POLITECNICO DI MILANO

FACOLTA' DI INGEGNERIA DEI SISTEMI

Corso di Laurea Magistrale in Ingegneria Biomedica

Analysis of Respiratory Parameters in COPD Patients Monitored at Home by a Novel Telemedicine System

Relatori: Prof. Raffaele Dellacà

              Prof. Marco Parvis

Correlatori: Ing. Alessandro Gobbi, Ph.D.

              Ing. Pasquale Pio Pompilio, Ph.D.

Tesi di Laurea di:

Ivan Cenci   Matr. 751565

Anno Accademico 2012-2013

Contents

Summary

Chronic obstructive pulmonary disease (COPD) represents a major public health problem. It is the fourth leading cause of chronic morbidity and mortality in the United States and by 2020, according to recent studies, it is projected to rank third among the leading causes of death and fifth in economical burden caused worldwide [1–3]. The increased mortality is driven by the expanding epidemic of smoking and the changing demographics in most countries, with an increase in population's life expectancy. Of these two forces, demographics is the stronger driver of the trend. Furthermore, although COPD has received increasing attention from the medical community in recent years, it is still relatively unknown or ignored by the public as well as public health and government officials.

From a pathophysiologic perspective, COPD is characterized by chronic airflow limitation and a range of pathological changes in the lung, some significant extra-pulmonary effects, and important comorbidities which may contribute to the severity of the disease in individual patients. Thus, COPD should be regarded as a pulmonary disease, but these significant comorbidities must be taken into account in a comprehensive diagnostic assessment of severity and in determining appropriate treatment.

The progression of the pathological condition is described by stable periods disrupted by sudden exacerbations of the symptoms during which a severe inflammatory process occurs. Such aggravations recur in a periodic fashion and often require hospitalization. A speculation proposed is that a long-term observation of objective parameters may help to characterize the progression of the disease over time and, possibly, early detect the onset of extreme future events.

Unfortunately, the diagnosis of this syndrome still relies on the measurements of spirometric indices that, given the effort-dependent maneuvers required, are not suitable for self-testing of airway obstruction. Therefore, a long-term and accurate observation of respiratory parameters is still missing.

A different and innovative approach for the characterization of the progression of COPD over time, for the association of clinical symptoms to quantitative parameters and for the early detection of acute events could be based on the measurement by

forced oscillation technique (FOT) of indices related to the mechanical properties of the respiratory system.

Accordingly, the present work aims at advancing knowledge about the pathology by employing a novel commercial telemedicine system based on FOT. To do so, the accuracy of this device is firstly evaluated in order to validate its functionality. Secondly, clinical data is acquired and analyzed thanks to the employment of the device in the largest and most detailed clinical trial ever defined for the observation of COPD. The work is presented in three chapters and their content is summarized as follows.

Chapter 1 - Chronic Obstructive Pulmonary Disease and assessment of lung function

As the present work is based on experimental studies involving COPD patients, the majority of Chapter 1 focuses on the pathophysiologic and epidemiologic characteristics of COPD. This is to provide a comprehensive background for the interpretation of the validation process of Chapter 2 and the data analysis of Chapter 3.

The disease is first defined, then processes characterizing pathophysiology, epidemiology and exacerbations of COPD are discerned in their main components and phases. Finally, clinical methods for the diagnosis of obstructive diseases and the assessment of their temporal variability are discussed. There is now a good understanding of how the underlying disease process in COPD leads to the characteristic physiologic abnormalities and symptoms. For example, decreased FEV_1 primarily results from inflammation and narrowing of peripheral airways, while decreased gas transfer arises from the parenchymal destruction of emphysema. Also, the peak expiratory flow (PEF) is the clinical tool recommended in most modern guidelines as a monitoring instrument for the measurement of variability of symptoms over time [4–6].

Since the just mentioned spirometric indexes are effort dependent and provide data with poor reproducibility when performed without supervision, it is speculated that more meaningful clinical outcomes could be obtained by leaving the spirometric approach and concentrating on the analysis of the temporal fluctuations of reliable and accurate parameters that do not require the execution of forced respiratory maneuvers, as the respiratory mechanical indices measured by FOT.

Chapter 2 - Validation of a novel telemedicine system for chronic respiratory diseases' home monitoring

In Chapter 2 a novel telemedicine system (RESMON® PRO) developed and commercialized by a spin-off society of the Politecnico di Milano (Restech Srl), in collaboration

with the Laboratory of Biomedical Research (TBM Lab) of the Politecnico di Milano, is presented and analyzed. The device detects flow and pressure signals at the airway openings and calculates the value of impedance of the tracheobronchial tree thanks to the application of the FOT. In addition, it assesses many other respiratory parameters and provides additional information about each test performed.

The analysis of the accuracy of this novel FOT device is based on comparisons between measurements performed by it and measurements performed by a reference FOT system for laboratory use which is taken as a gold standard. Through properly designed experimental protocols the accuracy of the macro-components of the device (i.e, front-end hardware and real-time processing algorithm) is investigated. Also, analyses are provided for assessing calibration stability and efficiency of the aspirator needed to expel the exhaled carbon dioxide. Every experimental protocol is supported by the employment of numerical tools of statistical significance.

Results of calibration stability indicate that the front-end maintains the same characteristics over a time span of at least two months, and that calibration is expected to remain valid for much longer periods. The aspirator, on the contrary, does not show very positive performances as it is not sufficient to completely extract the carbon dioxide present in the high-inertance path after exhalation. Therefore an alternative solution able to guarantee a greater flow rate should be taken into account.

Results of accuracy analysis fully validate the signal processing software installed on the device, whereas the accuracy of the front-end hardware cannot be assessed in absolute terms. This is due to the physiologic inter-test variability characterizing every subject tested by the device. In other words, physiologic inter-test variability "masks" the actual difference in accuracy between the two systems used for making comparisons. Several clues suggest that the actual difference in precision is rather small, but this is not directly provable.

Chapter 3 - Home monitoring of COPD and data analysis

Chapter 3 offers an analysis of clinical data acquired by the validated FOT device. The observational clinical study utilized as source of data moves from the speculation that a long-term observation of objective parameters may help to characterize the progression of the COPD over time and, possibly, early detect the onset of extreme future events. The purpose of the clinical trial is to assess daily variability of FOT data measured at home of a group of COPD patients in order to identify possible correlations between symptoms change, breathing pattern, lung mechanical impedance and occurrence of exacerbation. Every day, patients are requested to make a test at the

same time, usually in the morning and two hours after taking medications. It consists in administering a preliminary questionnaire that is a version of the COPD Assessment Test (CAT) purposely adapted to be shown on the screen of the FOT device [99]. The CAT is designed to measure the impact of COPD on a person's life, and how this changes over time. Successively, the actual FOT assessment is performed by providing multifrequency stimuli for 3 minutes. The main outputs of every test are represented by the total score of the CAT-based questionnaire and 13 respiratory parameters. The estimated enrollment of 80 patients and the time frame consisting in daily assessments for 6 to 8 months make this study the largest and most detailed clinical trial ever defined for the observation of COPD.

