94.7%)	0(0%)
Mean time from IPF diagnosis, years	2.44(±1.88)	N/A
Comorbidities		
Cardiovascular	7(36.8%)	2(20%)
Diabetes	4(21.1%)	0(0%)
GERD	9(47.4%)	1(10%)
Renal	1(5.3%)	0(0%)
Neurologic	1(5.3%)	0(0%)
Other	7(36.8%)	4(40%)

On average, the IPF population in this study had a moderate functional impairment at baseline (mean predicted FVC 73.74%) together with a reduction of DL_{CO} (mean predicted DL_{CO} 43.47%). Pulmonary function and gas exchange were well preserved in the control population.

When stratified using the GAP (gender, age, physiology) staging system, patients with IPF were found to be equally distributed between the three stages of severity/risk of the disease (6 patients were stage I, 7 were stage II and 6 were stage III).

At exercise testing (6MWT) there was an increase in the mean scores of Borg index for fatigue and dyspnea, and a desaturation with an average loss of -9.9%.

Table 36 - Baseline physiology measurements of IPF and control group. Data are counts and percentages (%) or mean ± standard deviation (SD). FVC = Forced Vital Capacity. FEV1 = Forced Expiratory Volume in the first second. RV = Residual Volume. TLC = Total Lung Capacity. GAP = Gender Age Physiology. DL_{CO} = Diffusion Lung Capacity for Carbon Monoxide. UCSD-SOBQ = University California San Diego-Shortness Of Breath Questionnaire. SGRQ = Saint George's Respiratory Questionnaire. 6MWD = 6-Minute Walk Distance. SaO2 = Saturation of oxygen.

	IPF group (n=19)	Healthy group (n=10)
Mean FVC volume, L	2.62(±0.68)	3.82(±0.6)
Mean FVC % predicted	73.74(±20.88)	109.1(±8.5)
Mean FEV_1, L	2.16(±0.58)	2.91(±0.51)
Mean FEV_1 % predicted	79.45(±20.44)	104(±7.39)
RV/TLC %	79.57(±15.04)	85.22(±9.18)
GAP Index		
Stage I	6(31.6%)	N/A
Stage II	7(36.8%)	N/A
Stage III	6(31.6%)	N/A
Mean DL_{CO}, mmol/min/kPa	3.58(±0.97)	8.06(±1.81)
Mean DL_{CO} % predicted	43.47(±9.21)	95.3(±21.3)
Mean relative UCSD-SOBQ, score	0.17(±0.08)	N/A
Mean SGRQ total, score	35.02(±11.05)	N/A
Mean 6MWD, m	442.61(±120.56)	N/A
Mean Borg scale of perceived dyspnea, score		
Pre-walk	0.41(±0.58)	N/A
Post-walk	2.16(±1.41)	N/A
Mean Borg scale of perceived fatigue, score		
Pre-walk	0.25(±0.58)	N/A
Post-walk	1.47(±1.34)	N/A
SaO_2 % (room air)		
At rest	95.21(±1.69)	97.5(±1.35)
Post-walk	85.06(±6.81)	N/A

5.4.2 Lung function

The results of Forced Vital Capacity (FVC) and Forced Expiratory Volume in the first second (FEV_1) from the pulmonary function tests performed at the different time points during the study for the IPF group are reported in Table 37.

Table 37 - Longitudinal measurements of Forced Vital Capacity (FVC) and Forced Expiratory Volume in the first second (FEV_1) expressed as absolute values (L = litres) and percentage of predicted values for IPF group. Data presented as mean ± standard deviation (SD), minimum (Min) and maximum (Max) values.

IPF group

Months	Parameter	Observations	Mean (±SD)	Min	Max
0 (baseline)	FVC(L)	19	2.62(±0.68)	1.72	4.14
	FVC(%pred)		73.74(±20.86)	42	127
	FEV_1 (L)	19	2.17(±0.58)	1.38	3.52
	FEV_1 (%pred)		79.45(±20.44)	46	129
2	FVC(L)	19	2.6(±0.72)	1.71	4.02
	FVC(%pred)		72.21(±19.21)	42	110
	FEV_1 (L)	19	2.15(±0.56)	1.39	3.27
	FEV_1 (%pred)		78.16(±19.5)	46	120
4	FVC(L)	17	2.54(±0.66)	1.37	3.63
	FVC(%pred)		73.59(±19.9)	41	113
	FEV_1 (L)	17	2.26(±0.63)	1.5	4.08
	FEV_1 (%pred)		79.53(±18.58)	48	120
6	FVC(L)	16	2.52(±0.66)	1.73	4.12
	FVC(%pred)		72.13(±20.66)	42	114
	FEV_1 (L)	16	2.14(±0.55)	1.41	3.26
	FEV_1 (%pred)		77.75(±19.57)	48	121
8	FVC(L)	14	2.49(±0.68)	1.72	3.97
	FVC(%pred)		72.07(±20.98)	45	110
	FEV_1 (L)	14	2.13(±0.57)	1.35	3.2
	FEV_1 (%pred)		76.71(±18.44)	50	119
10	FVC(L)	13	2.53(±0.66)	1.67	3.91
	FVC(%pred)		72.85(±20.68)	41	109
	FEV_1 (L)	13	2.07(±0.54)	1.29	3.07
	FEV_1 (%pred)		77.08(±19.28)	47	114
12	FVC(L)	12	2.51(±0.78)	1.42	3.91
	FVC(%pred)		71.83(±24.08)	36	108
	FEV_1 (L)	12	2.15(±0.63)	1.32	3.14
	FEV_1 (%pred)		76.25(±20.72)	47	116

On average, patients staying in the cohort experienced a slight deterioration of pulmonary function. Mean FVC declined from 2.62 to 2.51 litres (-110 ml) while mean FEV_1 declined from 2.17% to 2.15 (-200 ml). However, the number of observations was different across study time points because of the patients who dropped out of the study or because patients were not fit to perform PFTs in a few occasions.

ANOVA for repeated measures was performed to determine whether there was a significant functional decline in this cohort, expressed as % predicted FVC. Since no data imputation for missing data was performed in this study, the analysis was performed over 11 available observations across the study time points. Estimated means of % predicted FVC are reported in **Error! Reference source not found.** Table 38 and shown in the plot in **Error! Reference source not found.**Figure 22. A mean change of 5.9% was found between the first and the last visit, although it was not significant as reported in the pairwise comparisons between estimated means of % predicted FVC (95% CI -4.261;16.080, p=0.862), and neither there were significant differences between other single study time points. Pairwise comparisons are reported in full in Appendix 2 in the accompanying material of this thesis. Nevertheless, the overall rate of decline of % predicted FVC over the 12 months of follow-up was significant at the ANOVA tests of within-subject effects (p=0.013), shown in **Error! Reference source not found.**Table 39.

Table 38 - Estimated means for % predicted FVC in the IPF group. Data presented as means with standard error (SE) and 95% confidence intervals (95% CI).

Visit	Mean	SE	95% CI	
			Lower Bound	Upper Bound
1	80.545	6.382	66.325	94.766
2	78.818	5.348	66.903	90.733
3	78.091	5.849	65.059	91.123
4	77.818	6.025	64.393	91.244
5	76.364	6.296	62.335	90.392
6	75.182	6.047	61.709	88.654
7	74.636	6.969	59.108	90.165

Table 39 - Tests of within-subjects effects for % predicted FVC in the IPF group

Tests of Within-Subjects Effects

Measure: FVC

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Visit	Sphericity Assumed	287.896	6	47.983	2.997	.013
	Greenhouse-Geisser	287.896	3.216	89.518	2.997	.042
	Huynh-Feldt	287.896	4.920	58.515	2.997	.020
	Lower-bound	287.896	1.000	287.896	2.997	.114
Error(Visit)	Sphericity Assumed	960.675	60	16.011		
	Greenhouse-Geisser	960.675	32.161	29.871		
	Huynh-Feldt	960.675	49.200	19.526		
	Lower-bound	960.675	10.000	96.068		

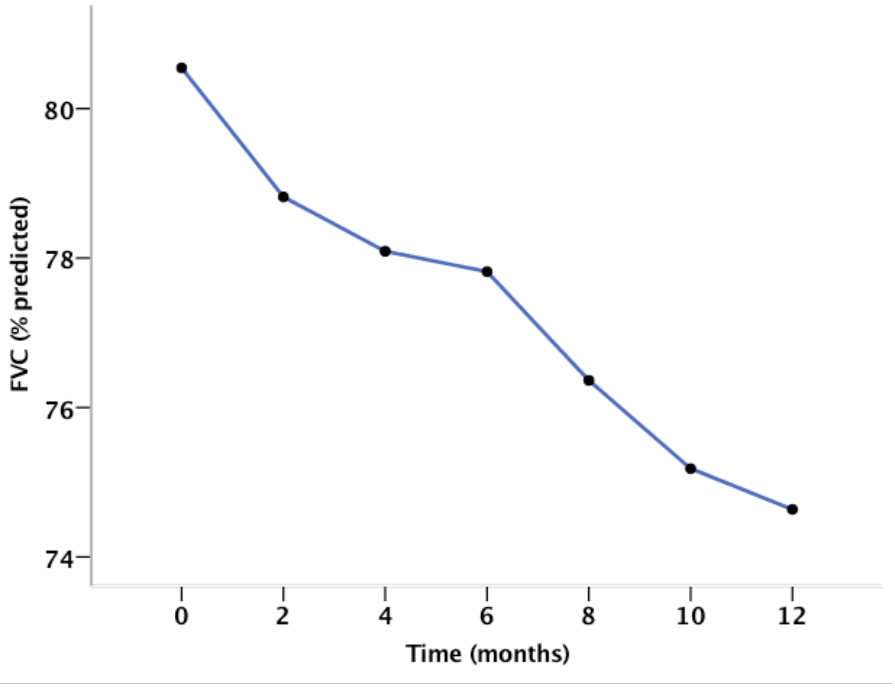


Figure 22 – Plot of estimated means of % predicted FVC in the IPF group.

The proportion of IPF patients who experienced disease progression was calculated as defined by the following: 1) absolute decline of FVC ≥ 10% at 12 months or 2) death from all causes or 3) drop-out for inability to perform study visits. 10 patients out of 19 (52.7%) progressed during the study (**Error! Reference source not found.**Figure 23).

Overall, these findings demonstrate that despite the small size of the population and its heterogeneity, there was significant disease progression expressed by functional decline.

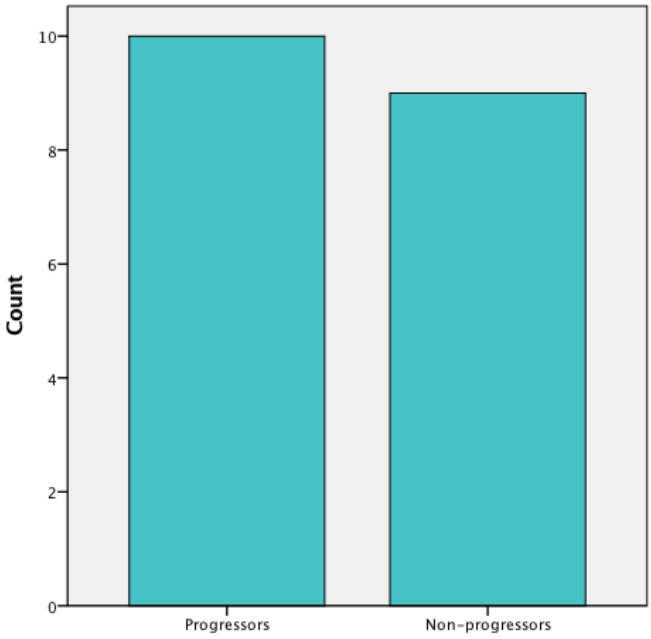


Figure 23 – "Progressors" and "non-progressors" in the IPF group. Disease progression was assessed via a composite endpoint defined by: 1) absolute decline of FVC ≥ 10% at 12 months or 2) death from all causes or 3) drop-out for inability to perform study visits.

5.4.3 Gas exchange

The results of the DL_{CO} test performed at the different time points during the study by the IPF group are reported in **Error! Reference source not found.**Table 40. On average, there was a reduction in DL_{CO} from 3.58 to 3.36 mmol/min/kPa (-0.22 mmol/min/kPa) between the first and the last visit.

Table 40 - Longitudinal measurements of Diffusion Capacity of CO (DL_{CO}) expressed as absolute and percentage of predicted values for the IPF group. Data presented as means ± standard deviations (SD), minimum (Min) and maximum (Max) values.

IPF group

Months	Parameter	Observations	Mean(±SD)	Min	Max
0 (baseline)	DL_{CO}(mmol/min/kPa) DL_{CO}(%pred)	19	3.58(±0.97) 43.47(±9.21)	2.26 30	5.49 59
2	DL_{CO}(mmol/min/kPa) DL_{CO}(%pred)	19	3.59(±1.16) 44.11(±13.6)	1.79 29	5.58 78
4	DL_{CO}(mmol/min/kPa) DL_{CO}(%pred)	17	3.34(±1.39) 41.18(±12.26)	2.42 24	6.88 73
6	DL_{CO}(mmol/min/kPa) DL_{CO}(%pred)	15	3.43(±1.06) 41.4(±10.8)	2.12 24	5.62 61
8	DL_{CO}(mmol/min/kPa) DL_{CO}(%pred)	14	3.25(±1.11) 39.71(±11.2)	1.41 22	4.86 58
10	DL_{CO}(mmol/min/kPa) DL_{CO}(%pred)	13	3.12(±1.19) 37.92(±12.79)	1.62 22	5.25 57
12	DL_{CO}(mmol/min/kPa) DL_{CO}(%pred)	11	3.36(±1.04) 40.36(±11.59)	1.8 20	5.25 57

Just as with FVC, ANOVA for repeated measures was performed over 10 available observations across all time points to check whether there was a significant worsening in gas exchange in this population. Estimated means of % predicted DL_{CO} are reported in **Error! Reference source not found.** Table 41 and shown in the plot in Figure 24**Error! Reference source not found.**. A mean difference of 5.5% was found between the first and the last visit over 10 available observations, although it was not significant (95% CI -2.336;13.336, p=0.35). Differences in estimated means of % predicted DL_{CO} between other single time points were not significant either, as shown in the pairwise comparisons in Appendix 2. The rate of decline of predicted DL_{CO} over 12 months of follow-up was however significant at the ANOVA tests of within-subject effects (p=0.015, Table 42**Error! Reference source not found.**).

Table 41 - Estimated means for % predicted FVC in the IPF group. Data presented as means with standard error (SE) and 95% confidence intervals (95% CI)

Visit	Mean	SE	95% CI	
			Lower Bound	Upper Bound
1	47.900	2.354	42.574	53.226
2	48.900	2.923	42.288	55.512
3	47.200	3.678	38.879	55.521
4	46.800	2.649	40.807	52.793
5	44.200	2.962	37.500	50.900
6	42.200	3.593	34.073	50.327
7	42.400	3.138	35.301	49.499

Table 42 - Tests of within-subjects effects for % predicted DLCO in the IPF group

Tests of Within-Subjects Effects

Measure: DLCO

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	439.171	6	73.195	2.924	.015
	Greenhouse-Geisser	439.171	2.836	154.876	2.924	.056
	Huynh-Feldt	439.171	4.276	102.716	2.924	.031
	Lower-bound	439.171	1.000	439.171	2.924	.121
Error(factor1)	Sphericity Assumed	1351.971	54	25.037		
	Greenhouse-Geisser	1351.971	25.521	52.975		
	Huynh-Feldt	1351.971	38.480	35.134		
	Lower-bound	1351.971	9.000	150.219		

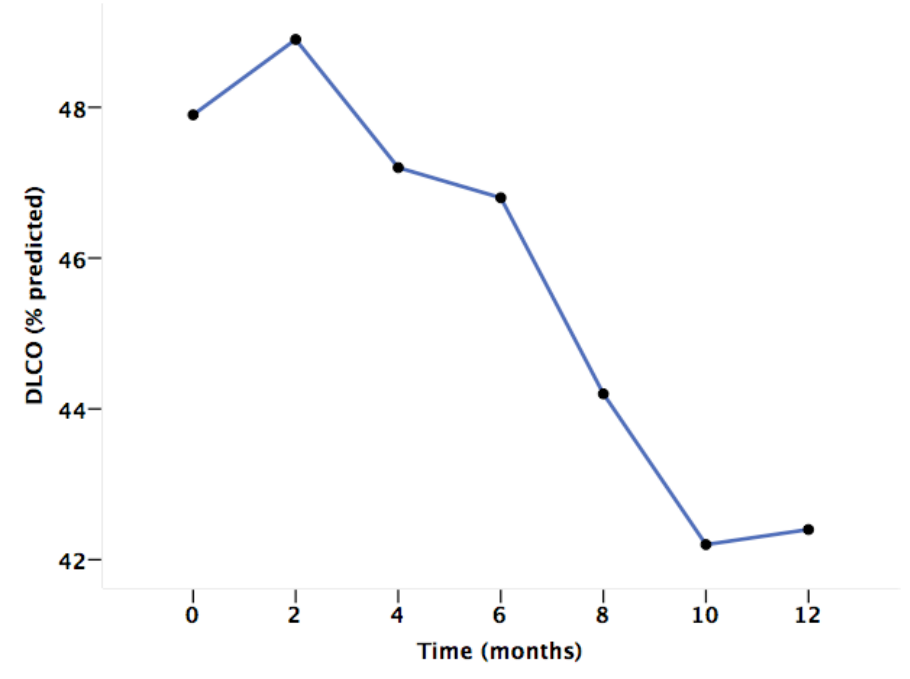


Figure 24 - Plot of estimated means of % predicted DLCO in the IPF group

5.4.4 Oxygen saturation

Measurements of resting oxygen saturation (SaO2) for the IPF group at different time points are reported in Table 43**Error! Reference source not found.**. On average, there was a reduction in SaO2 from 95.21% to 93.31% between the first and the last visit.

Table 43 - Longitudinal measurements of percentage of oxygen saturation (SaO2%) for the IPF group. Data presented as means ± standard deviations (SD), minimum (Min) and maximum (Max) values.

Visit	Observations	Mean(±SD)	Min	Max
0 (baseline)	19	95.21(±1.69)	92	97
2	18	95.22(±1.51)	93	98
4	15	95.47(±1.78)	91	98
6	15	94.8(±1.7)	90	97
8	14	94.07(±3.32)	88	99
10	11	94.82(±2.6)	88	97
12	13	93.31(±4.13)	84	97

ANOVA for repeated measures was performed over 9 available observations across all time points. Estimated means of SaO2% are reported in Table 44 and shown in Figure 25**Error! Reference source not found.**. Slight and non-significant differences in estimated means of SaO2% were found between the study time points (Appendix 2).

Mauchly's test was significant (p=0.028), as such sphericity couldn't be assumed; no one of the ANOVA tests of within-subject effects was significant (Table 45). As such, it can be concluded that SaO2% didn't change significantly in the study population over the observation period, and was not used in the subsequent correlation analysis with the acoustic features.

Table 44 - Estimated means for SaO2% in the IPF group. Data presented as means with standard error (SE) and 95% confidence intervals (95% CI)

Visit	Mean	SE	95% CI	
			Lower Bound	Upper Bound
1	95.44	0.5	94.28	96.6
2	95.22	0.52	94.02	96.42
3	96.11	0.39	95.21	97
4	94.78	0.68	93.2	96.36
5	95.44	0.78	93.64	97.25
6	94.89	0.96	92.67	97.11

7	93.56	1.75	90.16	96.95

Table 45 - Tests of within-subjects effects for SaO2% in the IPF group

Tests of Within-Subjects Effects

Measure: SaO2

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Visit	Sphericity Assumed	34.190	6	5.698	1.798	.119
	Greenhouse-Geisser	34.190	2.608	13.110	1.798	.183
	Huynh-Feldt	34.190	3.982	8.586	1.798	.154
	Lower-bound	34.190	1.000	34.190	1.798	.217
Error(Visit)	Sphericity Assumed	152.095	48	3.169		
	Greenhouse-Geisser	152.095	20.864	7.290		
	Huynh-Feldt	152.095	31.858	4.774		
	Lower-bound	152.095	8.000	19.012		

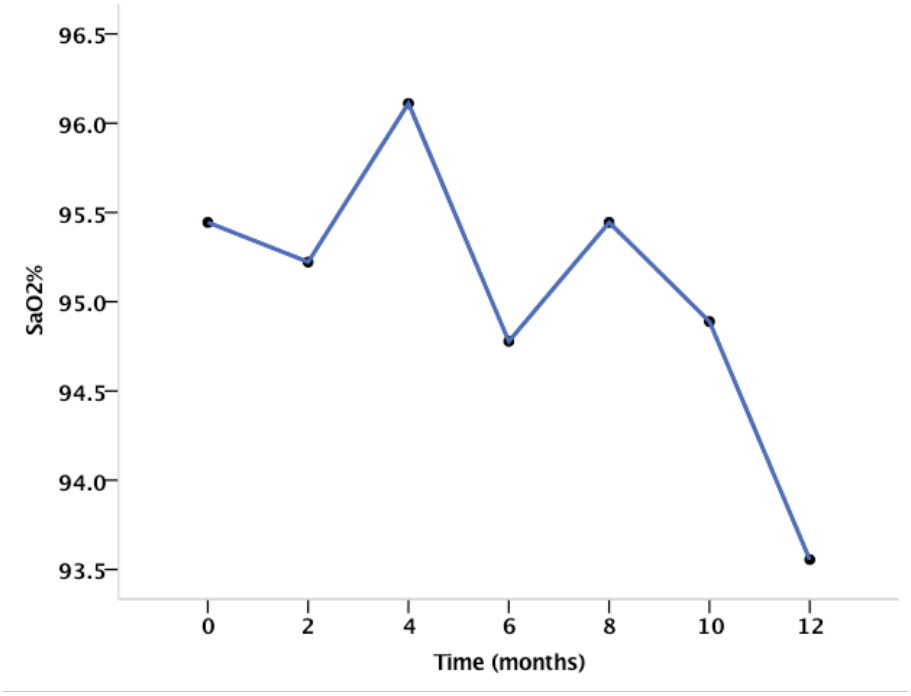


Figure 25 - Plot of estimated means of percentage of oxygen saturation (SaO2%) in the IPF group.

5.4.5 Tolerance to exercise

In the IPF group, tolerance to exercise was measured by calculating the distance covered at the 6-minute walk test (6-minute walk distance, 6MWD) and the average values for the 3 time points when 6MWT was performed (baseline, 6 and 12 months) are reported in Table 46 **Error! Reference source not found.**, together with the average scores at the BORG dyspnea and fatigue scales, and the percentage of saturation of oxygen (SaO2%), all recorded at rest and after the 6-minute walk test (6MWT) was performed.

At ANOVA for repeated measures performed over 11 available observations, a decrease in 6MWD was found between visit 4 and visit 6, indicating a decline over the last 6 months of the observation period (Figure 26). The decrease in 6MWD was 39 metres, slightly below the minimal clinically important difference value for 6MWD (>50 metres over 6 months) (du Bois et al., 2011c), and was not statistically significant (95% CI -43.267;136.904, p=0.5) (Appendix 2). The change in 6MWD observed over 12 months of follow-up, shown in the plot in **Error! Reference source not found.**, was not significant either at the tests of within-subjects (p=0.227, Table 47**Error! Reference source not found.**).

Mean differences at the scores at the BORG dyspnea and fatigue scale are reported separately in Table 48**Error! Reference source not found.**. The observed differences for both scales, compared using paired samples T test were significant at all study time points. Mean differences in percentages of oxygen saturation recorded at rest and after the 6-minute walk test are reported in Table 49**Error! Reference source not found.**. A significant desaturation (from -5.5% to -9.9%) was recorded at all study time points.

Table 46 - Longitudinal measurements of exercise tolerance for IPF group, including distance (metres) covered at the 6-minute walk test (6MWD), scores at BORG dyspnea and fatigue scales and percentage of saturation of oxygen (SaO2%) recorded before and after the test. Data presented as means ± standard deviation (SD), minimum (Min) and maximum (Max) values.

Months	Parameter	Observations	Mean(±SD)	Min	Max
0 (baseline)	6MWD	19	440.89(±117.4)	315	780
	BORG Dyspnea				
	Pre-walk	16	0.41(±0.58)	0	2
	Post-walk	16	2.16(±1.41)	0	5
	BORG Fatigue				
	Pre-walk	16	0.25(±0.58)	0	2
	Post-walk	16	1.47(±1.34)	0	5
	SaO2				
	Pre-walk	19	95.21(±1.69)	92	97
	Post-walk	19	85.32(±6.71)	72	97
6	6MWD	14	467.64(±110.29)	330	650
	BORG Dyspnea				
	Pre-walk	14	0.57(±0.55)	0	2
	Post-walk	14	2.82(±1.92)	0	7
	BORG Fatigue				
	Pre-walk	12	0.33(±0.62)	0	2
	Post-walk	12	1.7(±1.63)	0	5
	SaO2				
	Pre-walk	15	94.8(±1.7)	90	97
	Post-walk	14	85.21(±10.32)	60	97
12	6MWD	12	434.42(±117.97)	300	680
	BORG Dyspnea				
	Pre-walk	12	0.88(±0.86)	0	2
	Post-walk	12	2.63(±1.49)	0	5
	BORG Fatigue				
	Pre-walk	12	0.79(±1.18)	0	3
	Post-walk	11	2.27(±1.42)	0	5
	SaO2				
	Pre-walk	13	93.31(±4.13)	84	97
	Post-walk	12	87.92(±6.6)	77	97

Table 47 - Tests of within-subjects effects for 6-minute walk distance (6MWD) in the IPF group

Tests of Within-Subjects Effects

Measure: Distance

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	13838.242	2	6919.121	1.597	.227
	Greenhouse-Geisser	13838.242	1.359	10181.078	1.597	.235
	Huynh-Feldt	13838.242	1.499	9232.512	1.597	.234
	Lower-bound	13838.242	1.000	13838.242	1.597	.235
Error(factor1)	Sphericity Assumed	86642.424	20	4332.121		
	Greenhouse-Geisser	86642.424	13.592	6374.460		
	Huynh-Feldt	86642.424	14.989	5780.555		
	Lower-bound	86642.424	10.000	8664.242		

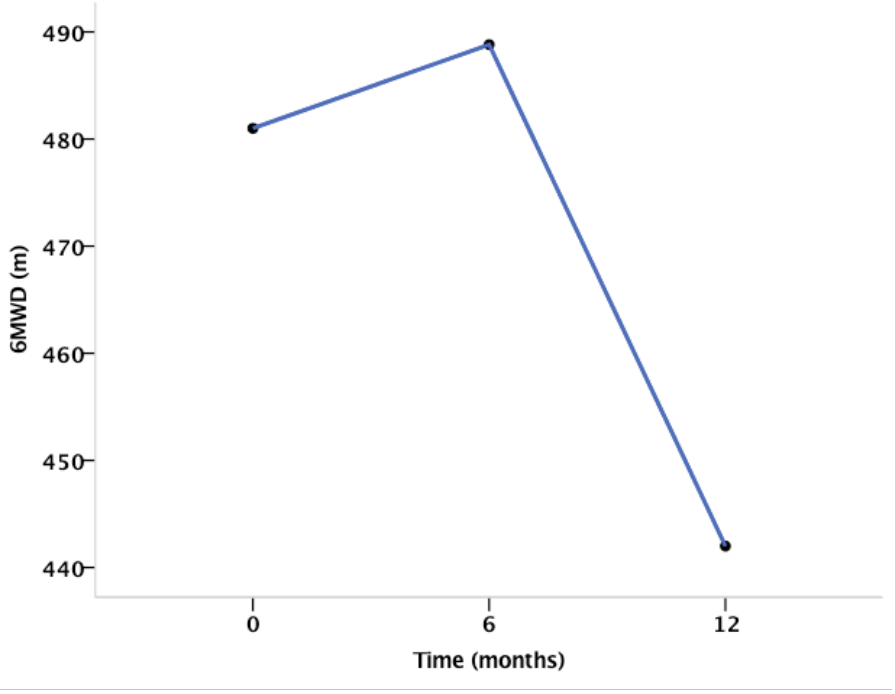


Figure 26 - Plot of estimated means of 6-minute walk distance (6MWD) in the IPF group. Values reported in metres (m).

Table 48 - Comparisons of mean scores at BORG dyspnea and fatigue scales before and after 6-minute walk test for different study time points in the IPF group, tested with. Data presented as mean differences with standard error (SE) and 95% confidence intervals (95% CI), and p value for paired samples T test.

Months	Parameter	Observations	Mean difference	SE	p	95% CI	
						Lower bound	Upper bound
0 (Baseline)	BORG Dyspnea Post-Pre	16	1.75	0.34	0.000	1.04	2.46
	BORG Fatigue Post-Pre	16	1.22	0.3	0.001	0.58	1.9
6	BORG Dyspnea Post-Pre	14	2.25	0.5	0.001	1.17	2.33
	BORG Fatigue Post-Pre	12	1.38	0.41	0.006	0.48	2.27
12	BORG Dyspnea Post-Pre	12	1.75	0.37	0.001	0.94	2.56
	BORG Fatigue Post-Pre	11	1.41	0.36	0.003	0.62	2.2

Table 49 - Comparisons of mean percentage of saturation of oxygen (SaO2%) before and after 6-minute walk test for different study time points in the IPF group, tested with. Data presented as mean differences with standard error (SE) and 95% confidence intervals (95% CI), and p value for paired samples T test.

Months	Parameter	Observations	Mean difference	Std. Error	p	95% CI	
						Lower bound	Upper bound
0 (Baseline)	SaO2	19	-9.9	1.28	0.000	-12.57	-7.22
6	SaO2	14	-9.93	2.64	0.002	-15.62	-4.23
12	SaO2	11	-5.64	1.17	0.001	-8.24	-3.03

5.4.6 Breathlessness and quality of life

In the IPF group, breathlessness was measured via the University of California San Diego – Shortness of Breath (UCSD-SOB) questionnaire. In this study, a "relative" score was calculated as fraction of the total score achievable based on the number of questions answered by the participant. This was necessary since some answers to the questions were missing and no method for imputing missing data was found in the literature for this questionnaire. The average relative scores for each time point are reported in Table 50**Error! Reference source not found.**. At ANOVA over 14 available observations, there were no significant differences between estimated means of scores at the UCSD-SOB between single time points, as shown in Appendix 2. The largest mean difference in the relative scores (0.15) was found between baseline and visit 7 over 14 available observations, but was not significant (95% CI –0.363;0.063, p=0.423). Nevertheless, a pretty steady and statistically significant increase in average scores at UCSD-SOB over 12 months of follow-up was found, as reported by the ANOVA tests of within-subject effects (p=0.001, **Error! Reference source not found.**) and shown in the plot in Figure 27**Error! Reference source not found.**.

Table 50 - Longitudinal measurements of breathlessness and quality of life expressed as relative scores (fraction of highest possible score) at the University of California San Diego – Shortness of Breath Questionnaire (UCSD-SOB) and total scores at the Saint George's Respiratory Questionnaire (SGRQ) for the IPF group. Data presented as means ± standard deviations (SD), minimum (Min) and maximum (Max) values.

Months	Parameter	Observations	Mean(±SD)	Min	Max
0 (baseline)	UCSD-SOB (relative) SGRQ (total)	19 19	0.17(±0.09) 35.02(±11.05)	0 15.12	0.33 51.14
2	UCSD-SOB (relative) SGRQ (total)	19 19	0.24(±0.16) 38.96(±14.97)	0.01 10.9	0.6 62.45
4	UCSD-SOB (relative) SGRQ (total)	17 16	0.25(±0.14) 39.9(±10.99)	0.04 16.29	0.55 53.9
6	UCSD-SOB (relative) SGRQ (total)	16 16	0.27(±0.17) 39.75(±12.94)	0.04 11.57	0.53 56.77
8	UCSD-SOB (relative) SGRQ (total)	16 15	0.32(±0.2) 47.39(±15.12)	0.03 17.25	0.75 68.23
10	UCSD-SOB (relative) SGRQ (total)	14 14	0.27(±0.19) 41.48(±11.09)	0.02 16.35	0.64 53.82
12	UCSD-SOB (relative) SGRQ (total)	14 14	0.32(±0.23) 43.45(±13.32)	0.03 11.71	0.89 61.58

Table 51 - Estimated means of relative scores at University of California San Diego - Shortness Of Breath Questionnaire (UCSD-SOBQ) in the IPF group. Data presented as means with standard error (SE) and 95% confidence intervals (95% CI).

Visit	Mean	SE	95% CI	
			Lower Bound	Upper Bound
1	0.17	0.02	0.12	0.21
2	0.22	0.04	0.13	0.3
3	0.25	0.04	0.16	0.34
4	0.26	0.05	0.16	0.36
5	0.31	0.06	0.18	0.43
6	0.27	0.05	0.16	0.38
7	0.32	0.06	0.18	0.45

Table 52 - Tests of within-subjects effects for relative scores at University of California San Diego - Shortness Of Breath Questionnaire (UCSD-SOBQ) in the IPF group

Tests of Within-Subjects Effects

Measure: SGRQ

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	1027.642	6	171.274	4.125	.001
	Greenhouse-Geisser	1027.642	3.152	326.022	4.125	.012
	Huynh-Feldt	1027.642	4.405	233.280	4.125	.004
	Lower-bound	1027.642	1.000	1027.642	4.125	.065
Error(factor1)	Sphericity Assumed	2989.813	72	41.525		
	Greenhouse-Geisser	2989.813	37.825	79.044		
	Huynh-Feldt	2989.813	52.862	56.559		
	Lower-bound	2989.813	12.000	249.151		

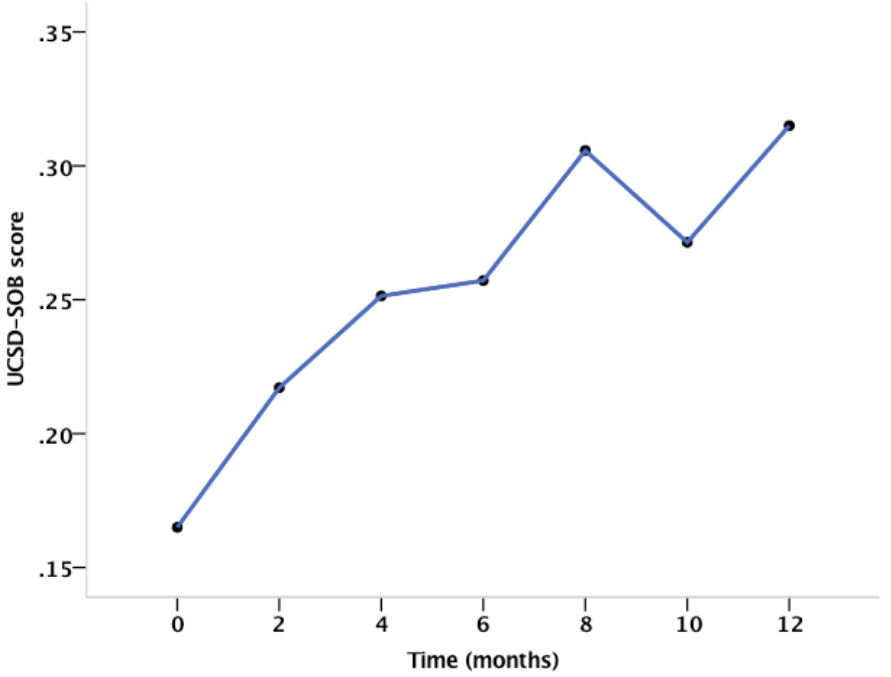


Figure 27 - Plot of estimated means of relative scores at the University of California San Diego – Shortness of Breath (UCSD-SOB) questionnaire for the IPF group.

Health status (quality of life) was investigated using the Saint George's Respiratory Questionnaire. The scores were calculated using a Microsoft Excel-based calculator provided with the questionnaire, which can automatically impute missing data to a certain extent. For this study, the total scores, were used in the analysis and are reported for all time points in Table 50**Error! Reference source not found.**. ANOVA was performed over 13 available measurements across all study time points. The estimated means of total scores are reported in Table 53**Error! Reference source not found.**; several statistically significant differences were found between single study time points, as shown in Appendix 2. The largest mean difference was found between baseline and visit 5 (corresponding to 8 months of observation) over 13 available observations (95% CI – 21.01; -0.52, p=0.035). A trend towards increment in total SGRQ scores was registered as shown in the plot in Figure 28**Error! Reference source not found.**, and was statistically significant at the ANOVA tests of within-subjects effects (p=0.01, Table 54**Error! Reference source not found.**).

Overall, these findings showed that the IPF population in this study experienced a significant deterioration in symptoms and quality of life, reflecting progression of the disease over the observation period.

Table 53 - Estimated means of relative scores at Saint George's Respiratory Questionnaire (SGRQ) in the IPF group. Data presented as means with standard error (SE) and 95% confidence intervals (95% CI).

Visit	Mean	SE	95% CI	
			Lower Bound	Upper Bound
1	34.84	3.17	27.93	41.76
2	38.07	4.3	28.71	47.44
3	39.08	3.2	32.1	46.06
4	37.73	3.57	29.95	45.5
5	45.6	4.18	36.51	54.71
6	41.22	3.19	34.28	48.18
7	43.13	3.83	34.79	51.48

Table 54 - Tests of within-subjects effects for total scores at Saint George's Respiratory Questionnaire (SGRQ) in the IPF group.

Tests of Within-Subjects Effects

Measure: SGRQ

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	1027.642	6	171.274	4.125	.001
	Greenhouse-Geisser	1027.642	3.152	326.022	4.125	.012
	Huynh-Feldt	1027.642	4.405	233.280	4.125	.004
	Lower-bound	1027.642	1.000	1027.642	4.125	.065
Error(factor1)	Sphericity Assumed	2989.813	72	41.525		
	Greenhouse-Geisser	2989.813	37.825	79.044		
	Huynh-Feldt	2989.813	52.862	56.559		
	Lower-bound	2989.813	12.000	249.151		

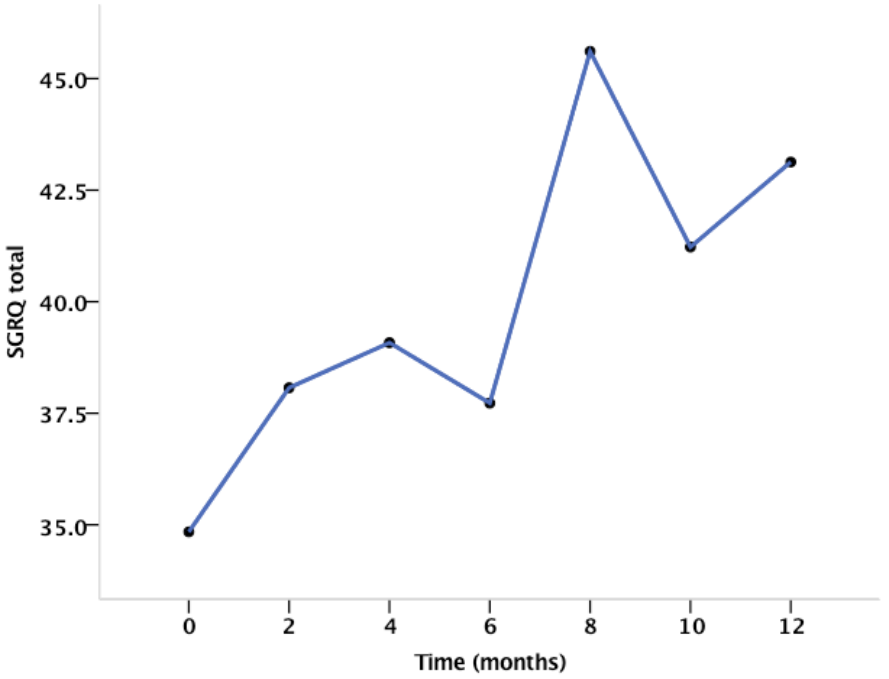


Figure 28 - Plot of estimated means of relative scores at the Saint George's Respiratory Questionnaire (SGRQ) for the IPF group.

5.4.7 Lung sounds data

1261 sound files for the IPF group and 557 files for the control group were analysed using the algorithms written in Matlab (version R2015a) as described in the Methodology chapter.

For each recording, 481 acoustic features were generated by the algorithms and automatically extracted onto a Microsoft Excel spreadsheet. For those sound files that were not processed by EMD at the first run of the algorithms, edited shorter bits were analysed separately. For each original sound file, one of these edited recordings was chosen at random and included in the analyses.

The files not processed via EMD in their full length were approximately 20% of the original data sets: 246 for the IPF group (19.5% of the original data set) and 146 for the healthy volunteers group (26.2 %). Table 55**Error! Reference source not found.** and Table 56**Error! Reference source not found.** show the distribution of these files across patients and recording sites for both study groups. The processing errors occurred more frequently in specific subjects, especially for the IPF group, although some patients withdrew from the study earlier and had therefore less recordings to be analysed. Among the patients who terminated the study (i.e. with a full data set of recordings available), no correlation was found between the frequency of "faulty" files and a more severe functional impairment at baseline (Spearman's correlation coefficient = 0.04, p=0.89), suggesting there was no relationship between the difficulty in processing the files sounds produced in a more advanced stage of the disease. When the frequency of faulty recordings was stratified per recording site, it was found higher for those sites corresponding to the lower lobes and lateral chest, in both IPF and control groups.
The potential issue of a consistent amount of missing data due to faulty processing was overcome by editing these files into shorter pieces of recording (approximately between 3 and 5 seconds in length), that the algorithms were able to process in the large majority of cases.

Table 55 - Recordings not processed via EMD per patient and per recording site in the IPF group. Data presented as counts (Number) and percentages (%). *=patient withdrawn from study. RUL=right upper lobe; RML=right middle lobe; RLL=right lower lobe; LUL=left upper lobe; LML=left middle lobe; LLL=left lower lobe; RLat=right lateral; LLat=left lateral; RAA=right apex anterior; LAA=left apex anterior.

Patient	Number	%	Recording site	Number	%
S-001	21	8.5	RUL	12	4.9
S-002	21	8.5	RML	14	5.7
S-003	8	3.3	RLL	41	16.7
S-004	12	4.9	LUL	13	5.3
S-005	5	2.0	LML	13	5.3
S-006	21	8.5	LLL	45	18.3
S-007	6	2.4	RLat	42	17.1
S-008*	3	1.2	LLat	45	18.3
S-009*	1	.4	RAA	9	3.7
S-010*	9	3.7	LAA	12	4.9
S-011*	0	0			
S-012	9	3.7			
S-013	40	16.3			
S-014*	6	2.4			
S-015	10	4.1			
S-016	22	8.9			
S-017	9	3.7			
S-018	9	3.7			
S-019	34	13.8			
Total	246	100.0	Total	246	100.0

Table 56 - Recordings not processed via EMD per patient and recording site in the control group (healthy volunteers). Data presented as counts (Number) and percentages (%). *=patient withdrawn from study. RUL=right upper lobe; RML=right middle lobe; RLL=right lower lobe; LUL=left upper lobe; LML=left middle lobe; LLL=left lower lobe; RLat=right lateral; LLat=left lateral; RAA=right apex anterior; LAA=left apex anterior.

Patient ID	Number	%	Recording site	Number	%
S-001	19	13.0	RUL	8	5.5
S-002	7	4.8	RML	11	7.5
S-003*	3	2.1	RLL	26	17.8
S-004	28	19.2	LUL	6	4.1
S-005	21	14.4	LML	6	4.1
S-006*	0	0	LLL	25	17.1
S-007	18	12.3	RLat	25	17.1
S-008	30	20.5	LLat	25	17.1
S-009*	5	3.4	RAA	5	3.4
S-010	15	10.3	LAA	9	6.2
Total	146	100.0	Total	146	100.0

The Excel spreadsheets reporting in full the values of the features extracted from both original and edited recordings have been included in the accompanying material of this thesis in appendix 2 (folder “Spreadsheets”).

The descriptive statistics for the 481 variables representing the acoustic features, stratified per study time points and recording sites, are presented in full in the accompanying material of this thesis in Appendix 3, for IPF and control group in separate subfolders. **Error! Reference source not found.** Table 57 below reports an example of descriptives for a set of 19 features (those found to be the most repeatable, see section) of lung sounds recorded over the right lower lobe at the 7 study time points in the IPF group.

Table 57 – Example of descriptives of selected acoustic features recorded over the right lower lobe in the IPF group. Data presented as means with standard deviations (SD) and minimum (Min and maximum (Max) values.