Methods designed for managing and checking the quality of the large amount of data produced by the clinical trial, such as tools for automatic outlier detection and missing data estimate, provide very positive performances. The user can quickly navigate through the acquisitions of the various patients and effectively visualize and correct acquisition suspected to be outliers. Whenever a test is missing, an algorithm of locally weighted regression estimates values for the 13 respiratory parameters in a way that leaves unaltered the original statistical properties of the time series. Despite the very good functioning of the mentioned methods, possible future improvements are identified and elucidated.

Results of the analyses performed on the parameters, namely a study of significance of the CAT-based questionnaire, the principal component analysis (PCA) and an analysis of variability at different temporal scales, supply interesting insights into the pathology despite the modest amount of patients currently available.

First, results firmly demonstrate that the CAT-based questionnaire cannot be used as a tool for the punctual assessment of the degree of lung obstruction in COPD. Indeed, linear regression ascertains that the questionnaire is not systematically correlated neither with any of the respiratory parameters taken singularly, nor with any of the 8192 different sets drawn from them. This confirms the typical criticism made about questionnaire-based measures consisting in the fact that the gathered information is by its nature subjective.

Second, PCA indicates that two variables, opportunely defined as linear combinations of the original set of 14 parameters, suffice to explain a fair percentage (64.8%) of the total information content represented by the complete set of parameters. Furthermore, the first principal component is very much related to the parameters of respiratory mechanics[1] whereas the second principal component is somewhat aligned with breath-

---

[1] Parameters describing the lung mechanical impedance are referred to as parameters of *respiratory mechanics*

ing pattern parameters[2]. This means that the parameters of respiratory mechanics are the most sensitive to any deterioration or improvement of the health status and that parameters of respiratory mechanics and parameters of breathing pattern might be used to represent different aspects of the pathology.

Third, the variability analysis at different temporal scales illustrates how short-term fluctuations of parameters of respiratory mechanics might constitute a measure of the severity of the disease and, at the same time, represent the effects of pathophysiologic mechanisms on both the central and peripheral airways. In fact, parameters of respiratory mechanics show fluctuations that are minimum in healthy subjects, modest in mild asthmatic patients and very strong in severe COPD patients. Differently, the long-term variability confirms that asthma and COPD are characterized by similar pathological changes in the central airways, whereas inflammatory processes in the peripheral airways and in lung parenchyma are much more pronounced in COPD patients. This is suggested by the fact that fluctuations for temporal scales greater than one day are comparable in asthma and COPD only for parameters of resistance. Parameters of reactance, on the contrary, are characterized by much larger fluctuations in COPD patients.

In conclusion, this work indicates how a home daily monitoring of respiratory mechanics may contribute to advance the understanding of COPD and to improve the clinical management of patients suffering from this disease.

---

[2]Parameters describing respiratory volumes and respiratory frequencies are referred to as parameters of *breathing pattern*

Sommario

La broncopneumopatia cronica ostruttiva (COPD) rappresenta uno dei principali problemi di salute pubblica. Essa è la quarta causa causa di morbilità e mortalità negli USA e, secondo recenti studi, nel 2020 diventerà la terza causa di morte e il quinto onere economico sanitario a livello mondiale [1–3]. L'aumentata mortalità è guidata dalla sempre maggiore diffusione del fumo di tabacco a da cambiamenti nella demografia di molti paesi, con associato un incremento dell'aspettativa di vita delle popolazioni. Tra questi due aspetti, la demografia è la causa più forte di questo trend. In aggiunta, sebbene la COPD ha ricevuto sempre maggiore attenzione da parte della comunità medica negli anni recenti, questa è ancora relativamente sconosciuta o ignorata sia dalla sanità pubblica che da quella privata.

Da un punto di vista patofisiologico, la COPD è caratterizzata da una limitazione cronica del flusso respiratorio e una serie di cambiamenti patologici del tessuto polmonare, alcuni significativi effetti extra-polmonari e importanti comorbilità che possono contribuire alla gravità della malattia. Perciò la COPD dovrebbe essere considerata una malattia polmonare con comorbilità che devono essere tenute in considerazione in una valutazione diagnostica globale della sua gravità, e nella determinazione di un trattamento appropriato

L'evoluzione della condizione patologica è descritta da periodi stabili interrotti da improvvise riacutizzazioni dei sintomi durante le quali si hanno seri processi infiammatori. Queste ricadute ricorrono periodicamente e spesso richiedono una ospedalizzazione. Una speculazione proposta indica che una osservazione a lungo-termine di parametri oggettivi potrebbe aiutare a caratterizzare la progressione temporale della malattia e, possibilmente, rilevare con il dovuto anticipo l'insorgenza di eventi estremi.

Sfortunatamente la diagnosi di questa sindrome fa ancora affidamento sulla misurazione di indici spirometrici che, date le manovre sforzo-dipendenti richieste, non sono adatti ad un testing non supervisionato del livello di ostruzione delle vie aeree. Di conseguenza un'accurata osservazione a lungo-termine dei parametri respiratori non è ancora disponibile nella pratica clinica.

Un approccio innovativo per la caratterizzazione della progressione temporale della

COPD, l'associazione dei sintomi clinici con dei parametri quantitativi e l'individuazione anticipata dell'insorgenza di eventi acuti potrebbe essere basato sulla misurazione di indici legati alle proprietà meccaniche del sistema respiratorio tramite la tecnica delle oscillazioni forzate (FOT).

Lo scopo di questo lavoro è incrementare la conoscenza della patologia tramite l'utilizzo di un nuovo sistema commerciale di telemedicina basato sulla FOT. A questo fine l'accuratezza del dispositivo è inizialmente valutata e validata. In seconda istanza dati clinici sono acquisiti e analizzati grazie all'impiego del sistema nel più grande e più dettagliato studio clinico mai definito per l'osservazione della COPD. La tesi è divisa in tre capitoli il cui contenuto è riassunto nelle seguenti Sezioni.

Chapter 1 - Chronic Obstructive Pulmonary Disease and assessment of lung function

Siccome questo lavoro è basato su studi sperimentali riguardanti pazienti COPD, la maggior parte del Capitolo 1 è focalizzato sulle caratteristiche patofisiologiche ed epidemiologiche della COPD. In questo modo è possibile fornire un ampio background per l'interpretazione del processo di validazione affrontato nel Capitolo 2 e l'analisi dati del Capitolo 3.

La malattia è inizialmente definita. Successivamente, i processi che caratterizzano patofisiologia, epidemiologia e riacutizzazioni della COPD sono distinti nelle loro fasi e componenti principali. Infine, metodi clinici per la diagnosi delle malattie ostruttive e per la valutazione della loro variabilità temporale sono discussi. Attualmente si ha una buona comprensione del modo in cui i processi fondanti la COPD portano ai caratteristici sintomi e anormalità fisiologiche. Per esempio, un ridotto FEV_1 risulta causato principalmente da una infiammazione e un restringimento delle vie aeree periferiche, mentre un ridotto scambio gassoso è generato dalla distruzione del parenchima per via dell'enfisema. In aggiunta, il picco di flusso espiratorio (PEF) è il tool clinico maggiormente raccomandato nelle moderne linee guida come strumento di monitoraggio della variabilità temporale dei sintomi [4–6].

Poichè gli indici spirometrici appena introdotti sono sforzo-dipendenti e forniscono dati con una povera riproducibilità quando questi sono eseguiti senza supervisione, si specula che risultati clinici più significativi sarebbero ottenibili tramite la sostituzione dell'approccio spirometrico con l'analisi delle fluttuazioni temporali di parametri affidabili e accurati che non richiedono l'esecuzione di manovre respiratorie forzate, come gli indici di meccanica respiratoria ottenuti tramite FOT.

Chapter 2 - Validation of a novel telemedicine system for chronic respiratory diseases' home monitoring

Nel Capitolo 2 un nuovo sistema di telemedicina (RESMON® PRO) sviluppato e commercializzato da una società spin-off del Politecnico di Milano (Restech Srl), in collaborazione con il Laboratorio di Tecnologie Biomediche (TBM Lab) del Politecnico di Milano, è presentato e analizzato. Il dispositivo rileva segnali di flusso e pressione all'apertura delle vie aeree e calcola il valore di impedenza dell'albero tracheobronchiale grazie all'applicazione della FOT. In aggiunta, esso valuta molti altri parametri respiratory e fornisce informazioni aggiuntive su ogni test effettuato.

L'analisi di accuratezza di questo dispositivo è basata su confronti tra misure effettuate da esso e misure effettuate da un sistema FOT di riferimento per uso di laboratorio che è considerato come gold standard. Tramite protocolli sperimentali opportunamente definiti, l'accuratezza dei macro-componenti del dispositivo (front-end hardware e algoritmo di elaborazione real-time) è valutata. Inoltre si propongono analisi per stimare la stabilità della calibrazione e l'efficienza dell'aspiratore che ha il compito di espellere il diossido di carbonio esalato. Ogni protocollo sperimentale è supportato dall'applicazione di strumenti numerici di significatività statistica.

I risultati di stabilità della calibrazione indicano che il front-end mantiene le stesse caratteristiche per un lasso di tempo di almeno due mesi, e che verosimilmente la calibrazione rimanga valida per periodi molto più lunghi. L'aspiratore, al contrario, non mostra prestazioni molto positive in quanto non è sufficiente ad eliminare completamente il diossido di carbonio presente a fine espirazione nel tubo ad alta inertanza. Di conseguenza una soluzione alternativa in grado di garantire una portata maggiore dovrebbe essere presa in considerazione.

I risultati dell'analisi di accuratezza validano pienamente il software di elaborazione dei segnali installato sul dispositivo, mentre l'accuratezza del front-end hardware non può essere quantificata in termini assoluti. Ciò è dovuto alla variabilità fisiologica inter-test caratterizzante tutti i soggetti acquisiti tramite il dispositivo. In altre parole, la variabilità fisiologica inter-test "maschera" la vera differenza in accuratezza tra i due sistemi usati per fare confronti. Diversi indizi portano ad affermare che la vera differenza in precisione è piuttosto piccola, ma questo non è direttamente provabile.

Chapter 3 - Home monitoring of COPD and data analysis

Il Capitolo 3 offre una analisi di dati clinici acquisiti tramite il dispositivo FOT validato. Lo studio clinico osservazionale usato come sorgente di dati nasce dalla speculazione che un'osservazione a lungo-termine di parametri oggettivi potrebbe aiutare nel carat-

terizzare la progressione temporale della COPD e, possibilmente, rilevare con anticipo l'insorgenza di futuri eventi estremi. Lo scopo dello studio clinico è quello di misurare la variabilità giornaliera di dati FOT acquisiti a casa di un gruppo di pazienti COPD al fine di identificare possibili correlazioni tra cambiamenti nei sintomi, pattern respiratorio, impedenza meccanica polmonare e occorrenza delle riacutizzazioni. Ogni giorno, ai pazienti è richiesto di effettuare un test alla stessa ora, possibilmente di mattina e due ore dopo l'assunzione dei farmaci. Il test consiste nella somministrazione di un questionario che è una versione del COPD Assessment Test (CAT) opportunamente adattato per essere visualizzato sullo schermo del dispositivo FOT [99]. Il CAT è ideato per misurare l'impatto della COPD sulla vita di una persona, e come questa cambia col passare del tempo. Successivamente il vero test FOT è eseguito tramite l'invio di stimoli a multifrequenza per 3 minuti. Gli output principali di ogni test sono rappresentati dal total score del questionario e da 13 parametri respiratori. Il reclutamento stimato di 80 pazienti e la valutazione giornaliera da 6 a 8 mesi fa di questo studio il più esteso e dettagliato trial clinico mai definito per l'osservazione della COPD.

I metodi progettati per gestire e verificare la qualità della grande mole di dati prodotta dallo studio clinico, come gli strumenti per la individuazione automatica degli outliers e per la stima dei dati mancanti, forniscono prestazioni molto elevate. L'utente può velocemente navigare tra le acquisizioni dei vari pazienti e può efficacemente visualizzare e correggere le acquisizioni sospette di essere outliers. Ogni qual volta un test è mancante, un algoritmo di regressione locale pesata stima i valori dei 13 parametri respiratori in modo da lasciare inalterate le proprietà statistiche delle serie temporali originali. Nonostante i metodi presentati dimostrano un funzionamento molto valido, possibili miglioramenti futuri sono identificati ed elucidati.

I risultati delle analisi effettuate sui parametri, ovvero uno studio di significatività del questionario, l'analisi delle componenti principali (PCA) e una analisi di variabilità a diverse scale temporali, forniscono interessanti spunti sulla patologia nonostante il modesto numero di pazienti attualmente disponibile.

Innanzi tutto, i risultati dimostrano fermamente come il questionario non può essere usato come strumento per la valutazione puntuale del grado di ostruzione polmonare nella COPD. Infatti, la regressione lineare accerta che il questionario non è sistematicamente correlato nè con nessun parametro respiratorio preso singolarmente, nè con nessuno dei 8192 diversi gruppi estratti dal set totale di parametri. Ciò conferma la comune critica mossa verso le misure basate su questionario che consiste nel fatto che l'informazione ottenuta è per sua natura soggettiva.

Inoltre la PCA indica che due variabili, opportunamente definite come combinazioni lineari dell'insieme originale di 14 parametri, sono sufficienti a spiegare una buona per-

centuale (64.8%) del contenuto informativo totale rappresentato dal set completo di parametri. La prima componente principale è molto legata ai parametri di meccanica respiratoria[3], mentre la seconda componente principale è piuttosto allineata con i parametri di pattern respiratorio[4]. Ciò significa che i parametri di meccanica respiratoria sono i più sensibili a qualsiasi deterioramento o miglioramento dello stato di salute e che i parametri di meccanica respiratoria e quelli di pattern respiratorio potrebbero essere usati per rappresentare diversi aspetti della patologia.