Visit	Feature	N	Min	Max	Mean	SD
	C3 EW_200_500Hz	19	.0558748300000000	.4622684000000000	.264131814000000	.125359462000000
	C4 EW_75_200Hz	19	.3126714780000000	.6905769020000000	.517346239000000	.100912663000000
	C4 EW_200_500Hz	19	.0208150600000000	.4399643700000001	.213872330000000	.111416912000000
	sig_zerocross	19	.0248483200000000	.0563750000000000	.041838035900000	.009428008720000
	sig_mfcc02	19	-11.046915900000000	-8.919270457000001	-9.891270587000001	.467572087000000
	sig_75_200Hz_zerocross	19	.0240625000000000	.0298375000000000	.026672384800000	.001512926200000
	sig_75_200Hz_centroid	19	108.99286529999999	132.25523479999998	122.497370300000000	6.515853403000000
	sig_200_500Hz_rms	19	.0673488680000000	.5479223450000000	.298695050000000	.129713276000000
	sig_200_500Hz_lowenergy	19	.5395683450000001	.8160000000000001	.686462907000000	.063102887900000
1	sig_200_500Hz_lowenergyASR	19	.5395683450000001	.8800000000000000	.706088951000000	.087448470100000
	sig_200_500Hz_zerocross	19	.0689890240000000	.0853125000000000	.077513576800000	.004644161830000
	sig_200_500Hz_std_meanframes	19	.0053201956600000	.0422821930000000	.014983318100000	.008008481530000
	sig_200_500Hz_std_medianframes	19	.0013972596900000	.0092824532000000	.003580338160000	.002099336130000
	sig_500_1000Hz_zerocross	19	.1420500000000000	.1513428810000000	.147513589000000	.002281964060000
	sig_500_1000Hz_rolloff85	19	685.180664100000	791.259765600000	750.36942330000000	27.884610080000000
	sig_500_1000Hz_centroid	19	597.9415931999999	644.2011201000000	626.350560099999900	12.867643210000000
	C3_mfcc02	19	-11.047493640000000	-8.946565627000000	-9.900612639000000	.465554061000000
	C4_zerocross	19	.0251891740000000	.0556250000000000	.042665444700000	.009335766280000
	C4_mfcc02	19	-11.046928200000000	-8.922181400000001	-9.891363439000001	.466940330000000

Visit	Feature	N	Min	Max	Mean	SD
	C3 EW_200_500Hz	19	.0235505950000000	.4218277230000000	.267235804000000	.122173392000000
	C4 EW_75_200Hz	19	.2868267290000000	.6290060540000001	.510942265000000	.094908743200000
	C4 EW_200_500Hz	19	.0149006710000000	.4366222930000001	.212884631000000	.109992198000000
	sig_zerocross	19	.0222693210000000	.0619625499000000	.042082790600000	.010117082400000
	sig_mfcc02	19	-10.915512330000000	-9.441148041000000	-10.000937960000000	.392155119000000
	sig_75_200Hz_zerocross	19	.0227607810000000	.0277763844000000	.026369558500000	.001284713560000
	sig_75_200Hz_centroid	19	103.71003970000000	131.98031010000000	122.796790899999990	7.130455741000001
	sig_200_500Hz_rms	19	.0508887810000000	.5752950960000001	.306621737000000	.126038047000000
	sig_200_500Hz_lowenergy	19	.6149068320000001	.7707006370000000	.701717368000000	.041329049800000
2	sig_200_500Hz_lowenergyASR	19	.6086956520000001	.9312500000000000	.751212460000000	.085078882600000
	sig_200_500Hz_zerocross	19	.0727362080000000	.0883521311000000	.079714683800000	.004334656250000
	sig_200_500Hz_std_meanframes	19	.0045396040000000	.0576097720000000	.015912036100000	.011883170300000
	sig_200_500Hz_std_medianframes	19	.0008994290620000	.0064782017600000	.002680748210000	.001431953040000
	sig_500_1000Hz_zerocross	19	.1430557550000000	.1509310940000000	.147394485000000	.001932489550000
	sig_500_1000Hz_rolloff85	19	693.115234400000	796.386718800000	748.13360919999990	32.014782320000000
	sig_500_1000Hz_centroid	19	595.7417177000000	650.6656254999999	627.660501900000000	16.302002100000000
	C3_mfcc02	19	-10.933743370000000	-9.452896643000000	-10.015216990000000	.390584408000000
	C4_zerocross	19	.0219621580000000	.0770980000000000	.045095689600000	.012870638700000
	C4_mfcc02	19	-10.916020880000000	-9.442221805000000	-10.003067890000000	.391373903000000
	C3 EW_200_500Hz	17	.1302929620000000	.4596395130000001	.259215777000000	.095957129300000
	C4 EW_75_200Hz	17	.1656896140000000	.6530177840000000	.483883397000000	.128900304000000
	C4 EW_200_500Hz	17	.0879709590000000	.4870029410000000	.220528953000000	.100426689000000
	sig_zerocross	17	.0329500000000000	.0604293699000000	.042153907600000	.007868800410000
	sig_mfcc02	17	-11.575471410000000	-9.539971130000001	-10.379230720000000	.534657646000000
	sig_75_200Hz_zerocross	17	.0252808989000000	.0275990770000000	.026667749600000	.000798924957000
	sig_75_200Hz_centroid	17	116.09417730000000	135.29847270000000	124.608465700000000	5.913992212000000
	sig_200_500Hz_rms	17	.1570694200000000	.5337867250000000	.307662470000000	.104384660000000
	sig_200_500Hz_lowenergy	17	.6814621410000000	.8316326530000000	.731857357000000	.043312379000000
3	sig_200_500Hz_lowenergyASR	17	.6890547260000001	.9336734690000000	.775538351000000	.063601653100000
	sig_200_500Hz_zerocross	17	.0691273070000000	.0893050945000000	.078737149500000	.005353248220000
	sig_200_500Hz_std_meanframes	17	.0066646472600000	.0421581890000000	.018499332300000	.009652880380000
	sig_200_500Hz_std_medianframes	17	.0012544433000000	.0044252510000000	.002569489080000	.000966945752000
	sig_500_1000Hz_zerocross	17	.1425875000000000	.1509860700000000	.147184587000000	.002545290290000
	sig_500_1000Hz_rolloff85	17	694.061279300000	837.646484400000	751.89208990000000	31.557553420000000
	sig_500_1000Hz_centroid	17	603.1510150000000	678.6299101999999	632.659317300000000	18.465716900000000
	C3_mfcc02	17	-11.576651450000000	-9.543175397000000	-10.400597120000000	.524010930000000
	C4_zerocross	17	.0332750000000000	.0830857050000000	.047927487400000	.012573179100000
	C4_mfcc02	17	-11.575495650000000	-9.540258481000000	-10.380219770000000	.532430100000000

Visit	Feature	N	Min	Max	Mean	SD
4	C3 EW_200_500Hz	16	.0447356296000000	.4049891950000000	.252179668000000	.105865784000000
	C4 EW_75_200Hz	16	.3070032130000000	.7007776320000001	.538186459000000	.103936539000000
	C4 EW_200_500Hz	16	.0385494122000000	.3650172110000000	.190419424000000	.098251231100000
	sig_zerocross	16	.0262250000000000	.0598528530000000	.041783877000000	.009965169110000
	sig_mfcc02	16	-11.497758420000000	-9.339541119000000	-10.559275320000000	.568254537000000
	sig_75_200Hz_zerocross	16	.0251031637000000	.0285785288000000	.026774653400000	.000952728731000
	sig_75_200Hz_centroid	16	115.17505460000000	132.43916389999998	124.601769100000000	5.105373511000001
	sig_200_500Hz_rms	16	.1507591310000000	.4252701190000000	.301028538000000	.085128889900000
	sig_200_500Hz_lowenergy	16	.5507246380000000	.9276859500000001	.725029929000000	.095015171800000
	sig_200_500Hz_lowenergyASR	16	.5507246380000000	.9545454550000001	.756849878000000	.117544770000000
	sig_200_500Hz_zerocross	16	.0572000000000000	.0843148020000000	.076185257000000	.006906662300000
	sig_200_500Hz_std_meanframes	16	.0066840914300000	.0325731080000000	.017326714500000	.009053019870000
	sig_200_500Hz_std_medianframes	16	.0007340327910000	.0128523830000000	.003396530790000	.002785296490000
	sig_500_1000Hz_zerocross	16	.1455987350000000	.1527178810000000	.148802092000000	.002068548070000
	sig_500_1000Hz_rolloff85	16	659.423828100000	830.566406300000	758.45909120000000	38.431043890000000
	sig_500_1000Hz_centroid	16	586.1564245000000	668.5845164000000	633.578641600000000	19.502063890000000
	C3_mfcc02	16	-11.513722600000000	-9.356406248000000	-10.570412000000001	.573240828000000
	C4_zerocross	16	.0259750000000000	.0615679060000000	.042992337300000	.010046782800000
	C4_mfcc02	16	-11.497604170000000	-9.341829402000000	-10.558437930000000	.568855359000000
5	C3 EW_200_500Hz	16	.1125574910000000	.3944220210000000	.264930731000000	.087188833500000
	C4 EW_75_200Hz	16	.2868355400000000	.7402653090000001	.506247980000000	.119098791000000
	C4 EW_200_500Hz	16	.0512453782000000	.3585245360000000	.221177733000000	.091088330100000
	sig_zerocross	16	.0274312317000000	.0579916667000000	.043504248700000	.008861060020000
	sig_mfcc02	16	-11.240351430000000	-9.334481973000000	-10.450674560000000	.463036528000000
	sig_75_200Hz_zerocross	16	.0253288120000000	.0301512320000000	.027093472100000	.001178699200000
	sig_75_200Hz_centroid	16	111.68684089999999	129.84368360000000	121.650312000000000	6.057675585000000
	sig_200_500Hz_rms	16	.1046244530000000	.3897484620000000	.268332203000000	.092613602600000
	sig_200_500Hz_lowenergy	16	.5815972220000001	.8606896550000000	.715340843000000	.082643481100000
	sig_200_500Hz_lowenergyASR	16	.5572916670000000	.9558620690000000	.756795076000000	.113708617000000
	sig_200_500Hz_zerocross	16	.0646664110000000	.0844709268000000	.076808935700000	.005708646190000
	sig_200_500Hz_std_meanframes	16	.0046175828900000	.0220504499000000	.014201026000000	.006287769090000
	sig_200_500Hz_std_medianframes	16	.0007601266890000	.0056611987200000	.002469795730000	.001480817280000
	sig_500_1000Hz_zerocross	16	.1453315340000000	.1520801380000000	.149246561000000	.002144420110000
	sig_500_1000Hz_rolloff85	16	682.861328100000	809.234619100000	758.26644900000000	36.862163760000000
	sig_500_1000Hz_centroid	16	603.0460223000000	656.8779191000000	637.713060699999900	15.386013860000000
	C3_mfcc02	16	-11.298155360000000	-9.306110241000000	-10.462740850000000	.472875622000000
	C4_zerocross	16	.0281995497000000	.0592624705000000	.045523346800000	.009372030270000
	C4_mfcc02	16	-11.241132000000000	-9.322974075000001	-10.448631620000000	.466197297000000

Visit	Feature	N	Min	Max	Mean	SD
	C3 EW_200_500Hz	14	.0431345411000000	.3965871410000000	.238014926000000	.096448394700000
	C4 EW_75_200Hz	14	.2584321240000000	.6906578460000000	.544076498000000	.113559159000000
	C4 EW_200_500Hz	14	.0480702459000000	.3331924290000000	.176353656000000	.093152261600000
	sig_zerocross	14	.0268179658000000	.0519559443000000	.037157005500000	.006817763400000
	sig_mfcc02	14	-11.564647050000000	-9.741830401000001	-10.419649280000000	.545104122000000
	sig_75_200Hz_zerocross	14	.0239437936000000	.0281570254000000	.026407549100000	.001300793030000
	sig_75_200Hz_centroid	14	114.42835009999999	134.20412869999998	122.768138999999990	6.069467901000000
	sig_200_500Hz_rms	14	.0958863181000000	.5163249080000001	.246866230000000	.111377659000000
	sig_200_500Hz_lowenergy	14	.6030150750000001	.8481806780000001	.701958796000000	.073607486700000
6	sig_200_500Hz_lowenergyASR	14	.5841584160000001	.9461966600000000	.747429039000000	.101122529000000
	sig_200_500Hz_zerocross	14	.0715598031000000	.0833592454000000	.076472500100000	.003597491570000
	sig_200_500Hz_std_meanframes	14	.0032133237400000	.0289143350000000	.013597461400000	.007096586870000
	sig_200_500Hz_std_medianframes	14	.0008644800290000	.0054202040000000	.002490112670000	.001490880850000
	sig_500_1000Hz_zerocross	14	.1454290220000000	.1526210310000000	.149575498000000	.002421434630000
	sig_500_1000Hz_rolloff85	14	716.125488300000	809.143066400000	771.09854560000000	25.682910260000000
	sig_500_1000Hz_centroid	14	616.9382405000000	663.6303846000000	641.994138400000000	13.889547770000000
	C3_mfcc02	14	-11.577691700000000	-9.798231501000000	-10.429241070000000	.542785724000000
	C4_zerocross	14	.0272694867000000	.0534595529000000	.039862294500000	.007987799550000
	C4_mfcc02	14	-11.564418090000000	-9.752126229000000	-10.420358950000000	.544090707000000
	C3 EW_200_500Hz	14	.0106196957000000	.3828882220000000	.204975664000000	.113059791000000
	C4 EW_75_200Hz	14	.3997073040000000	.7001299260000000	.571747896000000	.079232131400000
	C4 EW_200_500Hz	14	.0120021050000000	.3321810940000000	.157213136000000	.091727152500000
	sig_zerocross	14	.0223072862000000	.0495369430000000	.036188430100000	.008058919390000
	sig_mfcc02	14	-12.172846640000000	-9.986966139000000	-10.467198500000000	.571184116000000
	sig_75_200Hz_zerocross	14	.0237844610000000	.0284044833000000	.026269993200000	.001457492050000
	sig_75_200Hz_centroid	14	112.46990219999999	126.50485219999999	119.443412099999990	4.265655077000000
	sig_200_500Hz_rms	14	.0752193140000000	.4648535440000000	.250295072000000	.098840048100000
	sig_200_500Hz_lowenergy	14	.6089108910000001	.8132231400000001	.701204400000000	.066895787500000
7	sig_200_500Hz_lowenergyASR	14	.4752475250000000	.9239669420000001	.725878129000000	.120542657000000
	sig_200_500Hz_zerocross	14	.0563044483000000	.0776018722000000	.073679913700000	.005524378670000
	sig_200_500Hz_std_meanframes	14	.0049942272800000	.0340532480000000	.014603189700000	.008486951470000
	sig_200_500Hz_std_medianframes	14	.0008027102890000	.0066474890000000	.003351096810000	.002099134340000
	sig_500_1000Hz_zerocross	14	.1448241160000000	.1504219720000000	.147673827000000	.001370552660000
	sig_500_1000Hz_rolloff85	14	697.357177700000	799.682617200000	766.34434290000000	30.618569080000000
	sig_500_1000Hz_centroid	14	608.0697474000000	659.3555372000000	636.107928100000000	14.736685760000000
	C3_mfcc02	14	-12.192118730000000	-9.990508057000001	-10.474084880000000	.572462409000000
	C4_zerocross	14	.0222082893000000	.0611897010000000	.038229675300000	.010157203600000
	C4_mfcc02	14	-12.173034500000000	-9.987750256000000	-10.467198200000000	.571062238000000

5.4.7.1 Intra-subject reliability

The intra-subject reliability analysis was performed to identify those features that were most stable when recording lung sounds using a digital stethoscope. 34 observations were evaluated, consisting of recordings taken from different sites on the chest of 4 IPF subjects at 3 repeated measurements. 19 features were found to have ICC > 0.5 and therefore deemed to have acceptable repeatability, and are shown in Table 58. Most of these features were derived from the original signal, while a few were among those originated from Empirical Mode Decomposition (EMD). The full list of features with the corresponding ICC values are reported in the Appendix 3 of the accompanying material of this thesis.

Table 58 - Acoustic features showing reliability at 3 repeated measurements. Data presented as ICC (Intra Class Correlation) with 95% confidence intervals (95% CI) and relative p value. Hz = Hertz.

Feature name	Source	Description	ICC	95% CI		p value
				Lower bound	Upper bound	
C3_EW_200_500Hz	EMD	energy weight (EW) of the crackle component (N=3) in the specified frequency bands	0.66	0.50	0.79	0.0000
C4_EW_75_200Hz	EMD	energy weight (EW) of the crackle component (N=4) in the specified frequency bands	0.54	0.35	0.70	0.0000
C4_EW_200_500Hz	EMD		0.72	§§§	0.83	0.0000
sig_zerocross	Original	Zero-cross of the original signal	0.73	0.58	0.84	0.0000
sig_mfcc02	Original	2nd from the 13 MFCC of the original signal	0.59	0.41	0.74	0.0000
sig_75_200Hz_zerocross	Original	zero-cross of the original signal in the [75-200] Hz frequency range	0.59	0.42	0.75	0.0000
sig_75_200Hz_centroid	Original	Centroid of the original signal in the [75-200] Hz frequency range	0.54	0.35	0.71	0.0000
sig_200_500Hz_rms	Original	RMS of the original signal in the [200-500] Hz frequency range	0.54	0.35	0.70	0.0000
sig_200_500Hz_lowenergy	Original	Low Energy of the original signal in the [200-500] Hz frequency range	0.55	0.36	0.71	0.0000
sig_200_500Hz_lowenergyASR	Original	Average Silence Ratio of the original signal in the [200-500] Hz frequency range	0.52	0.33	0.69	0.0000
sig_200_500Hz_zerocross	Original	Zero-cross of the original signal in the [200-500] Hz frequency range	0.60	0.42	0.75	0.0000
sig_200_500Hz_std_meanframes	Original	mean of the frames of the original signal in the [200-500] Hz frequency range (mean, std, median)	0.53	0.34	0.70	0.0000
sig_200_500Hz_std_medianframes	Original	median of the frames of the original signal in the [200-500] Hz frequency range (mean, std, median)	0.54	0.35	0.70	0.0000
sig_500_1000Hz_zerocross	Original	Zero-cross of the original signal in the [500-1000] Hz frequency range	0.50	0.31	0.68	0.0000
sig_500_1000Hz_rolloff85	Original	Roll-off (threshold=85%) of the original signal in the [500-1000] Hz frequency range	0.52	0.33	0.69	0.0000
sig_500_1000Hz_centroid	Original	Centroid of the original signal in the [500-1000] Hz frequency range	0.59	0.41	0.74	0.0000
C3_mfcc02	EMD	2nd of the 13 MFCC of the crackle component (N=3)	0.59	0.41	0.75	0.0000
C4_zerocross	EMD	Zero-cross of the crackle component (N=4)	0.61	0.44	0.76	0.0000
C4_mfcc02	EMD	13 MFCC of the crackle component (N=4)	0.59	0.41	0.74	0.0000

5.4.7.2 Discrimination between IPF and controls

IPF patients and healthy controls were compared based on the 19 acoustic features that showed good reliability. Such analysis was important to understand whether these were features could be considered distinctive of IPF. The measurements taken from the 10 sites of recording at all study time points were entered in this analysis, for a total of 1261 observations for the IPF group and 557 observations for the control group.

None of the acoustic features followed a normal distribution, as reported by the Shapiro-Wilk test which was significant for all variables (**Error! Reference source not found.**Table 59). Nevertheless, the distribution of the residuals in the ANCOVA model could be considered as normal, which justified the choice of this parametric test.

Table 59 - Shapiro-Wilk test performed for the 19 repeatable acoustic features.

	Statistic	P value
C3 EW_200_500Hz	0.982	0.000
C4 EW_75_200Hz	0.950	0.000
C4 EW_200_500Hz	0.949	0.000
sig_zerocross	0.933	0.000
sig_mfcc02	0.988	0.000
sig_75_200Hz_zerocross	0.998	0.049
sig_75_200Hz_centroid	0.973	0.000
sig_200_500Hz_rms	0.959	0.000
sig_200_500Hz_lowenergy	0.994	0.000
sig_200_500Hz_lowenergyASR	0.981	0.000
sig_200_500Hz_zerocross	0.991	0.000
sig_200_500Hz_std_meanframes	0.871	0.000
sig_200_500Hz_std_medianframes	0.886	0.000
sig_500_1000Hz_zerocross	0.964	0.000
sig_500_1000Hz_rolloff85	0.992	0.000
sig_500_1000Hz_centroid	0.994	0.000
C3_mfcc02	0.988	0.000
C4_zerocross	0.909	0.000
C4_mfcc02	0.988	0.000

At the multivariate analysis of covariance (ANCOVA), the group showed to have a significant influence on the values of the examined features, indicating that there was a significant effect of presence of the fibrotic disease on the combination of the 19 acoustic features ($F=11.837$, $p=0.000$, Wilks' lambda=0.889).

When each of the dependent variables was examined individually, the model (corrected per age, sex, BMI, time point and recording site entered as random covariates) showed a significant effect of the study group factor on each of the features, as reported in the tests of between-subjects shown below (**Error! Reference source not found.**Table 60). Cepstral features such as Sig_mfcc02, C3_mfcc02 and C4mfcc02 and zero-cross rate of the signal (sig_zerocross) and its 500-1000 Hz component (sig_500_1000Hz_zerocross) were the most influenced by the independent factor (study group), as demonstrated by the highest R squared values, which represents the amount of variation of each dependent variable that can be accounted for the independent factor.

Table 60 - Tests of between-subjects effects for the 19 repeatable acoustic features. F= ANOVA test statistic.

Feature name	F	p	R squared
C3 EW_200_500Hz	13.723	0.000	0.04
C4 EW_75_200Hz	32.795	0.000	0.095
C4 EW_200_500Hz	37.728	0.000	0.108
sig_zerocross	43.457	0.000	0.123
sig_mfcc02	69.283	0.000	0.184
sig_75_200Hz_zerocross	29.750	0.000	0.087
sig_75_200Hz_centroid	10.905	0.000	0.032
sig_200_500Hz_rms	17.591	0.000	0.052
sig_200_500Hz_lowenergy	32.521	0.000	0.094
sig_200_500Hz_lowenergyASR	28.893	0.000	0.084
sig_200_500Hz_zerocross	35.863	0.000	0.103
sig_200_500Hz_std_meanframes	12.625	0.000	0.037
sig_200_500Hz_std_medianframes	17.166	0.000	0.051
sig_500_1000Hz_zerocross	40.412	0.000	0.115
sig_500_1000Hz_rolloff85	29.328	0.000	0.086
sig_500_1000Hz_centroid	23.383	0.000	0.069
C3_mfcc02	68.973	0.000	0.183
C4_zerocross	20.972	0.000	0.062
C4_mfcc02	69.093	0.000	0.184

Estimated means of the variables were also calculated for both groups based on the multivariate ANCOVA model, and the mean difference values between groups are reported in Table 61**Error! Reference source not found.**. The estimated mean differences were found significant for several acoustic features (indicated with an asterisk in the table). Overall, these findings point out that most of the features that showed to be repeatable could effectively discriminate IPF subjects from healthy controls.

Table 61 - Estimated mean differences between IPF and control group for the examined acoustic features. Data reported together with standard error (SE), p value and 95% Confidence Intervals (CI) for difference. * = the mean difference is significant at the 0.05 level.

Feature	Mean Difference (IPF-Controls)	SE	p	95% CI for Difference	
				Lower Bound	Upper Bound
C3 EW_200_500Hz	0.006	0.008	0.454	-0.009	0.021
C4 EW_75_200Hz	-0.044*	0.007	0.000	-0.057	-0.031
C4 EW_200_500Hz	0.024*	0.006	0.000	0.012	0.036
sig_zerocross	0.003*	0.001	0.000	0.002	0.004
sig_mfcc02	0.193*	0.040	0.000	0.114	0.272
sig_75_200Hz_zerocross	0.000*	0.000	0.000	0.000	0.000
sig_75_200Hz_centroid	-0.701	0.393	0.075	-1.471	0.070
sig_200_500Hz_rms	0.014*	0.006	0.022	0.002	0.027
sig_200_500Hz_lowenergy	-0.013*	0.005	0.017	-0.024	-0.002
sig_200_500Hz_lowenergyASR	-0.015	0.009	0.088	-0.032	0.002
sig_200_500Hz_zerocross	0.002*	0.000	0.000	0.002	0.003
sig_200_500Hz_std_meanframes	0.002*	0.000	0.001	0.001	0.002
sig_200_500Hz_std_medianframes	0.000	0.000	0.327	0.000	0.000
sig_500_1000Hz_zerocross	0.000	0.000	0.293	0.000	0.001
sig_500_1000Hz_rolloff85	4.584	2.665	0.086	-0.643	9.811
sig_500_1000Hz_centroid	3.987*	1.308	0.002	1.421	6.553
C3_mfcc02	0.195*	0.040	0.000	0.116	0.274
C4_zerocross	0.002*	0.001	0.014	0.000	0.003
C4_mfcc02	0.194*	0.040	0.000	0.115	0.273

5.4.7.3 Longitudinal changes in acoustic features in IPF

The next step consisted in investigating which acoustic features, among the most repeatable, significantly changed in the IPF group over the 12-month observation period of the study. This analysis could provide insights into the type of features that might be reflect disease progression.

ANOVA for repeated measures was used to assess the influence of time on the sound features, used as dependent variables. Recording sites were used as a random covariate in the model, together with age, sex and BMI of the patients, also entered for further adjustment. Just as for the comparison between IPF patients and healthy controls, the choice of a parametric test such as ANOVA was justified by the normal distribution of the residuals in the model.

Since no data imputation was performed, the analysis was performed over 132 available observations across the 7 study time points. **Error! Reference source not found.**Table 62 reports the F statistic values and p values of ANOVA tests of within-subjects effects for each variable. Only 4 features (sig_75_200Hz_centroid, sig_200_500Hz_std_meanframes, sig_200_500Hz_std_medianframes and sig_500_1000Hz_rolloff85), did not present significant change over time in the IPF group.

Table 62 - Tests of within-subjects effects for repeatable acoustic features in the IPF group.

Feature	F	P value
C3 EW_200_500Hz	5.803	0.000
C4 EW_75_200Hz	2.877	0.013
C4 EW_200_500Hz	3.853	0.001
sig_zerocross	4.552	0.000
sig_mfcc02	4.579	0.000
sig_75_200Hz_zerocross	2.973	0.009
sig_75_200Hz_centroid	1.362	0.233
sig_200_500Hz_rms	2.974	0.007
sig_200_500Hz_lowenergy	3.009	0.007
sig_200_500Hz_lowenergyASR	2.366	0.028
sig_200_500Hz_zerocross	3.813	0.001
sig_200_500Hz_std_meanframes	1.102	0.359
sig_200_500Hz_std_medianframes	1.035	0.398
sig_500_1000Hz_zerocross	4.964	0.000
sig_500_1000Hz_rolloff85	1.318	0.247
sig_500_1000Hz_centroid	2.391	0.027
C3_mfcc02	4.661	0.000
C4_zerocross	3.934	0.001
C4_mfcc02	4.569	0.000

Further information was obtained by calculating the estimated means of the variables for each time point and the difference of such means between baseline (visit 1) and end of study (visit 7), which are reported in Table 63. Actually, only 6 features significantly changed between the first and the last visit: Sig_zerocross (p=0.001), Sig_mfcc002 (p=0.000), sig_500_1000Hz_zerocross (p=0.000), sig_500_1000Hz_centroid (p=0.000), C3_mfcc02 (p=0.000) and C4_mfcc02 (p=0.000). sig_500_1000Hz_rolloff85 also changed significantly between beginning and end of study (p=0.000) but its overall change was not found significant, as reported above.

The plots in Figure 29 further show that not all the features showing a significant change over the 12-month period followed a clear trend towards increase or decrease of their values, as in the case of C3_EW_200_500 Hz, C4_EW_75_200 Hz, sig_200_500Hz_zerocross, sig_200_500Hz_lowenergy and sig_200_500Hz_lowenergyASR. The features showing significant mean differences between the beginning and the end of the study followed quite a steady change, however their values still showed an inversion at a certain point, usually occurring between 8 and 10 months of observation.

Looking at pairwise comparisons of estimated means of features between single time points (reported in Appendix 3 of the accompanying material of this thesis), a few other features demonstrated to change significantly between baseline and visit 6 (C4 EW_200_500Hz, Sig_75_200_ Hz_zerocross, sig_200_500Hz_rms) or between baseline and visit 5 (sig_75_200Hz_centroid).

Table 63 - Estimated means of acoustic features in the IPF group across study time points. Means are reported together with standard error (SE) and 95% confidence intervals (95% CI). Estimated means difference between baseline (visit 1) and end of study (visit 7) are reported together with p value and 95% confidence intervals (95% CI). * = mean difference significant at p<0.05.

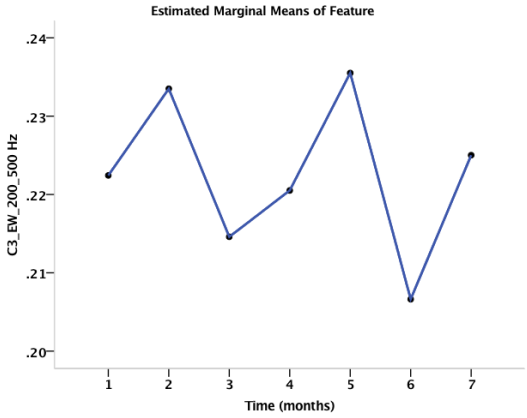
Feature	Visit	Mean	SE	95% CI		Mean difference 1-7	p	95% CI	
				Lower Bound	Upper Bound			Lower Bound	Upper Bound
C3_EW_200_500 Hz	1	0.222	0.009	0.204	0.241	-0.003	1	-0.037	0.032
	2	0.234	0.009	0.215	0.252				
	3	0.215	0.008	0.198	0.231				
	4	0.221	0.009	0.203	0.238				
	5	0.236	0.008	0.220	0.251				
	6	0.207	0.009	0.189	0.224				
	7	0.225	0.009	0.208	0.242				
C4_EW_75_200 Hz	1	0.572	0.009	0.554	0.591	-0.020	0.808	-0.050	0.010
	2	0.578	0.008	0.562	0.593				
	3	0.573	0.009	0.555	0.592				
	4	0.562	0.009	0.545	0.580				
	5	0.550	0.008	0.535	0.566				
	6	0.596	0.008	0.580	0.611				
	7	0.593	0.007	0.580	0.606				
C4_EW_200_500 Hz	1	0.171	0.008	0.154	0.187	0.023	0.268	-0.005	0.051
	2	0.162	0.008	0.147	0.177				
	3	0.155	0.007	0.141	0.170				
	4	0.161	0.008	0.146	0.176				
	5	0.164	0.007	0.150	0.177				
	6	0.137	0.007	0.124	0.150				
	7	0.148[a]	0.007	0.135	0.161				

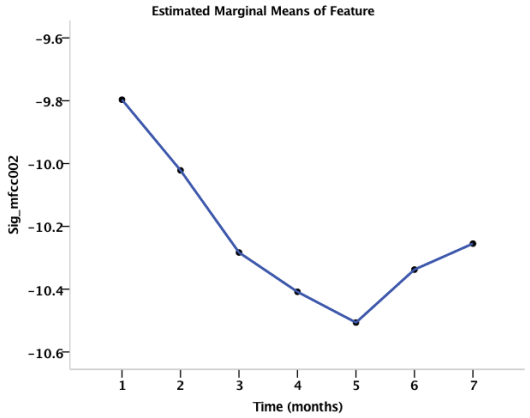
Sig_zerocross	1	0.171[a]	0.008	0.154	0.187	0.004*	0.001	0.001	0.006
	2	0.162[a]	0.008	0.147	0.177				
	3	0.155[a]	0.007	0.141	0.170				
	4	0.161[a]	0.008	0.146	0.176				
	5	0.164[a]	0.007	0.150	0.177				
	6	0.137[a]	0.007	0.124	0.150				
	7	0.148[a]	0.007	0.135	0.161				
Sig_mfcc002	1	-9.797[a]	0.039	-9.874	-9.720	0.458*	0.000	0.288	0.628
	2	-10.022	0.042	-10.105	-9.938				
	3	-10.283	0.047	-10.376	-10.190				
	4	-10.408	0.053	-10.513	-10.304				
	5	-10.506	0.044	-10.593	-10.420				
	6	-10.338	0.044	-10.425	-10.251				
	7	-10.255	0.042	-10.337	-10.173				
Sig_75_200_ Hz_zerocross	1	0.027	0.000	0.027	0.027	0.000	0.282	-7.039E-5	0.001
	2	0.027	0.000	0.027	0.027				
	3	0.027	0.000	0.026	0.027				
	4	0.027	0.000	0.026	0.027				
	5	0.026	0.000	0.026	0.027				
	6	0.026	0.000	0.026	0.027				
	7	0.027	0.000	0.026	0.027				
sig_75_200Hz_centroid	1	122.461	0.417	121.636	123.286	1.217	0.462	-0.410	2.845
	2	123.346	0.523	122.311	124.380				
	3	121.851	0.488	120.886	122.816				
	4	121.120	0.562	120.007	122.233				
	5	120.060	0.486	119.098	121.023				
	6	121.428	0.461	120.516	122.339				
	7	121.244	0.414	120.425	122.063				

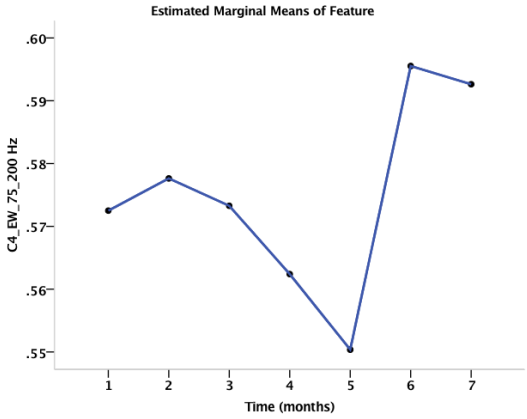
sig_200_500Hz_rms	1	0.261	0.009	0.243	0.280	0.015	1	-0.018	0.047
	2	0.264	0.009	0.246	0.282				
	3	0.246	0.008	0.230	0.262				
	4	0.233	0.008	0.217	0.249				
	5	0.226	0.008	0.211	0.241				
	6	0.226	0.007	0.213	0.239				
	7	0.247	0.008	0.232	0.262				
sig_200_500Hz_lowenergy	1	0.681	0.005	0.670	0.691	-0.011	1	-0.035	0.014
	2	0.691	0.006	0.680	0.703				
	3	0.704	0.006	0.691	0.716				
	4	0.689	0.007	0.676	0.702				
	5	0.712	0.007	0.699	0.725				
	6	0.698	0.006	0.685	0.711				
	7	0.691	0.006	0.678	0.704				
sig_200_500Hz_lowenergyASR	1	0.705	0.009	0.687	0.723	-0.026	1	-0.066	0.015
	2	0.732	0.010	0.713	0.751				
	3	0.745	0.010	0.726	0.764				
	4	0.722	0.011	0.701	0.743				
	5	0.754	0.010	0.735	0.773				
	6	0.737	0.010	0.717	0.757				
	7	0.731	0.011	0.710	0.752				
sig_200_500Hz_zerocross	1	0.077	0.000	0.076	0.078	0.002	0.068	-5.705E-5	0.003
	2	0.076	0.000	0.075	0.077				
	3	0.077	0.000	0.076	0.078				
	4	0.075	0.001	0.074	0.076				
	5	0.076	0.000	0.076	0.077				
	6	0.075	0.000	0.074	0.076				
	7	0.075	0.000	0.074	0.076				

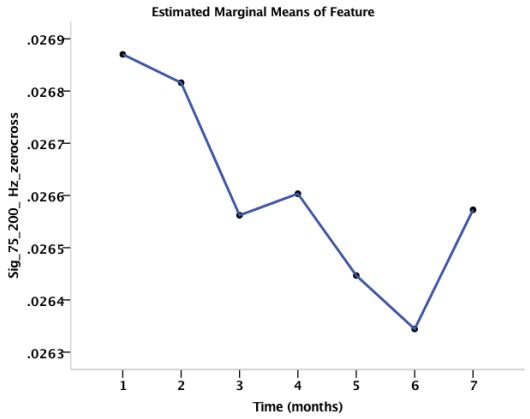
sig_200_500Hz_std_meanframes	1	0.013	0.001	0.012	0.014	0.001	1	-0.002	0.004
	2	0.013	0.001	0.012	0.014				
	3	0.012	0.001	0.011	0.014				
	4	0.012	0.001	0.011	0.013				
	5	0.011	0.001	0.010	0.013				
	6	0.011	0.001	0.010	0.012				
	7	0.012	0.001	0.011	0.013				
sig_200_500Hz_std_medianframes	1	0.003	0.000	0.003	0.004	0.000	1	0.000	0.001
	2	0.003	0.000	0.003	0.003				
	3	0.003	0.000	0.003	0.003				
	4	0.003	0.000	0.003	0.004				
	5	0.003	0.000	0.002	0.003				
	6	0.003	0.000	0.003	0.003				
	7	0.003	0.000	0.003	0.003				
sig_500_1000Hz_zerocross	1	0.148	0.000	0.147	0.148	-0.001*	0.000	-0.002	0.000
	2	0.148	0.000	0.147	0.148				
	3	0.148	0.000	0.148	0.148				
	4	0.150	0.000	0.149	0.150				
	5	0.149	0.000	0.149	0.150				
	6	0.150	0.000	0.149	0.150				
	7	0.149	0.000	0.148	0.149				
sig_500_1000Hz_rolloff85	1	740.409	3.031	734.410	746.407	-23.336*	0.000	-36.4	-10.273
	2	747.987	3.082	741.888	754.085				
	3	754.035	2.945	748.208	759.862				
	4	759.085	3.078	752.994	765.176				
	5	767.961	2.863	762.296	773.626				
	6	765.361	2.812	759.797	770.926				
	7	763.745	3.483	756.853	770.637				

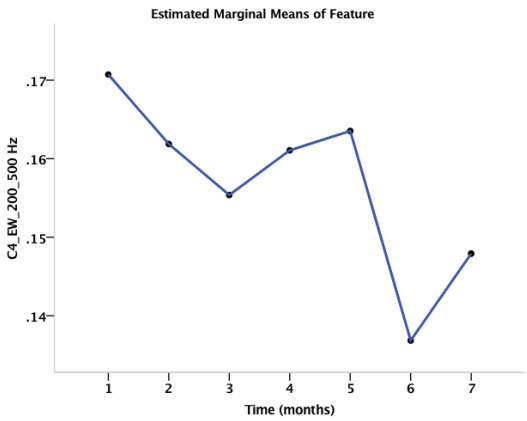
sig_500_1000Hz_centroid	1	625.106	1.378	622.379	627.833	-10.507*	0.000	-16.582	-4.431
	2	628.289	1.623	625.076	631.502				
	3	631.057	1.449	628.190	633.924				
	4	635.114	1.453	632.240	637.989				
	5	640.355	1.429	637.527	643.183				
	6	635.771	1.365	633.070	638.472				
	7	635.613	1.588	632.470	638.756				
C3_mfcc02	1	-9.805	0.039	-9.881	-9.728	0.458*	0.000	0.289	0.627
	2	-10.030	0.043	-10.114	-9.946				
	3	-10.293	0.047	-10.386	-10.200				
	4	-10.416	0.053	-10.521	-10.311				
	5	-10.514	0.044	-10.601	-10.427				
	6	-10.343	0.044	-10.430	-10.257				
	7	-10.262	0.041	-10.344	-10.180				
C4_zerocross	1	0.041	0.001	0.039	0.043	0.003	0.081	0.000	0.006
	2	0.039	0.001	0.037	0.040				
	3	0.040	0.001	0.038	0.041				
	4	0.040	0.001	0.038	0.042				
	5	0.041	0.001	0.039	0.043				
	6	0.037	0.001	0.035	0.038				
	7	0.038	0.001	0.036	0.039				
C4_mfcc02	1	-9.797	0.039	-9.874	-9.720	0.458*	0.000	0.288	0.628
	2	-10.022	0.042	-10.106	-9.938				
	3	-10.284	0.047	-10.376	-10.191				
	4	-10.407	0.053	-10.512	-10.302				
	5	-10.505	0.044	-10.592	-10.418				
	6	-10.337	0.044	-10.424	-10.250				
	7	-10.255	0.042	-10.337	-10.173				

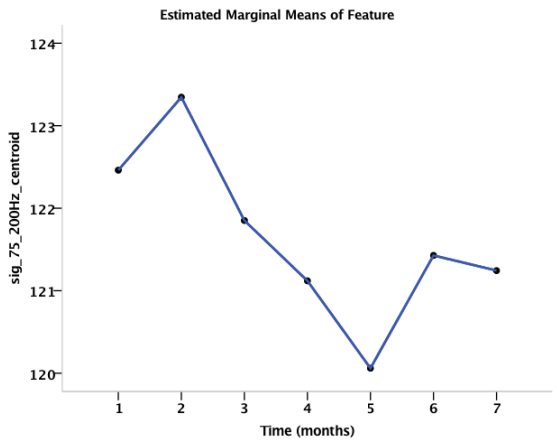

Estimated Marginal Means of Feature
C3_EW_200_500 Hz
.24
.23
.22
.21
.20
1 2 3 4 5 6 7
Time (months)

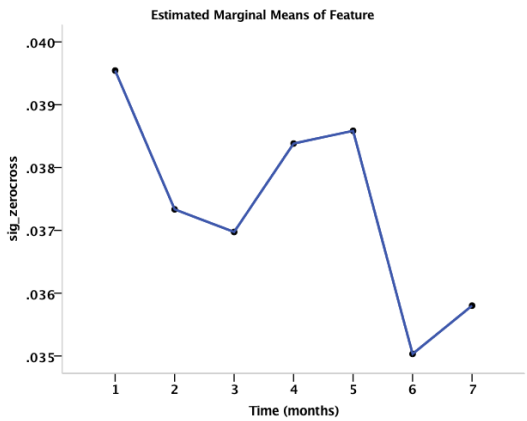

Estimated Marginal Means of Feature
Sig_mfcc002
-9.6
-9.8
-10.0
-10.2
-10.4
-10.6
1 2 3 4 5 6 7
Time (months)

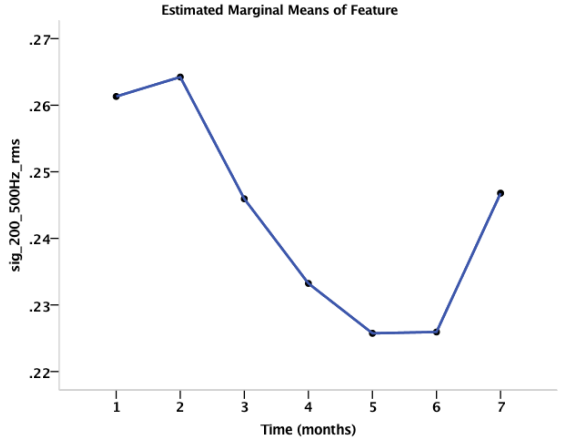

Estimated Marginal Means of Feature
C4_EW_75_200 Hz
.60
.59
.58
.57
.56
.55
1 2 3 4 5 6 7
Time (months)

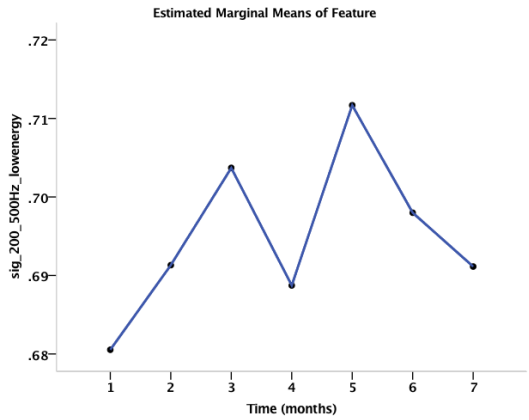

Estimated Marginal Means of Feature
Sig_75_200_ Hz_zerocross
.0269
.0268
.0267
.0266
.0265
.0264
.0263
1 2 3 4 5 6 7
Time (months)

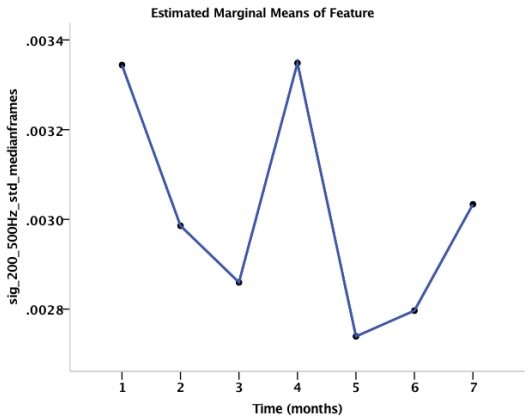

Estimated Marginal Means of Feature
C4_EW_200_500 Hz
.17
.16
.15
.14
1 2 3 4 5 6 7
Time (months)

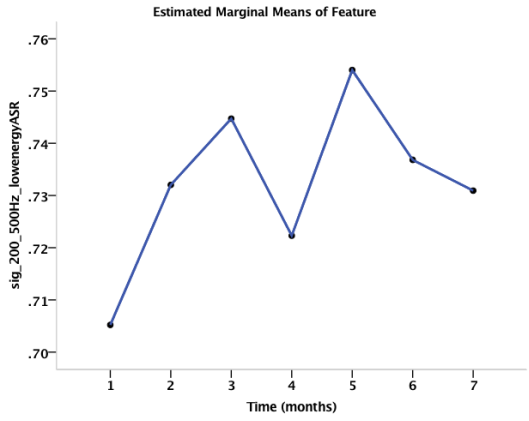

Estimated Marginal Means of Feature
sig_75_200Hz_centroid
124
123
122
121
120
1 2 3 4 5 6 7
Time (months)

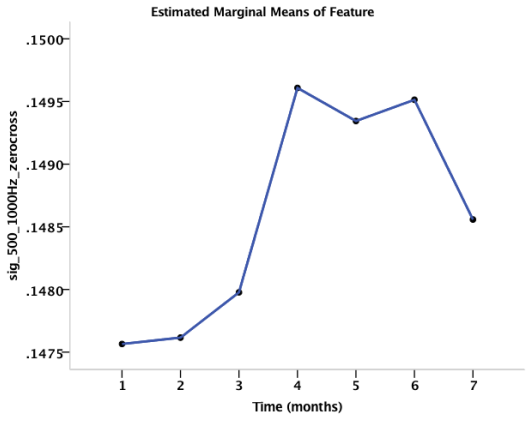

Estimated Marginal Means of Feature
sig_zerocross
.040
.039
.038
.037
.036
.035
1 2 3 4 5 6 7
Time (months)

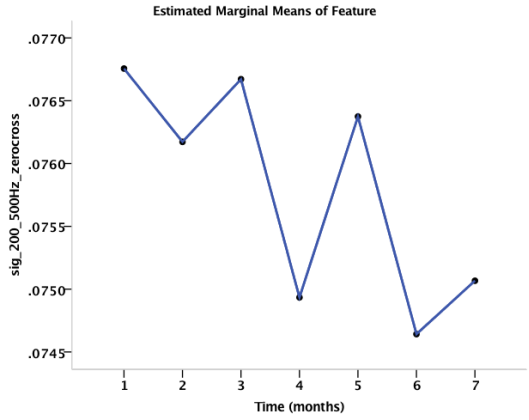

Estimated Marginal Means of Feature
sig_200_500Hz_rms
.27
.26
.25
.24
.23
.22
1 2 3 4 5 6 7
Time (months)

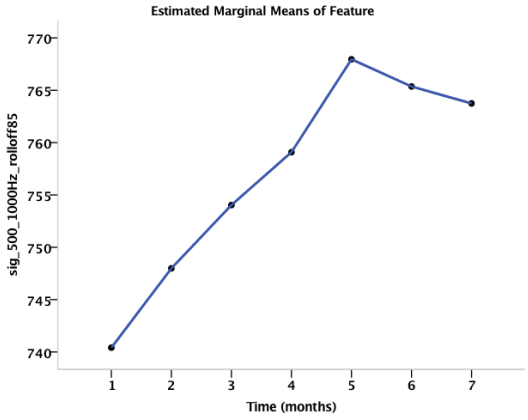

Estimated Marginal Means of Feature
sig_200_500Hz_lowenergy
.72
.71
.70
.69
.68
1
2
3
4
5
6
7
Time (months)

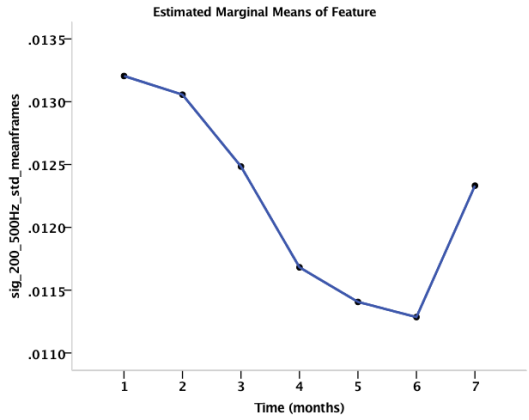

Estimated Marginal Means of Feature
sig_200_500Hz_std_medianframes
.0034
.0032
.0030
.0028
1
2
3
4
5
6
7
Time (months)

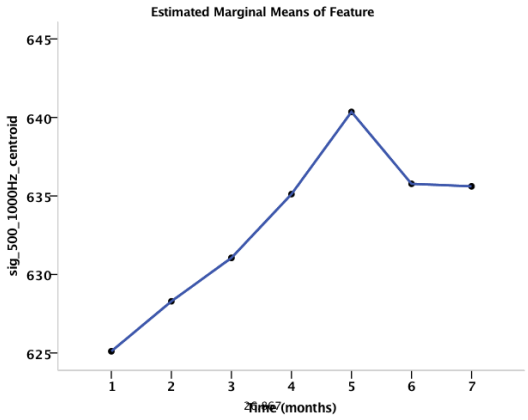

Estimated Marginal Means of Feature
sig_200_500Hz_lowenergyASR
.76
.75
.74
.73
.72
.71
.70
1
2
3
4
5
6
7
Time (months)


Estimated Marginal Means of Feature
sig_500_1000Hz_zerocross
.1500
.1495
.1490
.1485
.1480
.1475
1
2
3
4
5
6
7
Time (months)


Estimated Marginal Means of Feature
sig_200_500Hz_zerocross
.0770
.0765
.0760
.0755
.0750
.0745
1
2
3
4
5
6
7
Time (months)


Estimated Marginal Means of Feature
sig_500_1000Hz_rolloff85
770
765
760
755
750
745
740
1
2
3
4
5
6
7
Time (months)


Estimated Marginal Means of Feature
sig_200_500Hz_std_meanframes
.0135
.0130
.0125
.0120
.0115
.0110
1
2
3
4
5
6
7
Time (months)


Estimated Marginal Means of Feature
sig_500_1000Hz_centroid
645
640
635
630
625
1
2
3
4
5
6
7
Time (months)

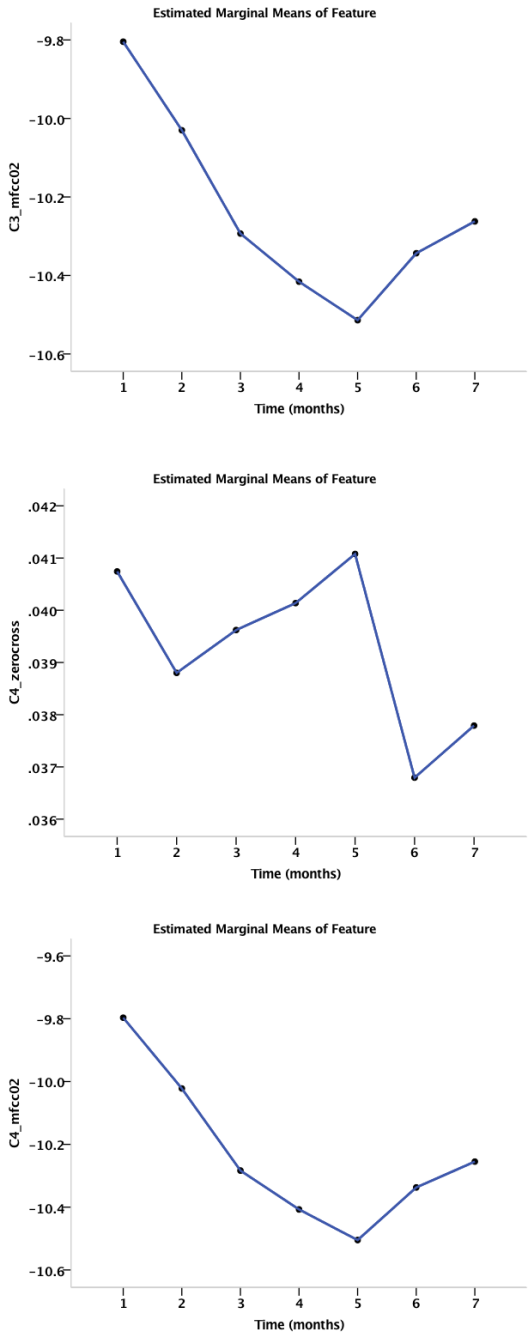


Figure 29 - Plots of estimated means of repeatable features in the IPF group.