Infine, l'analisi di variabilità a diverse scale temporali illustra come le fluttuazioni a breve-termine dei parametri di meccanica respiratoria potrebbero costituire una misura della gravità della malattia e, allo stesso tempo, rappresentare gli effetti di meccanismi patofisiologici su entrambe le via aeree centrali e periferiche. A conferma di ciò, i parametri di meccanica respiratoria mostrano fluttuazioni che sono minime in soggetti sani, modeste in pazienti lievemente asmatici e molto forti in pazienti COPD gravi. Al contrario, la variabilità lungo-termine conferma che asma e COPD sono caratterizzate da simili cambiamenti patologici nella vie aeree centrali, mentre i processi infiammatori delle vie aeree periferiche e nel parenchima sono molto più pronunciate nei pazienti COPD. Questo è suggerito dal fatto che le fluttuazioni per scale temporali maggiori di un giorno sono confrontabili in asma e in COPD solo per i parametri di resistenza. I parametri di reattanza, al contrario, sono caratterizzati da fluttuazioni molto più pronunciate nei pazienti COPD.

In conclusione, questo lavoro indica come un monitoraggio domiciliare della meccanica respiratoria potrebbe contribuire ad avanzare la comprensione della COPD e migliorare la gestione clinica dei pazienti che soffrono di questa malattia.

---

[3]I parametri che descrivono l'impedenza meccanica polmonare sono chiamati parametri di *meccanica respiratoria*

[4]I parametri che descrivono volumi polmonari e frequenze respiratorie sono chiamati parametri di *pattern respiratorio*

Introduction

Complex physiological rhythms and fluctuations characterize nearly all aspects of life [8] and appear at different spatial scales, ranging from the molecular dimension to the organ and organism level. Examples of the latter include rhythmic heart beat, daily cycle of sleep and wakefulness and respiration. Interestingly, most biological fluctuations are not strictly periodic but show irregular patterns over many time scales [9]. Moreover, they interact not only with each other but with the external environment, which itself also exhibits substantial irregular fluctuations.

During the entire course of life, each organism tries to maintain a constant internal state by regulating its physiological processes. This dynamic self control is usually called homeostasis. However, the maintenance of homeostasis requires some controlled parameters have a range of permitted values, and these parameters continuously oscillate within its boundaries. Therefore, a more appropriate term is homeokinesis, that is the ability of an organism interacting with a variable external environment to maintain an organized internal environment, which fluctuates within acceptable limits by dissipating energy in a state far from thermodynamic equilibrium [10]. This definition implies that, while continuous fluctuations are part of normal life, alterations in the fluctuations always signify abnormal physiology. The disruption of such rhythms often leads to collapse or even the death of the organism [11, 12].

If homeokinesis is healthy and pathology is lack or excess of fluctuations in some observed parameters, it is clear that variability itself contains an encoded message that needs to be deciphered, quantified, analyzed and understood. Understanding this message should shed light into what is considered health and which are the mechanisms that bring to illness.

Due to the complexity of natural and biological feedbacks and the presence of unknown environmental factors, it is becoming clear that variability can not be analyzed using a classic reductionist approach. This is noticeably true also for the altered variability observed in the principal chronic respiratory pathologies, bronchial asthma and chronic obstructive pulmonary disease (COPD).

Asthma and COPD affect millions of children and adults worldwide. It is estimated

that 300 millions of people have asthma (Global Initiative for Asthma 2009) and 80 millions suffer from COPD (World Health Organization 2006). Asthma has become more common in recent decades; its prevalence has been associated with an increase in atopic sensitization and is in parallel with the increase of other allergic disorders such as rhinitis. COPD is the fourth leading cause of death in the world and increases in its prevalence and mortality are expected in the next future [13].

Many features of these diseases, including pathogenesis and progression, are not fully understood. Indeed, COPD and asthma are characterized by high variability of symptoms over time and a proper and accurate evaluation of patients' condition at any point in time would require frequent, at least daily, assessment of their disease. However, new objective parameters for the follow up of their severity and variability were still needed. A strategy with more than one observed index or including statistical measures of objective parameters with time should provide a more comprehensive picture of the pathology and its progression and it would eventually help improving patients' quality of life and, potentially, their life expectancy.

Telemedicine and home-monitoring devices have been addressed as possible solutions for the chronic stabilization of patients suffering from chronic respiratory diseases. However, despite recent advances in medical technology, the assessment and management of asthma and COPD still relied on tests performed with spirometry or peakflow meters, which did not provide data consistent with the quality requirements of international guidelines when the measurements are performed in unsupervised environments [14, 15].

A different and innovative multidimensional approach for the characterization of the variability and the progression of asthma and COPD over time could be based on the measurement of parameters related to the mechanical properties of the respiratory system by forced oscillation technique (FOT). FOT constitutes a valid solution to develop a new model for a better management of such patients: as a tool for the investigation of respiratory mechanics in clinical practice, it is well supported theoretically and has the advantage of being a non-invasive, versatile method and demanding minimal cooperation of the patient. For these reasons, FOT has been indicated as a potential alternative to spirometry for home assessment of respiratory function [16, 17].

Accordingly, the purpose of the work here presented is to:

- validate a new strategy for the analysis of the temporal variability of asthma and COPD, based on the development of a novel FOT telemedicine system for chronic respiratory diseases' home monitoring;
- analyze clinical data from COPD patients at stage 3 and 4 of GOLD classification measured with the device.

The main outcomes are positive feedback about the reliability of the device and new insights into the pathology despite the modest amount of patients currently available.

Chapter 1

Chronic Obstructive Pulmonary Disease and assessment of lung function

1 Introduction

Chronic obstructive pulmonary disease (COPD) is a chronic respiratory disorder characterized by the presence of airflow limitation [13]. The pathology consists by a chronic inflammation of the respiratory tract that is mediated by increased expression of multiple inflammatory proteins, including cytokines, chemokines, adhesion molecules, inflammatory enzymes and receptors.

The progression of the pathological condition is described by stable periods disrupted by sudden exacerbations of the symptoms during which a severe inflammatory process occurs. Such aggravations recur in a periodic fashion and often require hospitalization.

Worldwide, cigarette smoking is the most commonly encountered risk factor for COPD, even though in many countries air pollution resulting from the burning of wood and other biomass fuels has also been identified as a COPD risk factor.

As the present work is based on experimental studies involving COPD patients, the majority of this Chapter will focus on the pathophysiologic and epidemiologic characteristics of COPD. Successively, a separate section will introduce a non-invasive methods useful to monitor the progression of the pathology on a daily, namely the forced oscillation technique (FOT).