In the primary analysis, the measurements taken from all over the chest were entered for the longitudinal assessment of acoustic features in order to provide a comprehensive evaluation of their behaviour. On the other hand, the fibrotic process in IPF is known to start in the lower regions of the lungs and progress upwards. As such, it is very likely that despite disease progression a few areas of lung parenchyma didn't become fibrotic over the 12 months of observation, and no "Velcro-type" crackles were present in the corresponding recordings. Hence, it cannot be excluded that the "normal" characteristics of the sounds recorded from these areas could possibly mask the acoustic changes occurring in the fibrotic regions.

A secondary analysis was therefore performed by repeating the assessment of longitudinal changes of acoustic features only for the sounds recorded over the lung bases and the lateral chest, chosen for being the lower recordings sites on the chest.

ANOVA was performed over a lower number of available observations (n=74) as compared to the primary analysis. **Error! Reference source not found.** Table 64 in the next page shows the comparison of the tests of within-subjects effects between the primary and the secondary analysis. A few acoustic features (sig_mfcc02, sig_200_500Hz_lowenergy, C3_mfcc02 and C4_mfcc02) significantly changed in the primary analysis, but they didn't in the ANOVA performed over the measurements obtained from the lower chest recording sites. Conversely, modifications of one feature (sig_75_200Hz_centroid) reached statistical significance in the secondary analysis only.

Table 64 - Tests of within-subjects effects for repeatable acoustic features in the IPF group – comparison between the analysis performed over all recording sites (n=10) or lower recording sites only (n=4). * = F value significant at $p < 0.05$. F values found significant in one of the two analyses but not in the other are highlighted in bold.

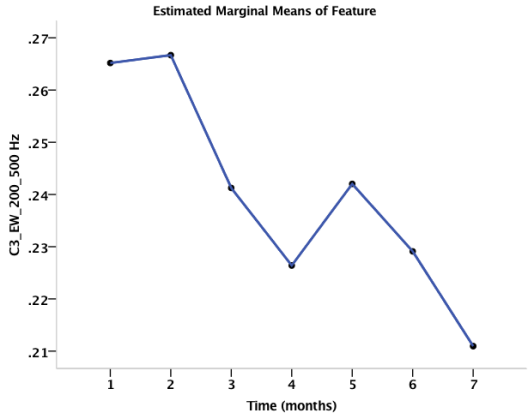
	All sites		**Lower sites**	
Feature	F	P	F	p
C3 EW_200_500Hz	5.803*	0.000	3.729*	0.001
C4 EW_75_200Hz	2.877*	0.013	3.543*	0.002
C4 EW_200_500Hz	3.853*	0.001	3.389*	0.003
sig_zerocross	4.552*	0.000	3.600*	0.002
sig_mfcc02	**4.579***	0.000	0.540	0.778
sig_75_200Hz_zerocross	2.973*	0.009	2.416*	0.026
sig_75_200Hz_centroid	1.362	0.233	**3.042***	0.006
sig_200_500Hz_rms	2.974*	0.007	2.971*	0.008
sig_200_500Hz_lowenergy	**3.009***	0.007	2.017	0.062
sig_200_500Hz_lowenergyASR	2.366*	0.028	1.778*	0.102
sig_200_500Hz_zerocross	3.813*	0.001	5.966*	0.000
sig_200_500Hz_std_meanframes	1.102	0.359	2.031	0.060
sig_200_500Hz_std_medianframes	1.035	0.398	0.846	0.535
sig_500_1000Hz_zerocross	4.964*	0.000	4.306*	0.000
sig_500_1000Hz_rolloff85	1.318	0.247	2.010	0.063
sig_500_1000Hz_centroid	2.391*	0.027	3.389*	0.003
C3_mfcc02	**4.661***	0.000	0.596	0.734
C4_zerocross	3.934*	0.001	4.591*	0.000
C4_mfcc02	**4.569***	0.000	0.550	0.770

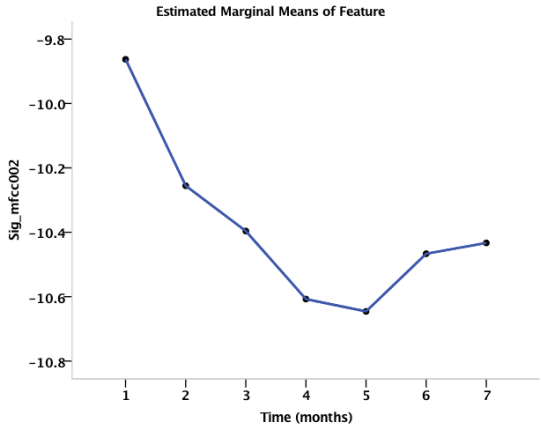
Table 65 **Error! Reference source not found.** reports the estimated mean differences between baseline (visit 1) and end of study (visit 7) calculated in both ANOVAs, the one including the measurements taken at all recording sites and the one including those taken at lower recording sites only. Overall, the sounds recorded over the lower regions of the lungs showed to undergo greater changes more over time, as demonstrated by the larger mean differences in most of the acoustic features after 12 months – in fact, only sig_500_1000Hz_centroid showed a larger difference in the primary analysis.

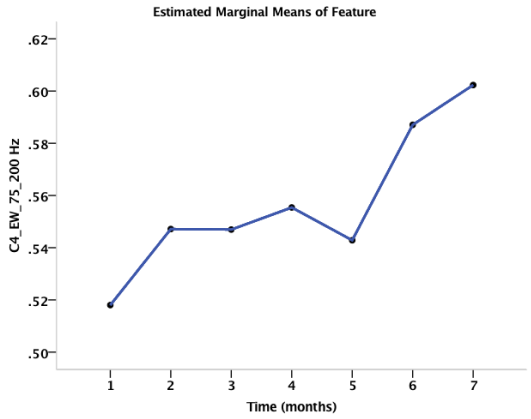
The plots in Figure 30 also show that the trends followed by the values of several features were more consistent in the secondary analysis, suggesting that the study of lung sounds recorded from lower lung regions might provide more relevant and reliable information as to the progression of the fibrotic process.

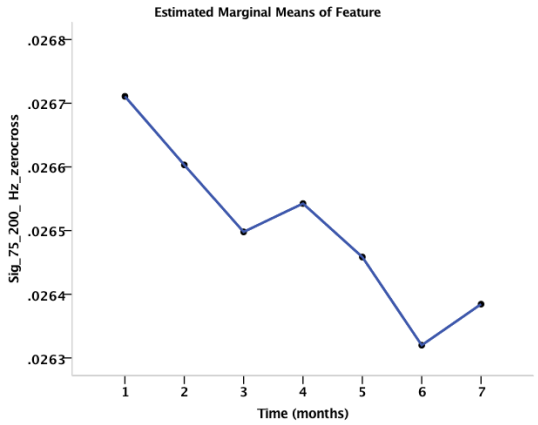
Table 65 – Estimated mean differences between baseline (visit 1) and end of study (visit 7) for measurements taken over all recording sites or lower chest recording sites only (lung bases and lateral chest). * = mean difference value significant at $p < 0.005$. The values found statistically significant in one of the two analyses of variance but not in the other are highlighted in bold.

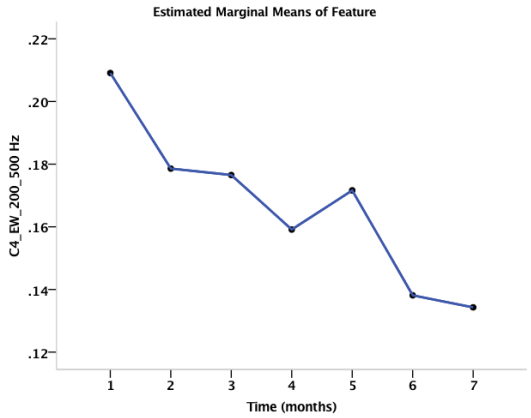
Feature	All sites	Lower sites
C3 EW_200_500Hz	-0.003	**0.054***
C4 EW_75_200Hz	-0.020	**-0.084***
C4 EW_200_500Hz	0.023	**0.075***
sig_zerocross	0.004*	**0.009***
sig_mfcc02	0.458*	0.570*
sig_75_200Hz_zerocross	0.000	0.000
sig_75_200Hz_centroid	1.217	**3.237***
sig_200_500Hz_rms	0.015	**0.064***
sig_200_500Hz_lowenergy	-0.011	-0.006
sig_200_500Hz_lowenergyASR	-0.026	0.029
sig_200_500Hz_zerocross	0.002	**0.005***
sig_200_500Hz_std_meanframes	0.001	**0.004***
sig_200_500Hz_std_medianframes	0.000	0.000
sig_500_1000Hz_zerocross	-0.001	-0.001
sig_500_1000Hz_rolloff85	-23.336*	-18.895*
sig_500_1000Hz_centroid	**-10.507***	-6.681
C3_mfcc02	0.458*	0.563*
C4_zerocross	0.003	**0.008***
C4_mfcc02	0.458*	0.569*

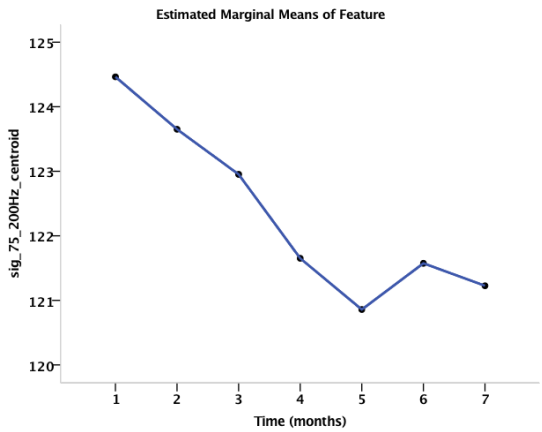

Estimated Marginal Means of Feature
C3_EW_200_500 Hz
.27
.26
.25
.24
.23
.22
.21
1 2 3 4 5 6 7
Time (months)

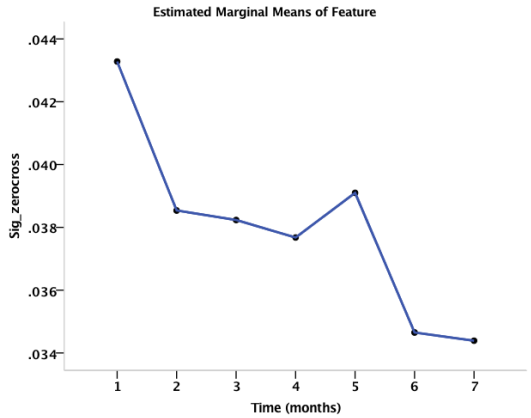

Estimated Marginal Means of Feature
Sig_mfcc002
-9.8
-10.0
-10.2
-10.4
-10.6
-10.8
1 2 3 4 5 6 7
Time (months)

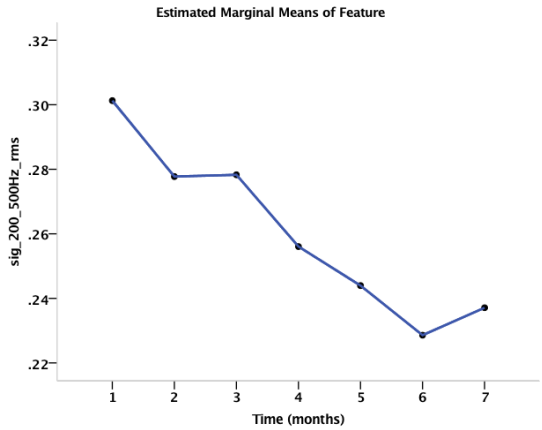

Estimated Marginal Means of Feature
C4_EW_75_200 Hz
.62
.60
.58
.56
.54
.52
.50
1 2 3 4 5 6 7
Time (months)

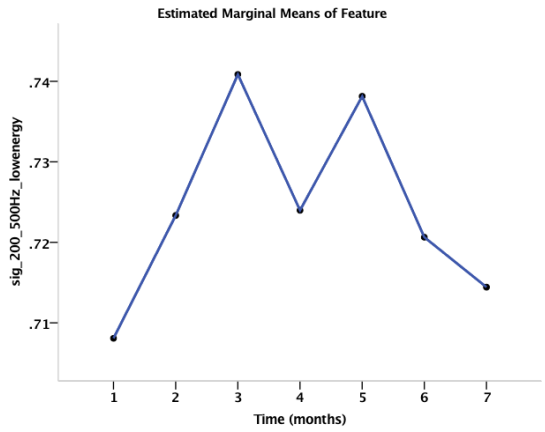

Estimated Marginal Means of Feature
Sig_75_200_ Hz_zerocross
.0268
.0267
.0266
.0265
.0264
.0263
1 2 3 4 5 6 7
Time (months)

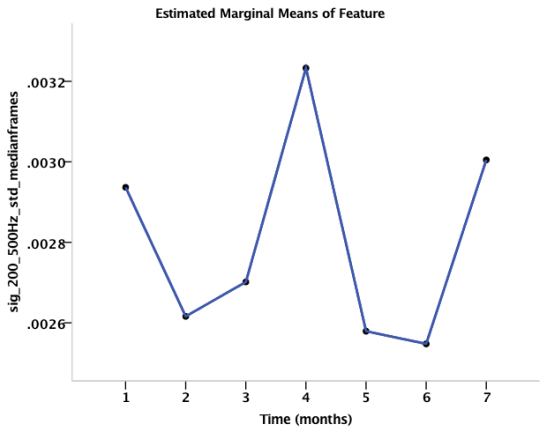

Estimated Marginal Means of Feature
C4_EW_200_500 Hz
.22
.20
.18
.16
.14
.12
1 2 3 4 5 6 7
Time (months)

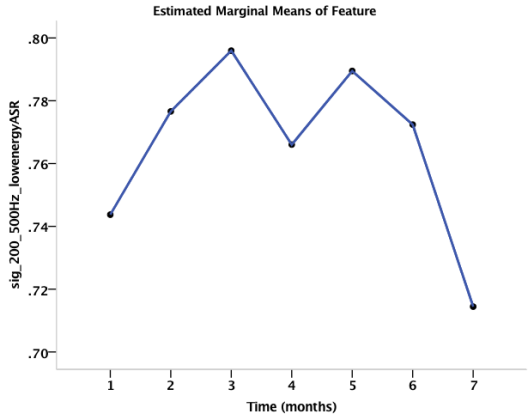

Estimated Marginal Means of Feature
sig_75_200Hz_centroid
125
124
123
122
121
120
1 2 3 4 5 6 7
Time (months)

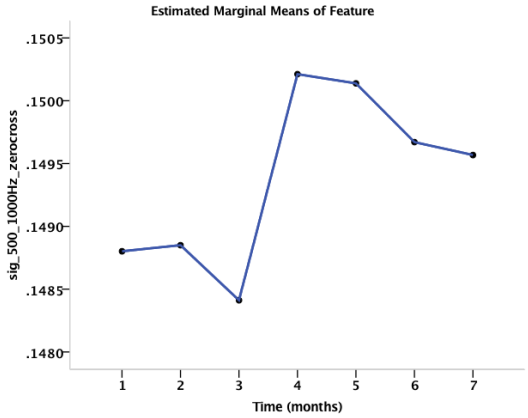

Estimated Marginal Means of Feature
Sig_zerocross
.044
.042
.040
.038
.036
.034
1 2 3 4 5 6 7
Time (months)

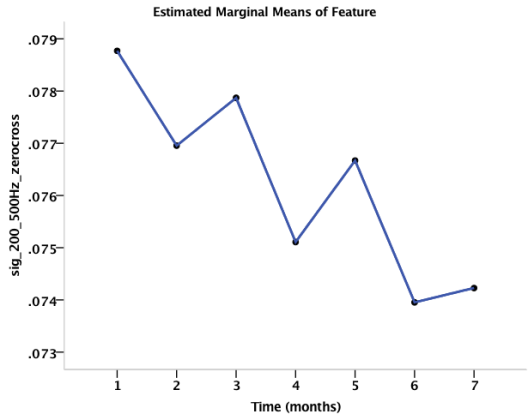

Estimated Marginal Means of Feature
sig_200_500Hz_rms
.32
.30
.28
.26
.24
.22
1 2 3 4 5 6 7
Time (months)

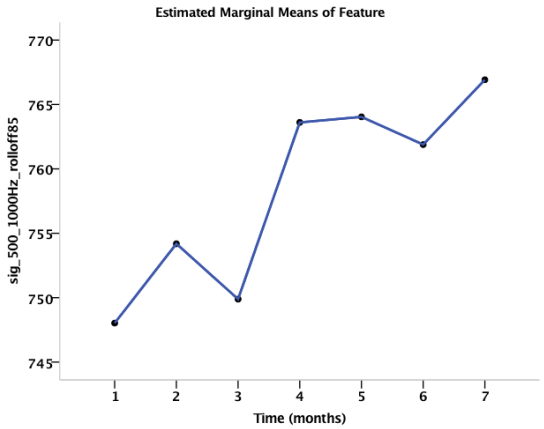

Estimated Marginal Means of Feature
sig_200_500Hz_lowenergy
.74
.73
.72
.71
1 2 3 4 5 6 7
Time (months)

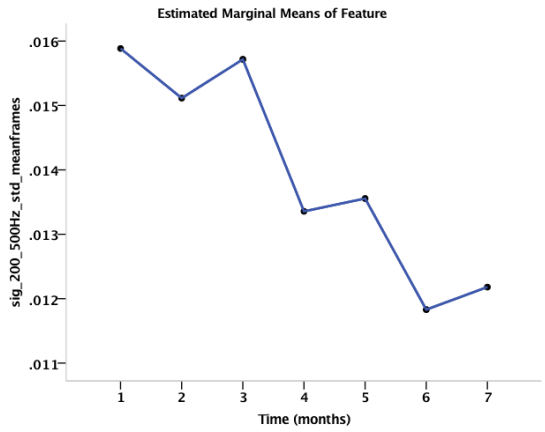

Estimated Marginal Means of Feature
sig_200_500Hz_std_medianframes
.0032
.0030
.0028
.0026
1 2 3 4 5 6 7
Time (months)

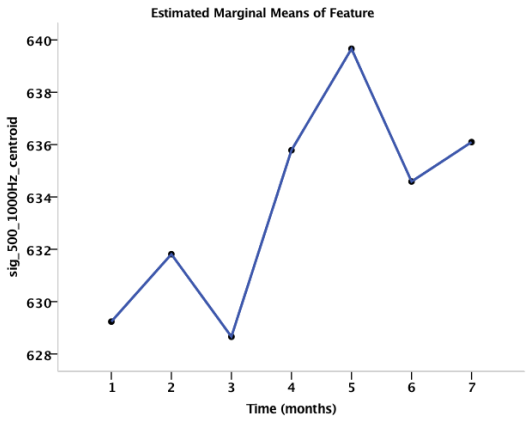

Estimated Marginal Means of Feature
sig_200_500Hz_lowenergyASR
.80
.78
.76
.74
.72
.70
1 2 3 4 5 6 7
Time (months)


Estimated Marginal Means of Feature
sig_500_1000Hz_zerocross
.1505
.1500
.1495
.1490
.1485
.1480
1 2 3 4 5 6 7
Time (months)


Estimated Marginal Means of Feature
sig_200_500Hz_zerocross
.079
.078
.077
.076
.075
.074
.073
1 2 3 4 5 6 7
Time (months)


Estimated Marginal Means of Feature
sig_500_1000Hz_rolloff85
770
765
760
755
750
745
1 2 3 4 5 6 7
Time (months)


Estimated Marginal Means of Feature
sig_200_500Hz_std_meanframes
.016
.015
.014
.013
.012
.011
1 2 3 4 5 6 7
Time (months)


Estimated Marginal Means of Feature
sig_500_1000Hz_centroid
640
638
636
634
632
630
628
1 2 3 4 5 6 7
Time (months)

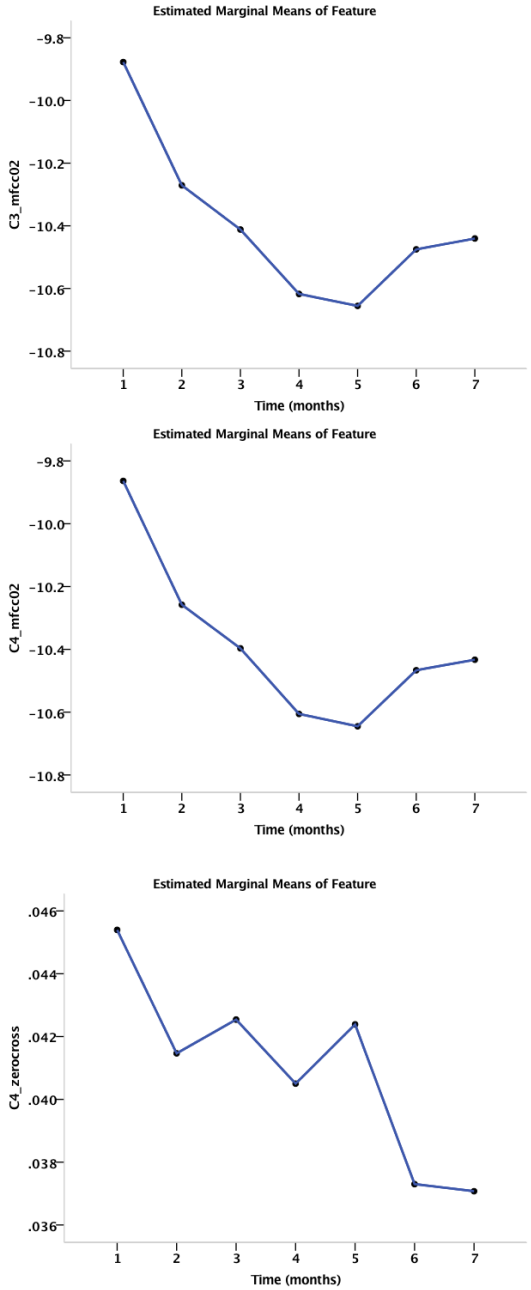


Figure 30 - Plots of estimated means of repeatable features in the IPF group for lower chest recording sites (lung bases and lateral chest)

5.4.7.4 Correlation of acoustic features with parameters of disease progression

The acoustic features that changed significantly over the observation period in the IPF group were correlated with the following clinical parameters: % predicted FVC, % predicted DL_{CO}, the distance walked at the 6MWT (6MWD) and the scores at UCSD- SOBQ and SGRQ.

At univariate analysis, several features showed statistically significant correlations (expressed as Pearson's correlation coefficients) with clinical parameters, especially with % predicted FVC and 6MWD, as shown in Table 66**Error! Reference source not found.**. The strongest associations were with C3 EW_75_200Hz, C4 EW_75_200Hz, C4 EW_200_500Hz, sig_zerocross, sig_200_500 Hz_rms, and C4_zerocross. Overall, the strength of the individual relationships between acoustic features and the other parameters was lower than the relationships between the different clinical measurements.

Table 66 - Univariate correlation analysis between clinical parameters and repeatable acoustic features, measured at different study time points in the IPF group. Data are expressed as Pearson's correlation coefficient. * = correlation was significant at the 0.05 level. ** = correlation was significant at the 0.01 level. FVC = Forced Vital Capacity. DLCO = Diffusion Lung capacity for CO. UCSD-SOBQ = University of California Sand Diego -Shortness Of Breath Questionnaire. SGRQ = Saint George's Respiratory Questionnaire. 6MWD = 6-Minute Walk Distance.

	Pred % FVC	**Pred % DL_{CO}**	**UCSD-SOBQ**	**SGRQ**	**6MWD**
Pred %FVC	**1**	**0.481****	**-0.381****	**-0.322****	**0.401****
Pred % DL_{CO}	**0.481***	**1**	**-0.378****	**-0.470****	**0.279****
UCSD-SOBQ	**-0.381****	**-0.378****	**1**	**0.744****	**-0.391****
SGRQ	**-0.322****	**-0.470****	**0.744****	**1**	**-0.245****
6MWD	**0.401****	**0.279****	**-0.391****	**-0.245****	**1**
C3 EW_200_500Hz	-0.070*	0.015	0.009	0.032	-0.240**
C4 EW_75_200Hz	0.117**	0.033	-0.023	-0.073**	0.200**
C4 EW_200_500Hz	-0.119**	-0.025	0.012	0.063*	-0.257**
sig_zerocross	-0.158**	-0.064*	0.023	0.074*	-0.227**
sig_mfcc02	-0.031	0.019	-0.034	-0.044	-0.164**
sig_75_200Hz_zerocross	-0.054	0.004	-0.007	-0.012	-0.183**
sig_200_500Hz_rms	-0.075**	-0.003	-0.030	0.048	-0.228**
sig_200_500Hz_lowenergy	0.059*	0.065*	-0.017	0.012	0.041
sig_200_500Hz_lowenergyASR	0.060*	0.087**	-0.005	-0.020	-0.024
sig_200_500Hz_zerocross	-0.092**	0.001	-0.013	0.017	-0.156**
sig_500_1000Hz_zerocross	0.043	-0.005	0.107**	0.019	-0.067
sig_500_1000Hz_centroid	0.040	0.020	0.102**	0.058*	0.064
C3_mfcc02	-0.031	0.019	-0.035	-0.041	-0.165**
C4_zerocross	-0.129**	-0.043	0.017	0.057*	-0.203**
C4_mfcc02	-0.031	0.019	-0.034	-0.043	-0.164**

Figure 31 and Figure 32 show the graphical plot of the values of the features with significant correlation coefficients at univariate analysis against the values of % predicted FVC and 6MWD. The univariate regression lines and equations are also reported over the graphs in the figures: for % predicted FVC the strongest relationships were found with zero cross rate properties of the signal and its crackle component (Sig_zerocross and C4_zerocross), that were shown to account respectively for 2.5% and 1.7% of the variability of pulmonary function. As for the distance walked at the 6-minute walk test, the strongest relationship was found with the energy content of the crackle components at different frequency bands (C3 EW_200_500Hz, C4 EW_75_200Hz and C4 EW_200_500Hz, who were shown to account for 5.7%, 4% and 6.6% of the variance of the dependent parameter), with the root mean square of the signal at the 200-500 Hz band and again with the zero-cross rate of the signal (both accounting for 5.2% of the variance of 6MWD).

Multivariable linear regression was then performed to investigate the relationships between the whole set of reproducible acoustic features and % predicted FVC and 6MWT – the parameters that showed more significant correlations with sound features on univariate analysis. Table 67 and

Table 68 report the R squared values indicating the amount of variability of the dependent clinical variable that can be explained by the model built using the combination of the acoustic features. The regression coefficients for the individual features included in the model are also reported. The model accounted for 3.9% of the variation of % predicted FVC and for 10% of the variation of the distance walked at the 6MWT. Only few acoustic features demonstrated to independently predict change in the dependent variables: among those, sig_zerocross and C4 EW_200_500Hz were the strongest determinants of % predicted FVC and 6MWD respectively, in adherence with the results of the univariate analysis.

Table 67 – Multivariable linear regression analysis demonstrating the relationship between repeatable acoustic features and % predicted FVC. Model's adjusted R squared (R^2) and the regression coefficients (ß) for the individual acoustic features are reported, together with p values and 95% Confidence Intervals (95% CI).

	R^2	ß	P value	95% CI	
				Lower bound	Upper bound
Model	0.039		0.000		
C3 EW_200_500Hz		0.051	0.397	-12.803	32.292
C4 EW_75_200Hz		0.040	0.608	-22.814	38.995
C4 EW_200_500Hz		-0.003	0.981	-48.973	47.820
sig_zerocross		-0.431	0.001	-1558.275	-426.909
sig_75_200Hz_zerocross		0.092	0.083	-220.941	3573.198
sig_75_200Hz_centroid		0.060	0.490	-24.687	51.463
sig_200_500Hz_rms		-0.102	0.687	-30.242	19.941
sig_200_500Hz_lowenergy		0.049	0.022	63.524	797.797
sig_200_500Hz_lowenergyASR		-0.030	0.043	9.724	580.084
sig_200_500Hz_zerocross		0.123	0.214	-1531.652	344.240
sig_200_500Hz_std_meanframes		0.101	0.455	-370.466	825.744
sig_200_500Hz_std_medianframes		-0.050	0.752	-0.087	0.063
sig_500_1000Hz_zerocross		0.030	0.583	-0.116	0.206
sig_500_1000Hz_rolloff85		-0.022	0.955	-79.208	83.854
sig_500_1000Hz_centroid		0.040	0.294	-205.528	678.283
C3_mfcc02		0.068	0.964	-83.150	79.437
C4_zerocross		0.116	0.397	-12.803	32.292
C4_mfcc02		-0.054	0.608	-22.814	38.995

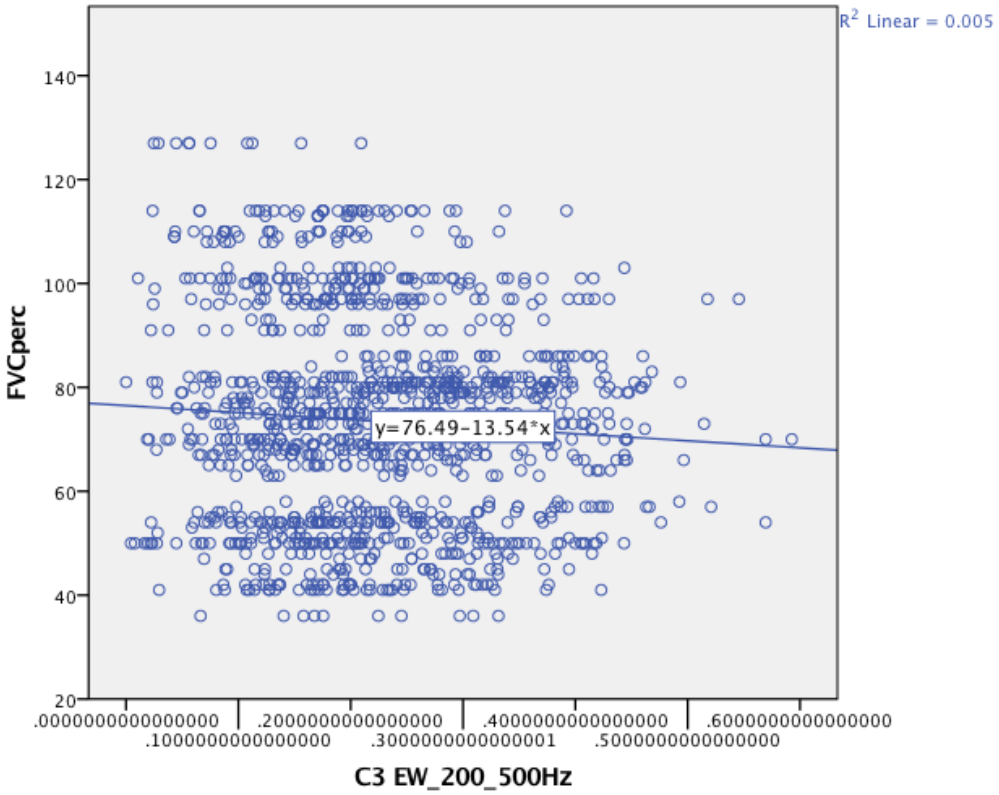
Table 68 - Multivariable linear regression analysis demonstrating the relationship between repeatable acoustic features and 6MWD. Model's adjusted R squared (R^2) and the regression coefficients (ß) for the individual acoustic features are reported, together with p values and 95% Confidence Intervals (95% CI).

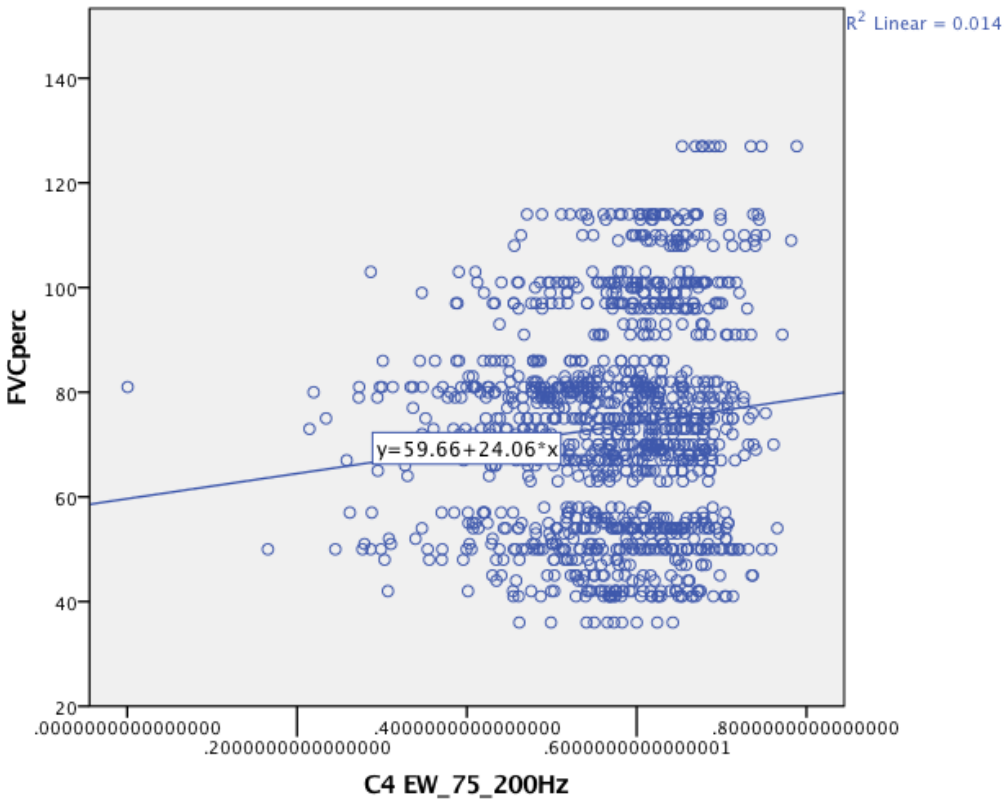
	R^2	ß	P value	95% CI	
				Lower bound	Upper bound
Model	0.104		0.000		
C3 EW_200_500Hz		-0.110	0.282	-322.825	94.252
C4 EW_75_200Hz		-0.143	0.368	-515.191	191.388
C4 EW_200_500Hz		-0.471	0.016	-1022.337	-106.413
sig_zerocross		-0.284	0.242	-9180.932	2321.648
sig_mfcc02		0.491	0.876	-1080.570	1266.849
sig_75_200Hz_zerocross		0.144	0.095	-2455.656	30443.047
sig_75_200Hz_centroid		-0.033	0.646	-3.623	2.251
sig_200_500Hz_rms		-0.132	0.210	-399.423	87.986
sig_200_500Hz_lowenergy		0.287	0.009	114.609	797.999
sig_200_500Hz_lowenergyASR		-0.268	0.029	-473.230	-25.850
sig_200_500Hz_zerocross		0.058	0.496	-2010.272	4145.296
sig_200_500Hz_std_meanframes		0.089	0.310	-1428.396	4494.349
sig_200_500Hz_std_medianframes		-0.076	0.236	-12687.084	3135.463
sig_500_1000Hz_zerocross		-0.175	0.004	-12167.345	-2289.389
sig_500_1000Hz_rolloff85		0.193	0.069	-0.045	1.210
sig_500_1000Hz_centroid		-0.174	0.123	-2.581	0.309
C3_mfcc02		0.231	0.911	-731.760	819.580
C4_zerocross		0.453	0.056	-120.630	10131.697
C4_mfcc02		-0.844	0.848	-1801.452	1481.327

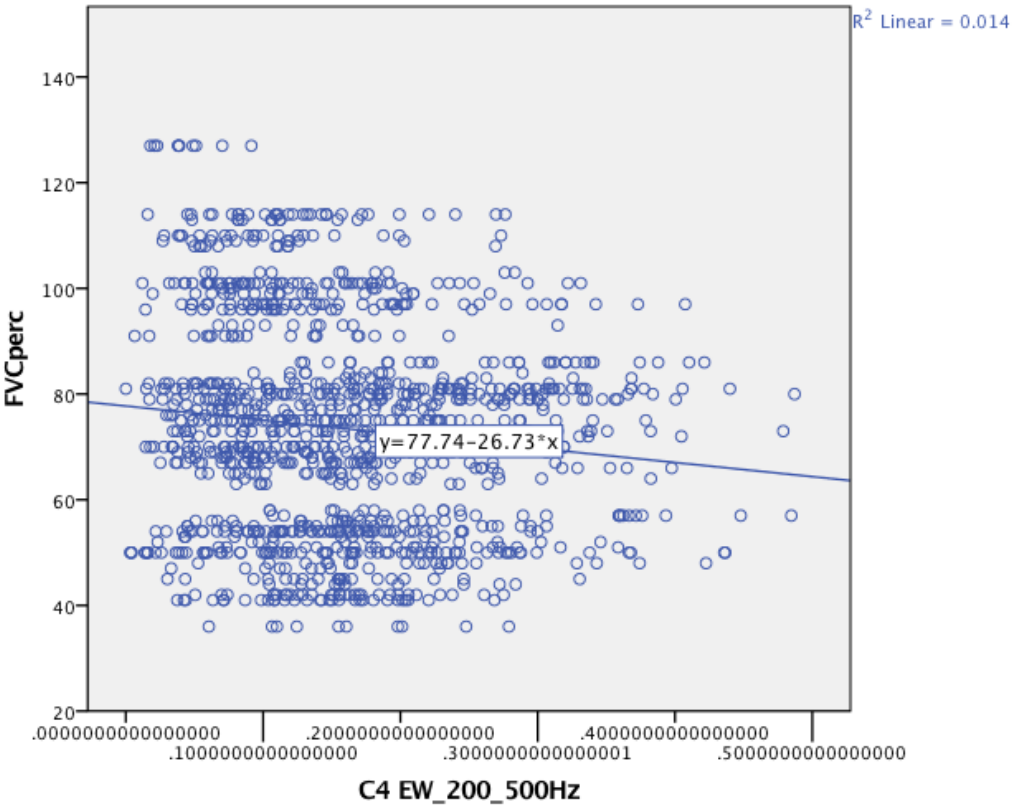
A univariate correlation analysis was finally performed entering only the measurements of the acoustic features from the lower lung regions, just as it was done for the secondary analysis of longitudinal changes (Table 69). Nevertheless, such analysis did not provide further relevant information, as only few significant, weak correlations were found, likely due to the fewer observations available.

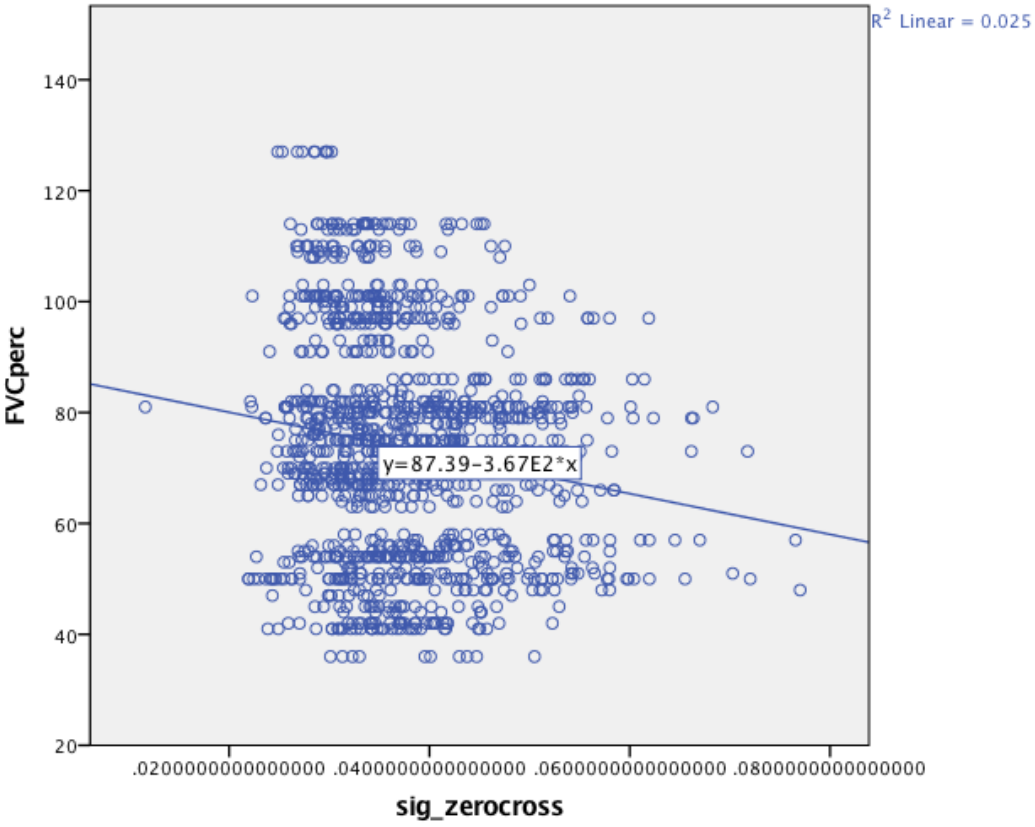
Table 69 - Univariate correlation analysis between clinical parameters (% pred FVC and 6MWD) and acoustic features measured from lung sounds recorded over the lower regions of the chest (lung bases and lateral chest in correspondence of the 4th or 5th intercostal space). Data are expressed as Pearson's correlation coefficient. * = correlation was significant at the 0.05 level. ** = correlation was significant at the 0.01 level.

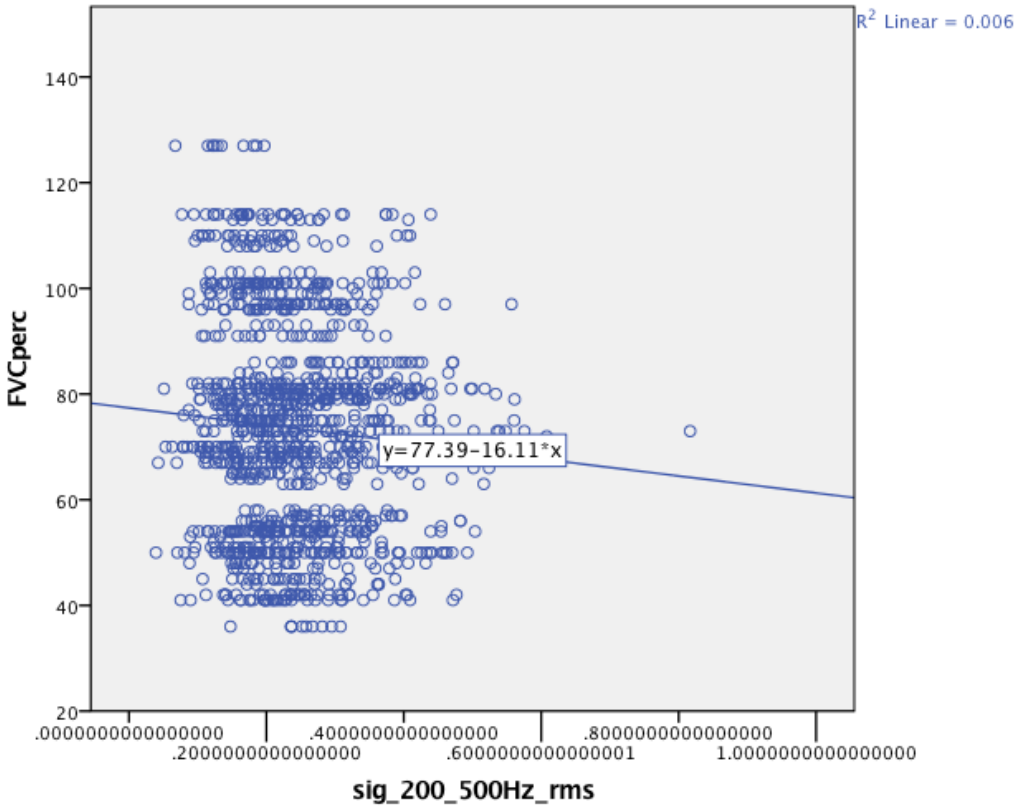
	Pred % FVC	**6MWD**
C3 EW_200_500Hz	-0.02	-0.161*
C4 EW_75_200Hz	0.150**	0.161*
C4 EW_200_500Hz	-0.077	-0.204**
sig_zerocross	-0.164**	-0.172*
sig_75_200Hz_zerocross	0.023	-0.098
sig_75_200Hz_centroid	0.022	-0.159*
sig_200_500Hz_rms	-0.038	-0.152*
sig_200_500Hz_lowenergyASR	0.071	-0.079
sig_200_500Hz_zerocross	-0.138**	-0.083
sig_500_1000Hz_zerocross	-0.031	-0.103
sig_500_1000Hz_centroid	-0.039	0.071
C4_zerocross	-0.131**	-0.137

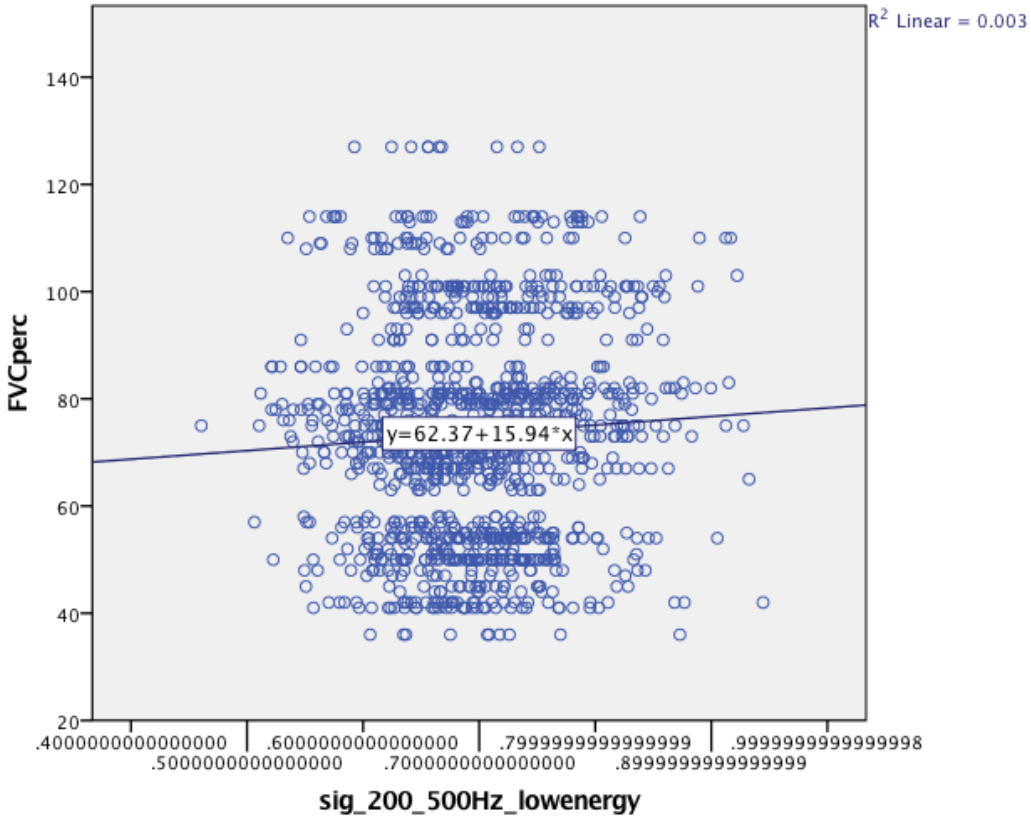

R² Linear = 0.005
FVCperc
y=76.49-13.54*x
C3 EW_200_500Hz

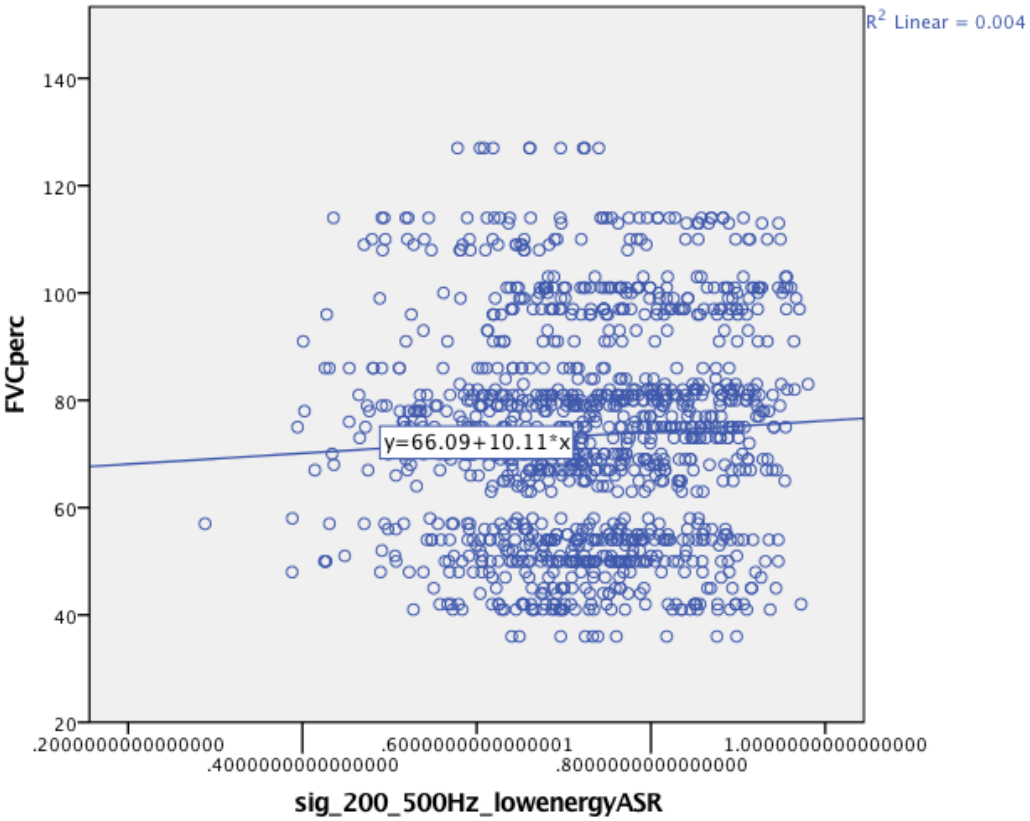

R² Linear = 0.014
FVCperc
y=59.66+24.06*x
C4 EW_75_200Hz

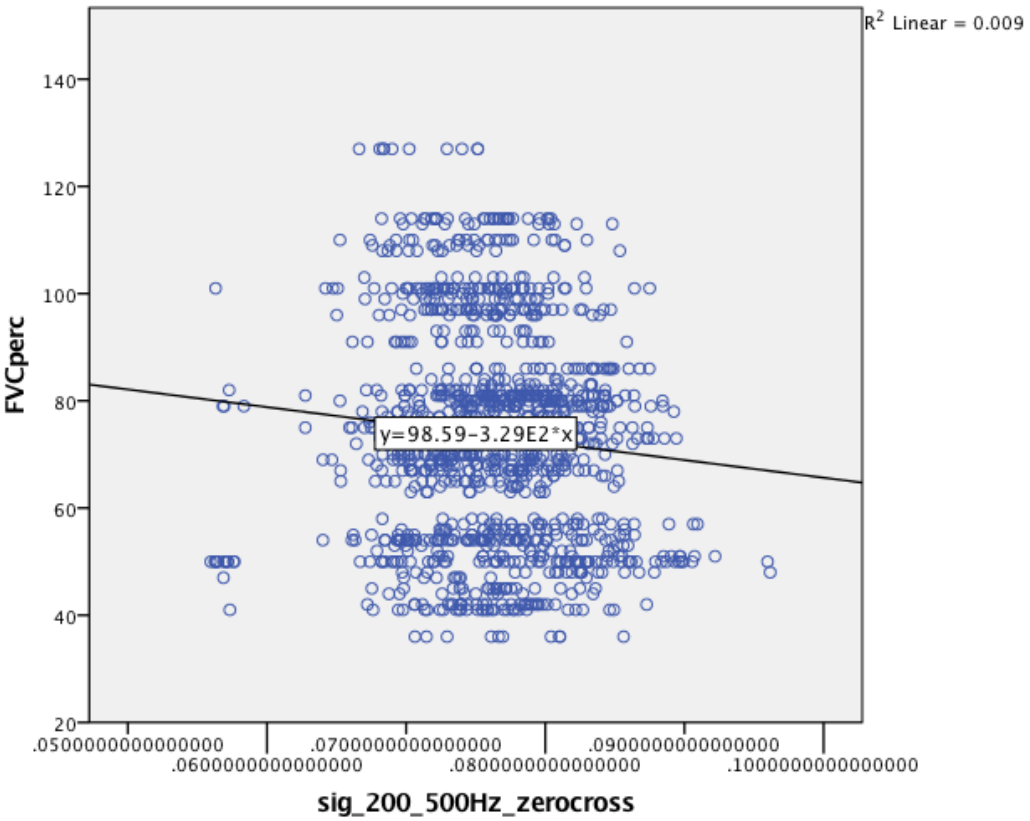

R² Linear = 0.014
FVCperc
140
120
100
80
60
40
20
y=77.74-26.73*x
.0000000000000000
.1000000000000000
.2000000000000000
.3000000000000001
.4000000000000000
.5000000000000000
C4 EW_200_500Hz


R² Linear = 0.025
FVCperc
140
120
100
80
60
40
20
y=87.39-3.67E2*x
.0200000000000000
.0400000000000000
.0600000000000000
.0800000000000000
sig_zerocross


R² Linear = 0.006
FVCperc
140
120
100
80
60
40
20
y=77.39-16.11*x
.0000000000000000
.2000000000000000
.4000000000000000
.6000000000000001
.8000000000000000
1.0000000000000000
sig_200_500Hz_rms


R² Linear = 0.003
FVCperc
140
120
100
80
60
40
20
y=62.37+15.94*x
.4000000000000000
.5000000000000000
.6000000000000000
.7000000000000000
.7999999999999999
.8999999999999999
.9999999999999998
sig_200_500Hz_lowenergy


R² Linear = 0.004
FVCperc
y=66.09+10.11*x
sig_200_500Hz_lowenergyASR


R² Linear = 0.009
FVCperc
y=98.59-3.29E2*x
sig_200_500Hz_zerocross

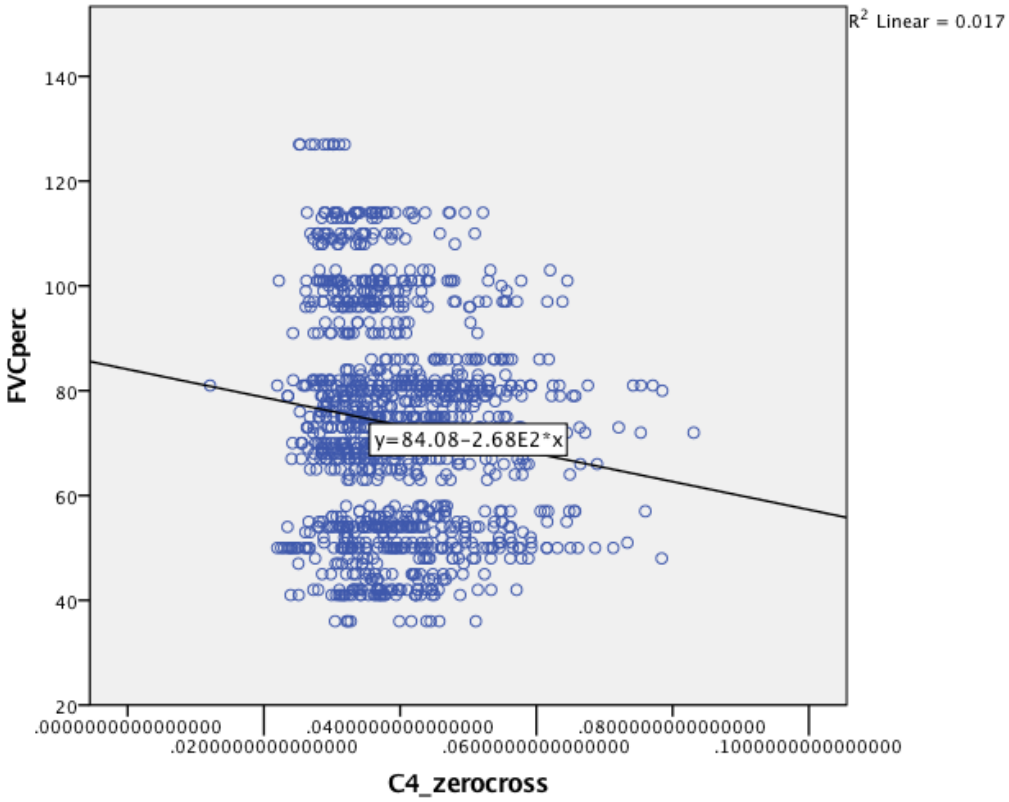
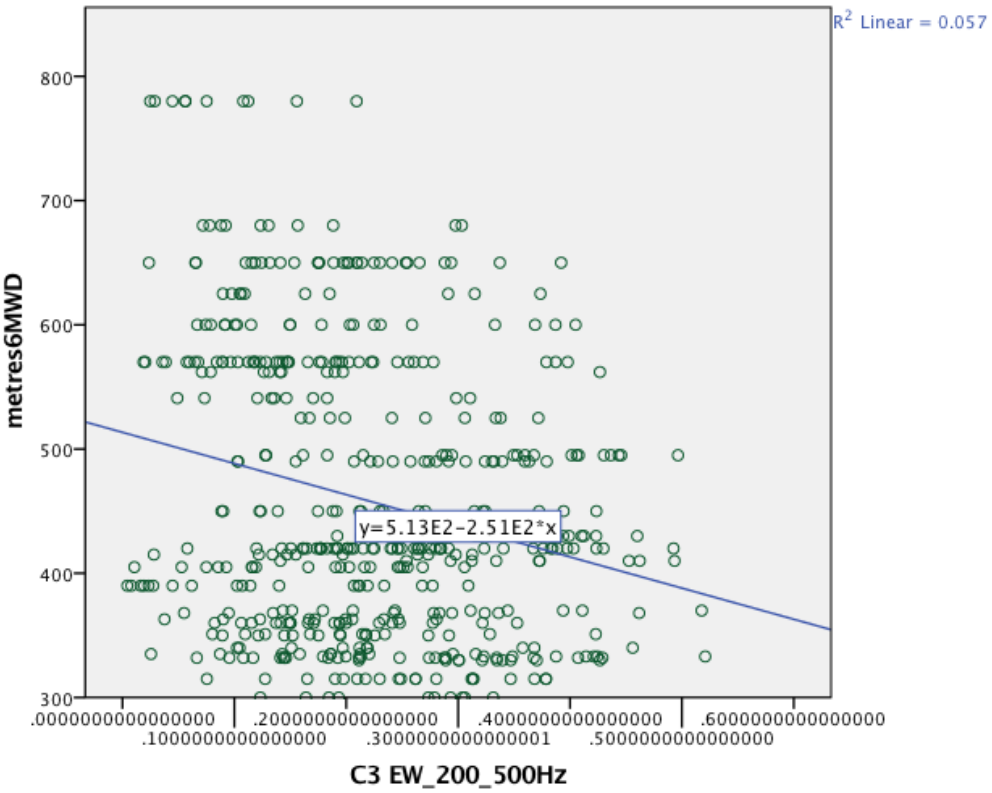
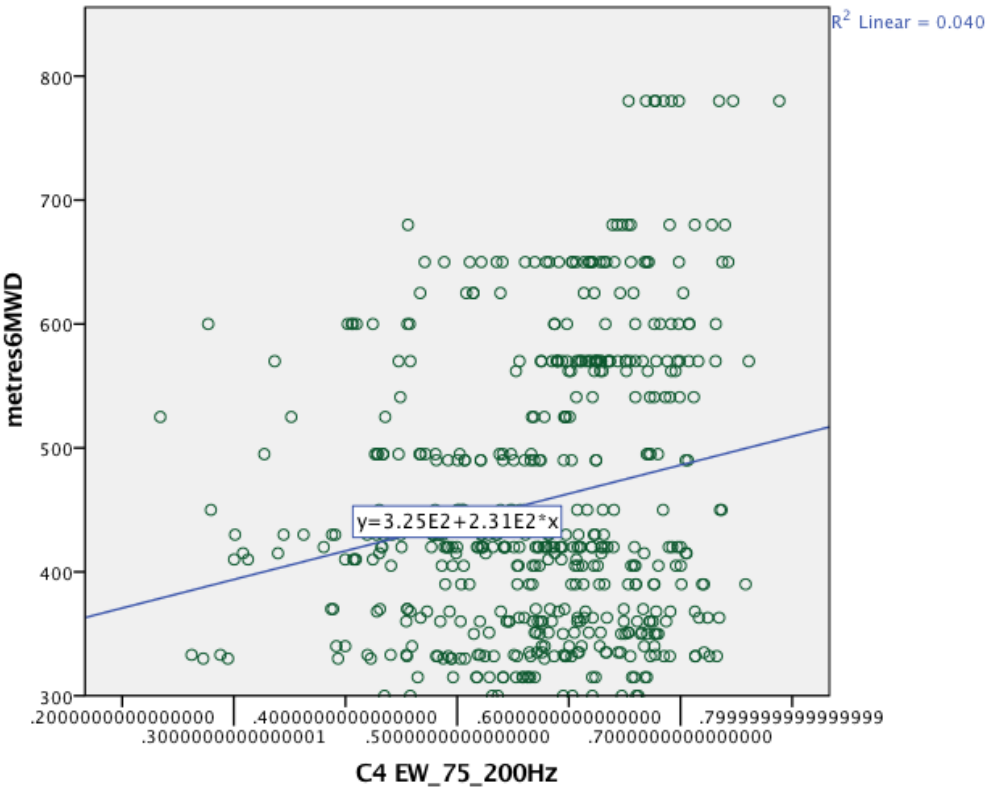
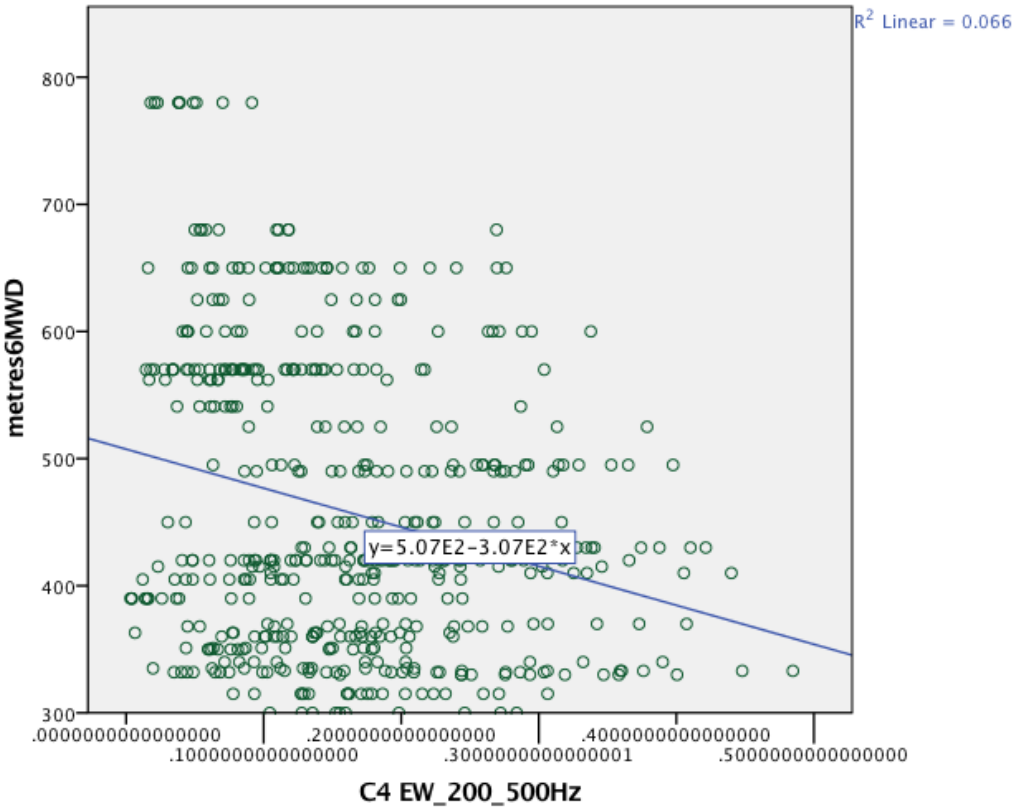
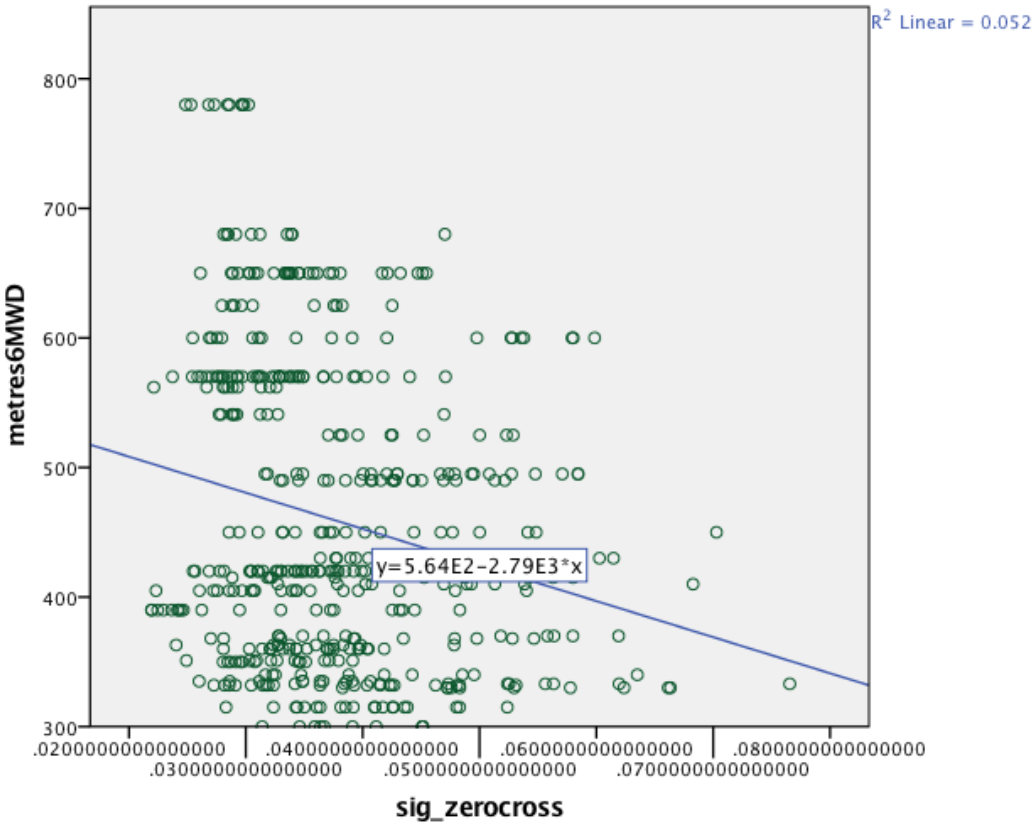


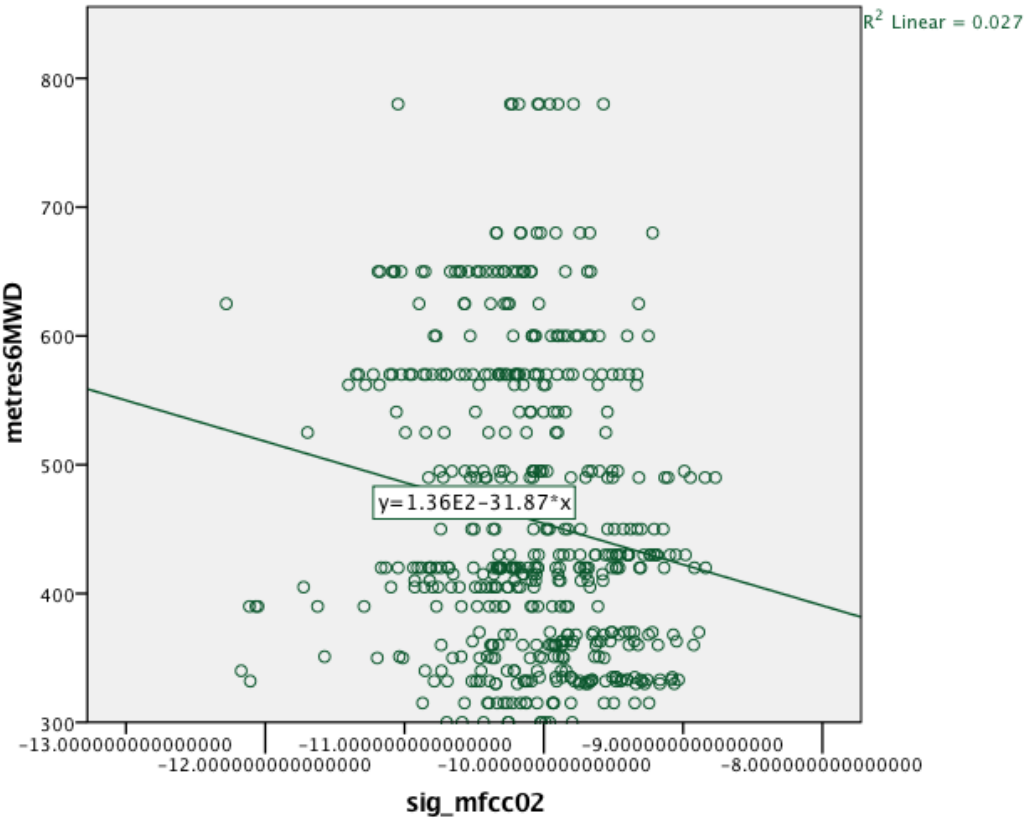
Figure 31 – Scatterplots showing correlation between acoustic features and % predicted Forced Vital Capacity (FVC). The regression line is shown together with the linear regression equation and R squared (top right).

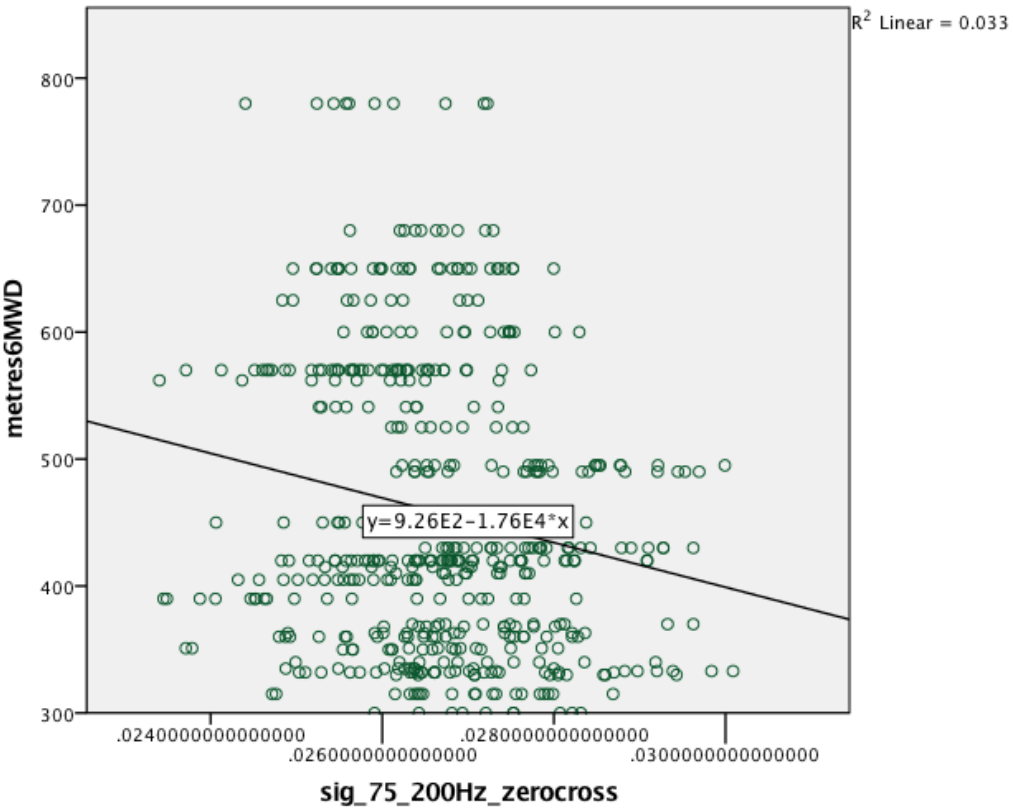

R² Linear = 0.057
metres6MWD
800
700
600
500
400
300
y=5.13E2-2.51E2*x
.0000000000000000
.1000000000000000
.2000000000000000
.3000000000000001
.4000000000000000
.5000000000000000
.6000000000000000
C3 EW_200_500Hz

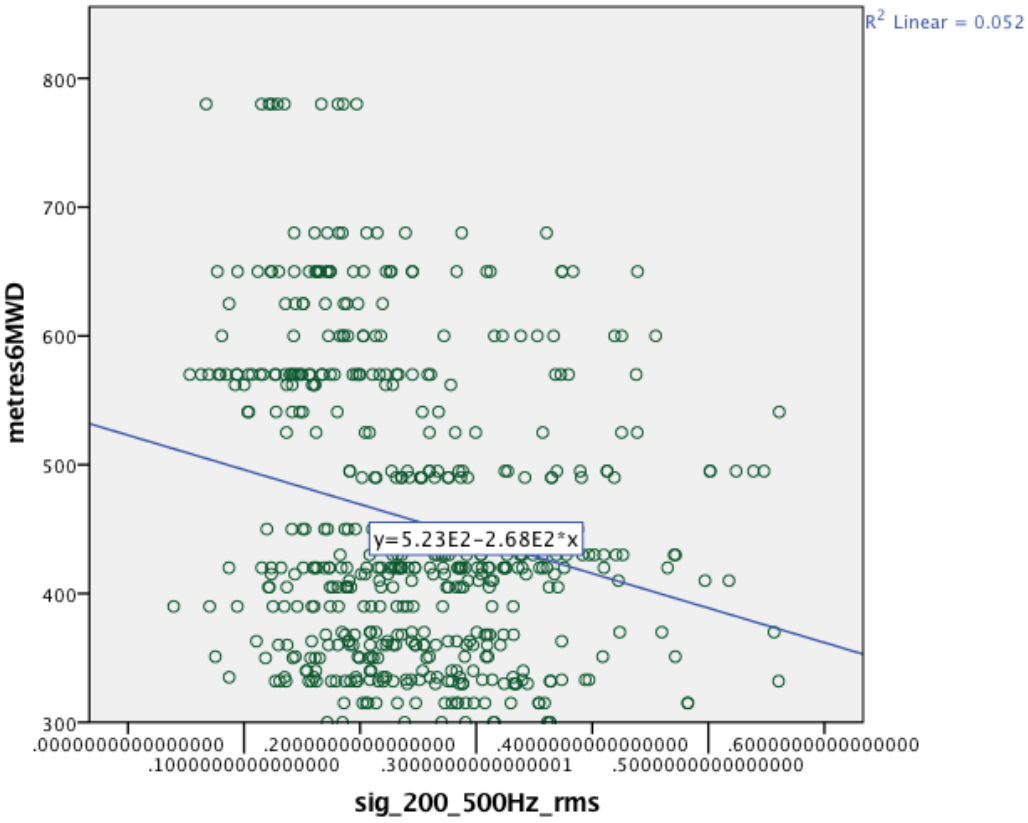

R² Linear = 0.040
metres6MWD
800
700
600
500
400
300
y=3.25E2+2.31E2*x
.2000000000000000
.3000000000000001
.4000000000000000
.5000000000000000
.6000000000000000
.7000000000000000
.7999999999999999
C4 EW_75_200Hz

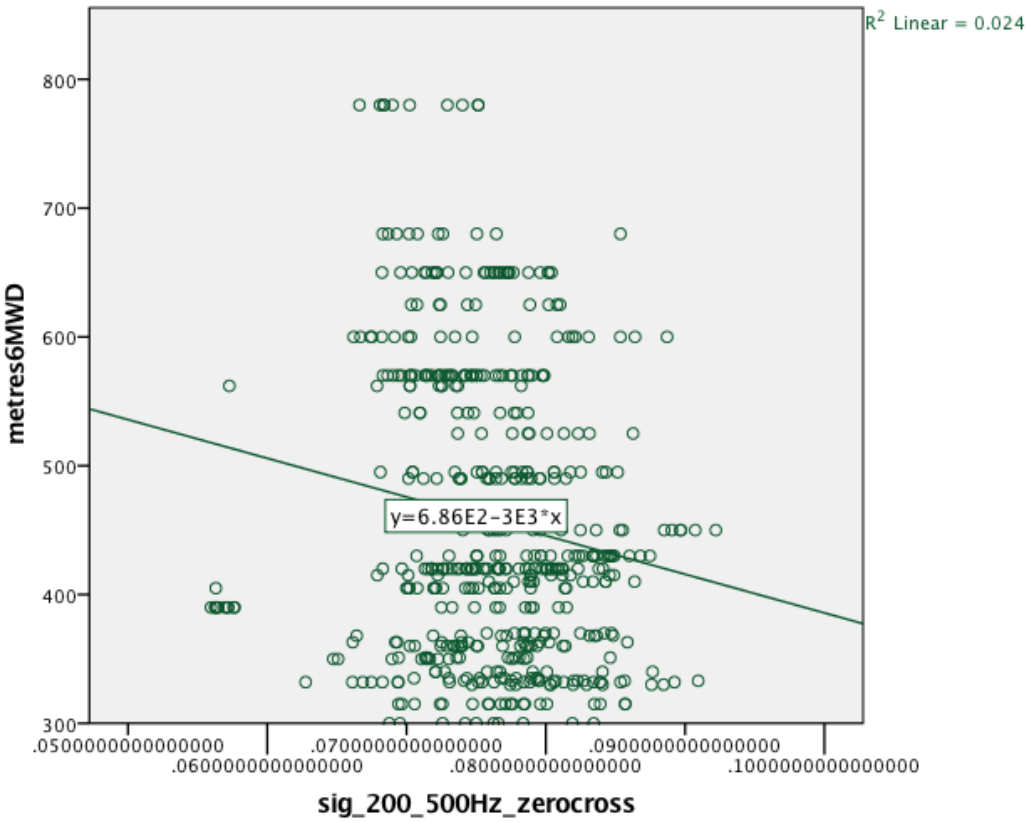

R² Linear = 0.066
metres6MWD
800
700
600
500
400
300
y=5.07E2-3.07E2*x
.0000000000000000
.1000000000000000
.2000000000000000
.3000000000000001
.4000000000000000
.5000000000000000
C4 EW_200_500Hz

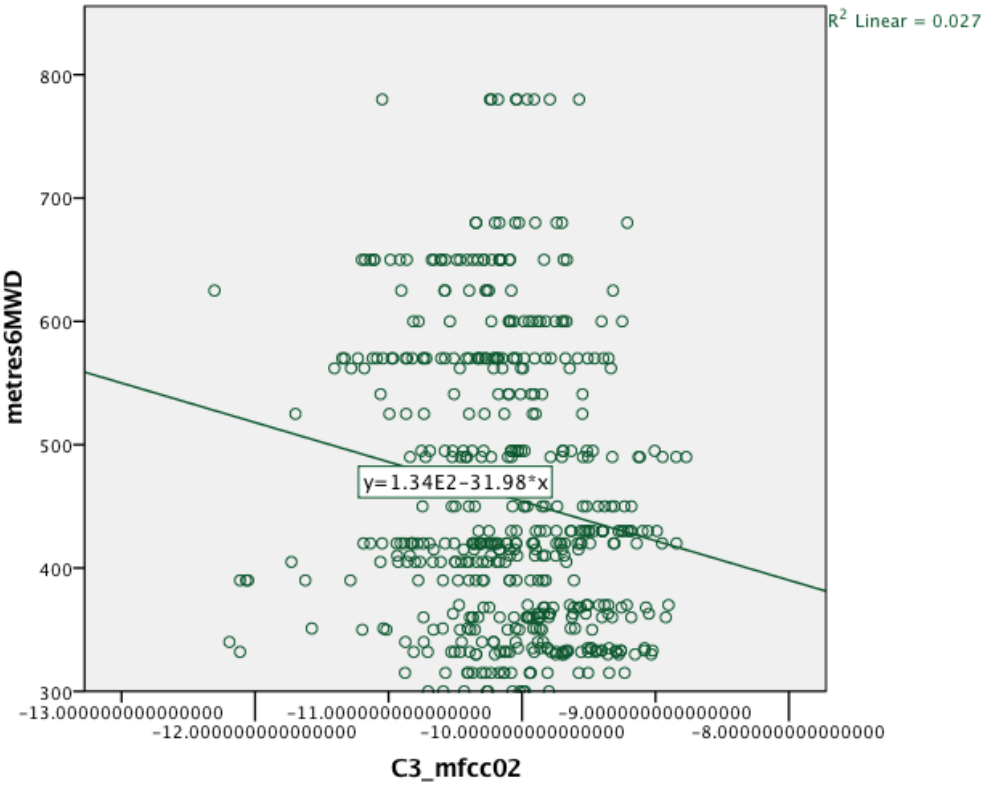

R² Linear = 0.052
metres6MWD
800
700
600
500
400
300
y=5.64E2-2.79E3*x
.0200000000000000
.0300000000000000
.0400000000000000
.0500000000000000
.0600000000000000
.0700000000000000
.0800000000000000
sig_zerocross

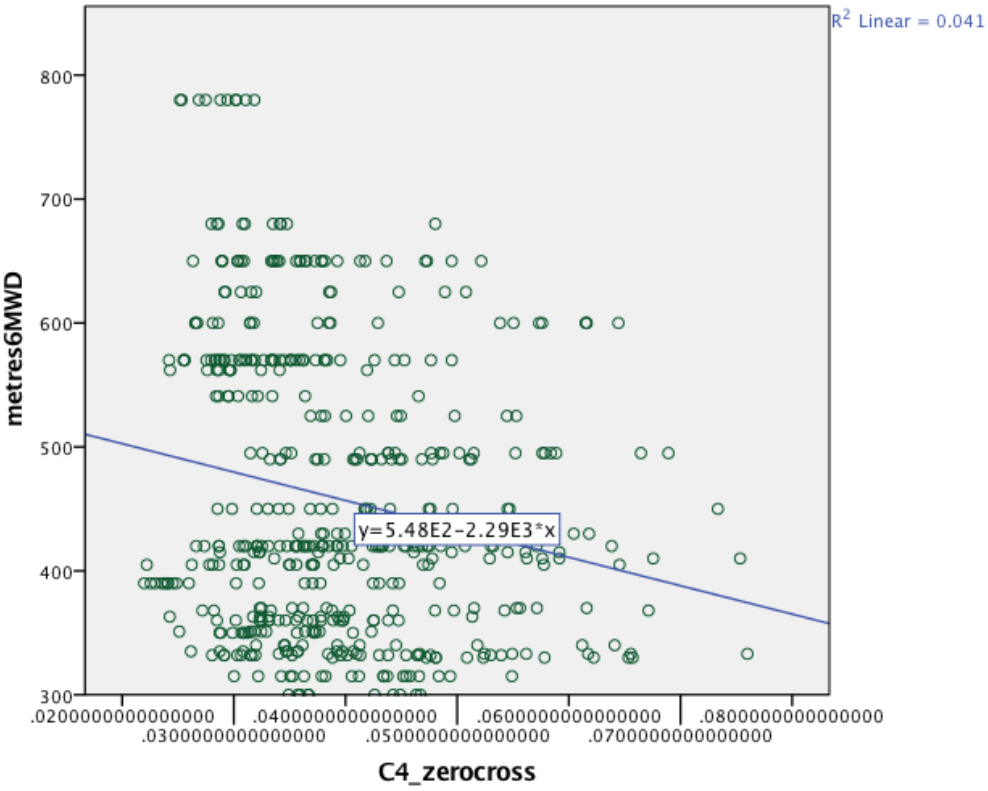

R² Linear = 0.027
metres6MWD
800
700
600
500
400
300
y=1.36E2-31.87*x
-13.000000000000000
-12.000000000000000
-11.000000000000000
-10.000000000000000
-9.000000000000000
-8.000000000000000
sig_mfcc02


R² Linear = 0.033
metres6MWD
800
700
600
500
400
300
y=9.26E2-1.76E4*x
.0240000000000000
.0260000000000000
.0280000000000000
.0300000000000000
sig_75_200Hz_zerocross


R² Linear = 0.052
metres6MWD
800
700
600
500
400
300
y=5.23E2-2.68E2*x
.0000000000000000
.1000000000000000
.2000000000000000
.3000000000000001
.4000000000000000
.5000000000000000
.6000000000000000
sig_200_500Hz_rms


R² Linear = 0.024
metres6MWD
800
700
600
500
400
300
y=6.86E2-3E3*x
.0500000000000000
.0600000000000000
.0700000000000000
.0800000000000000
.0900000000000000
.1000000000000000
sig_200_500Hz_zerocross


R² Linear = 0.027
metres6MWD
800
700
600
500
400
300
y=1.34E2-31.98*x
-13.000000000000000
-12.000000000000000
-11.000000000000000
-10.000000000000000
-9.000000000000000
-8.000000000000000
C3_mfcc02


R² Linear = 0.041
metres6MWD
800
700
600
500
400
300
y=5.48E2-2.29E3*x
.0200000000000000
.0300000000000000
.0400000000000000
.0500000000000000
.0600000000000000
.0700000000000000
.0800000000000000
C4_zerocross

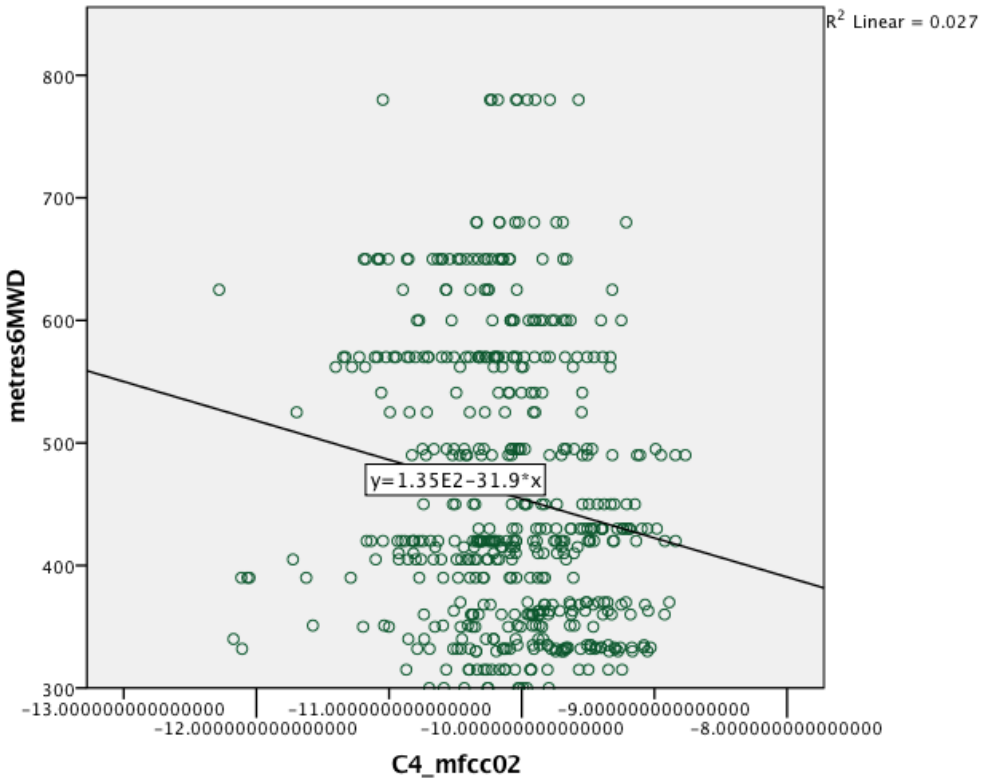


Figure 32 - Scatterplots showing correlations between acoustic features and 6-minute walk distance, expressed as metres walked (6MWD). The regression line is shown together with the linear regression equation and R squared (top right).