2   Definition of Chronic Obstructive Pulmonary Disease

Internationally accepted opinion has defined the Chronic Obstructive Pulmonary Disease (COPD) as a disease state characterized by chronic airflow limitation due to chronic bronchitis and emphysema [18]. Chronic bronchitis has been defined in clinical terms: the presence of chronic productive cough for at least 3 consecutive months in 2 consecutive years. Other causes of chronic productive cough must be ruled out. Emphysema, on the other hand, has been defined by its pathologic description: an abnormal enlargement of the air spaces distal to the terminal bronchioles accompanied by destruction of their walls and without obvious fibrosis. The latest ATS[1]/ERS[2] guidelines, like the GOLD[3] guidelines, have parted from this traditional description of COPD. Similar to the changes in the definition of asthma by the NHLBI[4], the definition of COPD has undergone major revision. COPD, like asthma, is now recognized as an inflammatory disease of the airways [22]. This is supported by extensive clinical and basic science research showing that asthma and COPD have different and distinct cellular and inflammatory mediator profiles. The current ATS/ERS definition reflects these scientific advances:

Chronic obstructive pulmonary disease (COPD) is a preventable and treatable disease characterized by airflow limitation that is not fully reversible. The airflow limitation is usually progressive and is associated with an abnormal inflammatory response of the lungs to noxious particles or gases, primarily caused by cigarette smoking. Although COPD affects the lungs, it also produces significant systemic consequences [20].

While the current guidelines do not specifically include chronic bronchitis and emphysema in the definition of COPD, it is made clear that they are considered the predominant causes of COPD.

---

[1]American Thoracic Society
[2]European Respiratory Society
[3]Global Initiative for Chronic Obstructive lung Disease
[4]National Heart, Lung, and Blood Institute

3 Pathogenesis, Pathology and Pathophysiology of COPD

The chronic airflow limitation characteristic of COPD is caused by a mixture of small airway disease (obstructive bronchiolitis) and parenchymal destruction (emphysema), the relative contributions of which vary from subject to subject (see Figure 1.1). Chronic inflammation causes structural changes and narrowing of the small airways. Destruction of the lung parenchyma, also by inflammatory processes, leads to the loss of alveolar attachments to the small airways and decreases lung elastic recoil. In turn, these changes diminish the ability of the airways to remain open during expiration.
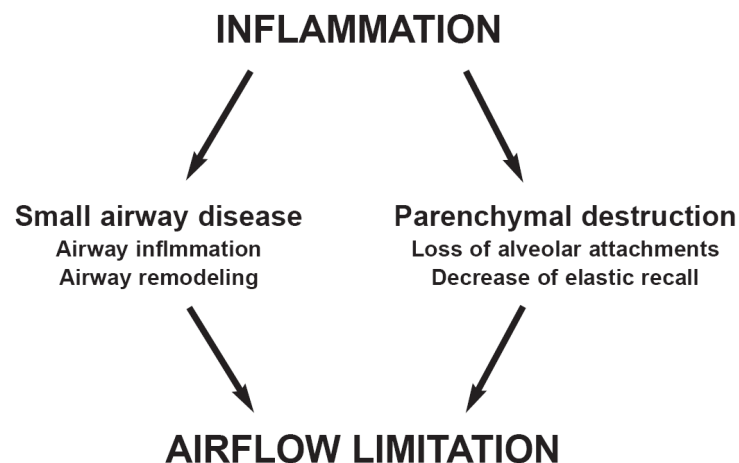
INFLAMMATION

Small airway disease
Airway inflammation
Airway remodeling

Parenchymal destruction
Loss of alveolar attachments
Decrease of elastic recall

AIRFLOW LIMITATION

Figure 1.1: Mechanisms underlying airflow limitation in COPD.

Emphysema, or destruction of the gasexchanging surfaces of the lung (alveoli), is a pathological term that is often (but incorrectly) used clinically and describes only one of several structural abnormalities present in patients with COPD. Chronic bronchitis, or the presence of cough and sputum production for at least 3 months in each of two consecutive years, remains a clinically and epidemiologically useful term as well. However, it does not reflect the major impact of airflow limitation on morbidity and mortality in COPD patients. It is also important to recognize that cough and sputum production may precede the development of airflow limitation but, conversely, some patients develop significant airflow limitation without chronic cough and sputum production.

3.1 Pathogenesis

The inflammation in the respiratory tract of COPD patients appears to be an amplification of the normal inflammatory response of the respiratory tract to chronic irritants such as cigarette smoke. The mechanisms for this amplification are not yet understood but may be genetically determined. Some patients develop COPD without smoking, but the nature of the inflammatory response in these patients is unknown [21]. Lung inflammation is further amplified by oxidative stress and an excess of proteinases in the lung. Together, these mechanisms lead to the characteristic pathological changes in COPD exhibited in Figure 1.2.
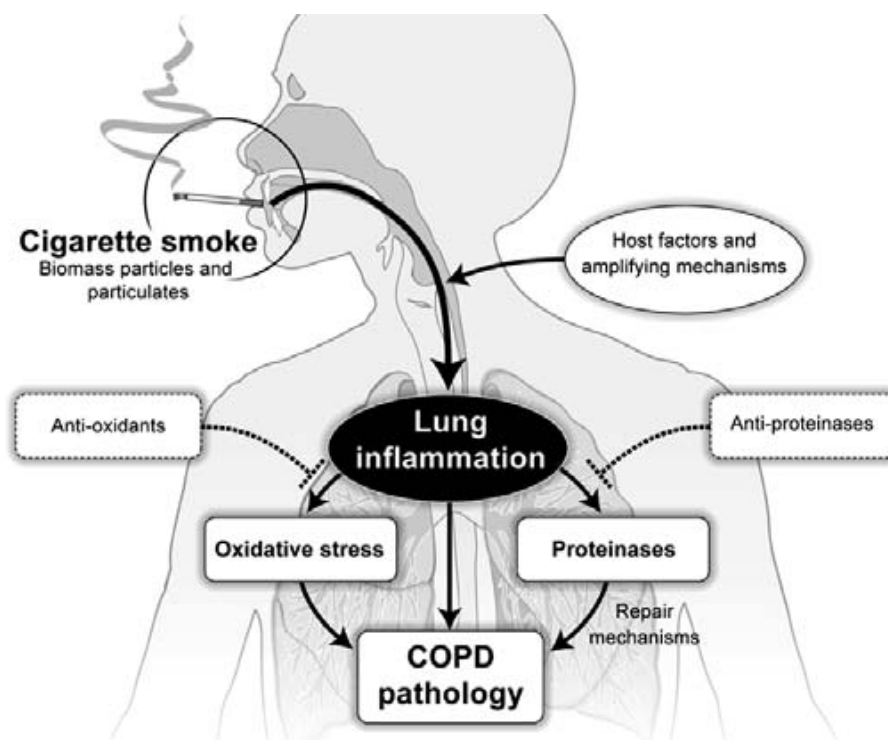
Figure 1.2: Pathogenesis of COPD.

Inflammation

COPD is characterized by a specific pattern of inflammation involving neutrophils, macrophages and lymphocytes [22]. These cells release inflammatory mediators and interact with structural cells in the airways and lung parenchyma.