5.5 Discussion

5.5.1 Demographic data

Despite being a small sample, the IPF population enrolled in the prospective study showed characteristics quite consistent to those of the populations recruited in several clinical trials in IPF.

The average age and the male predominance reflected the epidemiology of IPF, which occurs primarily in males with advanced age, usually after 60 years. In terms of disease severity at baseline (expressed as functional impairment), this population was pretty similar to the large populations enrolled in large randomised clinical trials (King et al., 2014b,Richeldi et al., 2014). Also, it was heterogeneous as confirmed by the calculation of the GAP stage. This was a positive aspect for study as it assured the inclusion of patients at different stages of severity, likely to progress over time at different rates. The odds of having a good proportion of patients progressing could have been limited by having most patients already in treatment with one of the two available anti-fibrotic agents, pirfenidone and nintedanib. Nevertheless, this could not be avoided as these were consecutive patients enrolled from the ILD clinic at the Southampton General Hospital, which is a prescription centre with many referrals also from other centres. Including patients in treatment with one of the available drugs for IPF is now a compelling issue that has to be faced by the trials coming up in the next years. Most patients in the IPF cohort were non-smokers, in contrast with smoking being a well-known risk factor for developing IPF. Whilst this might find an explanation into the eligibility criteria of the study, excluding those patients with a significant emphysema - which are more likely to be current or former smokers - there were no screen failures in this study, meaning that the prevalence of non-smokers was not due to this factor.

Because of the difference in average age, sex and BMI between the IPF and the control group – which was younger, with a female predominance and a higher BMI - the analyses conducted in this study were adjusted for these variables. In particular, the effect of weight and therefore of a possibly thicker thoracic wall could potentially affect transmission of lung sounds, although this has never been properly studied or clarified. As such, it was deemed as important to enter BMI as a covariate in the analyses involving acoustic features in this study.

5.5.2 Lung function and other clinical parameters

The longitudinal assessment of physiology and other clinical parameters collected from the IPF patients enrolled in the cohort study was crucial to determine whether this sample population

experienced disease progression over the 12 months of observation. This could not be given for granted due to the small size of the population and the possible effect of anti-fibrotic treatments (at baseline, two thirds of the IPF patients were in treatment with either pirfenidone or nintedanib). In line with the clinical practice and the available evidence in the literature, pulmonary function expressed as % predicted Forced Vital Capacity (FVC) was chosen as the main endpoint to express disease progression in this study. An absolute mean change of -110 ml was found between the beginning and the end of the study, a slighter decline than that experienced by populations enrolled in recent large clinical trials (Richeldi et al., 2014) (King et al., 2014b). Nevertheless, this value was calculated from the patients who remained in the study, and does not take into account those who died or dropped-out since no imputation of missing data was made. Solid proof of disease progression however came from ANOVA, which showed a significant rate of decline in % predicted FVC over 11 available observations, with an estimated mean change of 5.9% between baseline and end of study, well inside the range of minimal clinically important difference (MCID) values reported in the literature (du Bois et al., 2011a)). Further demonstration of disease progression in this population also came from the substantial proportion of patients (56.5%) who met the composite endpoint defined as absolute decline of FVC $\geq$ 10% at 12 months, death from all causes or drop-out for inability to perform study visits.

DL_{CO} also showed to decrease significantly in the study population, although the entity of such deterioration was small - mean change of % predicted DL_{CO} was 5.5%, below the MCID reported in the literature.

Low oxygen levels at rest and ultimately respiratory failure (defined as SaO2 < 90%) are signs of severe impairment of gas exchange and are likely to appear only in the most advanced stages of the disease. Resting SaO2 was quite preserved in most of the subjects enrolled in this study, and its values didn't change significantly in the population of IPF patients within the study time frame.

Conversely, a significant desaturation (defined as SaO2 falling below 90% or as a pre- and post-exercise difference > 4%) was found at all time points when 6MWT was performed. This was evident even at the end of the study when the most severe patients already died, dropped out or were unable to perform the test. This was also reflected by the need of oxygen supplemental therapy: at baseline, only one patient reported using ambulatory oxygen during exercise, but during the observation period other patients started using ambulatory oxygen. Despite the average change in the distance covered at the 6MWT didn't reach statistical significance (although a decrease was recorded over the last 6 months), the worsening of breathlessness and fatigue after the 6MWT, measured via the BORG scale, further confirmed the reduced tolerance to exercise of these patients. Overall, the findings pre- and post- exercise are consistent with a

typical IPF population; since concomitant conditions such as chronic heart failure or pulmonary hypertension were excluded at baseline pulmonary fibrosis might be indicated as primarily responsible, although it cannot be excluded some cardiovascular component appeared during the observation concurring to reduce the tolerance to exercise.

Finally, dyspnea and quality of life, measured as patient-reported outcomes via validated questionnaires (UCSD-SOBQ and SGRQ), further proved the significant clinical worsening of these patients over 12 months.

5.5.3 Definition of an "acoustic signature" of IPF

Most studies conducted so far in the field of computerised analysis of lung sounds adopted complex systems enabling simultaneous recording from different areas of the chest, such as multichannel sound analysers. Indeed, such approach is not cost-effective and hardly reproducible in the everyday clinical practice. The novelty of this research consisted in the application of methods for quantitative sound analysis on recordings consecutively collected from different regions on the patients' chest using a simple tool such as an electronic stethoscope. Due to the less controlled conditions, the most correct approach was to select the acoustic properties for characterisation of "Velcro-type" crackles based on a reliability analysis, in the perspective of a "data-driven" study. If validated, the definition of an "acoustic signature" of IPF obtained from a real-life clinical setting using a digital stethoscope for signal acquisition would provide more robust evidence of the clinical usefulness of lung sound analysis for both diagnostic and management purposes.

Lung sounds processing presented a few issues in this study. Since EMD is an empirical technique by definition, the reasons behind these issues are not clearly identifiable. Indeed, it's likely that there was some signal in these recordings impeding the correct decomposition of sounds. Despite a patient-related factor cannot be ruled out – a few patients in the IPF group presented a higher proportion of unprocessed sound files – this didn't seem to be related to the severity of the disease. On the other hand, unprocessed data was clustered at specific recording sites, in particular the recordings taken at the lower lobes and lateral chest sites were more prone to fail EMD. This occurred in both study groups, as such it can be hypothesised it wasn't related to the presence of adventitious lung sounds. Conversely, it's possible that the large range of motion of the chest wall in these regions could generate more acoustic artefacts, such as those produced by the involuntary rubbing of the stethoscope's diaphragm on the skin. This hypothesis is further supported by the successful processing after editing the original files into shorter bits of recordings, that possibly reduced the complexity of the signal and excluded those artefacts from

some edited recordings. Currently, there is no indication as to the optimal length of recording of lung sounds for quantitative analysis. Several studies analysing the characteristics of individual crackles focused on single inspiration or expiration phases, however the isolation of portions of the breathing cycle requires further steps of signal processing. Such approach didn't quite fit the design of this study, meant to provide readily available information based on the characteristics of the sound signal in its original length. As such, taking recordings including from two to three full breathing cycles might be a good compromise to ensure getting an acceptable amount of information and guarantee the successful processing with automated algorithms.

5.5.3.1 Reliability of acoustic features in IPF

To date, the only evidence on the test-retest reliability of the characteristics of lung sounds recorded using a digital stethoscope and processed with CALSA was provided by Marques and co-workers, who demonstrated good to excellent reliability of time-domain features of crackles (such as IDW and 2CD) in a population of patients with bronchiectasis and cystic fibrosis (Marques et al., 2009). On the other hand, the reliability of acoustic features of crackles in ILD has never been assessed. In this research, a reliability analysis was carried out on a large number of acoustic properties. A group of 19 features demonstrated acceptable repeatability (with a cut-off set at ICC > 0.5) when measured three times using a digital stethoscope. The statistical, energy and frequency domains were represented in the set of the reliable features. 13 features were derived from the original signal, whilst 6 were derived from the crackle component generated by EMD. Conversely, no individual intrinsic mode function (IMF) properties were found to be repeatable. Since IMFs represent distinct empirically-defined components of the original signal, it is possible they are less stable at consecutive measurements as compared to the original signal or the crackle component (which consists of the first 3 or 4 IMFs). Interestingly, most of the repeatable features were originated at the frequency bands 200-500 Hz and 500-1000 Hz, suggesting that the signal is more stable at these frequencies.

5.5.3.2 Discrimination between IPF and controls

The comparison between the study groups showed that the combination of the 19 reliable features can effectively discriminate between IPF patients from healthy controls. Furthermore, majority showed to be independently influenced by the presence of pulmonary fibrosis. The features with the largest amount of variation between the groups included the features derived from the cepstral domain, which refers to the analysis of the spectrum of a signal in terms of fineness (see the Glossary at the end of this thesis for more detailed information). These were cepstral coefficients extracted both from the original signal (Sig_mfcc02) and from the crackle components made of the first 3 and 4 IMFs (C3_mfcc02 and C4mfcc02, respectively). This finding

suggests that the properties related to the fineness of the signal and its "pathological" components are the most relevant as to the discrimination of the presence of a fibrotic disease.

The zero-cross rate (ZCR) of the original signal (sig_zerocross) and its component at the frequency band 200-500 Hz (sig_200_500Hz_zerocross) also showed to be different between IPF and controls. The zero-cross rate represents the number of times the signal changes from positive to negative polarity or back, which tends to be high in signals containing crackles due to their characteristic signal waveform.

As described in Chapter 3, different combinations of acoustic properties – among which the zero-cross rate - of "Velcro-type" crackles have already been used for automatic classification purposes, with good results in terms of accuracy (Bhattacharyya et al., 2015,Flietstra et al., 2011). The evidence collected in this study supports the rationale of incorporating a set of acoustic features in automatic classification algorithms with the goal of assisting the diagnostic process of fibrotic ILD, but also provide further useful information. First, the use of the digital stethoscope and the robust characterisation of the study population supports the feasibility of applying these methods in a clinical setting. Secondly, the features assessed in this study for their discriminating properties were not selected *a priori*, but they were screened instead for repeatability. Thirdly, the study design itself, including serial measurement of lung sounds, allowed enter a further level of discrimination consisting in the longitudinal change. As such, the outputs of the comparison analysis between the two groups also relied on the variability of the features over time.

On the other hand, this study had the important limitation of not comparing IPF patients with other potentially competent cardio-pulmonary conditions. Another limitation was represented by the eligibility criteria of the study, excluding patients with comorbidities such as significant emphysema and heart failure. As such, these results might not be applicable to the general IPF population, often presenting concomitant conditions that can alter the transmission of lung sounds. Further studies could clarify the real diagnostic value of this approach towards the identification of ILD/IPF by recruiting an appropriate target population made of ILD/IPF subjects, healthy controls and subjects with other respiratory disorders.

5.5.4 Longitudinal assessment of acoustic features and correlation with clinical parameters

So far, the only study proposing lung sounds as a potential clinical outcome in respiratory disorders is the recent one by Marques and colleagues (Marques et al., 2013), which demonstrated substantial changes before and after intervention (respiratory physiotherapy aimed to clearance of secretions) in time domain-related features of crackles in patients with bronchiectasis and cystic fibrosis. The longitudinal assessment of acoustic features in the IPF group provided novel evidence on the changes of lung sounds related to the progression of the fibrotic process. The cohort study presented in this thesis collected for the first time longitudinal data of lung sounds from a small population of IPF patients. Due to the small size of the population and the lack of similar studies for comparison, the results of this study should be interpreted with caution and further, larger studies are required for their validation.

The first important finding of this analysis was that most of the selected features (15 out of 19) underwent a significant change over the 12-month period in the IPF population. Most of the features not changing significantly were statistical properties of the signal (sig_75_200Hz_centroid, sig_200_500Hz_std_meanframes, and sig_200_500Hz_std_medianframes). Whilst properties belonging to the statistical domain have previously shown good accuracy as to automatic classification (Bhattacharyya et al., 2015), this evidence seems to suggest they may not be fit as indicators of disease progression.

On the other hand, not all the features that changed significantly followed a constant increasing or decreasing trend across the study time points. This observation was important to understand which properties could really reflect the progressive nature of the disease. 6 features showed a consistent trend from beginning till end of the study: Sig_zerocross, sig_500_1000Hz_zerocross, Sig_mfcc002, C3_mfcc02 and C4_mfcc02, and sig_500_1000Hz_centroid. Interestingly, Sig_zerocross, Sig_mfcc002, C3_mfcc02 and C4_mfcc02 were the same that best discriminated IPF subjects from healthy controls in the comparison analysis. This evidence indicates that the zero-cross rate of “Velcro-type” crackles and their frequency-related cepstral properties might be the most clinically relevant in IPF with regards both to discrimination and responsiveness to disease progression.

Nevertheless, an inversion in the trend of the values of these features could be noticed and started occurring around the last period of observation, between the 8th and the 10th month. Although such phenomenon is not easily interpretable, the physiopathology of the fibrotic lung might provide a plausible explanation. The progressive, aberrant deposition of fibrotic tissue in IPF makes lungs increasingly stiff, with consequent decrease of the trans-thoracic differential pressure and the reduction of lung volumes. Consequently, patients in the advanced stages of the disease tend to adapt to the limited pulmonary compliance by increasing the breathing rate, which also affects the velocity of air flowing into the lungs and thus its turbulence. It can be therefore hypothesised that several acoustic properties of lung sounds intensify up to a certain point due to the progression of the fibrotic alterations, and then increase or decrease in the opposite direction due to some change in these system-related, non-stationary components.

The secondary analysis, selectively including the measurements taken at the recording sites corresponding to the lower regions of the lungs (lung bases identified by chest percussion, and lateral chest regions in correspondence of the 4th or 5th intercostal space), provided further insights into the possible role of lung sounds for monitoring IPF. The greater change shown by the values of the acoustic features over time suggests that sounds recorded over the lower areas of the chest might be more efficient in reflecting the progression of the fibrotic process. Furthermore, the trends followed by the values of the features were more consistent than those shown when all recording sites were included in the analysis. Should these results be validated in the future, this evidence might have important implications as to the choice of the best sites to use as a tool for monitoring or predicting disease progression. In fact, since digital stethoscopes allow consecutive recording of sounds, the acquired signals must be analysed separately and it would be impractical to use recordings from all over the chest, other than potentially difficult to interpret. On the other hand, the potential issues related to the processing of signals acquired from lower regions - discussed in the previous section - should be taken into account and properly addressed in future studies.

Whilst the significant increase or decrease shown by acoustic features possibly reflected the disease progression that occurred in the IPF population, proved via the observed decline in pulmonary function, no strong relationships were found between individual acoustic properties of “Velcro-type” crackles and the other clinical parameters. Overall, there seemed to be no relationship at all between sounds and worsening of symptoms or health status, whilst some degree of correlation was shown between several acoustic features and % predicted FVC or the distance walked at 6MWT, indicating that the properties of the acoustic signals were responsive to the functional decline and to exercise tolerance.

On multivariable analysis however, while the set of acoustic features examined accounted for only 3.9% of the change in % predicted FVC, it showed to be a much better predictor of exercise tolerance, as it explained 10% of variability of the distance covered at the 6MWT. As such, these findings seem to suggest that lung sounds might be a valid parameter to monitor the overall physical performance of individuals affected by IPF.

Several reasons may explain the overall weak relationship between acoustic properties of lung sounds and most of the other parameters measured in this study. First, it is recognised that the approach chosen for the correlation analyses might have partly biased the results: the single values of clinical and functional parameters have been repeatedly associated with different measurements of acoustic features obtained from different sites. An alternative approach could consist in averaging the measurements of the acoustic features obtained from the same patient, however this would have significantly reduced the number of observations. A second explanation lies in the different domains these parameters belong to: lung sounds might reflect a broad range of pathobiologic mechanisms occurring in the lung parenchyma that not necessarily affect the overall pulmonary performance of an individual. Finally, modifiers might influence these parameters in different (and possibly opposite) ways. Although clinically significant comorbidities such as pulmonary hypertension, chronic heart failure and significant emphysema were excluded at baseline, some patients might have developed such conditions during the period of clinical observation. Emphysematous alterations tend to normalise pulmonary volumes and are believed to attenuate the transmission of pathological lung sounds. On the other hand, the presence of fluid in the peripheral airways or alveoli from congestive heart failure might influence the acoustic properties of lung sounds without affecting the overall pulmonary function. Interestingly, the hypothesis of fluid retention in the lungs of these patients might explain the stronger relationships found between acoustic features and the distance covered at the 6MWT, since exercise tolerance is clearly multifactorial, and its impairment may result from respiratory conditions as wells as other components such as concomitant cardiovascular disease.

Among the features that changed significantly during the observation period, the zero-cross rate of the original signal (sig_zerocross) was the only feature also responsive to changes in both pulmonary function and exercise tolerance. On univariate analysis, the other features with the strongest relationships with % predicted FVC and 6MWD were the zero-cross rate of the crackle component (C4_zerocross) and those related to the energy content of the crackle components of the signal at different frequency bands - C3 EW_75_200Hz, C4 EW_75_200Hz, C4 EW_200_500Hz. These findings were confirmed on multivariable regression analysis, where sig_zerocross and C4 EW_200_500Hz showed to be the best independent predictors of change in % predicted FVC and 6MWD, respectively. Such evidence supports the hypothesis that sig_zerocross might be the most

valid individual acoustic feature to assess disease progression in IPF, as it is highly reproducible, shows a constant trend of change over time and independently reflects changes in validated clinical parameters. On the other hand, energy-related features of the crackles might be good indicators of exercise tolerance: whilst these properties did not follow a steady trend during the observation period in the primary analysis of the longitudinal change of features, they actually did in the secondary analysis including only the measurements from lower recording sites, which indicate that the energy content of the crackle components of the signal might represent another relevant property for the assessment of disease progression.

5.6 Conclusions and future work

5.6.1 Main conclusions

This pilot cohort study provided the first evidence on the longitudinal changes of lung sounds in patients with fibrotic ILD, and explored the potential role of quantitative lung sounds analysis as a novel tool for monitoring disease progression. A set of highly reproducible acoustic features were screened for discriminating IPF patients from healthy controls; in a functionally declining population of IPF patients, some of the selected features demonstrated a significant and constant trend of change over time, that can be related to the underlying progression of the fibrotic process. Furthermore, the identified combination of acoustic features also showed a certain degree of responsiveness to other validated measurements of disease progression, including functional parameters and tolerance to exercise. As such, there is a set of acoustic properties of lung sounds which may be recognised as a clinically relevant “acoustic signature” of IPF.

These results were obtained using a longitudinal design with serial measurements in a small but heterogeneous cohort of patients with IPF. A large acoustic data set was serially collected from patients with IPF and healthy controls using a digital stethoscope. Standardised methods were used for computerised analysis of lung sounds, and a consequential, robust approach was followed to evaluate the reproducibility, the discriminating value and the changes in a broad set of acoustic features.

In conclusion, the quantitative analysis of lung sounds represents a new, appealing approach to monitor disease course in IPF. Further research is needed to validate the findings of this pilot study and determine the real prognostic yield of computerised analysis of lung sounds, thus clarifying whether it might improve the management of these patients in the clinical practice.

5.6.2 Summary of study limitations

In this pilot, single-centre study, the change in acoustic features of lung sounds was chosen as the primary outcome. However, with no sample size calculations nor prognostic enrichment strategies feasible, the risk was to fail in capturing a significant signal of change if patients with relatively stable disease were included. The high proportion of patients in treatment with anti-fibrotic agents was another potential limitation to meeting the proposed objectives, since these patients were more likely to experience disease progression at a slower rate. On the other hand, excluding such patients would have proved to be highly impractical, since only a minority of patients with IPF are not being treated nowadays. Moreover, most of these patients are not treated because they have advanced disease, as such this would have raised the risk of a high drop-out rate, with fewer observations available for analysis. Therefore, the only available strategy to maximise the efficiency of the study would have consisted in enlarging the study population base by allocating a longer time for recruitment: however, this was not feasible either due to the pre-specified overall study duration. More importantly, despite the prospective cohort design and the serial measurement of various parameters fit very well a prognostic research study, the investigation of the value of lung sounds as a marker of disease progression in this study was also clearly limited by the small population size: for example, a survival analysis was not feasible due to the low number of events recorded. The period of clinical observation was also relatively short to register an adequate number of events to determine the predictive value lung sounds toward mortality. The 12-month duration was decided based on most randomised clinical trials in IPF - usually powered to detect an agent's efficacy in terms of reduction of the rate of functional decline - since the primary objective included to determine the relationships between lung sounds and other parameters of disease progression.

The eligibility criteria used in this pilot study also had limitations. Patients with IPF and evidence of other pulmonary or cardiologic conditions were excluded at baseline when these were supposed to affect the transmission of lung sounds strictly related to the progression of the fibrotic disease. This was meant to avoid that these factors could mask the change in the acoustic features produced by the fibrotic process. Generally, such approach limits the applicability and generalisability of findings, and future studies should be designed to include these patients and quantify the influence of such modifiers. Furthermore, the inclusion of a population consisting of subjects with pulmonary conditions other than ILD and cardiovascular disease would represent a far better control population as compared to healthy volunteers to determine which acoustic feature are specifically distinctive of a pulmonary fibrosis.

On the other hand, the presence of other effect modifiers can be recognised but these were not fully addressed in the data analysis. While the comparison between IPF patients and healthy volunteers was adjusted for BMI, the specific impact of body composition on the strength of the relationships between the acoustic features and the other parameters was not measured. As such, future studies should consider obtain measurements of parameters such as the thickness of the thoracic wall on CT in order to understand whether and how they interfere with the transmission of the pathological sounds investigated. Flow rates and tidal volumes were not monitored or targeted in this research either, as such it cannot be excluded that the measurement of acoustic features might have been influenced by a certain degree of variability of such parameters in the examined population. On the other hand, while in healthy human volunteers lung sound intensity is known to be highly dependent on airflow at the mouth (Kraman, 1984), the frequency spectrum of lung sounds does not seem to be affected by airflow (Kraman, 1986). Moreover, it must be reminded that the standardisation of lung sounds recording through pneumotachography techniques is not easily applicable in the everyday clinical practice.

With regard to the data analysis, only the 19 most stable features have been evaluated longitudinally. However, the cut-off of ICC > 0.5 was chosen arbitrarily: whether more features could undergo relevant modifications using a less strict cut-off, it remains unknown. Finally, in the analyses of correlation single values of clinical and functional parameters have been repeatedly associated with measurements of acoustic properties obtained from different sites: as discussed in the previous section, this approach allowed maximise the number of observations in the study but might have biased the results of the analyses performed.

5.6.3 Predicting progression in IPF using lung sounds: future steps

Building on the insights provided by this pilot study, a practical proposal for future research to determine the value of the quantitative analysis of lung sounds toward prediction of disease progression in IPF and develop a prognostic tool useful in the clinical practice is described below.

Just as diagnostic research, prognostic research should be performed in close adherence to daily clinical practice to ensure wide applicability of the findings. The object of any prognostic process is inherently longitudinal, as such a longitudinal design should be followed where the candidate predictors are measured before the outcome is observed. The data should be preferably collected prospectively rather than retrospectively, as it allows the optimal measurement of predictors and outcomes as well as adequate (possibly complete) follow up. The study population domain should consist of patients presenting with a certain disease in a specific setting. When evaluating a new potential tool for monitoring disease progression, as in the case of lung sounds, a secondary or

tertiary care setting would be the ideal choice, since the follow up of these patients takes place in specialised centres, such as hospitals qualified for prescribing anti-fibrotic treatments. In the past, large randomised trials of IPF candidate therapies offered the opportunity to access a large amount of longitudinal data for a wide spectrum of parameters, leading to the development of the currently available models for risk prediction in IPF. Nevertheless, the applicability of these models in the clinical practice has been debated, mostly due to the strict eligibility criteria applied by these trials, excluding patients with more advanced disease and significant comorbidities (Wells, 2015). As such, large, prospective cohort studies including a broader, more representative population of IPF patients represent a more desirable approach. The downside of including a heterogeneous population is that the nature and the strength of the associations between the prognostic markers and the outcome is prone to the influence of modifiers. Some factors potentially influencing the measurement of lung sounds produced by pulmonary fibrosis have been recognised and discussed in this pilot study. These include coexisting pulmonary and cardiac disorders, the patient's own body composition and the airflow rates obtained during tidal breathing. These should be monitored and their impact properly assessed in future research: the relationships between the acoustic features of lung sounds and other parameters could be studied across different categories of patients, for example those with or without a concomitant chronic heart disease, or across different BMI ranges. The measurement of the thickness of the chest wall in correspondence to the sites of recording could be incorporated as covariate in the correlation analyses to assess whether (and to what extent) it may influence the strength of association between acoustic features and other parameters. In order to maximise the study efficiency in capturing a broad range of outcomes, a multicentre design should be preferred as it would allow increase the rate of outcome events in a significant way. With this regard, the strategy of including the quantitative assessment of lung sounds among the exploratory endpoints of future clinical trials should be considered.

Generally, when establishing whether a new tool is a suitable marker of disease progression or bad outcome, the first step consists in the assessment of its validity against other parameters of disease severity and progression, performed by determining their relationships through the calculation of correlation parameters or linear regression modelling. While this study demonstrated that several acoustic properties underwent a significant change over time in IPF patients, the strength of the relationships with validated parameters of disease progression such as FVC was pretty low. Indeed, larger cohorts (and thus larger data sets) might help clarify the relationships of the quantitative assessment of lung sounds with measurements of functional decline or tolerance to exercise. Notably, more robust demonstration of the validity of lung sounds as a marker of disease progression could be obtained through the direct association of

acoustic features with radiologic indices measured via visual scoring or via quantitative CT assessment of corresponding regions or volumes of parenchyma. In fact, building on the evidence available in this field (Jacob et al., 2016,Jacob et al., 2017,Maldonado et al., 2014a), the quantitative estimates of fibrosis on CT using a validated software such as CALIPER will allow determine the responsiveness of the acoustic properties to modifications in the extent of different interstitial abnormalities.

The final aim of a prognostic study is to determine the predictive value of one or more determinants toward the occurrence of an outcome of interest, i.e., which is clinically meaningful to patients. Mortality, incidence of hospitalisations or acute exacerbations, and future functional decline are all valid outcomes to explore in IPF. As discussed in section 5.2, prediction of future functional decline would be incredibly useful in clinical practice, but has proven to be extremely difficult to investigate, and IPF remains an unpredictable disease in individual patients. Since physicians use multiple predictors to estimate a patient's prognosis, including a variety of demographical and clinical characteristics together with tests results, prognostic research should also involve a combination of determinants, rather than a single one. As such, a multivariable approach should be followed in the design and analysis to determine which predictors independently contribute, and to what extent, to the outcome prediction. As described above, several clinical prediction models and risk scores have been already developed in IPF starting from different combinations of (formerly individually) validated parameters, which allowed estimate the risk of developing the disease-related event (e.g. mortality) in the daily practice. It must be kept in mind though that the choice of the candidate prognostic variables should always reflect clinical thinking, and should not be solely based on their individual predictive value in statistical models. Also, such choice should take into account the feasibility of the measurements and tests in the everyday practice.

Instead of selecting a narrow set of acoustic features empirically (an approach followed by many studies in the field of lung sound analysis so far), this pilot study importantly defined a highly reproducible combination of clinically relevant features that a future study might incorporate together with other simple parameters to build a valid prognostic tool for IPF patients. Short-term changes in lung sounds features might be evaluated in combination with other parameters such as age, measures of burden of symptoms, changes in lung function measurements and CT indices toward the occurrence of a composite endpoint representative of disease progression, e.g. including mortality, hospitalisation and categorical decline in predicted FVC >10%. The longer the time available to evaluate the occurrence of such outcome (and thus the period of clinical follow up of the study population), the greater the applicability of the results will be. On the other hand, the follow-up time between participants will vary substantially, and prediction will become more

problematic. 1- or 2-year follow up seems a good compromise to guarantee study feasibility and applicability of the findings. For time to event outcomes, univariable analysis can be performed using the Kaplan-Meyer method, while multivariable analysis can be performed using Cox proportional hazard modelling. Despite many studies include hundreds of patients that develop the outcome event, for Cox proportional regression analysis at least 10 subjects in the smallest of the outcome categories are needed for proper statistical modelling (Concato et al., 1995). In order to guide decision making in the individual patients, the reporting of a prognostic study should concentrate on period-specific absolute risk estimates of the outcome rather than reporting the relative risks of developing the outcome, which is not informative in clinical practice. When a novel, potentially valuable prognostic marker is being tested, it might be directly compared to other individual predictors in univariable analysis. Nevertheless, the best approach would consist in the development of a multivariable prediction model including the candidate predictors that retain independent association with the outcome. This should be followed by comparing the prognostic accuracy of two models – one including and one not including the proposed combination of acoustic features - to quantify the added value of the new predictor to more conventional parameters. Methods to determine and compare the discriminatory power (as such, the performance) of prediction models with time to event outcomes include the c-index statistic (which is numerically equivalent to a ROC curve) and more advanced metrics such as the net reclassification improvement (NRI), which quantifies the extent to which a model with a newly introduced predictor improves the classification of participants with and without the outcome as compared with the basic model without that predictor (Pencina et al., 2008); this is done by calculating the number of patients correctly reclassified into meaningful low or high risk categories when the new predictor is added.

After validating the findings in external cohorts and settings, the final step will require the development of a software that can automatically calculate and report the probability of a certain outcome for a patient after a series of parameters (including lung sounds recordings) are entered. This should proceed in parallel with the design and the validation of a simple, effective graphical interface that could be easily run by healthcare professionals and lay users.

Appendices

Appendix A Ethics documentation (cohort study)

A.1 Letter of favourable opinion

Health Research Authority
NRES Committee South Central - Berkshire
Bristol REC Centre
Whitefriars
Level 3, Block B
Lewins Mead
Bristol
BS1 2NT
Telephone: 0117 342 1389

22 January 2015

Dr. Giacomo Sgalla
Southampton Centre for Biomedical Research
Level D, South Block
Southampton General Hospital,
Tremona Road, Southampton
SO16 6YD

Dear Dr. Sgalla

Study title: **Electronic recording of lung sounds as clinical endpoint in Idiopathic Pulmonary Fibrosis: a longitudinal pilot study**
REC reference: **14/SC/1429**
IRAS project ID: **143584**

Thank you for your letter of 20 January 2015, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact the REC Manager, Libby Watson, at: nrescommittee.southcentral-berkshire@nhs.net.

Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

NHS
Health Research Authority

Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Sponsor's insurance]		20 November 2014
GP/consultant information sheets or letters [Letter to GP]	1	19 November 2014
IRAS Checklist XML [Checklist_16012015]		16 January 2015
Letters of invitation to participant [Letter of invitation to participants]	1	19 November 2014
Other [Study protocol - revised (tracked changes)]	2	14 January 2015
Other [Response letter to provisional opinion]		16 January 2015
Participant consent form [IParticipant consent form]	2	14 January 2015
Participant consent form [Participant consent form - revised]	2	14 January 2015
Participant information sheet (PIS) [PIS - IPF patients]	2	14 January 2015
Participant information sheet (PIS) [PIS - IPF patients - revised]	2	14 January 2015
Participant information sheet (PIS) [PIS - Healthy volunteers]	2	14 January 2015
Participant information sheet (PIS) [PIS - healthy volunteers - revised]	2	14 January 2015
REC Application Form [REC_Form_15012015]		15 January 2015
Referee's report or other scientific critique report [SUHT peer review]	1	11 September 2014
Research protocol or project proposal [Study protocol]	2	14 January 2015
Summary CV for Chief Investigator (CI) [Chief Investigator's CV]		
Summary CV for supervisor (student research) [Supervisor's CV]		
Validated questionnaire [SGRQ questionnaire]		
Validated questionnaire [SOBQ questionnaire]		

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document *"After ethical review – guidance for researchers"* gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

NHS
Health Research Authority

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:
http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at
http://www.hra.nhs.uk/hra-training/

14/SC/1429	Please quote this number on all correspondence

With the Committee's best wishes for the success of this project.

Yours sincerely

pp.

Mr David Carpenter
Chair

Email: nrescommittee.southcentral-berkshire@nhs.net

Enclosures: "After ethical review – guidance for researchers" *[SL-AR2]*

Copy to: *Ms Diana Galpin*
Mrs Penny Silsbury, University Hospitals Southampton NHS Foundation Trust

A Research Ethics Committee established by the Health Research Authority

NHS
Health Research Authority

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database. This should be before the first participant is recruited but no later than 6 weeks after recruitment of the first participant.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra.studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non-registration may be permissible with prior agreement from NRES. Guidance on where to register is provided on the HRA website.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

NHS sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Copies of advertisement materials for research participants [Poster]	2	12 January 2015

A.2 Information sheets

A.2.1 IPF group

University of Southampton

University Hospital Southampton NHS
NHS Foundation Trust

Participant Information Sheet

Electronic recording of lung sounds as a clinical endpoint in Idiopathic Pulmonary Fibrosis: a longitudinal pilot study

Dear prospective participant,

Many thanks for taking the time to consider participating in our research study. Before you decide, we would like you to understand why this research is being done and what it would involve for you. One of our team will go through the information sheet with you and answer any questions that you might have. We would suggest this should take about ten to fifteen minutes.

Talk to others about the study if you wish.

Part 1 tells you the purpose of this study and what will happen if you take part.

Part 2 gives you more detailed information about the conduct of the study.

Please ask us if there is anything that is not clear.

Local Contacts:

Your Study Doctor:
Dr Giacomo Sgalla
Tel: 02381205399
Email: giacomo.sgalla@uhs.nhs.uk

Your research nurses:
Ms Louise Stanley
Ms Kate Howard
Tel: 02381204479
Emails: louise.stanley@uhs.nhs.uk ; kate.howard@uhs.nhs.uk

IPF cohort

PIS Role of Lung sounds in IPF
REC ref: 14SC1429
Version 3

Page 1 of 12

Part 1

1. What is the purpose of the study?

Idiopathic Pulmonary Fibrosis (IPF) is a chronic disease causing a progressive scarring into the lungs. This scarring leads to the development of typical sounds produced by lungs, which are easily audible by doctors when listening to the chest of a patient through the stethoscope.

This is an educational study aiming to evaluate if lung sounds in patients with IPF may provide useful information about the course of the disease, i.e. how fast it will progress in the future. This would inform physicians about the management of IPF patients, as to making decisions about the optimal therapy to choose.

2. Why have I been invited?

You have Idiopathic Pulmonary Fibrosis and your doctors think you might be suitable for this study. Around 30 patients will take part in the study.

3. Do I have to take part?

No.

It is up to you to decide whether to join the study. We will describe the study and go through this information sheet. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any time, without giving a reason. This would not affect the standard of care you receive from your doctor.

4. What will happen to me if I take part?

You will be in the study for 12 months. Visits will take place every 2 months at the Wellcome Trust research facilities - Southampton Centre for Biomedical Research at the University Hospital Southampton NHS Foundation Trust. This means you will be attending 7 visits (1 screening/baseline visit + 6 follow up visits) over this period.
The study will cover 2 years, a time necessary to collect all the data and analyse them.
If you are eligible following screening, you will enter the study.

Screening/Baseline (Visit 1) assessments

Screening visit is needed to determine whether you can take part to the study. If so, you will go through some tests and procedures described below. Some of these tests or procedures may be part of your regular medical care and may be done even if you do not take part in the study. If you have had some of them recently, they may not need to be repeated.

IPF cohort

- Discussion of this study and review and signing of this Informed Consent Form

  You will discuss study details with a member of the research staff (the study doctor or a research nurse). If you are happy to take part in the study you will be asked to sign a consent form. You will be given a signed copy of the consent form. Your demographic information, including your age, sex, and race/ethnicity will be recorded.

- Review of your medical history and any medications you are taking.

- Complete physical examination, measurement of your vital signs (breathing rate, heart rate, blood pressure, and body temperature), resting oxygen saturation.

  Pulse oximetry is a simple, non-invasive technique used to monitor the oxygen in your blood. It monitors the percentage of hemoglobin that has absorbed oxygen. For the test, a sensor is placed on a thin part of the body, usually a fingertip or earlobe.

- Recording of lung sounds.

  Lung sounds will be recorded from 6 sites on the back of your chest through an electronic stethoscope. This stethoscope is no different from a regular one, but it contains a microphone capable of recording sounds when it is applied on the chest.

- Measurement of your height and weight

- High-Resolution Computed Tomography

  High-resolution computed tomography (HRCT) is a type of X-ray used to look at the structure of your lungs and to evaluate the presence of fibrosis in your lung tissue. You may not need to undergo HRCT if you have this assessment performed recently (within 12 months prior to the screening visit).

- Spirometry/Pulmonary Function Testing

  Spirometry is a breathing test that requires that you breathe in and out through a tube you place in your mouth. This requires you to take deep breaths and blow out hard and fast. A spirometry test is used in the evaluation and diagnosis of your breathing difficulty. The purpose of this test is to help your doctor to find out how

much the lung fibrosis has limited your ability to breathe in and out.

- Gas Exchange-Diffusion Capacity of the Lung for Carbon Monoxide Assessment (DL_{CO})

  A DL_{CO} test is used in the evaluation and diagnosis of your breathing difficulty. The purpose of this test is to help your doctor in determining how much your IPF impairs the exchange of oxygen between your lungs and your blood. The test is typically performed as part of spirometry testing.

- 6-Minute Walk Test (6MWT)

  The 6-minute walk test (6MWT) measures your exercise tolerance. The 6MWT measures how far you can walk over a total of 6 minutes on a hard, flat surface. The goal is for the individual to walk as far as possible in 6 minutes. You will be required to fill out a brief questionnaire before and at the end of the 6MWT.

- Blood Samples

  About 2-3 tablespoons (15–20 mL) of blood will be obtained from a vein in your arm to test the serum levels of specific substances (called biomarkers) that might be related to IPF and its progression. Other tests may be performed to look at specific genes in your DNA that are related to IPF and its progression.
  With your specific consent, any remaining samples may be used for future research after approval by a Research Ethics Committee. Otherwise, they will be destroyed at the end of the study.
  The samples provided will be stored in secured facilities at the University Hospital of Southampton for the whole duration of the study.

- Completion of questionnaires about your IPF and your overall health.

The visit 1 assessments will be repeated at each of the following 6 visits, except for:

- Complete physical examination, measurement of vital signs, pulse oximetry: only at baseline (visit 1).

- Blood samples: performed at visits 1, 4 and 7 (End of study visit), i.e. a total of 3 times throughout the study.

- 6 minute walking test: performed at visits 1, 4 and 7 (End of study visit), i.e. a total of 3 times throughout the study.

- HRCT scan: performed at visit 1 (if not performed within previous 12 months) and visit 7 (end of study) - a maximum of 2 tests performed throughout the study. This serves to evaluate if the fibrosis (the scarring in your lungs) has progressed and how much.

Visit 1, 4 and 7 would be slightly longer than the others (approximately 2.5 hours each). It is possible that you will be asked to undergo the HRCT scan on different days from the scheduled dates for visit 1 and 7, depending upon when it is possible to book an appointment for you.

5. Will I receive expenses and/or payments?

You will receive a payment of £ 70 per each visit, to compensate you for your time and inconvenience and for reimbursement of travel/parking expenses.
You will be provided with drinks and meals during the visits.

6. What will I have to do?

<u>During the study you should:</u>

- Attend all your study appointments. If you know that you will miss an appointment, contact the study staff to reschedule it as soon as possible.
- Follow the study staff's instructions with regards to study procedures.
- Complete questionnaires about your health / disease
- Tell the study staff about any adverse events or symptoms whether related or unrelated to the study, GP or hospital visits that you may have, as well as any medicines or supplements you might be taking.
- Ask questions as you think of them.
- Tell the study staff if you change your mind about staying in the study.
- Continue your current IPF medications.
- Tell the study staff if you think you have become pregnant.

- Tell the study staff if you are taking part or if you intend to take part in any other studies. Since this study does not involve drugs or invasive procedures, taking part to other studies is permitted, but it is advisable to check with the study doctor the sustainability of participating to more studies at once, which will depend on your preferences and health status.

7. What are the possible disadvantages and risks of taking part?

There are possible risks, disadvantages and inconveniences with any research study. Consider these carefully before agreeing to participate in this study.

You will have more tests and procedures if you take part in the study, compared to standard hospital visits. Study visits could take more time than standard hospital visits and you are likely to have more blood taken.

Risks associated with recording lung sounds
Recording lung sounds through an electronic stethoscope is a completely non-invasive and safe procedure.

Possible risks and discomfort associated with drawing blood

During this study, blood will be drawn from a vein and used for tests. Drawing blood may cause pain where the needle is inserted, and there is a small risk of bruising and/or infection at the place where the needle is inserted. Some people experience dizziness, upset stomach, or fainting when their blood is drawn.
The maximum amount of blood that will be taken is approximately 1-2 tablespoons per each visit in which this test is due. The total amount taken over the entire study is not considered to be a health risk for you.

Spirometry risks

Breathing tests will be done using machines that measure how much air you breathe in and out. A machine to measure your breathing (a spirometer) will be provided to your doctor for use during this study. Your doctor and the staff at your doctor's office will be specially trained on how to use and maintain the machines to measure your breathing. You will wear nose clips and be instructed to blow into a machine. You will be expected to blow as hard and as fast as you can for as long as you can. You may be asked to do this up to 3 or 4 times.
A second type of test called diffusion capacity (DL_{CO}) will also be performed with this machine. The purpose of these tests is to estimate how much your IPF impairs the

Insurance

Before taking part you should consider if this will affect any insurance you have, including travel insurance, and seek advice if necessary.

8. What are the possible benefits of taking part?

There are no direct benefits from taking part in this study, however the information gathered from your participation may help others suffering with the same condition in the future.
You may benefit from more frequent medical supervision.

9. What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

10. Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. The details are included in Part 2.

This completes part 1.

If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.

exchange of oxygen between your lungs and your blood. You will be instructed to exhale then inhale rapidly and hold your breath for approximately 8 seconds. After the 8 seconds have passed you will be asked to blow out again. The test might be repeated two or more times. Sometimes these procedures can cause you to cough, experience shortness of breath, or feel lightheaded, but there is no pain expected with these tests.

Possible risks and discomfort associated with HRCT scans

High Resolution Computerised Tomography (HRCT) scans are special X-ray tests used to study the internal organs and bones of your body. This is necessary for the diagnosis of IPF, and useful to see how the scarring in the lungs progresses over time.
You will be asked to undergo 1 or 2 HRCT scans throughout the study:

- You will undergo a HRCT scan at baseline (visit 1) if you are found eligible for the study. You don't need to have this scan if you have already had one within 12 months prior to screening.

- You will undergo a HRCT scan at the end of the study (visit 7), approximately 12 months after the start of the study.

To undergo a HRCT scan means you will be exposed to radiation, that can cause cell damage which may, after many years or decades, turn cancerous. The total radiation dose to your whole body per single test is considered equivalent to about 3 years of natural background radiation in the UK. The total maximum dose of radiation for this study relates to a 1 in 1400 risk of a potentially fatal induced cancer in a healthy population.
If you are a woman and are pregnant when you enter the study, or should you become pregnant at any moment during the study, you will be kept in the study but HRCT scans won't be performed if they fall into the gestational period, in order not to expose the foetus to radiation.

Possible risks and discomfort associated with the 6-minute walk test.

The 6-minute walk test (6MWT) measures your exercise tolerance. The 6MWT measures how far you can walk over a total of 6 minutes on a hard, flat surface. The goal is for the individual to walk as far as possible in 6 minutes. If you are receiving oxygen at home, the amount you are receiving may temporarily be changed during the study. You will be required to fill out a brief questionnaire before and at the end of the 6MWT.
During the 6MWT you will be asked to walk as far as you can in 6 minutes, but you may slow down or stop if you feel tired or have other symptoms. Symptoms that you may experience include: chest pain, severe shortness of breath, leg cramps, sweating, mental confusion, and/or headache.

Insurance

Before taking part you should consider if this will affect any insurance you have, including travel insurance, and seek advice if necessary.

8. What are the possible benefits of taking part?

There are no direct benefits from taking part in this study, however the information gathered from your participation may help others suffering with the same condition in the future.
You may benefit from more frequent medical supervision.

9. What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

10. Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. The details are included in Part 2.

This completes part 1.

If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.

University of Southampton

University Hospital Southampton NHS
NHS Foundation Trust

Part 2

11. What if relevant new information becomes available?

If the study is stopped for any other reason, we will tell you and arrange your continuing care.

12. What will happen if I don't want to carry on with the study?

You can withdraw from the study at any time. You don't have to give a reason. If you withdraw from the study you will not be able to rejoin the study at a future date.
Your study doctor may withdraw you from the study without your consent if:

- They decide that continuing could be harmful to you
- The study is cancelled
- You fail to follow the instructions of the study staff

If you withdraw from the study, data and samples already collected up to the point of your withdrawal may still be used, but they will be destroyed if you wish. In the unfortunate event of you suffering a loss of capacity to consent during the study, you will be withdrawn but data and samples already collected will be used confidentially in connection with the purposes for which consent was originally sought.

13. What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions - please raise your concerns in the first instance with Dr Giacomo Sgalla; his/her contact details are at the end of this information sheet.

Independent from the study team you can also contact the University of Southampton's Research Integrity & Governance Office on 0238059 5058 or rgoinfo@soton.ac.uk who acts as the sponsor representative.

If you remain unhappy and wish to complain formally, you can do this via the NHS complaints procedure. You can get help from Patient Advice and Liaison Service 023 8079 8498 (available 9am to 4.30 pm Monday to Friday, out of hours there is an answer phone, or from

PALS
C Level Centre Block

Email: PALS@suht.swest.nhs.uk
Tel: 023 8079 8498

Mailpoint 81
Southampton General Hospital
Tremona Road
Southampton
SO16 6YD

14. Harm

If something does go wrong and you are harmed during the research and this is due to someone's carelessness, then you may have grounds for legal action to compensation against University Hospital Southampton NHS Foundation Trust but you may have to pay your legal costs.

15. Will my taking part in this study be kept confidential?

Once on the trial you will be identified only by a unique code number and information about the code will be kept in a secure location and access limited to research study personnel. The data will be coded, stored and protected by Southampton University Hospitals NHS Trust for 15 years from the end of the study.
All information which is collected about you during the course of the research will be kept strictly confidential, and any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised.
Your identifiable data will be accessible by authorised persons such as the study doctor/chief investigator and other research staff (co-investigators, study nurses, respiratory technicians). They may also be looked at by authorised people to check that the study is being carried out correctly, such as regulatory authorities and R&D audit.
All will have a duty of confidentiality to you as a research participant and we will do our best to meet this duty.
It is possible that data from this study will be re-analysed at a later date and it may needed to perform further statistical tests on the data. The results of this study may be used for future medical research.