The wide variety of inflammatory mediators that have been shown to be increased in COPD patients attract inflammatory cells from the circulation (chemotactic factors), amplify the inflammatory process (proinflammatory cytokines), and induce structural changes (growth factors) [23].

Oxidative stress

Oxidative stress may be an important amplifying mechanism in COPD [24]. Biomarkers of oxidative stress (e.g., hydrogen peroxide, 8-isoprostane) are increased in the exhaled breath condensate, sputum, and systemic circulation of COPD patients. Oxidative stress is further increased in exacerbations. Oxidants are generated by cigarette smoke and other inhaled particulates, and released from activated inflammatory cells such as macrophages and neutrophils. There may also be a reduction in endogenous antioxidants in COPD patients. Oxidative stress has several adverse consequences in the lungs, including activation of inflammatory genes, inactivation of antiproteases, stimulation of mucus secretion, and stimulation of increased plasma exudation. Many of these adverse effects are mediated by peroxynitrite, which is formed via an interaction between superoxide anions and nitric oxide. In turn, the nitric oxide is generated by inducible nitric oxide synthase, which is expressed in the peripheral airways and lung parenchyma of COPD patients. Oxidative stress may also account for a reduction in histone deacetylase activity in lung tissue from COPD patients, which may lead to enhanced expression of inflammatory genes and also a reduction in the antiinflammatory action of glucocorticosteroids [25].

Protease-antiprotease imbalance

There is compelling evidence for an imbalance in the lungs of COPD patients between proteases that break down connective tissue components and antiproteases that protect against this. Several proteases, derived from inflammatory cells and epithelial cells, are increased in COPD patients. There is increasing evidence that they may interact with each other. Protease-mediated destruction of elastin, a major connective tissue component in lung parenchyma, is an important feature of emphysema and is likely to be irreversible.

3.2 Pathology

Pathological changes characteristic of COPD are found in the central airways, peripheral airways, lung parenchyma, and pulmonary vasculature [26]. In patients with chronic bronchitis, an inflammatory exudate of fluid and cells infiltrates the epithelium lining the central airways and the associated glands and ducts.

Figure 1.3 shows how the early decline in lung function in COPD is correlated with inflammatory changes in the peripheral airways similar to those that occur in the central airways [27]: exudate of fluid and cells in the airway wall and lumen, goblet

and squamous cell metaplasia of the epithelium, edema of the airway mucosa due to inflammation and excess mucus in the airways due to goblet cell metaplasia. However, the most characteristic change in the peripheral airways of patients with COPD is airway narrowing.
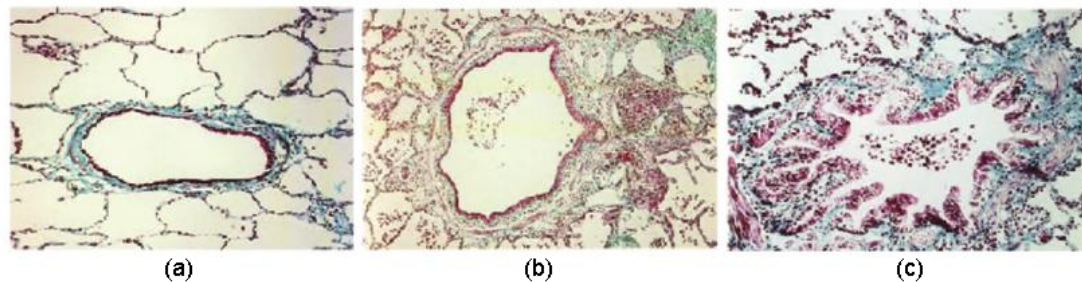

Figure 1.3: Histological section of peripheral airways. (a) Section from a sigarette smoker with normal lung function showing a nearly normal airway with small number of inflammatory cells. (b) Section from a patient with small airway disease showing inflammatory exudate in the wall and lumen of the airway. (c) Section showing more advanced small airway disease, with reduced lumen causing structural reorganization of the airway wall, increased smooth muscle and deposition of peribronchial connective tissue.

Inflammation leads to repeated cycles of injury and repair of the walls of the peripheral airways. This injury then initiates repair processes. Although airway repair is only partly understood, it seems likely that disordered repair processes can lead to tissue remodeling with altered structure and function [28]. This injury-and-repair process results in a structural remodeling of the airway wall, with increasing collagen content and scar tissue formation that narrows the lumen and produces fixed airways obstruction. The peripheral airways become the major site of airways obstruction in COPD [29].

The most common type of parenchymal destruction in COPD patients is the centriacinar form of emphysema (see Figure 1.4), which involves dilatation and destruction of the respiratory bronchioles [30]. These lesions occur more frequently in the upper lung regions in milder cases. Panacinar emphysema, which extends throughout the acinus, involves dilatation and destruction of the alveolar ducts and sacs as well as the respiratory bronchioles. It tends to affect the lower more than the upper lung regions.

Pulmonary vascular changes in COPD are characterized by a thickening of the vessel wall that begins early in the natural history of the disease, followed by an increase in vascular smooth muscle and the infiltration of the vessel wall by inflammatory cells. Since endothelium plays an important role in regulating vascular tone and cell prolif-
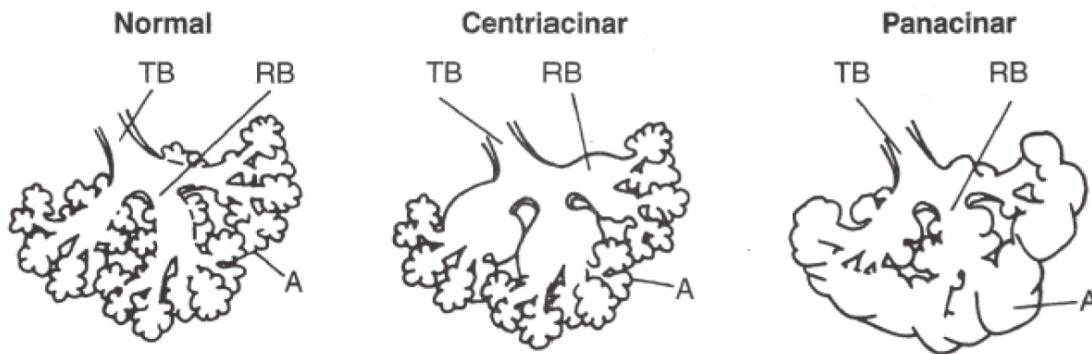
Figure 1.4: Centriacinar and panacinar emphysema. Notice that in centriacinar emphysema the destruction is confined to the terminal and respiratory bronchioles (TB and RB). In panacinar emphysema the peripheral alveoli (A) are also involved.

eration, it is likely that endothelial dysfunction might initiate the sequence of events that results ultimately in structural changes. These structural changes are correlated with an increase in pulmonary vascular pressure that develops first with exercise and then at rest. As COPD worsens, greater amounts of smooth muscle, proteoglycans, and collagen further thicken the vessel wall. In advanced disease, the changes in the muscular arteries may be associated with emphysematous destruction of the pulmonary capillary bed.