16. Involvement of the General Practitioner/Family doctor (GP)

If you agree, your GP will be contacted in the case of clinically significant findings during the course of the study.

17. What will happen to any samples I give?

Data resulting from the samples is treated with the same confidentiality as the data that

you discuss with the study doctor or nurse. The samples will be stored by Dr. Giacomo Sgalla in a secured facility at the University Hospital of Southampton. They will be used for the purposes of this study. The remaining samples, with your consent, will be stored for use in future undefined projects that will be approved by a Research Ethics Committee.
Should a third party outside the UK provide research analysis that is superior or not available to that at the Trust/University, the samples will be anonymised prior to their export.

If you decide to withdraw at any point during the study, no more samples will be collected from you. However, the samples already collected will be retained for analysis. If you request it, such samples will be also destroyed.

18. Sample and data sharing

In the future, your anonymised data may be shared with researchers outside the UK and outside the European Union. We would like to do this to ensure the best and widest use of the time you helped us conduct this project. You will not have any financial gain from your data being shared with other researchers.

If you do not wish for your data to be shared, please let us know and indicate this on your consent form.

19. What will happen to the results of the research study?

The results of the research will be published in a report and shared with the other research doctors on the study and shown to other doctors at meetings. You will not be identified in any report or publication. Your individual results will not be available to you. You will be informed through a letter or email summarising the results of the study on completion of th e study and eventual publication.

20. Who is organising and funding the research?

This research study is sponsored by the University of Southampton.
The study is being funded by the NIHR Respiratory Biomedical research Unit of the Southampton University Hospital NHS Trust.

21. Who has reviewed the study?

All research in the NHS is looked at by an independent group of people, called a Research

Ethics Committee, which has given favourable opinion for the conduction of the study.

22. Further information and contact details

If you have any questions about this research study, the procedures, risks or benefits or alternative treatments, please call the study staff using the information on the front page of this booklet. You can also discuss your participation with your regular doctor.

If you have any questions about your rights as a subject in this study you may want to contact your local Patients Advice and Liaison Service Office or www.pals.nhs.uk. You will be given a copy of this information sheet and a copy of your signed consent form to keep. You should keep this information in your possession for as long as you are in the study.

During the study, if there is an emergency please contact your doctor on the telephone number given. Should you have to visit another doctor tell him/her that you are taking part in this study so that he/she can contact your study doctor if necessary.

This is the end of Part 2.

Thank you for considering taking part in this research study.

A.2.2 Healthy volunteers

University of Southampton

University Hospital Southampton NHS
NHS Foundation Trust

Participant Information Sheet

Electronic recording of lung sounds as a clinical endpoint in Idiopathic Pulmonary Fibrosis: a longitudinal pilot study

Dear prospective participant,

Many thanks for taking the time to consider participating in our research study. Before you decide, we would like you to understand why this research is being done and what it would involve for you. One of our team will go through the information sheet with you and answer any questions that you might have. We would suggest this should take about ten to fifteen minutes.

Talk to others about the study if you wish.

Part 1 tells you the purpose of this study and what will happen if you take part.

Part 2 gives you more detailed information about the conduct of the study.

Please ask us if there is anything that is not clear.

Local Contacts:

Your Study Doctor:

Dr Giacomo Sgalla

Tel: 02381205399

Email: giacomo.sgalla@uhs.nhs.uk

Your research nurses:

Ms Louise Stanley

Ms Kate Howard

Tel: 02381204479

Emails: louise.stanley@uhs.nhs.uk ; kate.howard@uhs.nhs.uk

University of Southampton

University Hospital Southampton NHS
NHS Foundation Trust

Part 1

1. **What is the purpose of the study?**

Idiopathic Pulmonary Fibrosis (IPF) is a chronic disease causing a progressive scarring into the lungs. This scarring leads to the development of typical sounds produced by lungs, which are easily audible by doctors when listening to the chest of a patient through the stethoscope.

This is an educational study aiming to evaluate if lung sounds in patients with IPF may provide useful information about the course of the disease, i.e. how fast it will progress in the future. This would inform physicians about the management of IPF patients, as to making decisions about the optimal therapy to choose.

2. Why have I been invited?

The inclusion of a group of healthy volunteers in the study is justified by the need to obtain information about the normal variability of sounds produced by "healthy" lungs over time. Variations in lung sounds features will be also compared to the results of the spirometry (breathing test). This will help the analysis of pathological sounds in the patients affected by IPF.

3. Do I have to take part?

NO.

It is up to you to decide whether to join the study. We will describe the study and go through this information sheet. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any time, without giving a reason.

4. What will happen to me if I take part?

You will be in the study for 12 months. Visits will take place every 2 months at the Wellcome Trust research facilities - Southampton Centre for Biomedical Research at the University Hospital Southampton NHS Foundation Trust. This means you will be attending 7 visits (1 screening/baseline visit + 6 follow up visits) over this period.
The study will cover 2 years, a time necessary to collect all the data and analyse them.
If you result eligible at screening, you will enter the study.

Screening/Baseline (Visit 1) assessments

Screening visit is needed to determine whether you can take part to the study. If so, you will go through some tests and procedures described below.

Blood will be obtained from a vein in your arm to test the serum levels of specific substances (called biomarkers) that might be related to IPF and its progression. Other tests may be performed to look at specific genes in your DNA that are related to IPF and its progression.
The reason why these samples are being taken from you is to compare the results of the analyses described above with people afftected by IPF.
With your specific consent, any remaining samples may be used for future research after approval by a Research Ethics Committee.
The samples provided will be stored in secured facilities at the University Hospital of Southampton for the whole duration of the study.
The total amount of blood taken each time won't exceed 80 ml (about 8-10 tablespoons).

Follow-up visits (visit 2, 3, 4, 5, 6, 7) assessments

- At each of the following 6 visits, lung sound recording and spirometry will be repeated.
- Blood samples: repeated at visit 4 and 7, i.e. a total of 3 times throughout the study

5. Will I receive expenses and/or payments?

You will receive a payment of £ 70 per each visit, to compensate you for your time and inconvenience and for reimbursement of travel/parking expenses.
You will be provided with drinks and meals during the visits.

6. What will I have to do?

During the study you should:

- Attend all your study appointments. If you know that you will miss an appointment, contact the study staff to reschedule it as soon as possible.
- Follow the study staff's instructions with regards to study procedures.
- Ask questions as you think of them.
- Tell the study staff if you change your mind about staying in the study.

7. What are the possible disadvantages and risks of taking part?

- Discussion of this study and review and signing of this Informed Consent Form

  You will discuss study details with a member of the research staff (the study doctor or a research nurse). If you are happy to take part to the study you will be asked to sign a consent form. You will be given a signed copy of the consent form. Your demographic information, including your age, sex, and race/ethnicity will be recorded.

- Review of your medical history and any medications you are taking.

- Complete physical examination and recording of lung sounds.

  Lung sounds will be recorded from 6 sites on the back of your chest through an electronic stethoscope. This stethoscope is no different from a regular one, but it contains a microphone capable of recording sounds when it is applied on the chest.

- Measurement of your height and weight

- Measurement of your vital signs (breathing rate, heart rate, blood pressure, and body temperature)

- Resting Pulse Oximetry

  Pulse oximetry is a simple, non-invasive technique used to monitor the oxygen in your blood. It monitors the percentage of hemoglobin that has absorbed oxygen. For the test, a sensor is placed on a thin part of the body, usually a fingertip or earlobe.

- Spirometry/Pulmonary Function Testing

  Spirometry is a breathing test that requires that you breathe in and out through a tube you place in your mouth. This requires that you take deep breaths and blow out hard and fast. A spirometry test is used in the evaluation of how lungs function.

- Gas Exchange-Diffusion Capacity of the Lung for Carbon Monoxide Assessment (DL_{CO})

  A DL_{CO} test is used to determine how well the oxygen is exchanged between your lungs and your blood. The test is typically performed as part of spirometry testing.

- Blood Samples

There are possible risks, disadvantages and inconveniences with any research study. Consider these carefully before agreeing to participate in this study.

Risks associated with recording lung sounds
Recording lung sounds through an electronic stethoscope is a completely non-invasive and safe procedure.

Spirometry risks

Breathing tests will be done using machines that measure how much air you breathe in and out. A machine to measure your breathing (a spirometer) will be provided to your doctor for use during this study. Your doctor and the staff at your doctor's office will be specially trained on how to use and maintain the machines to measure your breathing. You will wear nose clips and be instructed to blow into a machine. You will be expected to blow as hard and as fast as you can for as long as you can. You may be asked to do this up to 3 or 4 times.
A second type of test called diffusion capacity (DL_{CO}) will also be performed with this machine. You will be instructed to exhale then inhale rapidly and hold your breath for approximately 8 seconds. After the 8 seconds have passed you will be asked to blow out again. The test might be repeated two or more times. Sometimes these procedures can cause you to cough, experience shortness of breath, or feel lightheaded, but there is no pain expected with these tests.

Insurance

Before taking part you should consider if this will affect any insurance you have, including travel insurance, and seek advice if necessary.

8. What are the possible benefits of taking part?

There are no direct benefits from taking part in this study, however the information gathered from your participation may help others suffering from IPF in the future. .

9. What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

10. Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. The details are included in Part 2.

This completes part 1.

If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.

University of Southampton

University Hospital Southampton NHS
NHS Foundation Trust

Part 2

11. What if relevant new information becomes available?

If the study is stopped for any other reason, we will tell you.

12. What will happen if I don't want to carry on with the study?

You can withdraw from the study at any time. You don't have to give a reason. If you withdraw from the study you will not be able to rejoin the study at a future date.
Your study doctor may withdraw you from the study without your consent if:

- They decide that continuing could be harmful to you
- The study is cancelled
- You fail to follow the instructions of the study staff

If you withdraw from the study, data already collected up to the point of your withdrawal may still be used, but they will be destroyed if you wish. In the unfortunate event of you suffering a loss of capacity to consent during the study, you will be withdrawn but data already collected will be used confidentially in connection with the purposes for which consent was originally sought.

13. What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions - please raise your concerns in the first instance with Dr Giacomo Sgalla; his/her contact details are at the end of this information sheet.

Independent from the study team you can also contact the University of Southampton's Research Integrity & Governance Office on 0238059 5058 or rgoinfo@soton.ac.uk who acts as the sponsor representative.

If you remain unhappy and wish to complain formally, you can do this via the NHS complaints procedure. You can get help from Patient Advice and Liaison Service 023 8079 8498 (available 9am to 4.30 pm Monday to Friday, out of hours there is an answer phone), or from

PALS
C Level Centre Block
Mailpoint 81

Email: PALS@suht.swest.nhs.uk
Tel: 023 8079 8498

Sgalla in a secured facility at the University Hospital of Southampton. They will be used for the purposes of this study. The remaining samples, with your consent, will be stored for use in future undefined projects that will be approved by a Research Ethics Committee.
Should a third party outside the UK provide research analysis that is superior or not available to that at the Trust/University, the samples will be anonymised prior to their export.

18. Sample and data sharing

In the future, your anonymised data may be shared with researchers outside the UK and outside the European Union. We would like to do this to ensure the best and widest use of the time you helped us conduct this project. You will not have any financial gain from your data being shared with other researchers.

If you do not wish for your data to be shared, please let us know and indicate this on your consent form.

19. What will happen to the results of the research study?

The results of the research will be published in a report and shared with the other research doctors on the study and shown to other doctors at meetings. You will not be identified in any report or publication. Your individual results will not be available to you. You will be informed through a letter or email summarising the results of the study on completion of th e study and eventual publication.

20. Who is organising and funding the research?

This research study is sponsored by the University of Southampton.
The study is being funded by the NIHR Respiratory Biomedical research Unit of the Southampton University Hospital NHS Trust.

21. Who has reviewed the study?

All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee, which has given favourable opinion for the conduction of the study.

22. Further information and contact details

If you have any questions about this research study, the procedures, risks or benefits,

Sgalla in a secured facility at the University Hospital of Southampton. They will be used for the purposes of this study. The remaining samples, with your consent, will be stored for use in future undefined projects that will be approved by a Research Ethics Committee. Should a third party outside the UK provide research analysis that is superior or not available to that at the Trust/University, the samples will be anonymised prior to their export.

18. Sample and data sharing

In the future, your anonymised data may be shared with researchers outside the UK and outside the European Union. We would like to do this to ensure the best and widest use of the time you helped us conduct this project. You will not have any financial gain from your data being shared with other researchers.

If you do not wish for your data to be shared, please let us know and indicate this on your consent form.

19. What will happen to the results of the research study?

The results of the research will be published in a report and shared with the other research doctors on the study and shown to other doctors at meetings. You will not be identified in any report or publication. Your individual results will not be available to you. You will be informed through a letter or email summarising the results of the study on completion of th e study and eventual publication.

20. Who is organising and funding the research?

This research study is sponsored by the University of Southampton.
The study is being funded by the NIHR Respiratory Biomedical research Unit of the Southampton University Hospital NHS Trust.

21. Who has reviewed the study?

All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee, which has given favourable opinion for the conduction of the study.

22. Further information and contact details

If you have any questions about this research study, the procedures, risks or benefits,

Southampton General Hospital
Tremona Road
Southampton
SO16 6YD

14. Harm

If something does go wrong and you are harmed during the research and this is due to someone's carelessness, then you may have grounds for a legal action and compensation against University Hospital Southampton NHS Foundation Trust but you may have to pay your legal costs.

15. Will my taking part in this study be kept confidential?

Once on the trial you will be identified only by a unique code number and information about the code will be kept in a secure location and access limited to research study personnel. The data will be coded, stored and protected by Southampton University Hospitals NHS Trust for 15 years from the end of the study.
All information which is collected about you during the course of the research will be kept strictly confidential, and any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised.
Your identifiable data will be accessible by authorised persons such as the study doctor/chief investigator and other research staff (academic supervisor, study nurses, respiratory technicians). They may also be looked at by authorised people to check that the study is being carried out correctly, such as regulatory authorities and R&D audit.
All will have a duty of confidentiality to you as a research participant and we will do our best to meet this duty.
It is possible that data from this study will be re-analysed at a later date and it may needed to perform further statistical tests on the data. The results of this study may be used for future medical research.

16. Involvement of the General Practitioner/Family doctor (GP)

If you agree, your GP will be contacted in the case of clinically significant findings during the course of the study.

17. What will happen to any samples I give?

Data resulting from the samples is treated with the same confidentiality as the data that you discuss with the study doctor or nurse. The samples will be stored by Dr. Giacomo

please call the study staff using the information on the front page of this booklet. You can also discuss your participation with your regular doctor.

If you have any questions about your rights as a subject in this study you may want to contact your local Patients Advice and Liaison Service Office or www.pals.nhs.uk. You will be given a copy of this information sheet and a copy of your signed consent form to keep. You should keep this information in your possession for as long as you are in the study.

During the study, if there is an emergency please contact your doctor on the telephone number given. Should you have to visit another doctor tell him/her that you are taking part in this study so that he/she can contact your study doctor if necessary.

This is the end of Part 2.

Thank you for considering taking part in this research study.

A.3 Informed consent

Respiratory Biomedical Research Unit,
Southampton Centre for Biomedical Research,
MP218
Southampton General Hospital
Tremona Road
SO16 6YD

***Investigator**: Dr. Giacomo Sgalla*
Contact for queries
If you have any questions about this study, you can contact:
Louise Stanley / Kate Howard on 0231204479
Dr Giacomo Sgalla on 02381205399

INFORMED CONSENT FORM

Title of study:	Role of lung sounds in IPF
Name of Chief Investigator:	Dr. Giacomo Sgalla
Centre/Site:	University Hospital of Southampton
Study number:	
REC approval number:	
Participant ID:	

Thank you for reading the information about our research project. If you would like to take part, please read and sign this form.

PART A: Consent for the current study
PLEASE INITIAL THE BOXES IF YOU AGREE WITH EACH SECTION:

1. I have read and understand the information sheet version 4 dated 25/01/2016 for the above study and have been given a copy to keep. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐

3. Should a third party outside the UK provide research analysis that is superior or not available to that at the Trust/University, I give consent for my sample of blood to be exported outside the UK. I understand that the samples will be linked or linked anonymised prior to their export. ☐

4. I agree to give a sample of (blood/ tissue/ other as appropriate) for research in this study. I understand how the sample will be collected, that giving a sample for this research is voluntary and that I am free to withdraw my approval for use of the sample at any time. ☐

5. I understand that University of Southampton is sponsor of the study, and that the SCBR Respiratory BRU is funding the study. I understand that researchers (Doctors, Nurses and Scientists) are working on the study under the instruction of Dr Giacomo Sgalla. ☐

6. I understand that relevant sections of my medical notes and data collected during the study, may be looked at by individuals from University Hospital of Southampton, staff working in the clinical research facility and from regulatory authorities, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. I understand that the information will be kept confidential. ☐

7. I agree that my GP may be informed of my participation if any of the results of tests done as part of the research are important for my health. ☐

Respiratory Biomedical Research Unit,
Southampton Centre for Biomedical Research,
MP218
Southampton General Hospital
Tremona Road
SO16 6YD

***Investigator**: Dr. Giacomo Sgalla*
Contact for queries
If you have any questions about this study, you can contact:
Louise Stanley / Kate Howard on 0231204479
Dr Giacomo Sgalla on 02381205399

8.	I understand that I will not benefit financially if this research leads to the development of a new treatment or test.	☐
9.	I have informed the researchers of my participation in any other research study	☐
10.	I know how to contact the research team if I need to.	☐
11.	I agree to participate in this study.	☐
12.	I understand that information that could lead to my identification will not be disclosed in any reports on the project. No identifiable personal data will be published	☐
13.	I am interested in future research studies and give permission to be contacted. I understand that there will be a separate information sheet and separate consent form for my details to be kept on a research participant database.	☐

Participant: name surname	Date	Signature

Researcher taking consent: name surname	Date	Signature

Original for Investigator Site File , 1 copy for participant, 1 copy for medical record/hospital notes

Respiratory Biomedical Research Unit,
Southampton Centre for Biomedical Research,
MP218
Southampton General Hospital
Tremona Road
SO16 6YD

***Investigator**: Dr. Giacomo Sgalla*
Contact for queries
If you have any questions about this study, you can contact:
Louise Stanley / Kate Howard on 0231204479
Dr Giacomo Sgalla on 02381205399

PART B : Consent for future studies
Linked or linked anonymised samples
PLEASE INITIAL THE BOXES IF YOU AGREE WITH EACH SECTION:

1. I have read the information sheet version 4 dated 25/01/2016 for the above study and have been given a copy to keep. I give permission for my samples and the information gathered about me to be stored by Dr. Giacomo Sgalla at the University of Southampton for possible use in future undefined projects. I understand that some of these projects may be carried out by other researchers, including researchers working for commercial companies. ☐

2. I understand that future studies will be reviewed and approved by a Research Ethics Committee prior to my sample being used, and that I can alter these decisions at any time by letting the research team know. ☐

3. I give permission for the sample to be used only for research about IPF. ☐

4. I give permission for the sample to be used for other undefined research, the precise nature of which will depend upon future scientific advances. ☐

5. I am aware that I can request at any time in the future that my samples are destroyed without giving any reason as to my decision. ☐

6. I give permission for the samples to be exported outside the UK to a third party research collaborator. I understand that these samples will be linked or linked anonymised. ☐

7. I want / do not want (delete as applicable) to be told the results of any future test which may have health implications for me. ☐

8. I give permission for sections of my medical notes to be looked at by responsible individuals where it is relevant to such future study. I expect that my medical notes will be treated confidentially at all times. ☐

Respiratory Biomedical Research Unit,
Southampton Centre for Biomedical Research,
MP218
Southampton General Hospital
Tremona Road
SO16 6YD

***Investigator**: Dr. Giacomo Sgalla*
Contact for queries
If you have any questions about this study, you can contact:
Louise Stanley / Kate Howard on 0231204479
Dr Giacomo Sgalla on 02381205399

PART C: Consent for future studies
Unlinked anonymised samples
PLEASE INITIAL THE BOXES IF YOU AGREE WITH EACH SECTION:

1. I have read the information sheet version 4 dated 26/01/2016 for the above study and have been given a copy to keep. ☐

2. I give permission for my sample to be stored for possible use in future undefined projects, provided that the projects have been reviewed and approved by a Research Ethics Committee. I understand that my name or identifying details will not be linked to this sample. ☐

3. I give permission for the sample/tissue/images to be exported outside the UK to a third party research collaborator. I understand that these samples will be anonymised. ☐

Participant first name surname	Date	Signature

Person taking consent first name surname	Date	Signature

Original for Investigator Site File/Researcher, 1 copy for participant, 1 copy for medical record/hospital notes

Respiratory Biomedical Research Unit,
Southampton Centre for Biomedical Research,
MP218
Southampton General Hospital
Tremona Road
SO16 6YD

***Investigator**: Dr. Giacomo Sgalla*
Contact for queries
If you have any questions about this study, you can contact:
Louise Stanley / Kate Howard on 0231204479

Dr Giacomo Sgalla on 02381205399

Participant first name surname	Date	Signature

Person taking consent first name surname	Date	Signature

Original for Investigator Site File, 1 copy for participant, 1 copy for medical record/hospital notes

Appendix B Questionnaires and scales (cohort study)

B.1 Saint George's Respiratory Questionnaire (SGRQ)

ST. GEORGE'S RESPIRATORY QUESTIONNAIRE
ORIGINAL ENGLISH VERSION

ST. GEORGE'S RESPIRATORY QUESTIONNAIRE (SGRQ)

This questionnaire is designed to help us learn much more about how your breathing is troubling you and how it affects your life. We are using it to find out which aspects of your illness cause you most problems, rather than what the doctors and nurses think your problems are.

Please read the instructions carefully and ask if you do not understand anything. Do not spend too long deciding about your answers.

Before completing the rest of the questionnaire:

Please tick in one box to show how you describe your current health:

Very good	Good	Fair	Poor	Very poor
☐	☐	☐	☐	☐

P.W. Jones, PhD FRCP
Professor of Respiratory Medicine,
St. George's University of London,
Jenner Wing,
Cranmer Terrace,
London SW17 ORE, UK.

Tel. +44 (0) 20 8725 5371
Fax +44 (0) 20 8725 5955

continued...

St. George's Respiratory Questionnaire
PART 1

Questions about how much chest trouble you have had over the past 3 months.

Please tick (✓) *one box for each question*:

	most days a week	several days a week	a few days a month	only with chest infections	not at all
1. Over the past 3 months, I have coughed:	☐	☐	☐	☐	☐
2. Over the past 3 months, I have brought up phlegm (sputum):	☐	☐	☐	☐	☐
3. Over the past 3 months, I have had shortness of breath:	☐	☐	☐	☐	☐
4. Over the past 3 months, I have had attacks of wheezing:	☐	☐	☐	☐	☐

5. During the past 3 months how many severe or very unpleasant attacks of chest trouble have you had?

Please tick (✓) *one*:

- more than 3 attacks ☐
- 3 attacks ☐
- 2 attacks ☐
- 1 attack ☐
- no attacks ☐

6. How long did the worst attack of chest trouble last?
(Go to question 7 if you had no severe attacks)

Please tick (✓) *one*:

- a week or more ☐
- 3 or more days ☐
- 1 or 2 days ☐
- less than a day ☐

7. Over the past 3 months, in an average week, how many good days (with little chest trouble) have you had?

Please tick (✓) *one*:

- No good days ☐
- 1 or 2 good days ☐
- 3 or 4 good days ☐
- nearly every day is good ☐
- every day is good ☐

8. If you have a wheeze, is it worse in the morning?

Please tick (✓) *one*:

- No ☐
- Yes ☐

St. George's Respiratory Questionnaire
PART 2

Section 1

How would you describe your chest condition?

Please tick (✓) *one*:

- The most important problem I have ☐
- Causes me quite a lot of problems ☐
- Causes me a few problems ☐
- Causes no problem ☐

If you have ever had paid employment.

Please tick (✓) *one*:

- My chest trouble made me stop work altogether ☐
- My chest trouble interferes with my work or made me change my work ☐
- My chest trouble does not affect my work ☐

Section 2

Questions about what activities usually make you feel breathless <u>these days</u>.

Please tick (✓) in ***each box*** that applies to you ***these days***:

	True	False
Sitting or lying still	☐	☐
Getting washed or dressed	☐	☐
Walking around the home	☐	☐
Walking outside on the level	☐	☐
Walking up a flight of stairs	☐	☐
Walking up hills	☐	☐
Playing sports or games	☐	☐

continued…

St. George's Respiratory Questionnaire
PART 2

Section 3

Some more questions about your cough and breathlessness <u>these days</u>.

	Please tick (✓) in ***each box*** that applies to you ***these days***: True	False
My cough hurts	☐	☐
My cough makes me tired	☐	☐
I am breathless when I talk	☐	☐
I am breathless when I bend over	☐	☐
My cough or breathing disturbs my sleep	☐	☐
I get exhausted easily	☐	☐

Section 4

Questions about other effects that your chest trouble may have on you <u>these days</u>.

	Please tick (✓) in ***each box*** that applies to you ***these days***: True	False
My cough or breathing is embarrassing in public	☐	☐
My chest trouble is a nuisance to my family, friends or neighbours	☐	☐
I get afraid or panic when I cannot get my breath	☐	☐
I feel that I am not in control of my chest problem	☐	☐
I do not expect my chest to get any better	☐	☐
I have become frail or an invalid because of my chest	☐	☐
Exercise is not safe for me	☐	☐
Everything seems too much of an effort	☐	☐

Section 5

Questions about your medication, if you are receiving no medication go straight to section 6.

	Please tick (✓) in ***each box*** that applies to you ***these days***: True	False
My medication does not help me very much	☐	☐
I get embarrassed using my medication in public	☐	☐
I have unpleasant side effects from my medication	☐	☐
My medication interferes with my life a lot	☐	☐

continued…

f:\institut\cultadap\project\gsk1881\question\final versions\sgrqoriq.doc 14/03/03

St. George's Respiratory Questionnaire
PART 2

Section 6

These are questions about how your activities might be affected by your breathing.

	Please tick (✔) in ***each box*** that applies to you ***because of your breathing***: True	False
I take a long time to get washed or dressed	☐	☐
I cannot take a bath or shower, or I take a long time	☐	☐
I walk slower than other people, or I stop for rests	☐	☐
Jobs such as housework take a long time, or I have to stop for rests	☐	☐
If I walk up one flight of stairs, I have to go slowly or stop	☐	☐
If I hurry or walk fast, I have to stop or slow down	☐	☐
My breathing makes it difficult to do things such as walk up hills, carrying things up stairs, light gardening such as weeding, dance, play bowls or play golf	☐	☐
My breathing makes it difficult to do things such as carry heavy loads, dig the garden or shovel snow, jog or walk at 5 miles per hour, play tennis or swim	☐	☐
My breathing makes it difficult to do things such as very heavy manual work, run, cycle, swim fast or play competitive sports	☐	☐

Section 7

We would like to know how your chest <u>usually</u> affects your daily life.

	Please tick (✔) in ***each box*** that applies to you ***because of your chest trouble***: True	False
I cannot play sports or games	☐	☐
I cannot go out for entertainment or recreation	☐	☐
I cannot go out of the house to do the shopping	☐	☐
I cannot do housework	☐	☐
I cannot move far from my bed or chair	☐	☐

continued…

St. George's Respiratory Questionnaire

Here is a list of other activities that your chest trouble may prevent you doing. (You do not have to tick these, they are just to remind you of ways in which your breathlessness may affect you):

Going for walks or walking the dog
Doing things at home or in the garden
Sexual intercourse
Going out to church, pub, club or place of entertainment
Going out in bad weather or into smoky rooms
Visiting family or friends or playing with children

Please write in any other important activities that your chest trouble may stop you doing:

. . . . . .
. . . . . .
. . . . . .
. . . . . .

Now would you tick in the box (one only) which you think best describes how your chest affects you:

It does not stop me doing anything I would like to do ☐
It stops me doing one or two things I would like to do ☐
It stops me doing most of the things I would like to do ☐
It stops me doing everything I would like to do ☐

Thank you for filling in this questionnaire. Before you finish would you please check to see that you have answered all the questions.

B.2 University of California San Diego Shortness of Breath Questionnaire (UCSD-SOB)

436

Saint Agnes Medical Center
Fresno, California

SHORTNESS OF BREATH QUESTIONNAIRE

Page 1 of 2

106500-436 (5/01)

NAME_____

MR# _____

DATE _____

UCSD MEDICAL CENTER
PULMONARY REHABILITATION PROGRAM
SHORTNESS-OF-BREATH QUESTIONNAIRE

Instructions: For each activity listed below, please rate your breathlessness on a scale between zero and five where 0 is not at all breathless and 5 is maximally breathless or too breathless to do the activity. If the activity is one which you do not perform, please give your best estimate of breathlessness. Your responses should be for an "average" day during the past week. Please respond to all items. Read the two examples below then turn the page to begin the questionnaire.

0 Not at all

1

2

3

4 Severely

5 Maximally or unable to do because of breathlessness

Example 1:

How short of breath do you get while:

1. Brushing teeth 0 1 2 3 4 5

Harry has felt moderately short of breath during the past week while brushing his teeth and so circles a three for this activity.

Example 2:

How short of breath do you get while:

1. Mowing the lawn 0 1 2 3 4 5

Anne has never mowed the lawn before but estimates that she would have been too breathless to do this activity during the past week. She circles a five for this activity.

436

Saint Agnes Medical Center
Fresno, California

SHORTNESS OF BREATH QUESTIONNAIRE

Page 2 of 2

106500-436 (5/01)

How short of breath do you get:

1.	At rest	0	1	2	3	4	5
2.	Walking on a level at your own pace	0	1	2	3	4	5
3.	Walking on a level with others your age	0	1	2	3	4	5
4.	Walking up a hill	0	1	2	3	4	5
5.	Walking up stairs	0	1	2	3	4	5
6.	While eating	0	1	2	3	4	5
7.	Standing up from a chair	0	1	2	3	4	5
8.	Brushing teeth	0	1	2	3	4	5
9.	Shaving and/or brushing hair	0	1	2	3	4	5
10.	Showering / bathing	0	1	2	3	4	5
11.	Dressing	0	1	2	3	4	5
12.	Picking up and straightening	0	1	2	3	4	5
13.	Doing dishes	0	1	2	3	4	5
14.	Sweeping / vacuuming	0	1	2	3	4	5
15.	Making a bed	0	1	2	3	4	5
16.	Shopping	0	1	2	3	4	5
17.	Doing laundry	0	1	2	3	4	5
18.	Washing car	0	1	2	3	4	5
19.	Mowing lawn	0	1	2	3	4	5
20.	Watering lawn	0	1	2	3	4	5
21.	Sexual activities	0	1	2	3	4	5

How much do these limit you in your daily life?

22.	Shortness of breath	0	1	2	3	4	5
23.	Fear of "hurting myself" by overexerting	0	1	2	3	4	5
24.	Fear of shortness of breath	0	1	2	3	4	5

B.3 Borg scale – dyspnea and fatigue index

Post-walk Borg· Dyspnea Index

Time: _____:_____ *(0:00-23:59)*

☐	0.0	Nothing at all
☐	0.5	Very, very slight (just noticeable}
☐	1	Very slight
☐	2	Slight
☐	3	Moderate
☐	4	Somewhat severe
☐	5	Severe
☐	6	
☐	7	Very severe
☐	8	
☐	9	
☐	10	Very, very severe (Maximal)
☐		Not Done

Post-walk Borg· Fatigue Index

☐	0.0	Nothing at all
☐	0.5	Very, very slight (just noticeable}
☐	1	Very slight
☐	2	Slight
☐	3	Moderate
☐	4	Somewhat severe
☐	5	Severe
☐	6	
☐	7	Very severe
☐	8	
☐	9	
☐	10	Very, very severe (Maximal)
☐		Not Done

Pre-walk Borg· Dyspnea Index

Time: _____:_____ *(0:00-23:59)*

- ☐ 0.0 Nothing at all
- ☐ 0.5 Very, very slight (just noticeable)
- ☐ 1 Very slight
- ☐ 2 Slight
- ☐ 3 Moderate
- ☐ 4 Somewhat severe
- ☐ 5 Severe
- ☐ 6
- ☐ 7 Very severe
- ☐ 8
- ☐ 9
- ☐ 10 Very, very severe (Maximal)
- ☐ Not Done

Pre-walk Borg· Fatigue Index

- ☐ 0.0 Nothing at all
- ☐ 0.5 Very, very slight (just noticeable)
- ☐ 1 Very slight
- ☐ 2 Slight
- ☐ 3 Moderate
- ☐ 4 Somewhat severe
- ☐ 5 Severe
- ☐ 6
- ☐ 7 Very severe
- ☐ 8
- ☐ 9
- ☐ 10 Very, very severe (Maximal)
- ☐ Not Done

Appendix C Example of chart for subjective assessment of lung sounds (case-control study)

Sound file	Do you hear crackles ? (Y/N)	If YES, do you think crackles are fibrotic in nature? (Y/N)
16		
61		
84		
119		
187		
240		
299		
358		
397		
453		
476		
505		
561		
28		
68		
91		
155		
216		
247		
306		
375		
409		
460		
489		
518		
573		
32		
72		
95		
159		
220		
281		
316		
379		
…		

Appendix D Description of acoustic features (cohort study)

N	EMD	Feature name	Feature description	Notes
1	No	total_energy	total energy of the original signal	
2	Yes	IMF01_energy	energy of each IMF	
3	Yes	IMF02_energy		
4	Yes	IMF03_energy		
5	Yes	IMF04_energy		
6	Yes	IMF05_energy		
7	Yes	IMF06_energy		
8	Yes	IMF07_energy		
9	Yes	IMF08_energy		
10	Yes	IMF09_energy		
11	Yes	IMF10_energy		
12	Yes	IMF01_norm_energy	energy of each IMF normalised by the total energy of the original signal	
13	Yes	IMF02_norm_energy		
14	Yes	IMF03_norm_energy		
15	Yes	IMF04_norm_energy		
16	Yes	IMF05_norm_energy		
17	Yes	IMF06_norm_energy		
18	Yes	IMF07_norm_energy		
19	Yes	IMF08_norm_energy		
20	Yes	IMF09_norm_energy		
21	Yes	IMF10_norm_energy		
22	Yes	Ec3	energy of the crackle component (N=3)	
23	Yes	Er3	energy of the respiratory component (N=3)	
24	Yes	Ec3Er3	ratio of the energies of the crackle and respiratory components (N=3)	
25	Yes	Ec3n	energy of the crackle component (N=3) normalised by the total energy of the original signal	
26	Yes	Er3n	energy of the respiratory component (N=3) normalised by the total energy of the original signal	
27	Yes	Ec3nEr3n	ratio of the energies of the crackle and respiratory components (N=3) normalised by the total energy of the original signal	
28	Yes	Ec4	energy of the crackle component (N=4)	
29	Yes	Er4	energy of the respiratory component (N=4)	
30	Yes	Ec4Er4	ratio of the energies of the crackle and respiratory components (N=4)	
31	Yes	Ec4n	energy of the crackle component (N=4) normalised by the total energy of the original signal	

32	Yes	Er4n	energy of the respiratory component (N=4) normalised by the total energy of the original signal	
33	Yes	Ec4nEr4n	ratio of the energies of the crackle and respiratory components (N=4) normalised by the total energy of the original signal	
34	Yes	IMF01_avP150_450Hz	average power of each IMF in the [150-450] Hz frequency band	
35	Yes	IMF02_avP150_450Hz		
36	Yes	IMF03_avP150_450Hz		
37	Yes	IMF04_avP150_450Hz		
38	Yes	IMF05_avP150_450Hz		
39	Yes	IMF06_avP150_450Hz		
40	Yes	IMF07_avP150_450Hz		
41	Yes	IMF08_avP150_450Hz		
42	Yes	IMF09_avP150_450Hz		
43	Yes	IMF10_avP150_450Hz		
44	Yes	avPc3_150_450Hz	average power of the crackle component (N=3) in the [150-450] Hz frequency band	
45	Yes	avPr3_150_450Hz	average power of the respiratory component (N=3) in the [150-450] Hz frequency band	
46	Yes	avPc3avPr3_150_450Hz	ratio of the average powers of the crackle and respiratory components (N=3) in the [150-450] Hz frequency band	
47	Yes	avPc4_150_450Hz	average power of the crackle component (N=4) in the [150-450] Hz frequency band	
48	Yes	avPr4_150_450Hz	average power of the respiratory component (N=4) in the [150-450] Hz frequency band	
49	Yes	avPc4avPr4_150_450Hz	ratio of the average powers of the crackle and respiratory components (N=4) in the [150-450] Hz frequency band	
50	Yes	IMF01_MH_0_1500Hz	marginal HHT spectrum (MH) of the IMF1 in the specified frequency bands	
51	Yes	IMF01_MH_75_200Hz		
52	Yes	IMF01_MH_200_500Hz		
53	Yes	IMF01_MH_500_1000Hz		
54	Yes	IMF01_MH_1000_1500Hz		
55	Yes	IMF01_EW_75_200Hz	energy weight (EW) of the IMF1 in the specified frequency bands	
56	Yes	IMF01_EW_200_500Hz		
57	Yes	IMF01_EW_500_1000Hz		
58	Yes	IMF01_EW_1000_1500Hz		
59	Yes	IMF02_MH_0_1500Hz	marginal HHT spectrum (MH) of the IMF2 in the specified frequency bands	
60	Yes	IMF02_MH_75_200Hz		
61	Yes	IMF02_MH_200_500Hz		
62	Yes	IMF02_MH_500_1000Hz		
63	Yes	IMF02_MH_1000_1500Hz		
64	Yes	IMF02_EW_75_200Hz	energy weight (EW) of the IMF2 in the specified frequency bands	
65	Yes	IMF02_EW_200_500Hz		
66	Yes	IMF02_EW_500_1000Hz		
67	Yes	IMF02_EW_1000_1500Hz		
68	Yes	IMF03_MH_0_1500Hz		

69	Yes	IMF03_MH_75_200Hz	marginal HHT spectrum (MH) of the IMF3 in the specified frequency bands	
70	Yes	IMF03_MH_200_500Hz		
71	Yes	IMF03_MH_500_1000Hz		
72	Yes	IMF03_MH_1000_1500Hz		
73	Yes	IMF03_EW_75_200Hz	energy weight (EW) of the IMF3 in the specified frequency bands	
74	Yes	IMF03_EW_200_500Hz		
75	Yes	IMF03_EW_500_1000Hz		
76	Yes	IMF03_EW_1000_1500Hz		
77	Yes	IMF04_MH_0_1500Hz	marginal HHT spectrum (MH) of the IMF4 in the specified frequency bands	
78	Yes	IMF04_MH_75_200Hz		
79	Yes	IMF04_MH_200_500Hz		
80	Yes	IMF04_MH_500_1000Hz		
81	Yes	IMF04_MH_1000_1500Hz		
82	Yes	IMF04_EW_75_200Hz	energy weight (EW) of the IMF4 in the specified frequency bands	
83	Yes	IMF04_EW_200_500Hz		
84	Yes	IMF04_EW_500_1000Hz		
85	Yes	IMF04_EW_1000_1500Hz		
86	Yes	IMF05_MH_0_1500Hz	marginal HHT spectrum (MH) of the IMF5 in the specified frequency bands	
87	Yes	IMF05_MH_75_200Hz		
88	Yes	IMF05_MH_200_500Hz		
89	Yes	IMF05_MH_500_1000Hz		
90	Yes	IMF05_MH_1000_1500Hz		
91	Yes	IMF05_EW_75_200Hz	energy weight (EW) of the IMF5 in the specified frequency bands	
92	Yes	IMF05_EW_200_500Hz		
93	Yes	IMF05_EW_500_1000Hz		
94	Yes	IMF05_EW_1000_1500Hz		
95	Yes	IMF06_MH_0_1500Hz	marginal HHT spectrum (MH) of the IMF6 in the specified frequency bands	
96	Yes	IMF06_MH_75_200Hz		
97	Yes	IMF06_MH_200_500Hz		
98	Yes	IMF06_MH_500_1000Hz		
99	Yes	IMF06_MH_1000_1500Hz		
100	Yes	IMF06_EW_75_200Hz	energy weight (EW) of the IMF6 in the specified frequency bands	
101	Yes	IMF06_EW_200_500Hz		
102	Yes	IMF06_EW_500_1000Hz		
103	Yes	IMF06_EW_1000_1500Hz		
104	Yes	IMF07_MH_0_1500Hz	marginal HHT spectrum (MH) of the IMF7 in the specified frequency bands	
105	Yes	IMF07_MH_75_200Hz		
106	Yes	IMF07_MH_200_500Hz		
107	Yes	IMF07_MH_500_1000Hz		
108	Yes	IMF07_MH_1000_1500Hz		
109	Yes	IMF07_EW_75_200Hz	energy weight (EW) of the IMF7 in the specified frequency bands	
110	Yes	IMF07_EW_200_500Hz		
111	Yes	IMF07_EW_500_1000Hz		
112	Yes	IMF07_EW_1000_1500Hz		
113	Yes	IMF08_MH_0_1500Hz	marginal HHT spectrum (MH) of the IMF8 in the specified frequency bands	
114	Yes	IMF08_MH_75_200Hz		

115	Yes	IMF08_MH_200_500Hz		
116	Yes	IMF08_MH_500_1000Hz		
117	Yes	IMF08_MH_1000_1500Hz		
118	Yes	IMF08_EW_75_200Hz	energy weight (EW) of the IMF8 in the specified frequency bands	
119	Yes	IMF08_EW_200_500Hz		
120	Yes	IMF08_EW_500_1000Hz		
121	Yes	IMF08_EW_1000_1500Hz		
122	Yes	IMF09_MH_0_1500Hz	marginal HHT spectrum (MH) of the IMF9 in the specified frequency bands	
123	Yes	IMF09_MH_75_200Hz		
124	Yes	IMF09_MH_200_500Hz		
125	Yes	IMF09_MH_500_1000Hz		
126	Yes	IMF09_MH_1000_1500Hz		
127	Yes	IMF09_EW_75_200Hz	energy weight (EW) of the IMF9 in the specified frequency bands	
128	Yes	IMF09_EW_200_500Hz		
129	Yes	IMF09_EW_500_1000Hz		
130	Yes	IMF09_EW_1000_1500Hz		
131	Yes	IMF10_MH_0_1500Hz	marginal HHT spectrum (MH) of the IMF10 in the specified frequency bands	
132	Yes	IMF10_MH_75_200Hz		
133	Yes	IMF10_MH_200_500Hz		
134	Yes	IMF10_MH_500_1000Hz		
135	Yes	IMF10_MH_1000_1500Hz		
136	Yes	IMF10_EW_75_200Hz	energy weight (EW) of the IMF10 in the specified frequency bands	
137	Yes	IMF10_EW_200_500Hz		
138	Yes	IMF10_EW_500_1000Hz		
139	Yes	IMF10_EW_1000_1500Hz		
140	Yes	C3_MS0_0_1500Hz	marginal HHT spectrum (MH) of the crackle component (N=3) in the specified frequency bands	
141	Yes	C3_MSc_75_200Hz		
142	Yes	C3_MSc_200_500Hz		
143	Yes	C3_MSc_500_1000Hz		
144	Yes	C3_MSc_1000_1500Hz		
145	Yes	C3_EW_75_200Hz	energy weight (EW) of the crackle component (N=3) in the specified frequency bands	
146	Yes	C3_EW_200_500Hz		
147	Yes	C3_EW_500_1000Hz		
148	Yes	C3_EW_1000_1500Hz		
149	Yes	R3_MS0_0_1500Hz	marginal HHT spectrum (MH) of the respiratory component (N=3) in the specified frequency bands	
150	Yes	R3_MSc_75_200Hz		
151	Yes	R3_MSc_200_500Hz		
152	Yes	R3_MSc_500_1000Hz		
153	Yes	R3_MSc_1000_1500Hz		
154	Yes	R3_EW_75_200Hz	energy weight (EW) of the respiratory component (N=3) in the specified frequency bands	
155	Yes	R3_EW_200_500Hz		
156	Yes	R3_EW_500_1000Hz		
157	Yes	R3_EW_1000_1500Hz		
158	Yes	C4_MS0_0_1500Hz	marginal HHT spectrum (MH) of the crackle component (N=4) in the specified frequency bands	
159	Yes	C4_MSc_75_200Hz		
160	Yes	C4_MSc_200_500Hz		

161	Yes	C4_MSc_500_1000Hz		
162	Yes	C4_MSc_1000_1500Hz		
163	Yes	C4_EW_75_200Hz	energy weight (EW) of the crackle component (N=4) in the specified frequency bands	
164	Yes	C4_EW_200_500Hz		
165	Yes	C4_EW_500_1000Hz		
166	Yes	C4_EW_1000_1500Hz		
167	Yes	R4_MS0_0_1500Hz	marginal HHT spectrum (MH) of the respiratory component (N=4) in the specified frequency bands	
168	Yes	R4_MSc_75_200Hz		
169	Yes	R4_MSc_200_500Hz		
170	Yes	R4_MSc_500_1000Hz		
171	Yes	R4_MSc_1000_1500Hz		
172	Yes	R4_EW_75_200Hz	energy weight (EW) of the respiratory component (N=4) in the specified frequency bands	
173	Yes	R4_EW_200_500Hz		
174	Yes	R4_EW_500_1000Hz		
175	Yes	R4_EW_1000_1500Hz		
176	No	sig_rms	RMS of the original signal	
177	No	sig_lowenergy	Low Energy of the original signal	
178	No	sig_lowenergyASR	Average Silence Ratio of the original signal	
179	No	sig_zerocross	Zero-cross of the original signal	
180	No	sig_rolloff85	Roll-off (threshold=85%) of the original signal	
181	No	sig_rolloff95	Roll-off (threshold=95%) of the original signal	
182	No	sig_brightness_1000Hz	Brightness (cutt-off frequency=1000 Hz) of the original signal	
183	No	sig_brightness_1500Hz	Brightness (cutt-off frequency=1500 Hz) of the original signal	
184	No	sig_centroid	Centroid of the original signal	
185	No	sig_spread	Spread of the original signal	
186	No	sig_skewness	Skewness of the original signal	
187	No	sig_kurtosis	Kurtosis of the original signal	
188	No	sig_flatness	Flatness of the original signal	
189	No	sig_entropy	Entropy of the original signal	
190	No	sig_regularity	Regularity of the original signal	
191	No	sig_mfcc01	13 MFCC of the original signal	
192	No	sig_mfcc02		
193	No	sig_mfcc03		
194	No	sig_mfcc04		
195	No	sig_mfcc05		
196	No	sig_mfcc06		
197	No	sig_mfcc07		
198	No	sig_mfcc08		
199	No	sig_mfcc09		
200	No	sig_mfcc10		
201	No	sig_mfcc11		
202	No	sig_mfcc12		
203	No	sig_mfcc13		
204	No	sig_mean_meanframes		

205	No	sig_std_meanframes	mean of the frames of the original signal (mean, std, median)	
206	No	sig_median_meanframes		
207	No	sig_mean_medianframes	median of the frames of the original signal (mean, std, median)	
208	No	sig_std_medianframes		
209	No	sig_median_medianframes		
210	No	sig_75_200Hz_rms	RMS of the original signal in the [75-200] Hz frequency range	same statistical features as those numbered 176-209 extracted from the original signal in the [75-200] Hz frequency range
211	No	sig_75_200Hz_lowenergy	Low Energy of the original signal in the [75-200] Hz frequency range	
212	No	sig_75_200Hz_lowenergyASR	Average Silence Ratio of the original signal in the [75-200] Hz frequency range	
213	No	sig_75_200Hz_zerocross	Zero-cross of the original signal in the [75-200] Hz frequency range	
214	No	sig_75_200Hz_rolloff85	Roll-off (threshold=85%) of the original signal in the [75-200] Hz frequency range	
215	No	sig_75_200Hz_rolloff95	Roll-off (threshold=95%) of the original signal in the [75-200] Hz frequency range	
216	No	sig_75_200Hz_brightness_1000Hz	Brightness (cutt-off frequency=1000 Hz) of the original signal in the [75-200] Hz frequency range	
217	No	sig_75_200Hz_brightness_1500Hz	Brightness (cutt-off frequency=1500 Hz) of the original signal in the [75-200] Hz frequency range	
218	No	sig_75_200Hz_centroid	Centroid of the original signal in the [75-200] Hz frequency range	
219	No	sig_75_200Hz_spread	Spread of the original signal in the [75-200] Hz frequency range	
220	No	sig_75_200Hz_skewness	Skewness of the original signal in the [75-200] Hz frequency range	
221	No	sig_75_200Hz_kurtosis	Kurtosis of the original signal in the [75-200] Hz frequency range	
222	No	sig_75_200Hz_flatness	Flatness of the original signal in the [75-200] Hz frequency range	
223	No	sig_75_200Hz_entropy	Entropy of the original signal in the [75-200] Hz frequency range	
224	No	sig_75_200Hz_regularity	Regularity of the original signal in the [75-200] Hz frequency range	
225	No	sig_75_200Hz_mfcc01	13 MFCC of the original signal in the [75-200] Hz frequency range	
226	No	sig_75_200Hz_mfcc02		
227	No	sig_75_200Hz_mfcc03		
228	No	sig_75_200Hz_mfcc04		
229	No	sig_75_200Hz_mfcc05		
230	No	sig_75_200Hz_mfcc06		
231	No	sig_75_200Hz_mfcc07		
232	No	sig_75_200Hz_mfcc08		
233	No	sig_75_200Hz_mfcc09		
234	No	sig_75_200Hz_mfcc10		
235	No	sig_75_200Hz_mfcc11		
236	No	sig_75_200Hz_mfcc12		
237	No	sig_75_200Hz_mfcc13		
238	No	sig_75_200Hz_mean_meanframes		

239	No	sig_75_200Hz_std_meanframes	mean of the frames of the original signal in the [75-200] Hz frequency range (mean, std, median)	
240	No	sig_75_200Hz_median_meanframes		
241	No	sig_75_200Hz_mean_medianframes	median of the frames of the original signal in the [75-200] Hz frequency range (mean, std, median)	
242	No	sig_75_200Hz_std_medianframes		
243	No	sig_75_200Hz_median_medianframes		
244	No	sig_200_500Hz_rms	RMS of the original signal in the [200-500] Hz frequency range	same statistical features as those numbered 176-209 extracted from the original signal in the [200-500] Hz frequency range
245	No	sig_200_500Hz_lowenergy	Low Energy of the original signal in the [200-500] Hz frequency range	
246	No	sig_200_500Hz_lowenergyASR	Average Silence Ratio of the original signal in the [200-500] Hz frequency range	
247	No	sig_200_500Hz_zerocross	Zero-cross of the original signal in the [200-500] Hz frequency range	
248	No	sig_200_500Hz_rolloff85	Roll-off (threshold=85%) of the original signal in the [200-500] Hz frequency range	
249	No	sig_200_500Hz_rolloff95	Roll-off (threshold=95%) of the original signal in the [200-500] Hz frequency range	
250	No	sig_200_500Hz_brightness_1000Hz	Brightness (cutt-off frequency=1000 Hz) of the original signal in the [200-500] Hz frequency range	
251	No	sig_200_500Hz_brightness_1500Hz	Brightness (cutt-off frequency=1500 Hz) of the original signal in the [200-500] Hz frequency range	
252	No	sig_200_500Hz_centroid	Centroid of the original signal in the [200-500] Hz frequency range	
253	No	sig_200_500Hz_spread	Spread of the original signal in the [200-500] Hz frequency range	
254	No	sig_200_500Hz_skewness	Skewness of the original signal in the [200-500] Hz frequency range	
255	No	sig_200_500Hz_kurtosis	Kurtosis of the original signal in the [200-500] Hz frequency range	
256	No	sig_200_500Hz_flatness	Flatness of the original signal in the [200-500] Hz frequency range	
257	No	sig_200_500Hz_entropy	Entropy of the original signal in the [200-500] Hz frequency range	
258	No	sig_200_500Hz_regularity	Regularity of the original signal in the [200-500] Hz frequency range	
259	No	sig_200_500Hz_mfcc01	13 MFCC of the original signal in the [200-500] Hz frequency range	
260	No	sig_200_500Hz_mfcc02		
261	No	sig_200_500Hz_mfcc03		
262	No	sig_200_500Hz_mfcc04		
263	No	sig_200_500Hz_mfcc05		
264	No	sig_200_500Hz_mfcc06		
265	No	sig_200_500Hz_mfcc07		
266	No	sig_200_500Hz_mfcc08		
267	No	sig_200_500Hz_mfcc09		
268	No	sig_200_500Hz_mfcc10		
269	No	sig_200_500Hz_mfcc11		
270	No	sig_200_500Hz_mfcc12		
271	No	sig_200_500Hz_mfcc13		

272	No	sig_200_500Hz_mean_meanframes	mean of the frames of the original signal in the [200-500] Hz frequency range (mean, std, median)	
273	No	sig_200_500Hz_std_meanframes		
274	No	sig_200_500Hz_median_meanframes		
275	No	sig_200_500Hz_mean_medianframes	median of the frames of the original signal in the [200-500] Hz frequency range (mean, std, median)	
276	No	sig_200_500Hz_std_medianframes		
277	No	sig_200_500Hz_median_medianframes		
278	No	sig_500_1000Hz_rms	RMS of the original signal in the [500-1000] Hz frequency range	same statistical features as those numbered 176-209 extracted from the original signal in the [500-1000] Hz frequency range
279	No	sig_500_1000Hz_lowenergy	Low Energy of the original signal in the [500-1000] Hz frequency range	
280	No	sig_500_1000Hz_lowenergyASR	Average Silence Ratio of the original signal in the [500-1000] Hz frequency range	
281	No	sig_500_1000Hz_zerocross	Zero-cross of the original signal in the [500-1000] Hz frequency range	
282	No	sig_500_1000Hz_rolloff85	Roll-off (threshold=85%) of the original signal in the [500-1000] Hz frequency range	
283	No	sig_500_1000Hz_rolloff95	Roll-off (threshold=95%) of the original signal in the [500-1000] Hz frequency range	
284	No	sig_500_1000Hz_brightness_1000Hz	Brightness (cutt-off frequency=1000 Hz) of the original signal in the [500-1000] Hz frequency range	
285	No	sig_500_1000Hz_brightness_1500Hz	Brightness (cutt-off frequency=1500 Hz) of the original signal in the [500-1000] Hz frequency range	
286	No	sig_500_1000Hz_centroid	Centroid of the original signal in the [500-1000] Hz frequency range	
287	No	sig_500_1000Hz_spread	Spread of the original signal in the [500-1000] Hz frequency range	
288	No	sig_500_1000Hz_skewness	Skewness of the original signal in the [500-1000] Hz frequency range	
289	No	sig_500_1000Hz_kurtosis	Kurtosis of the original signal in the [500-1000] Hz frequency range	
290	No	sig_500_1000Hz_flatness	Flatness of the original signal in the [500-1000] Hz frequency range	
291	No	sig_500_1000Hz_entropy	Entropy of the original signal in the [500-1000] Hz frequency range	
292	No	sig_500_1000Hz_regularity	Regularity of the original signal in the [500-1000] Hz frequency range	
293	No	sig_500_1000Hz_mfcc01	13 MFCC of the original signal in the [500-1000] Hz frequency range	
294	No	sig_500_1000Hz_mfcc02		
295	No	sig_500_1000Hz_mfcc03		
296	No	sig_500_1000Hz_mfcc04		
297	No	sig_500_1000Hz_mfcc05		
298	No	sig_500_1000Hz_mfcc06		
299	No	sig_500_1000Hz_mfcc07		
300	No	sig_500_1000Hz_mfcc08		
301	No	sig_500_1000Hz_mfcc09		
302	No	sig_500_1000Hz_mfcc10		
303	No	sig_500_1000Hz_mfcc11		

304	No	sig_500_1000Hz_mfcc12		
305	No	sig_500_1000Hz_mfcc13		
306	No	sig_500_1000Hz_mean_meanframes	mean of the frames of the original signal in the [500-1000] Hz frequency range (mean, std, median)	
307	No	sig_500_1000Hz_std_meanframes		
308	No	sig_500_1000Hz_median_meanframes		
309	No	sig_500_1000Hz_mean_medianframes	median of the frames of the original signal in the [500-1000] Hz frequency range (mean, std, median)	
310	No	sig_500_1000Hz_std_medianframes		
311	No	sig_500_1000Hz_median_medianframes		
312	No	sig_1000_1500Hz_rms	RMS of the original signal in the [1000-1500] Hz frequency range	same statistical features as those numbered 176-209 extracted from the original signal in the [1000-1500] Hz frequency range
313	No	sig_1000_1500Hz_lowenergy	Low Energy of the original signal in the [1000-1500] Hz frequency range	
314	No	sig_1000_1500Hz_lowenergyASR	Average Silence Ratio of the original signal in the [1000-1500] Hz frequency range	
315	No	sig_1000_1500Hz_zerocross	Zero-cross of the original signal in the [1000-1500] Hz frequency range	
316	No	sig_1000_1500Hz_rolloff85	Roll-off (threshold=85%) of the original signal in the [1000-1500] Hz frequency range	
317	No	sig_1000_1500Hz_rolloff95	Roll-off (threshold=95%) of the original signal in the [1000-1500] Hz frequency range	
318	No	sig_1000_1500Hz_brightness_1000Hz	Brightness (cutt-off frequency=1000 Hz) of the original signal in the [1000-1500] Hz frequency range	
319	No	sig_1000_1500Hz_brightness_1500Hz	Brightness (cutt-off frequency=1500 Hz) of the original signal in the [1000-1500] Hz frequency range	
320	No	sig_1000_1500Hz_centroid	Centroid of the original signal in the [1000-1500] Hz frequency range	
321	No	sig_1000_1500Hz_spread	Spread of the original signal in the [1000-1500] Hz frequency range	
322	No	sig_1000_1500Hz_skewness	Skewness of the original signal in the [1000-1500] Hz frequency range	
323	No	sig_1000_1500Hz_kurtosis	Kurtosis of the original signal in the [1000-1500] Hz frequency range	
324	No	sig_1000_1500Hz_flatness	Flatness of the original signal in the [1000-1500] Hz frequency range	
325	No	sig_1000_1500Hz_entropy	Entropy of the original signal in the [1000-1500] Hz frequency range	
326	No	sig_1000_1500Hz_regularity	Regularity of the original signal in the [1000-1500] Hz frequency range	
327	No	sig_1000_1500Hz_mfcc01	13 MFCC of the original signal in the [1000-1500] Hz frequency range	
328	No	sig_1000_1500Hz_mfcc02		
329	No	sig_1000_1500Hz_mfcc03		
330	No	sig_1000_1500Hz_mfcc04		
331	No	sig_1000_1500Hz_mfcc05		
332	No	sig_1000_1500Hz_mfcc06		
333	No	sig_1000_1500Hz_mfcc07		

334	No	sig_1000_1500Hz_mfcc08		
335	No	sig_1000_1500Hz_mfcc09		
336	No	sig_1000_1500Hz_mfcc10		
337	No	sig_1000_1500Hz_mfcc11		
338	No	sig_1000_1500Hz_mfcc12		
339	No	sig_1000_1500Hz_mfcc13		
340	No	sig_1000_1500Hz_mean_meanframes	mean of the frames of the original signal in the [1000-1500] Hz frequency range (mean, std, median)	
341	No	sig_1000_1500Hz_std_meanframes		
342	No	sig_1000_1500Hz_median_meanframes		
343	No	sig_1000_1500Hz_mean_medianframes	median of the frames of the original signal in the [1000-1500] Hz frequency range (mean, std, median)	
344	No	sig_1000_1500Hz_std_medianframes		
345	No	sig_1000_1500Hz_median_medianframes		
346	Yes	C3_rms	RMS of the crackle component (N=3)	same statistical features as those numbered 176-209 extracted from the crackle componen t (N=3)
347	Yes	C3_lowenergy	Low Energy of the crackle component (N=3)	
348	Yes	C3_lowenergyASR	Average Silence Ratio of the crackle component (N=3)	
349	Yes	C3_zerocross	Zero-cross of the crackle component (N=3)	
350	Yes	C3_rolloff85	Roll-off (threshold=85%) of the crackle component (N=3)	
351	Yes	C3_rolloff95	Roll-off (threshold=95%) of the crackle component (N=3)	
352	Yes	C3_brightness_1000Hz	Brightness (cutt-off frequency=1000 Hz) of the crackle component (N=3)	
353	Yes	C3_brightness_1500Hz	Brightness (cutt-off frequency=1500 Hz) of the crackle component (N=3)	
354	Yes	C3_centroid	Centroid of the crackle component (N=3)	
355	Yes	C3_spread	Spread of the crackle component (N=3)	
356	Yes	C3_skewness	Skewness of the crackle component (N=3)	
357	Yes	C3_kurtosis	Kurtosis of the crackle component (N=3)	
358	Yes	C3_flatness	Flatness of the crackle component (N=3)	
359	Yes	C3_entropy	Entropy of the crackle component (N=3)	
360	Yes	C3_regularity	Regularity of the crackle component (N=3)	
361	Yes	C3_mfcc01	13 MFCC of the crackle component (N=3)	
362	Yes	C3_mfcc02		
363	Yes	C3_mfcc03		
364	Yes	C3_mfcc04		
365	Yes	C3_mfcc05		
366	Yes	C3_mfcc06		
367	Yes	C3_mfcc07		
368	Yes	C3_mfcc08		
369	Yes	C3_mfcc09		
370	Yes	C3_mfcc10		
371	Yes	C3_mfcc11		
372	Yes	C3_mfcc12		
373	Yes	C3_mfcc13		

374	Yes	C3_mean_meanframes	mean of the frames of the crackle component (N=3) (mean, std, median)	
375	Yes	C3_std_meanframes		
376	Yes	C3_median_meanframes		
377	Yes	C3_mean_medianframes	median of the frames of the crackle component (N=3) (mean, std, median)	
378	Yes	C3_std_medianframes		
379	Yes	C3_median_medianframes		
380	Yes	R3_rms	RMS of the respiratory component (N=3)	same statistical features as those numbered 176-209 extracted from the respiratory componen t (N=3)
381	Yes	R3_lowenergy	Low Energy of the respiratory component (N=3)	
382	Yes	R3_lowenergyASR	Average Silence Ratio of the respiratory component (N=3)	
383	Yes	R3_zerocross	Zero-cross of the respiratory component (N=3)	
384	Yes	R3_rolloff85	Roll-off (threshold=85%) of the respiratory component (N=3)	
385	Yes	R3_rolloff95	Roll-off (threshold=95%) of the respiratory component (N=3)	
386	Yes	R3_brightness_1000Hz	Brightness (cutt-off frequency=1000 Hz) of the respiratory component (N=3)	
387	Yes	R3_brightness_1500Hz	Brightness (cutt-off frequency=1500 Hz) of the respiratory component (N=3)	
388	Yes	R3_centroid	Centroid of the respiratory component (N=3)	
389	Yes	R3_spread	Spread of the respiratory component (N=3)	
390	Yes	R3_skewness	Skewness of the respiratory component (N=3)	
391	Yes	R3_kurtosis	Kurtosis of the respiratory component (N=3)	
392	Yes	R3_flatness	Flatness of the respiratory component (N=3)	
393	Yes	R3_entropy	Entropy of the respiratory component (N=3)	
394	Yes	R3_regularity	Regularity of the respiratory component (N=3)	
395	Yes	R3_mfcc01	13 MFCC of the respiratory component (N=3)	
396	Yes	R3_mfcc02		
397	Yes	R3_mfcc03		
398	Yes	R3_mfcc04		
399	Yes	R3_mfcc05		
400	Yes	R3_mfcc06		
401	Yes	R3_mfcc07		
402	Yes	R3_mfcc08		
403	Yes	R3_mfcc09		
404	Yes	R3_mfcc10		
405	Yes	R3_mfcc11		
406	Yes	R3_mfcc12		
407	Yes	R3_mfcc13		
408	Yes	R3_mean_meanframes	mean of the frames of the respiratory component (N=3) (mean, std, median)	
409	Yes	R3_std_meanframes		
410	Yes	R3_median_meanframes		
411	Yes	R3_mean_medianframes	median of the frames of the respiratory component (N=3) (mean, std, median)	
412	Yes	R3_std_medianframes		
413	Yes	R3_median_medianframes		

414	Yes	C4_rms	RMS of the crackle component (N=4)	same statistical features as those numbered 176-209 extracted from the crackle componen t (N=4)
415	Yes	C4_lowenergy	Low Energy of the crackle component (N=4)	
416	Yes	C4_lowenergyASR	Average Silence Ratio of the crackle component (N=4)	
417	Yes	C4_zerocross	Zero-cross of the crackle component (N=4)	
418	Yes	C4_rolloff85	Roll-off (threshold=85%) of the crackle component (N=4)	
419	Yes	C4_rolloff95	Roll-off (threshold=95%) of the crackle component (N=4)	
420	Yes	C4_brightness_1000Hz	Brightness (cutt-off frequency=1000 Hz) of the crackle component (N=4)	
421	Yes	C4_brightness_1500Hz	Brightness (cutt-off frequency=1500 Hz) of the crackle component (N=4)	
422	Yes	C4_centroid	Centroid of the crackle component (N=4)	
423	Yes	C4_spread	Spread of the crackle component (N=4)	
424	Yes	C4_skewness	Skewness of the crackle component (N=4)	
425	Yes	C4_kurtosis	Kurtosis of the crackle component (N=4)	
426	Yes	C4_flatness	Flatness of the crackle component (N=4)	
427	Yes	C4_entropy	Entropy of the crackle component (N=4)	
428	Yes	C4_regularity	Regularity of the crackle component (N=4)	
429	Yes	C4_mfcc01	13 MFCC of the crackle component (N=4)	
430	Yes	C4_mfcc02		
431	Yes	C4_mfcc03		
432	Yes	C4_mfcc04		
433	Yes	C4_mfcc05		
434	Yes	C4_mfcc06		
435	Yes	C4_mfcc07		
436	Yes	C4_mfcc08		
437	Yes	C4_mfcc09		
438	Yes	C4_mfcc10		
439	Yes	C4_mfcc11		
440	Yes	C4_mfcc12		
441	Yes	C4_mfcc13		
442	Yes	C4_mean_meanframes	mean of the frames of the crackle component (N=4) (mean, std, median)	
443	Yes	C4_std_meanframes		
444	Yes	C4_median_meanframes		
445	Yes	C4_mean_medianframes	median of the frames of the crackle component (N=4) (mean, std, median)	
446	Yes	C4_std_medianframes		
447	Yes	C4_median_medianframes		
448	Yes	R4_rms	RMS of the respiratory component (N=4)	same statistical features as those numbered 176-209 extracted from the respiratory componen t (N=4)
449	Yes	R4_lowenergy	Low Energy of the respiratory component (N=4)	
450	Yes	R4_lowenergyASR	Average Silence Ratio of the respiratory component (N=4)	
451	Yes	R4_zerocross	Zero-cross of the respiratory component (N=4)	
452	Yes	R4_rolloff85	Roll-off (threshold=85%) of the respiratory component (N=4)	
453	Yes	R4_rolloff95	Roll-off (threshold=95%) of the respiratory component (N=4)	

454	Yes	R4_brightness_1000Hz	Brightness (cutt-off frequency=1000 Hz) of the respiratory component (N=4)	
455	Yes	R4_brightness_1500Hz	Brightness (cutt-off frequency=1500 Hz) of the respiratory component (N=4)	
456	Yes	R4_centroid	Centroid of the respiratory component (N=4)	
457	Yes	R4_spread	Spread of the respiratory component (N=4)	
458	Yes	R4_skewness	Skewness of the respiratory component (N=4)	
459	Yes	R4_kurtosis	Kurtosis of the respiratory component (N=4)	
460	Yes	R4_flatness	Flatness of the respiratory component (N=4)	
461	Yes	R4_entropy	Entropy of the respiratory component (N=4)	
462	Yes	R4_regularity	Regularity of the respiratory component (N=4)	
463	Yes	R4_mfcc01	13 MFCC of the respiratory component (N=4)	
464	Yes	R4_mfcc02		
465	Yes	R4_mfcc03		
466	Yes	R4_mfcc04		
467	Yes	R4_mfcc05		
468	Yes	R4_mfcc06		
469	Yes	R4_mfcc07		
470	Yes	R4_mfcc08		
471	Yes	R4_mfcc09		
472	Yes	R4_mfcc10		
473	Yes	R4_mfcc11		
474	Yes	R4_mfcc12		
475	Yes	R4_mfcc13		
476	Yes	R4_mean_meanframes	mean of the frames of the respiratory component (N=4) (mean, std, median)	
477	Yes	R4_std_meanframes		
478	Yes	R4_median_meanframes		
479	Yes	R4_mean_medianframes	median of the frames of the respiratory component (N=4) (mean, std, median)	
480	Yes	R4_std_medianframes		
481	Yes	R4_median_medianframes		

Glossary

MFCCs = Mel Frequency Cepstral Coefficients are features of the cepstral domain. Cepstrum is the result of taking the inverse Fourier transform (IFT) of the logarithm of the estimated spectrum of a signal. One main property of the cepstrum is the analysis of the spectrum regarding to its fineness. Coefficients with lower indices describe the coarse structure of the signal, coefficients with high indices carry information about its fine structure.

ZCR = the zero-crossing rate is the rate of sign-changes along a signal, i.e., the rate at which the signal changes from positive to negative polarity or back. It is used in two different ways: on the one hand, it is used as single feature itself, on the other hand it is used to enhance the evaluation of the Empirical Mode Decomposition (EMD). The zero-crossing rate is an important feature to distinguish between voiced and unvoiced signals. For unvoiced, noisy signals its value is high, for voiced signals the rate of zero-crossings is low.

RMS = root mean square is defined as the square root of mean square (the arithmetic mean of the squares of a set of numbers). RMS can also be defined for a continuously varying function in terms of an integral of the squares of the instantaneous values during a cycle.

Roll-off is the steepness of a transmission function with frequency. It is usual to measure roll-off as a function of logarithmic frequency, consequently, the units of roll-off are either decibels per decade (dB/decade), where a decade is a 10-times increase in frequency, or decibels per octave (dB/8ve), where an octave is 2-times increase in frequency.

The **spectral centroid** is a measure used in digital signal processing to characterise a spectrum. It indicates where the "center of mass" of the spectrum is. Perceptually, it has a robust connection with the impression of "brightness" of a sound, a term used as an indication of the amount of high-frequency content in a sound. Because the spectral centroid is a good predictor of the "brightness" of a sound, it is widely used in digital audio and music processing as an automatic measure of musical timbre.

Bibliography

Abe S, Azuma A, Mukae H *et al.* Polymyxin B-immobilized fiber column (PMX) treatment for idiopathic pulmonary fibrosis with acute exacerbation: a multicenter retrospective analysis. *Intern Med*, 51(12), 1487-1491 (2012).

Abe S, Hayashi H, Seo Y *et al.* Reduction in serum high mobility group box-1 level by polymyxin B-immobilized fiber column in patients with idiopathic pulmonary fibrosis with acute exacerbation. *Blood Purif*, 32(4), 310-316 (2011).

Abella M, Formolo J, Penney DG. Comparison of the acoustic properties of six popular stethoscopes. *J Acoust Soc Am*, 91(4 Pt 1), 2224-2228 (1992).

Akhtar AA, Ali MA, Smith RP. Depression in patients with idiopathic pulmonary fibrosis. *Chronic respiratory disease*, 10(3), 127-133 (2013).

Akira M, Hamada H, Sakatani M, Kobayashi C, Nishioka M, Yamamoto S. CT findings during phase of accelerated deterioration in patients with idiopathic pulmonary fibrosis. *AJR Am J Roentgenol*, 168(1), 79-83 (1997).

Akira M, Kozuka T, Yamamoto S, Sakatani M. Computed tomography findings in acute exacerbation of idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 178(4), 372-378 (2008).

al Jarad N, Strickland B, Bothamley G, Lock S, Logan-Sinclair R, Rudd RM. Diagnosis of asbestosis by a time expanded wave form analysis, auscultation and high resolution computed tomography: a comparative study. *Thorax*, 48(4), 347-353 (1993).

Al-Hameed FM, Sharma S. Outcome of patients admitted to the intensive care unit for acute exacerbation of idiopathic pulmonary fibrosis. *Can Respir J*, 11(2), 117-122 (2004).

Albera C, Costabel U, Fagan EA *et al.* Efficacy of pirfenidone in patients with idiopathic pulmonary fibrosis with more preserved lung function. *The European respiratory journal*, 48(3), 843-851 (2016).

Allen JT, Spiteri MA. Growth factors in idiopathic pulmonary fibrosis: relative roles. *Respir Res*, 3, 13 (2002).

Ando M, Miyazaki E, Ito T *et al.* Significance of serum vascular endothelial growth factor level in patients with idiopathic pulmonary fibrosis. *Lung*, 188(3), 247-252 (2010).

Antoniou KM, Hansell DM, Rubens MB *et al.* Idiopathic pulmonary fibrosis: outcome in relation to smoking status. *American journal of respiratory and critical care medicine*, 177(2), 190-194 (2008).

Arakawa H, Honma K. Honeycomb Lung: History and Current Concepts. *American Journal of Roentgenology*, 196(4), 773-782 (2011).

Atkins CP, Loke YK, Wilson AM. Outcomes in idiopathic pulmonary fibrosis: a meta-analysis from placebo controlled trials. *Respir Med*, 108(2), 376-387 (2014).

ATS. Report on pulmonary nomenclature. *ATS news*, (3), 5-6 (1977).

ATS. Guidelines for the six-minute walk test. *American journal of respiratory and critical care medicine*, 166(1), 111-117 (2002).

Bibliography

ATS, ERS. Idiopathic Pulmonary Fibrosis: diagnosis and treatment: international consensus statement. *American journal of respiratory and critical care medicine*, 161(2), 646-664 (2000).

Azuma A, Nukiwa T, Tsuboi E *et al.* Double-blind, placebo-controlled trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 171(9), 1040-1047 (2005).

Banaszak EF, Kory RC, Snider GL. Phonopneumography. *Am Rev Respir Dis*, 107(3), 449-455 (1973).

Bando M, Hosono T, Mato N *et al.* Long-term efficacy of inhaled N-acetylcysteine in patients with idiopathic pulmonary fibrosis. *Intern Med*, 49(21), 2289-2296 (2010).

Bando M, Ohno S, Hosono T *et al.* Risk of Acute Exacerbation After Video-assisted Thoracoscopic Lung Biopsy for Interstitial Lung Disease. *J Bronchology Interv Pulmonol*, 16(4), 229-235 (2009).

Bando M, Ohno S, Oshikawa K, Takahashi M, Okamoto H, Sugiyama Y. Infection of TT virus in patients with idiopathic pulmonary fibrosis. *Respir Med*, 95(12), 935-942 (2001).

Behr J, Bendstrup E, Crestani B *et al.* Safety and tolerability of acetylcysteine and pirfenidone combination therapy in idiopathic pulmonary fibrosis: a randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet Respir Med*, 4(6), 445-453 (2016).

Best AC, Lynch AM, Bozic CM, Miller D, Grunwald GK, Lynch DA. Quantitative CT indexes in idiopathic pulmonary fibrosis: relationship with physiologic impairment. *Radiology*, 228(2), 407-414 (2003).

Best AC, Meng J, Lynch AM *et al.* Idiopathic pulmonary fibrosis: physiologic tests, quantitative CT indexes, and CT visual scores as predictors of mortality. *Radiology*, 246(3), 935-940 (2008).

Bhattacharyya P, Mondal A, Dey R, Saha D, Saha G. Novel algorithm to identify and differentiate specific digital signature of breath sound in patients with diffuse parenchymal lung disease. *Respirology*, 20(4), 633-639 (2015).

Blackwell TS, Tager AM, Borok Z *et al.* Future directions in idiopathic pulmonary fibrosis research. An NHLBI workshop report. *American journal of respiratory and critical care medicine*, 189(2), 214-222 (2014).

Bonella F, Wessendorf TE, Costabel U. [Clinical experience with pirfenidone for the treatment of idiopathic pulmonary fibrosis]. *Dtsch Med Wochenschr*, 138(11), 518-523 (2013).

Boon K, Bailey NW, Yang J *et al.* Molecular phenotypes distinguish patients with relatively stable from progressive idiopathic pulmonary fibrosis (IPF). *PLoS One*, 4(4), e5134 (2009).

Borg GA. Psychophysical bases of perceived exertion. *Medicine and science in sports and exercise*, 14(5), 377-381 (1982).

Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ*, 304(6840), 1491-1494 (1992).

Brown AW, Fischer CP, Shlobin OA *et al.* Outcomes after hospitalization in idiopathic pulmonary fibrosis: a cohort study. *Chest*, 147(1), 173-179 (2015).

Brownell R, Moua T, Henry TS *et al.* The use of pretest probability increases the value of high-resolution CT in diagnosing usual interstitial pneumonia. *Thorax*, 72(5), 424-429 (2017).

Cantin AM, Hubbard RC, Crystal RG. Glutathione deficiency in the epithelial lining fluid of the lower respiratory tract in idiopathic pulmonary fibrosis. *Am Rev Respir Dis*, 139(2), 370-372 (1989).

Casoni GL, Tomassetti S, Cavazza A *et al.* Transbronchial lung cryobiopsy in the diagnosis of fibrotic interstitial lung diseases. *PLoS One*, 9(2), e86716 (2014).

Charleston-Villalobos S, Gonzalez-Camarena R, Chi-Lem G, Aljama-Corrales T. Crackle sounds analysis by empirical mode decomposition. Nonlinear and nonstationary signal analysis for distinction of crackles in lung sounds. *IEEE Eng Med Biol Mag*, 26(1), 40-47 (2007).

Chaudhary NI, Roth GJ, Hilberg F *et al.* Inhibition of PDGF, VEGF and FGF signalling attenuates fibrosis. *The European respiratory journal*, 29(5), 976-985 (2007).

Chen CR, Shu WY, Chang CW, Hsu IC. Identification of under-detected periodicity in time-series microarray data by using empirical mode decomposition. *PLoS One*, 9(11), e111719 (2014).

Choi SM, Lee J, Park YS *et al.* Postoperative pulmonary complications after surgery in patients with interstitial lung disease. *Respiration*, 87(4), 287-293 (2014).

Collard HR, Chen SY, Yeh WS *et al.* Health care utilization and costs of idiopathic pulmonary fibrosis in U.S. Medicare beneficiaries aged 65 years and older. *Ann Am Thorac Soc*, 12(7), 981-987 (2015).

Collard HR, King TE, Jr., Bartelson BB, Vourlekis JS, Schwarz MI, Brown KK. Changes in clinical and physiologic variables predict survival in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 168(5), 538-542 (2003).

Collard HR, Moore BB, Flaherty KR *et al.* Acute exacerbations of idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 176(7), 636-643 (2007).

Collard HR, Ryerson CJ, Corte TJ *et al.* Acute Exacerbation of Idiopathic Pulmonary Fibrosis: An International Working Group Report. *American journal of respiratory and critical care medicine*, (2016).

Collard HR, Yow E, Richeldi L, Anstrom KJ, Glazer C, investigators IP. Suspected acute exacerbation of idiopathic pulmonary fibrosis as an outcome measure in clinical trials. *Respir Res*, 14, 73 (2013).

Concato J, Peduzzi P, Holford TR, Feinstein AR. Importance of events per independent variable in proportional hazards analysis. I. Background, goals, and general strategy. *J Clin Epidemiol*, 48(12), 1495-1501 (1995).

Cordier JF, Cottin V. Neglected evidence in idiopathic pulmonary fibrosis: from history to earlier diagnosis. *The European respiratory journal*, 42(4), 916-923 (2013).

Corte TJ, Wort SJ, Wells AU. Pulmonary hypertension in idiopathic pulmonary fibrosis: a review. *Sarcoidosis, vasculitis, and diffuse lung diseases : official journal of WASOG / World Association of Sarcoidosis and Other Granulomatous Disorders*, 26(1), 7-19 (2009).

Costabel U, Albera C, Bradford WZ *et al.* Analysis of lung function and survival in RECAP: An open-label extension study of pirfenidone in patients with idiopathic pulmonary fibrosis. *Sarcoidosis, vasculitis, and diffuse lung diseases : official journal of WASOG / World Association of Sarcoidosis and Other Granulomatous Disorders*, 31(3), 198-205 (2014).

Costabel U, Inoue Y, Richeldi L *et al.* Efficacy of Nintedanib in Idiopathic Pulmonary Fibrosis across Prespecified Subgroups in INPULSIS. *American journal of respiratory and critical care medicine*, 193(2), 178-185 (2016).

Cottin V. The impact of emphysema in pulmonary fibrosis. *Eur Respir Rev*, 22(128), 153-157 (2013).

Cottin V, Cordier JF. Velcro crackles: the key for early diagnosis of idiopathic pulmonary fibrosis? *Eur Respir J*, 40(3), 519-521 (2012).

Cottin V, Hansell DM, Sverzellati N *et al.* Effect of Emphysema Extent on Serial Lung Function in Patients with Idiopathic Pulmonary Fibrosis. *American journal of respiratory and critical care medicine*, (2017).

Cottin V, Richeldi L. Neglected evidence in idiopathic pulmonary fibrosis and the importance of early diagnosis and treatment. *Eur Respir Rev*, 23(131), 106-110 (2014).

Coultas DB, Zumwalt RE, Black WC, Sobonya RE. The epidemiology of interstitial lung diseases. *American Journal of Respiratory and Critical Care Medicine*, 150(4), 967-972 (1994).

Coward WR, Saini G, Jenkins G. The pathogenesis of idiopathic pulmonary fibrosis. *Ther Adv Respir Dis*, 4(6), 367-388 (2010).

Daniels CE, Lasky JA, Limper AH *et al.* Imatinib treatment for idiopathic pulmonary fibrosis: Randomized placebo-controlled trial results. *Am J Respir Crit Care Med*, 181(6), 604-610 (2010).

Daniels CE, Yi ES, Ryu JH. Autopsy findings in 42 consecutive patients with idiopathic pulmonary fibrosis. *Eur Respir J*, 32(1), 170-174 (2008).

de Hemptinne Q, Remmelink M, Brimioulle S, Salmon I, Vincent JL. ARDS: a clinicopathological confrontation. *Chest*, 135(4), 944-949 (2009).

De Vries J, Kessels BL, Drent M. Quality of life of idiopathic pulmonary fibrosis patients. *The European respiratory journal*, 17(5), 954-961 (2001).

Demedts M, Behr J, Buhl R *et al.* High-dose acetylcysteine in idiopathic pulmonary fibrosis. *N Engl J Med*, 353(21), 2229-2242 (2005).

Ding J, Chen Z, Feng K. Procalcitonin-guided antibiotic use in acute exacerbations of idiopathic pulmonary fibrosis. *Int J Med Sci*, 10(7), 903-907 (2013).

Donahoe M, Valentine VG, Chien N *et al.* Autoantibody-Targeted Treatments for Acute Exacerbations of Idiopathic Pulmonary Fibrosis. *PLoS One*, 10(6), e0127771 (2015).

du Bois RM. An earlier and more confident diagnosis of idiopathic pulmonary fibrosis. *Eur Respir Rev*, 21(124), 141-146 (2012).

du Bois RM, Albera C, Bradford WZ *et al.* 6-Minute walk distance is an independent predictor of mortality in patients with idiopathic pulmonary fibrosis. *The European respiratory journal*, 43(5), 1421-1429 (2014).

du Bois RM, Nathan SD, Richeldi L, Schwarz MI, Noble PW. Idiopathic pulmonary fibrosis: lung function is a clinically meaningful endpoint for phase III trials. *American journal of respiratory and critical care medicine*, 186(8), 712-715 (2012).

Du Bois RM, Richeldi L. *Interstitial Lung Diseases* (European Respiratory Society Journals Ltd, 2009).

du Bois RM, Weycker D, Albera C *et al.* Forced vital capacity in patients with idiopathic pulmonary fibrosis: test properties and minimal clinically important difference. *American journal of respiratory and critical care medicine*, 184(12), 1382-1389 (2011a).

du Bois RM, Weycker D, Albera C *et al.* Ascertainment of individual risk of mortality for patients with idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 184(4), 459-466 (2011b).

du Bois RM, Weycker D, Albera C *et al.* Six-minute-walk test in idiopathic pulmonary fibrosis: test validation and minimal clinically important difference. *American journal of respiratory and critical care medicine*, 183(9), 1231-1237 (2011c).

Durheim MT, Collard HR, Roberts RS *et al.* Association of hospital admission and forced vital capacity endpoints with survival in patients with idiopathic pulmonary fibrosis: analysis of a pooled cohort from three clinical trials. *Lancet Respir Med*, 3(5), 388-396 (2015).

Eakin EG, Resnikoff PM, Prewitt LM, Ries AL, Kaplan RM. Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire. University of California, San Diego. *Chest*, 113(3), 619-624 (1998).

Fagan TJ. Nomogram for Bayes's Theorem. *New England Journal of Medicine*, 293(5), 257-257 (1975).

Falaschetti E, Laiho J, Primatesta P, Purdon S. Prediction equations for normal and low lung function from the Health Survey for England. *The European respiratory journal*, 23(3), 456-463 (2004).

Farkas L, Gauldie J, Voelkel NF, Kolb M. Pulmonary hypertension and idiopathic pulmonary fibrosis: a tale of angiogenesis, apoptosis, and growth factors. *Am J Respir Cell Mol Biol*, 45(1), 1-15 (2011).

Fell CD. Idiopathic pulmonary fibrosis: phenotypes and comorbidities. *Clin Chest Med*, 33(1), 51-57 (2012).

Fell CD, Martinez FJ, Liu LX *et al.* Clinical predictors of a diagnosis of idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 181(8), 832-837 (2010).

Ferguson ND, Frutos-Vivar F, Esteban A *et al.* Acute respiratory distress syndrome: underrecognition by clinicians and diagnostic accuracy of three clinical definitions. *Crit Care Med*, 33(10), 2228-2234 (2005).

Fernandez Perez ER, Daniels CE, Schroeder DR *et al.* Incidence, prevalence, and clinical course of idiopathic pulmonary fibrosis: a population-based study. *Chest*, 137(1), 129-137 (2010).

Ferreira A, Garvey C, Connors GL *et al.* Pulmonary rehabilitation in interstitial lung disease: benefits and predictors of response. *Chest*, 135(2), 442-447 (2009).

Flaherty K, Thwaite E, Kazerooni E *et al.* Radiological versus histological diagnosis in UIP and NSIP: survival implications. *Thorax*, 58(2), 143-148 (2003).

Flaherty KR, Toews GB, Travis WD *et al.* Clinical significance of histological classification of idiopathic interstitial pneumonia. *The European respiratory journal*, 19(2), 275-283 (2002).

Fleiss JL. *The design and analysis of clinical experiments* (John Wiley and Sons, New York, 1986).

Flietstra B, Markuzon N, Vyshedskiy A, Murphy R. Automated Analysis of Crackles in Patients with Interstitial Pulmonary Fibrosis. *Pulmonary Medicine*, 2011, 590506 (2011).

Forgacs P. Crackles and wheezes. *Lancet*, 2(7508), 203-205 (1967).

Fredberg JJ HS. Discrete lung sounds: crackles (rales) as stress-relaxation quadropoles. *J Acoust Soc Am*, (73), 1036-1046 (1983).

Friedman SL, Sheppard D, Duffield JS, Violette S. Therapy for fibrotic diseases: nearing the starting line. *Science translational medicine*, 5(167), 167sr161 (2013).

Fujimoto K, Taniguchi H, Johkoh T *et al.* Acute exacerbation of idiopathic pulmonary fibrosis: high-resolution CT scores predict mortality. *Eur Radiol*, 22(1), 83-92 (2012).

George TJ, Arnaoutakis GJ, Shah AS. LUng transplant in idiopathic pulmonary fibrosis. *Archives of Surgery*, 146(10), 1204-1209 (2011).

Ghatol A, Ruhl AP, Danoff SK. Exacerbations in idiopathic pulmonary fibrosis triggered by pulmonary and nonpulmonary surgery: a case series and comprehensive review of the literature. *Lung*, 190(4), 373-380 (2012).

Greene KE, King TE, Jr., Kuroki Y *et al.* Serum surfactant proteins-A and -D as biomarkers in idiopathic pulmonary fibrosis. *Eur Respir J*, 19(3), 439-446 (2002).

Gribbin J, Hubbard RB, Le Jeune I, Smith CJ, West J, Tata LJ. Incidence and mortality of idiopathic pulmonary fibrosis and sarcoidosis in the UK. *Thorax*, 61(11), 980-985 (2006).

Gruden JF, Panse PM, Gotway MB, Jensen EA, Wellnitz CV, Wesselius L. Diagnosis of Usual Interstitial Pneumonitis in the Absence of Honeycombing: Evaluation of Specific CT Criteria With Clinical Follow-Up in 38 Patients. *AJR Am J Roentgenol*, 1-9 (2015).

Gurung A, Scrafford CG, Tielsch JM, Levine OS, Checkley W. Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: a systematic review and meta-analysis. *Respir Med*, 105(9), 1396-1403 (2011).

Hagiwara SI, Ishii Y, Kitamura S. Aerosolized administration of N-acetylcysteine attenuates lung fibrosis induced by bleomycin in mice. *American journal of respiratory and critical care medicine*, 162(1), 225-231 (2000).

Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*, 148(3), 839-843 (1983).

Hansell DM, Bankier AA, MacMahon H, McLoud TC, Muller NL, Remy J. Fleischner Society: glossary of terms for thoracic imaging. *Radiology*, 246(3), 697-722 (2008a).

Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J. Fleischner Society: glossary of terms for thoracic imaging. *Radiology*, 246 (2008b).

Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*, 15(4), 361-387 (1996).

Hartley PG, Galvin JR, Hunninghake GW *et al.* High-resolution CT-derived measures of lung density are valid indexes of interstitial lung disease. *J Appl Physiol (1985)*, 76(1), 271-277 (1994).

Hilberg F, Roth GJ, Krssak M *et al.* BIBF 1120: triple angiokinase inhibitor with sustained receptor blockade and good antitumor efficacy. *Cancer Res*, 68(12), 4774-4782 (2008).

Hoevers J, Loudon RG. Measuring crackles. *Chest*, 98(5), 1240-1243 (1990).

Holford SK. Discontinuous adventitious lung sounds: measurement, classification and modeling. *PhD dissertation, Cambridge*, (1981).

Holland AE, Fiore JF, Jr., Bell EC *et al.* Dyspnoea and comorbidity contribute to anxiety and depression in interstitial lung disease. *Respirology*, 19(8), 1215-1221 (2014).

Homma S, Sakamoto S, Kawabata M *et al.* Cyclosporin treatment in steroid-resistant and acutely exacerbated interstitial pneumonia. *Intern Med*, 44(11), 1144-1150 (2005).

Horita N, Akahane M, Okada Y *et al.* Tacrolimus and steroid treatment for acute exacerbation of idiopathic pulmonary fibrosis. *Intern Med*, 50(3), 189-195 (2011).

Huie TJ, Olson AL, Cosgrove GP *et al.* A detailed evaluation of acute respiratory decline in patients with fibrotic lung disease: aetiology and outcomes. *Respirology*, 15(6), 909-917 (2010).

Hunninghake GM. A new hope for idiopathic pulmonary fibrosis. *N Engl J Med*, 370(22), 2142-2143 (2014).

Hunninghake GM, Hatabu H, Okajima Y *et al.* MUC5B promoter polymorphism and interstitial lung abnormalities. *N Engl J Med*, 368(23), 2192-2200 (2013).

Hunninghake GW, Lynch DA, Galvin JR *et al.* Radiologic findings are strongly associated with a pathologic diagnosis of usual interstitial pneumonia. *Chest*, 124(4), 1215-1223 (2003).

Inase N, Sawada M, Ohtani Y *et al.* Cyclosporin A followed by the treatment of acute exacerbation of idiopathic pulmonary fibrosis with corticosteroid. *Intern Med*, 42(7), 565-570 (2003).

Isshiki T, Sakamoto S, Kinoshita A, Sugino K, Kurosaki A, Homma S. Recombinant human soluble thrombomodulin treatment for acute exacerbation of idiopathic pulmonary fibrosis: a retrospective study. *Respiration*, 89(3), 201-207 (2015).

Iwata T, Yoshida S, Nagato K *et al.* Experience with perioperative pirfenidone for lung cancer surgery in patients with idiopathic pulmonary fibrosis. *Surg Today*, 45(10), 1263-1270 (2015).