3.3 Pathophysiology

There is now a good understanding of how the underlying disease process in COPD leads to the characteristic physiologic abnormalities and symptoms. For example, decreased FEV_1 primarily results from inflammation and narrowing of peripheral airways, while decreased gas transfer arises from the parenchymal destruction of emphysema.

Airflow limitation and air trapping

The extent of inflammation, fibrosis, and luminal exudates in small airways is correlated with the reduction in FEV_1 and FEV_1/FVC ratio, and probably with the accelerated decline in FEV_1 characteristic of COPD [31]. This peripheral airway obstruction progressively traps air during expiration, resulting in hyperinflation.

Although emphysema is more associated with gas exchange abnormalities than with reduced FEV_1, it does contribute to air trapping during expiration. This is especially so

as alveolar attachments to small airways are destroyed when the disease becomes more severe. Hyperinflation reduces inspiratory capacity such that functional residual capacity increases, particularly during exercise (when this abnormality is known as dynamic hyperinflation), and this results in dyspnea and limitation of exercise capacity. It is now thought that hyperinflation develops early in the disease and is the main mechanism for exertional dyspnea [32]. Bronchodilators acting on peripheral airways reduce air trapping, thereby reducing lung volumes and improving symptoms and exercise capacity.

Gas exchange abnormalities

Gas exchange abnormalities result in hypoxemia and hypercapnia, and have several mechanisms in COPD. In general, gas transfer worsens as the disease progresses. The severity of emphysema correlates with arterial PO_2 and other markers of ventilation-perfusion (V_A/Q) imbalance. Peripheral airway obstruction also results in V_A/Q imbalance, and combines with ventilatory muscle impaired function in severe disease to reduce ventilation, leading to carbon dioxide retention. The abnormalities in alveolar ventilation and a reduced pulmonary vascular bed further worsen the V_A/Q abnormalities.

Mucus hypersecretion

Mucus hypersecretion, resulting in a chronic productive cough, is a feature of chronic bronchitis and is not necessarily associated with airflow limitation. Conversely, not all patients with COPD have symptomatic mucus hypersecretion. When present, it is due to mucous metaplasia with increased numbers of goblet cells and enlarged submucosal glands in response to chronic airway irritation by cigarette smoke and other noxious agents. Several mediators and proteases stimulate mucus hypersecretion and many of them exert their effects through the activation of epidermal growth factor receptor (EGFR) [33].

Pulmonary hypertension

Mild to moderate pulmonary hypertension may develop late in the course of COPD and is due to hypoxic vasoconstriction of small pulmonary arteries, eventually resulting in structural changes that include intimal hyperplasia and later smooth muscle hypertrophy/hyperplasia [34]. There is an inflammatory response in vessels similar to that seen in the airways and evidence for endothelial cell dysfunction.

The loss of the pulmonary capillary bed in emphysema may also contribute to increased pressure in the pulmonary circulation. Progressive pulmonary hypertension

may lead to right ventricular hypertrophy and eventually to right-side cardiac failure (cor pulmonale).

Systemic features

It is increasingly recognized that COPD involves several systemic features, particularly in patients with severe disease, which have a major impact on survival and comorbid diseases [35, 36]. These are:

- Cachexia: loss of fat free mass.
- Skeletal muscle wasting: apoptosis, disuse atrophy.
- Osteoporosis.
- Depression.
- Normochromic normocytic anemia.
- Increased risk of cardiovascular disease: associated with increase in C-reactive protein (CRP).

4 Epidemiology of COPD

In the past, imprecise and variable definitions of COPD have made it difficult to quantify prevalence, morbidity and mortality. Furthermore, the underrecognition and underdiagnosis of COPD lead to significant underreporting. The extent of the underreporting varies across countries and depends on the level of awareness and understanding of COPD among health professionals, the organization of health care services to cope with chronic diseases, and the availability of medications for the treatment of COPD [37].

4.1 Prevalence

Existing COPD prevalence data show remarkable variation due to differences in survey methods, diagnostic criteria and analytic approaches [38, 39]. Despite these complexities, data are emerging that enable some conclusions to be drawn regarding COPD prevalence. A systematic review and meta-analysis of studies carried out in 28 countries between 1990 and 2004 [38], and an additional study from Japan [40], provide evidence that the prevalence of COPD (Stage I: Mild COPD and higher) is appreciably higher in smokers and ex-smokers than in nonsmokers, in those over 40 years than those under 40, and in men than in women.

The Latin American Project for the Investigation of Obstructive Lung Disease (PLATINO) examined the prevalence of post-bronchodilator airflow limitation (Stage

I: Mild COPD and higher) among persons over age 40 in five major Latin American cities each in a different country (Brazil, Chile, Mexico, Uruguay, and Venezuela). In each country, the prevalence of Stage I: Mild COPD and higher increased steeply with age, with the highest prevalence among those over 60 years, ranging from a low of 18.4% in Mexico City, Mexico to a high of 32.1% in Montevideo, Uruguay. In all cities/countries the prevalence was appreciably higher in men than in women. The reasons for the differences in prevalence across the five Latin American cities are still under investigation [41].

In 12 Asia-Pacific countries and regions a study based on a prevalence estimation model indicated a mean prevalence rate for moderate to severe COPD among individuals 30 years and older of 6.3% for the region. The rates varied twofold across the 12 Asian countries and ranged from a minimum of 3.5% (Hong Kong and Singapore) to a maximum of 6.7% (Vietnam) [42].

4.2 Morbidity

Morbidity measures traditionally include physician visits, emergency department visits, and hospitalizations. Although COPD databases for these outcome parameters are less readily available and usually less reliable than mortality databases, the limited data available indicate that morbidity due to COPD increases with age and is greater in men than in women [43–45]. In these data sets, however, COPD in its early stages (Stage I: Mild COPD and Stage 2: Moderate COPD) is usually not recognized, diagnosed, or treated, and therefore may not be included as a diagnosis in a patient's medical record.

Morbidity from COPD may be affected by other comorbid chronic conditions [46] (e.g., musculoskeletal disease, diabetes mellitus) that are not directly related to COPD but nevertheless may have an impact on the patient's health status, or may negatively interfere with COPD management. In patients with more advanced disease (Stage III: Severe COPD and Stage IV: Very Severe COPD), morbidity from COPD may be misattributed to another comorbid condition.

Another way of estimating the morbidity burden of disease is to calculate years of living with disability (YLD). The Global Burden of Disease Study estimates that COPD results in 1.68 YLD per 1,000 population, representing 1.8% of all YLDs, with a greater burden in men than in women (1.93% vs. 1.42%) [1–3].

4.3 Mortality

The World Health Organization (WHO) publishes mortality statistics for selected causes of death annually for all WHO regions. Data must be interpreted cautiously, however, because of inconsistent use of terminology for COPD. Nowadays the problem of labeling the disease has been partly solved, but underrecognition and underdiagnosis of COPD still affect the accuracy of mortality data. Although COPD is often a primary cause of death, it is more likely to be listed as a contributory cause of death or omitted from the death certificate entirely, and the death attributed to another condition such as cardiovascular disease.