Jacob J, Bartholmai BJ, Rajagopalan S *et al.* Automated Quantitative Computed Tomography Versus Visual Computed Tomography Scoring in Idiopathic Pulmonary Fibrosis: Validation Against Pulmonary Function. *J Thorac Imaging*, 31(5), 304-311 (2016).

Jacob J, Bartholmai BJ, Rajagopalan S *et al.* Mortality prediction in idiopathic pulmonary fibrosis: evaluation of computer-based CT analysis with conventional severity measures. *The European respiratory journal*, 49(1) (2017).

Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *Jama*, 271(5), 389-391 (1994).

Jankowich MD, Rounds S. Combined pulmonary fibrosis and emphysema alters physiology but has similar mortality to pulmonary fibrosis without emphysema. *Lung*, 188(5), 365-373 (2010).

Jegal Y, Kim DS, Shim TS *et al.* Physiology is a stronger predictor of survival than pathology in fibrotic interstitial pneumonia. *American journal of respiratory and critical care medicine*, 171(6), 639-644 (2005).

Jenkins RG, Simpson JK, Saini G *et al.* Longitudinal change in collagen degradation biomarkers in idiopathic pulmonary fibrosis: an analysis from the prospective, multicentre PROFILE study. *The Lancet Respiratory Medicine*, 3(6), 462-472 (2015).

Jeon K, Chung MP, Lee KS *et al.* Prognostic factors and causes of death in Korean patients with idiopathic pulmonary fibrosis. *Respir Med*, 100(3), 451-457 (2006).

Jin GY, Lynch D, Chawla A *et al.* Interstitial lung abnormalities in a CT lung cancer screening population: prevalence and progression rate. *Radiology*, 268(2), 563-571 (2013).

Johannson K, Collard HR. Acute Exacerbation of Idiopathic Pulmonary Fibrosis: A Proposal. *Curr Respir Care Rep*, 2(4) (2013).

Johannson KA, Vittinghoff E, Lee K *et al.* Acute exacerbation of idiopathic pulmonary fibrosis associated with air pollution exposure. *Eur Respir J*, 43(4), 1124-1131 (2014).

John AE, Luckett JC, Tatler AL *et al.* Preclinical SPECT/CT imaging of alphavbeta6 integrins for molecular stratification of idiopathic pulmonary fibrosis. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, 54(12), 2146-2152 (2013).

Johnston I BJ, Kinnear W, Logan R. Rising mortality from cryptogenic fibrosing alveolitis. *BMJ*, (01(6759)), 1017-1021 (1990).

Johnston I, Britton J, Kinnear W, Logan R. Rising mortality from cryptogenic fibrosing alveolitis. *BMJ*, 301(6759), 1017-1021 (1990).

Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. *Respir Med*, 85 Suppl B, 25-31; discussion 33-27 (1991).

Judge EP, Fabre A, Adamali HI, Egan JJ. Acute exacerbations and pulmonary hypertension in advanced idiopathic pulmonary fibrosis. *The European respiratory journal*, 40(1), 93-100 (2012).

Kahloon RA, Xue J, Bhargava A *et al.* Patients with idiopathic pulmonary fibrosis with antibodies to heat shock protein 70 have poor prognoses. *American journal of respiratory and critical care medicine*, 187(7), 768-775 (2013).

Kaisla T, Sovijarvi A, Piirila P, Rajala HM, Haltsonen S, Rosqvist T. Validated method for automatic detection of lung sound crackles. *Med Biol Eng Comput*, 29(5), 517-521 (1991).

Kandaswamy A KC, Ramanathan RP, Jaymaraman S, Malmurugan N. Neural classification of lung sounds using wavelet coefficients. *Comput Biol Med*, (34), 523-537 (2004).

Kataoka K, Taniguchi H, Kondoh Y *et al.* Recombinant Human Thrombomodulin in Acute Exacerbation of Idiopathic Pulmonary Fibrosis. *Chest*, 148(2), 436-443 (2015).

Kazerooni EA, Martinez FJ, Flint A *et al.* Thin-section CT obtained at 10-mm increments versus limited three-level thin-section CT for idiopathic pulmonary fibrosis: correlation with pathologic scoring. *American Journal of Roentgenology*, 169(4), 977-983 (1997).

Kim DS, Park JH, Park BK, Lee JS, Nicholson AG, Colby T. Acute exacerbation of idiopathic pulmonary fibrosis: frequency and clinical features. *The European respiratory journal*, 27(1), 143-150 (2006).

Kinder BW, Brown KK, McCormack FX *et al.* Serum surfactant protein-A is a strong predictor of early mortality in idiopathic pulmonary fibrosis. *Chest*, 135(6), 1557-1563 (2009).

Kinder BW, Collard HR, Koth L *et al.* Idiopathic nonspecific interstitial pneumonia: lung manifestation of undifferentiated connective tissue disease? *American journal of respiratory and critical care medicine*, 176(7), 691-697 (2007).

King TE, Jr., Albera C, Bradford WZ *et al.* All-cause mortality rate in patients with idiopathic pulmonary fibrosis. Implications for the design and execution of clinical trials. *Am J Respir Crit Care Med*, 189(7), 825-831 (2014a).

King TE, Jr., Albera C, Bradford WZ *et al.* Effect of interferon gamma-1b on survival in patients with idiopathic pulmonary fibrosis (INSPIRE): a multicentre, randomised, placebo-controlled trial. *Lancet*, 374(9685), 222-228 (2009).

King TE, Jr., Behr J, Brown KK *et al.* BUILD-1: a randomized placebo-controlled trial of bosentan in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 177(1), 75-81 (2008).

King TE, Jr., Bradford WZ, Castro-Bernardini S *et al.* A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N Engl J Med*, 370(22), 2083-2092 (2014b).

King TE, Jr., Brown KK, Raghu G *et al.* BUILD-3: a randomized, controlled trial of bosentan in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 184(1), 92-99 (2011a).

King TE, Jr., Tooze JA, Schwarz MI, Brown KR, Cherniack RM. Predicting survival in idiopathic pulmonary fibrosis: scoring system and survival model. *American journal of respiratory and critical care medicine*, 164(7), 1171-1181 (2001a).

King TE, Pardo A, Selman M. Idiopathic pulmonary fibrosis. *The Lancet*, 378(9807), 1949-1961 (2011b).

King TE, Schwarz MI, Brown K *et al.* Idiopathic Pulmonary Fibrosis. *American journal of respiratory and critical care medicine*, 164(6), 1025-1032 (2001b).

Kishaba T, Tamaki H, Shimaoka Y, Fukuyama H, Yamashiro S. Staging of acute exacerbation in patients with idiopathic pulmonary fibrosis. *Lung*, 192(1), 141-149 (2014).

Klingberg F, Chow ML, Koehler A *et al.* Prestress in the extracellular matrix sensitizes latent TGF-beta1 for activation. *The Journal of cell biology*, 207(2), 283-297 (2014).

Kolb M, Richeldi L, Behr J *et al.* Nintedanib in patients with idiopathic pulmonary fibrosis and preserved lung volume. *Thorax*, (2016).

Kompis M, Pasterkamp H, Wodicka GR. Acoustic imaging of the human chest. *Chest*, 120(4), 1309-1321 (2001).

Kondoh Y, Taniguchi H, Ebina M *et al.* Risk factors for acute exacerbation of idiopathic pulmonary fibrosis--Extended analysis of pirfenidone trial in Japan. *Respir Investig*, 53(6), 271-278 (2015).

Kondoh Y, Taniguchi H, Katsuta T *et al.* Risk factors of acute exacerbation of idiopathic pulmonary fibrosis. *Sarcoidosis, vasculitis, and diffuse lung diseases : official journal of WASOG / World Association of Sarcoidosis and Other Granulomatous Disorders*, 27(2), 103-110 (2010).

Konishi K, Gibson KF, Lindell KO *et al.* Gene expression profiles of acute exacerbations of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 180(2), 167-175 (2009).

Korthagen NM, van Moorsel CH, Barlo NP *et al.* Serum and BALF YKL-40 levels are predictors of survival in idiopathic pulmonary fibrosis. *Respir Med*, 105(1), 106-113 (2011).

Kraman SS. The relationship between airflow and lung sound amplitude in normal subjects. *Chest*, 86(2), 225-229 (1984).

Kraman SS. Effects of lung volume and airflow on the frequency spectrum of vesicular lung sounds. *Respir Physiol*, 66(1), 1-9 (1986).

Kubo H, Nakayama K, Yanai M *et al.* Anticoagulant therapy for idiopathic pulmonary fibrosis. *Chest*, 128(3), 1475-1482 (2005).

Kulkarni AA, Thatcher TH, Hsiao HM *et al.* The triterpenoid CDDO-Me inhibits bleomycin-induced lung inflammation and fibrosis. *PLoS One*, 8(5), e63798 (2013).

Kupferberg DH, Kaplan RM, Slymen DJ, Ries AL. Minimal clinically important difference for the UCSD Shortness of Breath Questionnaire. *J Cardiopulm Rehabil*, 25(6), 370-377 (2005).

Kurashima K, Takayanagi N, Tsuchiya N *et al.* The effect of emphysema on lung function and survival in patients with idiopathic pulmonary fibrosis. *Respirology*, 15(5), 843-848 (2010).

Lamas DJ, Kawut SM, Bagiella E, Philip N, Arcasoy SM, Lederer DJ. Delayed access and survival in idiopathic pulmonary fibrosis: a cohort study. *Am J Respir Crit Care Med*, 184(7), 842-847 (2011).

Larsen BT, Colby TV. Update for Pathologists on Idiopathic Interstitial Pneumonias. *Archives of Pathology & Laboratory Medicine*, 136(10), 1234-1241 (2012).

Lartillot O, Toiviainen P. A Matlab toolbox for musical feature extraction from audio. *International Conference on Digital Audio Effects*, 237-244 (2007).

Lederer DJ, Jelic S, Bhattacharya J, Basner RC. Is obstructive sleep apnea a cause of idiopathic pulmonary fibrosis? *Arch Pathol Lab Med*, 136(5), 470; author reply 470 (2012).

Lee AS, Mira-Avendano I, Ryu JH, Daniels CE. The burden of idiopathic pulmonary fibrosis: an unmet public health need. *Respir Med*, 108(7), 955-967 (2014).

Lee JS, Collard HR, Anstrom KJ *et al.* Anti-acid treatment and disease progression in idiopathic pulmonary fibrosis: an analysis of data from three randomised controlled trials. *The Lancet Respiratory Medicine*, 1(5), 369-376 (2013).

Lee JS, Ryu JH, Elicker BM *et al.* Gastroesophageal reflux therapy is associated with longer survival in patients with idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 184(12), 1390-1394 (2011).

Lee JS, Song JW, Wolters PJ *et al.* Bronchoalveolar lavage pepsin in acute exacerbation of idiopathic pulmonary fibrosis. *The European respiratory journal*, 39(2), 352-358 (2012).

Ley B, Bradford WZ, Vittinghoff E, Weycker D, du Bois RM, Collard HR. Predictors of Mortality Poorly Predict Common Measures of Disease Progression in Idiopathic Pulmonary Fibrosis. *American journal of respiratory and critical care medicine*, (2016).

Ley B, Bradford WZ, Weycker D, Vittinghoff E, du Bois RM, Collard HR. Unified baseline and longitudinal mortality prediction in idiopathic pulmonary fibrosis. *The European respiratory journal*, 45(5), 1374-1381 (2015).

Ley B, Collard HR, King TE, Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 183(4), 431-440 (2011).

Ley B, Elicker BM, Hartman TE *et al.* Idiopathic pulmonary fibrosis: CT and risk of death. *Radiology*, 273(2), 570-579 (2014).

Ley B, Ryerson CJ, Vittinghoff E *et al.* A multidimensional index and staging system for idiopathic pulmonary fibrosis. *Ann Intern Med*, 156(10), 684-691 (2012).

Li X, Peng S, Wei L, Li Z. Relevance analysis of clinical and lung function parameters changing and prognosis of idiopathic pulmonary fibrosis. *International journal of clinical and experimental medicine*, 7(12), 4759-4769 (2014).

Liu QX, Guan WJ, Xie YQ *et al.* Vibration response imaging in idiopathic pulmonary fibrosis: a pilot study. *Respir Care*, 59(7), 1071-1077 (2014).

Loeh B, Drakopanagiotakis F, Bandelli GP *et al.* Intraindividual response to treatment with pirfenidone in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 191(1), 110-113 (2015).

Lota HK, Wells AU. The evolving pharmacotherapy of pulmonary fibrosis. *Expert opinion on pharmacotherapy*, 14(1), 79-89 (2013).

Lynch DA, Godwin JD, Safrin S *et al.* High-resolution computed tomography in idiopathic pulmonary fibrosis: diagnosis and prognosis. *American journal of respiratory and critical care medicine*, 172(4), 488-493 (2005).

MacIntyre N, Crapo RO, Viegi G *et al.* Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *European Respiratory Journal*, 26(4), 720-735 (2005).

Maharaj S, Shimbori C, Kolb M. Fibrocytes in pulmonary fibrosis: a brief synopsis. *Eur Respir Rev*, 22(130), 552-557 (2013).

Maher T, Cottin V, Skoeld M *et al.* S11 Pirfenidone Post-authorisation Safety Registry (passport)–interim Analysis Of Ipf Treatment. *Thorax*, 69(Suppl 2), A8-A9 (2014).

Maher TM. PROFILEing idiopathic pulmonary fibrosis: rethinking biomarker discovery. *Eur Respir Rev*, 22(128), 148-152 (2013).

Maldonado F, Moua T, Rajagopalan S *et al.* Automated quantification of radiologic patterns predicts survival in idiopathic pulmonary fibrosis. *The European respiratory journal*, 43 (2014a).

Maldonado F, Moua T, Rajagopalan S *et al.* Automated quantification of radiological patterns predicts survival in idiopathic pulmonary fibrosis. *The European respiratory journal*, 43(1), 204-212 (2014b).

Manali ED, Stathopoulos GT, Kollintza A *et al.* The Medical Research Council chronic dyspnea score predicts the survival of patients with idiopathic pulmonary fibrosis. *Respir Med*, 102(4), 586-592 (2008).

Mannino DM, Etzel RA, Parrish RG. Pulmonary fibrosis deaths in the United States, 1979-1991. An analysis of multiple-cause mortality data. *American journal of respiratory and critical care medicine*, 153(5), 1548-1552 (1996).

Maoua M, El Maalel O, Abdelghani A *et al.* [Impact of COPD on quality of life and mental health among one hundred Tunisian patients]. *Revue de pneumologie clinique*, 70(4), 195-202 (2014).

Marques A, Bruton A, Barney A. The reliability of lung crackle characteristics in cystic fibrosis and bronchiectasis patients in a clinical setting. *Physiol Meas*, 30(9), 903-912 (2009).

Marques ASPD, Oliveira ALA, Jácome CIO. Computerized adventitious respiratory sounds as outcome measures for respiratory therapy: a systematic review. *Respiratory Care*, (2013).

Martinez F, al e. The clinical course of patients with idiopathic pulmonary fibrosis. *Ann Intern Med*, 142, 963-967 (2005).

Martinez FJ, Chisholm A, Collard HR *et al.* The diagnosis of idiopathic pulmonary fibrosis: current and future approaches. *Lancet Respir Med*, 5(1), 61-71 (2017).

Martinez FJ, de Andrade JA, Anstrom KJ, King TE, Jr., Raghu G. Randomized trial of acetylcysteine in idiopathic pulmonary fibrosis. *N Engl J Med*, 370(22), 2093-2101 (2014).

Matthay MA, Ware LB, Zimmerman GA. The acute respiratory distress syndrome. *J Clin Invest*, 122(8), 2731-2740 (2012).

Milger K, Kneidinger N, Neurohr C, Reichenberger F, Behr J. Switching to nintedanib after discontinuation of pirfenidone due to adverse events in IPF. *The European respiratory journal*, 46(4), 1217-1221 (2015).

Miller MR, Hankinson J, Brusasco V *et al.* Standardisation of spirometry. *The European respiratory journal*, 26(2), 319-338 (2005).

Mizuno Y, Iwata H, Shirahashi K *et al.* The importance of intraoperative fluid balance for the prevention of postoperative acute exacerbation of idiopathic pulmonary fibrosis after pulmonary resection for primary lung cancer. *Eur J Cardiothorac Surg*, 41(6), e161-165 (2012).

Moeller A, Ask K, Warburton D, Gauldie J, Kolb M. The bleomycin animal model: a useful tool to investigate treatment options for idiopathic pulmonary fibrosis? *Int J Biochem Cell Biol*, 40(3), 362-382 (2008).

Moeller A, Gilpin SE, Ask K *et al.* Circulating fibrocytes are an indicator of poor prognosis in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 179(7), 588-594 (2009).

Molyneaux PL, Cox MJ, Willis-Owen SA *et al.* The role of bacteria in the pathogenesis and progression of idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 190(8), 906-913 (2014).

Moore BB, Lawson WE, Oury TD, Sisson TH, Raghavendran K, Hogaboam CM. Animal models of fibrotic lung disease. *Am J Respir Cell Mol Biol*, 49(2), 167-179 (2013).

Mura M, Porretta MA, Bargagli E *et al.* Predicting survival in newly diagnosed idiopathic pulmonary fibrosis: a 3-year prospective study. *The European respiratory journal*, 40(1), 101-109 (2012).

Murphy RL Jr HS, Knowler WC. Visual lung-sound characterization by time expanded waveform analysis. *N Engl J Med*, (296), 968-971 (1977).

Murphy RL, Jr., Del Bono EA, Davidson F. Validation of an automatic crackle (rale) counter. *Am Rev Respir Dis*, 140(4), 1017-1020 (1989).

Murphy RL VA, Power-Charnitsky VA, Bana DS, Marinelli PM, Wong-Tse A, Paciej R. Automated lung sound analysis in patients with pneumonia. *Resp Care*, (49(12)), 1490-1497 (2004).

Naik PK, Bozyk PD, Bentley JK *et al.* Periostin promotes fibrosis and predicts progression in patients with idiopathic pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol*, 303(12), L1046-1056 (2012).

Nathan SD, Meyer KC. IPF clinical trial design and endpoints. *Curr Opin Pulm Med*, 20(5), 463-471 (2014).

Nathan SD, Shlobin OA, Weir N *et al.* Long-term course and prognosis of idiopathic pulmonary fibrosis in the new millennium. *Chest*, 140(1), 221-229 (2011).

Natsuizaka M, Chiba H, Kuronuma K *et al.* Epidemiologic survey of Japanese patients with idiopathic pulmonary fibrosis and investigation of ethnic differences. *Am J Respir Crit Care Med*, 190(7), 773-779 (2014).

Navaratnam V, Fleming KM, West J *et al.* The rising incidence of idiopathic pulmonary fibrosis in the U.K. *Thorax*, 66(6), 462-467 (2011).

Nicholson AG, Colby TV, du Bois RM, Hansell DM, Wells AU. The prognostic significance of the histologic pattern of interstitial pneumonia in patients presenting with the clinical entity of cryptogenic fibrosing alveolitis. *American journal of respiratory and critical care medicine*, 162(6), 2213-2217 (2000).

Nishiyama O, Taniguchi H, Kondoh Y *et al.* A simple assessment of dyspnoea as a prognostic indicator in idiopathic pulmonary fibrosis. *Eur Respir J*, 36(5), 1067-1072 (2010).

Noble PW, Albera C, Bradford WZ *et al.* Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. *Lancet*, 377(9779), 1760-1769 (2011a).

Noble PW, Albera C, Bradford WZ *et al.* Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. *The Lancet*, 377(9779), 1760-1769 (2011b).

Noth I, Anstrom KJ, Calvert SB *et al.* A placebo-controlled randomized trial of warfarin in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 186(1), 88-95 (2012).

Noth I, Zhang Y, Ma SF *et al.* Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med*, 1(4), 309-317 (2013).

O'Riordan TG, Smith V, Raghu G. Development of Novel Agents for Idiopathic Pulmonary Fibrosis: Progress in Target Selection and Clinical Trial Design. *Chest*, 148(4), 1083-1092 (2015).

Oda K, Ishimoto H, Yatera K *et al.* High-resolution CT scoring system-based grading scale predicts the clinical outcomes in patients with idiopathic pulmonary fibrosis. *Respiratory Research*, 15(1), 1-9 (2014a).

Oda K, Ishimoto H, Yatera K *et al.* High-resolution CT scoring system-based grading scale predicts the clinical outcomes in patients with idiopathic pulmonary fibrosis. *Respir Res*, 15, 10 (2014b).

Ogura T, Taniguchi H, Azuma A *et al.* Safety and pharmacokinetics of nintedanib and pirfenidone in idiopathic pulmonary fibrosis. *The European respiratory journal*, 45(5), 1382-1392 (2015).

Ohno S, Nakaya T, Bando M, Sugiyama Y. Idiopathic pulmonary fibrosis--results from a Japanese nationwide epidemiological survey using individual clinical records. *Respirology*, 13(6), 926-928 (2008).

Ohshimo S, Ishikawa N, Horimasu Y *et al.* Baseline KL-6 predicts increased risk for acute exacerbation of idiopathic pulmonary fibrosis. *Respir Med*, 108(7), 1031-1039 (2014).

Ohshimo S, Sadamori T, Tanigawa K. Innovation in Analysis of Respiratory Sounds. *Ann Intern Med*, (2016).

Oishi K, Mimura-Kimura Y, Miyasho T *et al.* Association between cytokine removal by polymyxin B hemoperfusion and improved pulmonary oxygenation in patients with acute exacerbation of idiopathic pulmonary fibrosis. *Cytokine*, 61(1), 84-89 (2013).

Okamoto M, Hoshino T, Kitasato Y *et al.* Periostin, a matrix protein, is a novel biomarker for idiopathic interstitial pneumonias. *The European respiratory journal*, 37(5), 1119-1127 (2011).

Okuda R, Hagiwara E, Baba T, Kitamura H, Kato T, Ogura T. Safety and efficacy of pirfenidone in idiopathic pulmonary fibrosis in clinical practice. *Respir Med*, 107(9), 1431-1437 (2013).

Oldham JM, Ma S-F, Martinez FJ *et al.* TOLLIP, MUC5B, and the Response to N-Acetylcysteine among Individuals with Idiopathic Pulmonary Fibrosis. *American journal of respiratory and critical care medicine*, 192(12), 1475-1482 (2015).

Olson AL, Swigris JJ, Lezotte DC, Norris JM, Wilson CG, Brown KK. Mortality from pulmonary fibrosis increased in the United States from 1992 to 2003. *American journal of respiratory and critical care medicine*, 176(3), 277-284 (2007).

Olson J, Colby TV, Elliott CG. Hamman-Rich syndrome revisited. *Mayo Clin Proc*, 65(12), 1538-1548 (1990).

Oltmanns U, Kahn N, Palmowski K *et al.* Pirfenidone in idiopathic pulmonary fibrosis: real-life experience from a German tertiary referral center for interstitial lung diseases. *Respiration; international review of thoracic diseases*, 88(3), 199-207 (2014).

Pajares V, Puzo C, Castillo D *et al.* Diagnostic yield of transbronchial cryobiopsy in interstitial lung disease: A randomized trial. *Respirology*, 19(6), 900-906 (2014).

Park JH, Kim DK, Kim DS *et al.* Mortality and risk factors for surgical lung biopsy in patients with idiopathic interstitial pneumonia. *Eur J Cardiothorac Surg*, 31(6), 1115-1119 (2007).

Pasterkamp H, Brand PLP, Everard M, Garcia-Marcos L, Melbye H, Priftis KN. Towards the standardisation of lung sound nomenclature. *European Respiratory Journal*, 47(3), 724-732 (2016).

Pasterkamp H, Patel S, Wodicka GR. Asymmetry of respiratory sounds and thoracic transmission. *Med Biol Eng Comput*, 35(2), 103-106 (1997).

Peljto AL, Zhang Y, Fingerlin TE *et al.* Association between the MUC5B promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *Jama*, 309(21), 2232-2239 (2013).

Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*, 27(2), 157-172; discussion 207-112 (2008).

Petrosyan F, Culver DA, Reddy AJ. Role of bronchoalveolar lavage in the diagnosis of acute exacerbations of idiopathic pulmonary fibrosis: a retrospective study. *BMC Pulm Med*, 15, 70 (2015).

Piirilä P. Changes in crackle characteristics during the clinical course of pneumonia. *Chest*, 102(1), 176-183 (1992).

Piirila P, Sovijarvi AR. Crackles: recording, analysis and clinical significance. *Eur Respir J*, 8(12), 2139-2148 (1995).

Piirila P, Sovijarvi AR, Kaisla T, Rajala HM, Katila T. Crackles in patients with fibrosing alveolitis, bronchiectasis, COPD, and heart failure. *Chest*, 99(5), 1076-1083 (1991).

Ponte DF, Moraes R, Hizume DC, Alencar AM. Characterization of crackles from patients with fibrosis, heart failure and pneumonia. *Medical engineering & physics*, 35(4), 448-456 (2013).

Prasse A, Probst C, Bargagli E *et al.* Serum CC-chemokine ligand 18 concentration predicts outcome in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 179(8), 717-723 (2009).

Putman RK, Rosas IO, Hunninghake GM. Genetics and early detection in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 189(7), 770-778 (2014).

Raghu G, Anstrom KJ, King TE, Lasky JA, Martinez FJ. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. *N Engl J Med*, 366 (2012a).

Raghu G, Behr J, Brown KK *et al.* Treatment of idiopathic pulmonary fibrosis with ambrisentan: a parallel, randomized trial. *Ann Intern Med*, 158(9), 641-649 (2013a).

Raghu G, Chen SY, Yeh WS *et al.* Idiopathic pulmonary fibrosis in US Medicare beneficiaries aged 65 years and older: incidence, prevalence, and survival, 2001-11. *Lancet Respir Med*, 2(7), 566-572 (2014a).

Raghu G, Collard HR, Anstrom KJ *et al.* Idiopathic pulmonary fibrosis: clinically meaningful primary endpoints in phase 3 clinical trials. *American journal of respiratory and critical care medicine*, 185(10), 1044-1048 (2012b).

Raghu G, Collard HR, Egan JJ *et al.* An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *American journal of respiratory and critical care medicine*, 183(6), 788-824 (2011).

Raghu G, Lynch D, Godwin JD *et al.* Diagnosis of idiopathic pulmonary fibrosis with high-resolution CT in patients with little or no radiological evidence of honeycombing: secondary analysis of a randomised, controlled trial. *The Lancet Respiratory Medicine*, 2(4), 277-284 (2014b).

Raghu G, Million-Rousseau R, Morganti A, Perchenet L, Behr J, Group MS. Macitentan for the treatment of idiopathic pulmonary fibrosis: the randomised controlled MUSIC trial. *The European respiratory journal*, 42(6), 1622-1632 (2013b).

Raghu G, Rochwerg B, Zhang Y *et al.* An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline. *American journal of respiratory and critical care medicine*, 192(2), e3-e19 (2015a).

Raghu G, Wells A, Nicholson A *et al.* Consistent effect of Nintedanib on decline in FVC in patients across subgroups based on HRCT diagnostic criteria: results from the INPULSIS trials in IPF. *Abstract ATS*, (2015b).

Raghu G, Wells AU, Nicholson AG *et al.* Effect of Nintedanib in Subgroups of Idiopathic Pulmonary Fibrosis by Diagnostic Criteria. *American journal of respiratory and critical care medicine*, 195(1), 78-85 (2017).

Raghu G, Weycker D, Edelsberg J, Bradford WZ, Oster G. Incidence and Prevalence of Idiopathic Pulmonary Fibrosis. *American journal of respiratory and critical care medicine*, 174(7), 810-816 (2006).

Reichert S, Gass R, Brandt C, Andres E. Analysis of respiratory sounds: state of the art. *Clin Med Circ Respirat Pulm Med*, 2, 45-58 (2008).

Reichmann WM, Yu YF, Macaulay D, Wu EQ, Nathan SD. Change in forced vital capacity and associated subsequent outcomes in patients with newly diagnosed idiopathic pulmonary fibrosis. *BMC Pulm Med*, 15(1), 167 (2015).

Richards TJ, Kaminski N, Baribaud F *et al.* Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 185(1), 67-76 (2012).

Richeldi L, Costabel U, Selman M *et al.* Efficacy of a Tyrosine Kinase Inhibitor in Idiopathic Pulmonary Fibrosis. *New England Journal of Medicine*, 365(12), 1079-1087 (2011).

Richeldi L, Davies HR, Ferrara G, Franco F. Corticosteroids for idiopathic pulmonary fibrosis. *Cochrane Database Syst Rev*, (3), CD002880. Edited; published in Issue 002882, 002010 (2003).

Richeldi L, Davies HR, Spagnolo P, Luppi F. Corticosteroids for idiopathic pulmonary fibrosis. *Cochrane Database Syst Rev*, (2), CD002880 (2010).

Richeldi L, du Bois RM, Raghu G *et al.* Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med*, 370(22), 2071-2082 (2014).

Richeldi L, Ryerson CJ, Lee JS *et al.* Relative versus absolute change in forced vital capacity in idiopathic pulmonary fibrosis. *Thorax*, 67(5), 407-411 (2012).

Rosas IO, Richards TJ, Konishi K *et al.* MMP1 and MMP7 as potential peripheral blood biomarkers in idiopathic pulmonary fibrosis. *PLoS Med*, 5(4), e93 (2008).

Rush B, Wiskar K, Berger L, Griesdale D. The use of mechanical ventilation in patients with idiopathic pulmonary fibrosis in the United States: A nationwide retrospective cohort analysis. *Respir Med*, 111, 72-76 (2016).

Ryerson CJ, Arean PA, Berkeley J *et al.* Depression is a common and chronic comorbidity in patients with interstitial lung disease. *Respirology*, 17(3), 525-532 (2012).

Ryerson CJ, Berkeley J, Carrieri-Kohlman VL, Pantilat SZ, Landefeld CS, Collard HR. Depression and functional status are strongly associated with dyspnea in interstitial lung disease. *Chest*, 139(3), 609-616 (2011).

Ryerson CJ, Collard HR. Acute exacerbations complicating interstitial lung disease. *Curr Opin Pulm Med*, 20(5), 436-441 (2014).

Ryerson CJ, Corte TJ, Lee JS *et al.* A Standardized Diagnostic Ontology for Fibrotic Interstitial Lung Disease: An International Working Group Perspective. *American journal of respiratory and critical care medicine*, (2017).

Ryerson CJ, Cottin V, Brown KK, Collard HR. Acute exacerbation of idiopathic pulmonary fibrosis: shifting the paradigm. *Eur Respir J*, 46(2), 512-520 (2015).

Ryerson CJ, Urbania TH, Richeldi L *et al.* Prevalence and prognosis of unclassifiable interstitial lung disease. *The European respiratory journal*, 42(3), 750-757 (2013).

Sackett DL, Haynes RB, Tugwell P. *Clinical Epidemiology: A Basic Science for Clinical Medicine* (Little, Brown, 1985).

Saini G, Porte J, Weinreb PH *et al.* alphavbeta6 integrin may be a potential prognostic biomarker in interstitial lung disease. *The European respiratory journal*, 46(2), 486-494 (2015).

Sakamoto K, Taniguchi H, Kondoh Y *et al.* Acute exacerbation of IPF following diagnostic bronchoalveolar lavage procedures. *Respir Med*, 106(3), 436-442 (2012).

Sakamoto S, Homma S, Miyamoto A, Kurosaki A, Fujii T, Yoshimura K. Cyclosporin A in the treatment of acute exacerbation of idiopathic pulmonary fibrosis. *Intern Med*, 49(2), 109-115 (2010).

Sakamoto S, Homma S, Mun M, Fujii T, Kurosaki A, Yoshimura K. Acute exacerbation of idiopathic interstitial pneumonia following lung surgery in 3 of 68 consecutive patients: a retrospective study. *Intern Med*, 50(2), 77-85 (2011).

Salisbury ML, Xia M, Murray S *et al.* Predictors of idiopathic pulmonary fibrosis in absence of radiologic honeycombing: A cross sectional analysis in ILD patients undergoing lung tissue sampling. *Respir Med*, 118, 88-95 (2016).

Samejima J, Tajiri M, Ogura T *et al.* Thoracoscopic lung biopsy in 285 patients with diffuse pulmonary disease. *Asian Cardiovasc Thorac Ann*, 23(2), 191-197 (2015).

Santos GC, Parra ER, Stegun FW, Cirqueira CS, Capelozzi VL. Immunohistochemical detection of virus through its nuclear cytopathic effect in idiopathic interstitial pneumonia other than acute exacerbation. *Braz J Med Biol Res*, 46(11), 985-992 (2013).

Satoh H, Kurishima K, Ishikawa H, Ohtsuka M. Increased levels of KL-6 and subsequent mortality in patients with interstitial lung diseases. *J Intern Med*, 260(5), 429-434 (2006).

Schmidt SL, Tayob N, Han MK *et al.* Predicting Pulmonary Fibrosis Disease Course From Past Trends in Pulmonary Function. *Chest*, 145(3), 579-585 (2014).

Schunemann HJ, Jaeschke R, Cook DJ *et al.* An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am J Respir Crit Care Med*, 174(5), 605-614 (2006).

Schupp JC, Binder H, Jager B *et al.* Macrophage activation in acute exacerbation of idiopathic pulmonary fibrosis. *PLoS One*, 10(1), e0116775 (2015).

Seibold MA, Wise AL, Speer MC *et al.* A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med*, 364(16), 1503-1512 (2011).

Sellarés J, Hernández-González F, Lucena CM *et al.* Auscultation of Velcro Crackles is Associated With Usual Interstitial Pneumonia. *Medicine*, 95(5), e2573 (2016).

Selman M, Carrillo G, Estrada A *et al.* Accelerated variant of idiopathic pulmonary fibrosis: clinical behavior and gene expression pattern. *PLoS One*, 2(5), e482 (2007).

Seo Y, Abe S, Kurahara M *et al.* Beneficial effect of polymyxin B-immobilized fiber column (PMX) hemoperfusion treatment on acute exacerbation of idiopathic pulmonary fibrosis. *Intern Med*, 45(18), 1033-1038 (2006).

Shulgina L, Cahn AP, Chilvers ER *et al.* Treating idiopathic pulmonary fibrosis with the addition of co-trimoxazole: a randomised controlled trial. *Thorax*, 68(2), 155-162 (2013).

Simon-Blancal V, Freynet O, Nunes H *et al.* Acute exacerbation of idiopathic pulmonary fibrosis: outcome and prognostic factors. *Respiration*, 83(1), 28-35 (2012).

Song JW, Do KH, Jang SJ, Colby TV, Han S, Kim DS. Blood biomarkers MMP-7 and SP-A: predictors of outcome in idiopathic pulmonary fibrosis. *Chest*, 143(5), 1422-1429 (2013).

Song JW, Hong SB, Lim CM, Koh Y, Kim DS. Acute exacerbation of idiopathic pulmonary fibrosis: incidence, risk factors and outcome. *Eur Respir J*, 37(2), 356-363 (2011).

Sovijärvi ARA, Vanderschoot J, Earis JE. *Computerized respiratory sound analysis (CORSA) : recommended standards for terms and techniques : ERS Task Force Report* (Munksgaard, Copenhagen, 2000).

Spagnolo P, Bonella F, Maher TM. New guideline on treatment of idiopathic pulmonary fibrosis. *Lancet Respir Med*, 3(9), e31-32 (2015a).

Spagnolo P, Del Giovane C, Luppi F *et al.* Non-steroid agents for idiopathic pulmonary fibrosis. *Cochrane Database Syst Rev*, (9), CD003134 (2010).

Spagnolo P, du Bois RM, Cottin V. Rare lung disease and orphan drug development. *Lancet Respir Med*, 1(6), 479-487 (2013).

Spagnolo P, Maher TM, Richeldi L. Idiopathic pulmonary fibrosis: Recent advances on pharmacological therapy. *Pharmacology & therapeutics*, 152, 18-27 (2015b).

Spagnolo P, Wells AU, Collard HR. Pharmacological treatment of idiopathic pulmonary fibrosis: an update. *Drug Discovery Today*, 20(5), 514-524 (2015c).

Steele MP, Schwartz DA. Molecular mechanisms in progressive idiopathic pulmonary fibrosis. *Annual review of medicine*, 64, 265-276 (2013).

Sugino K, Nakamura Y, Ito T, Isshiki T, Sakamoto S, Homma S. Comparison of clinical characteristics and outcomes between combined pulmonary fibrosis and emphysema associated with usual interstitial pneumonia pattern and non-usual interstitial pneumonia. *Sarcoidosis Vasc Diffuse Lung Dis*, 32(2), 129-137 (2015).

Sugiura H, Takeda A, Hoshi T *et al.* Acute exacerbation of usual interstitial pneumonia after resection of lung cancer. *Ann Thorac Surg*, 93(3), 937-943 (2012).

Sumikawa H, Johkoh T, Colby TV *et al.* Computed tomography findings in pathological usual interstitial pneumonia: relationship to survival. *Am J Respir Crit Care Med*, 177(4), 433-439 (2008).

Sundaram B, Gross BH, Martinez FJ *et al.* Accuracy of High-Resolution CT in the Diagnosis of Diffuse Lung Disease: Effect of Predominance and Distribution of Findings. *American Journal of Roentgenology*, 191(4), 1032-1039 (2008).

Suzuki H, Sekine Y, Yoshida S *et al.* Risk of acute exacerbation of interstitial pneumonia after pulmonary resection for lung cancer in patients with idiopathic pulmonary fibrosis based on preoperative high-resolution computed tomography. *Surg Today*, 41(7), 914-921 (2011).

Swigris J, Han M, Vij R *et al.* The UCSD shortness of breath questionnaire has longitudinal construct validity in idiopathic pulmonary fibrosis. *Respir Med*, 106(10), 1447-1455 (2012a).

Swigris JJ, Brown KK, Behr J *et al.* The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. *Respiratory medicine*, 104(2), 296-304 (2010).

Swigris JJ, Olson AL, Huie TJ *et al.* Ethnic and racial differences in the presence of idiopathic pulmonary fibrosis at death. *Respir Med*, 106(4), 588-593 (2012b).

Tachibana K, Inoue Y, Nishiyama A *et al.* Polymyxin-B hemoperfusion for acute exacerbation of idiopathic pulmonary fibrosis: serum IL-7 as a prognostic marker. *Sarcoidosis Vasc Diffuse Lung Dis*, 28(2), 113-122 (2011).

Tachikawa R, Tomii K, Ueda H *et al.* Clinical features and outcome of acute exacerbation of interstitial pneumonia: collagen vascular diseases-related versus idiopathic. *Respiration*, 83(1), 20-27 (2012).

Tajiri M, Okamoto M, Fujimoto K *et al.* Serum level of periostin can predict long-term outcome of idiopathic pulmonary fibrosis. *Respir Investig*, 53(2), 73-81 (2015).

Taniguchi H, Ebina M, Kondoh Y *et al.* Pirfenidone in idiopathic pulmonary fibrosis. *The European respiratory journal*, 35(4), 821-829 (2010).

Tcherakian C, Cottin V, Brillet PY *et al.* Progression of idiopathic pulmonary fibrosis: lessons from asymmetrical disease. *Thorax*, 66(3), 226-231 (2011).

Thabut G, Crestani B, Porcher R, Richeldi L. Missing data in IPF trials: do not let methodological issues undermine a major therapeutic breakthrough. *The European respiratory journal*, 46(3), 607-614 (2015).

Thomeer M, Demedts M, Vandeurzen K, Diseases VWGoIL. Registration of interstitial lung diseases by 20 centres of respiratory medicine in Flanders. *Acta Clin Belg*, 56(3), 163-172 (2001).

Tinelli C, De Silvestri A, Richeldi L, Oggionni T. The Italian register for diffuse infiltrative lung disorders (RIPID): a four-year report. *Sarcoidosis, vasculitis, and diffuse lung diseases : official journal of WASOG / World Association of Sarcoidosis and Other Granulomatous Disorders*, 22 Suppl 1, S4-8 (2005).

Tomassetti S, Cavazza A, Colby TV *et al.* Transbronchial biopsy is useful in predicting UIP pattern. *Respiratory Research*, 13(1), 96-96 (2012).

Tomassetti S, Gurioli C, Ryu JH *et al.* The impact of lung cancer on survival of idiopathic pulmonary fibrosis. *Chest*, 147(1), 157-164 (2015).

Tomassetti S, Ruy JH, Gurioli C *et al.* The effect of anticoagulant therapy for idiopathic pulmonary fibrosis in real life practice. *Sarcoidosis Vasc Diffuse Lung Dis*, 30(2), 121-127 (2013).

Tomioka H, Sakurai T, Hashimoto K, Iwasaki H. Acute exacerbation of idiopathic pulmonary fibrosis: role of Chlamydophila pneumoniae infection. *Respirology*, 12(5), 700-706 (2007).

Travis WD, Costabel U, Hansell DM *et al.* An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *American journal of respiratory and critical care medicine*, 188(6), 733-748 (2013).

Tsushima K, Yamaguchi K, Kono Y *et al.* Thrombomodulin for acute exacerbations of idiopathic pulmonary fibrosis: a proof of concept study. *Pulm Pharmacol Ther*, 29(2), 233-240 (2014).

Ushiki A, Yamazaki Y, Hama M, Yasuo M, Hanaoka M, Kubo K. Viral infections in patients with an acute exacerbation of idiopathic interstitial pneumonia. *Respir Investig*, 52(1), 65-70 (2014).

Usui Y, Kaga A, Sakai F *et al.* A cohort study of mortality predictors in patients with acute exacerbation of chronic fibrosing interstitial pneumonia. *BMJ Open*, 3(7) (2013).

Vancheri C, Failla M, Crimi N, Raghu G. Idiopathic pulmonary fibrosis: a disease with similarities and links to cancer biology. *Eur Respir J*, 35(3), 496-504 (2010).

Vannuccini L, Rossi M, Pasquali G. A new method to detect crackles in respiratory sounds. *Technol Health Care*, 6(1), 75-79 (1998).

Vij R, Noth I, Strek ME. Autoimmune-featured interstitial lung disease: A distinct entity. *Chest*, 140(5), 1292-1299 (2011).

von Plessen C, Grinde O, Gulsvik A. Incidence and prevalence of cryptogenic fibrosing alveolitis in a Norwegian community. *Respir Med*, 97(4), 428-435 (2003).

Vuga LJ, Tedrow JR, Pandit KV *et al.* C-X-C motif chemokine 13 (CXCL13) is a prognostic biomarker of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 189(8), 966-974 (2014).

Vyshedskiy A, Alhashem RM, Paciej R *et al.* Mechanism of inspiratory and expiratory crackles. *Chest*, 135(1), 156-164 (2009).

Vyshedskiy A, Bezares F, Paciej R, Ebril M, Shane J, Murphy R. Transmission of crackles in patients with interstitial pulmonary fibrosis, congestive heart failure, and pneumonia. *Chest*, 128(3), 1468-1474 (2005).

Walsh SF, Sverzellati N, Devaraj A, Wells A, Hansell D. Chronic hypersensitivity pneumonitis: high resolution computed tomography patterns and pulmonary function indices as prognostic determinants. *Eur Radiol*, 22(8), 1672-1679 (2012).

Walsh SL, Calandriello L, Sverzellati N, Wells AU, Hansell DM, Consort UIPO. Interobserver agreement for the ATS/ERS/JRS/ALAT criteria for a UIP pattern on CT. *Thorax*, 71(1), 45-51 (2016).

Walsh SL, Sverzellati N, Devaraj A, Keir GJ, Wells AU, Hansell DM. Connective tissue disease related fibrotic lung disease: high resolution computed tomographic and pulmonary function indices as prognostic determinants. *Thorax*, 69(3), 216-222 (2014).

Watadani T, Sakai F, Johkoh T *et al.* Interobserver variability in the CT assessment of honeycombing in the lungs. *Radiology*, 266(3), 936-944 (2013).

Watanabe A, Higami T, Ohori S, Koyanagi T, Nakashima S, Mawatari T. Is lung cancer resection indicated in patients with idiopathic pulmonary fibrosis? *J Thorac Cardiovasc Surg*, 136(5), 1357-1363, 1363 e1351-1352 (2008).

Wells A. Combination therapy in idiopathic pulmonary fibrosis: the way ahead will be hard. *European Respiratory Journal*, 45(5), 1208-1210 (2015).

Wells AU, Desai SR, Rubens MB *et al.* Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. *American journal of respiratory and critical care medicine*, 167(7), 962-969 (2003).

Wootton SC, Kim DS, Kondoh Y *et al.* Viral infection in acute exacerbation of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 183(12), 1698-1702 (2011).

Wuyts WA, Antoniou KM, Borensztajn K *et al.* Combination therapy: the future of management for idiopathic pulmonary fibrosis? *Lancet Respir Med*, 2(11), 933-942 (2014).

Xue J, Kass DJ, Bon J *et al.* Plasma B lymphocyte stimulator and B cell differentiation in idiopathic pulmonary fibrosis patients. *J Immunol*, 191(5), 2089-2095 (2013).

Yoko M, Keishi S, Naoshi K *et al.* Efficacy Of Inhaled N-Acetylcysteine Monotherapy On Lung Function And Redox Balance In Idiopathic Pulmonary Fibrosis. In: *C55. INTERSTITIAL LUNG DISEASE: DIAGNOSTIC AND THERAPEUTIC CONSIDERATIONS.* (American Thoracic Society, 2012) A4589-A4589.

Yokoyama T, Kondoh Y, Taniguchi H *et al.* Noninvasive ventilation in acute exacerbation of idiopathic pulmonary fibrosis. *Intern Med*, 49(15), 1509-1514 (2010).

Yoon RG, Seo JB, Kim N *et al.* Quantitative assessment of change in regional disease patterns on serial HRCT of fibrotic interstitial pneumonia with texture-based automated quantification system. *Eur Radiol*, 23(3), 692-701 (2013).

Yorke J, Jones PW, Swigris JJ. Development and validity testing of an IPF-specific version of the St George's Respiratory Questionnaire. *Thorax*, 65(10), 921-926 (2010).

Zappala CJ, Latsi PI, Nicholson AG *et al.* Marginal decline in forced vital capacity is associated with a poor outcome in idiopathic pulmonary fibrosis. *The European respiratory journal*, 35(4), 830-836 (2010).

Zavaletta VA, Bartholmai BJ, Robb RA. High resolution multidetector CT-aided tissue analysis and quantification of lung fibrosis. *Acad Radiol*, 14(7), 772-787 (2007).