Despite the problems with the accuracy of the COPD mortality data, it is clear that COPD is one of the most important causes of death in most countries. The Global Burden of Disease Study has projected that COPD, which ranked sixth as the cause of death in 1990, will become the third leading cause of death worldwide by 2020 [1–3]. This increased mortality is driven by the expanding epidemic of smoking and the changing demographics in most countries, with more of the population living longer. Of these two forces, demographics is the stronger driver of the trend.

Trends in mortality rates over time provide further important information but, again, these statistics are greatly affected by terminology, awareness of the disease, and potential gender bias in its diagnosis. COPD mortality trends generally track several decades behind smoking trends. Trends in age-standardized death rates for the six leading causes of death in the United States from 1970 through 2002 [47] indicates that while mortality from several of these chronic conditions declined over that period, COPD mortality increased (refer to Figure 1.5). Death rates for COPD in Canada, in both men and women, have also been increasing since 1997. In Europe, however, the trends are different, with decreasing mortality from COPD already being seen in many countries [48]. There is no obvious reason for the difference between trends in North America and Europe, although presumably factors such as awareness, changing terminology, and diagnostic bias contribute to these differences.

The mortality trends for COPD have been particularly striking for women. In Canada, the death rate from COPD among women accelerated in the 1990s and is expected to soon overtake the rate among men [45]. In the United States, COPD deaths among women have been rising steeply since the 1970s. In 2000, the number of deaths from COPD in the United States was greater among women than men (59,936 vs. 59,118), although the mortality rates among women remain somewhat lower than among men [49].

Worldwide, recent increases in COPD deaths are likely to continue. The Global
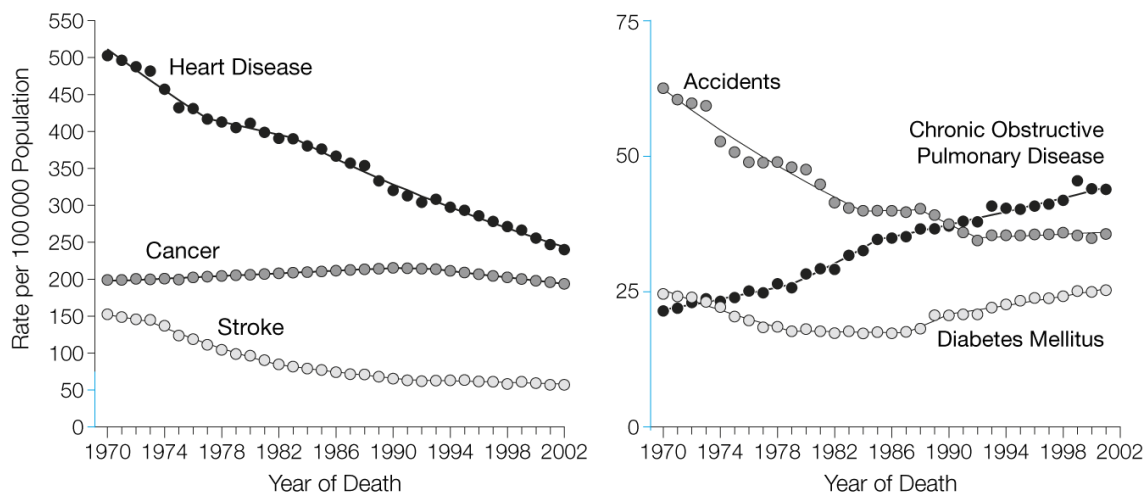
Figure 1.5: Trends in Age-standardized Death Rates for the 6 Leading Causes of Death in the United States, 1970-2002 [47].

Burden of Disease Study [1–3] projected baseline, optimistic, and pessimistic models for COPD mortality from 1990 to 2020 that take into account the expected aging of the world's population, projected increases in smoking rates, and projected declines in other causes of death such as diarrheal and HIV-related diseases.

4.4 Risk factors

The identification of risk factors is an important step towards the development of strategies for prevention and treatment of any disease. The division into *host factors* and *exposures* reflects the current understanding of COPD as resulting from an interaction between the two types of factors.

A list of host factors is:

- Genes (e.g., alpha-1 antitrypsin deficiency).
- Airway hyperresponsiveness.
- Lung growth.

A list of exposures is:

- Direct and environmental tobacco smoke.
- Occupational dust and chemicals.
- Indoor and outdoor air pollution.
- Infections.
- Socioeconomic status.

The best-documented host factor is a severe hereditary deficiency of alpha-1 antitrypsin, a major circulating inhibitor of serine proteases. This severe deficiency causes premature and accelerated development of panlobular emphysema and decline in lung function. Asthma and airway hyperresponsiveness, identified as risk factors that contribute to the development of COPD, are complex disorders related to a number of genetic and environmental factors. Reduced maximal attained lung function (as measured by spirometry) may identify individuals who are at increased risk for the development of COPD.

The major environmental risk factors are tobacco smoke [50], occupational dusts and chemicals (vapors, irritants, fumes), and indoor and outdoor air pollution. However, it is very difficult to demonstrate that a given risk factor is sufficient to cause the disease.

5 Exacerbations of COPD

As already mentioned COPD is characterized by a progressive and irreversible decline in lung function, breathlessness and other respiratory symptoms (e.g., cough and sputum production), and a deterioration of health status [51]. In addition to their chronic disease, patients with COPD often experience regular acute exacerbations (typically around 2-3 per year) that increase in frequency with increased disease severity [51,52].

These debilitating exacerbations have a substantial impact on patients and on healthcare systems. COPD exacerbations are also associated with considerable physiologic deterioration and increased airway inflammatory changes that are caused by a variety of factors such as viruses, bacteria, common pollutants and low temperatures [53]. COPD exacerbations are more common in the winter months and there may be important interactions between cold temperatures and exacerbations caused by viruses or pollutants.

Definition of a COPD exacerbation

Despite their importance, there is no standardized, universally accepted definition for COPD exacerbations. This situation reflects the multifactorial, heterogenous and poorly understood pathophysiology of exacerbations and the difficulty in differentiating true exacerbations from normal day-to-day variations of COPD. As a result, exacerbations remain poorly recognized and often poorly treated.

One of most widely known definitions of COPD exacerbations is that proposed by Anthonisen et al. in 1987 [54]. This definition is based on the presence of three specific symptoms in patients with COPD, namely increased dyspnoea, sputum volume

and sputum purulence. Three sub-types of exacerbations were defined: type 1, occurrence of all three symptoms; type 2, occurrence of two out of the three symptoms; and type 3, occurrence of one of the three symptoms in addition to at least one of the following: recent upper respiratory infection, fever, increased wheezing, increased cough, or increased respiratory or heart rate.

Epidemiology of COPD exacerbation

Exacerbations are an important cause of hospital admission and have a considerable impact on quality of life and activities of daily living of the patients.

Patients with COPD are accustomed to frequent symptom changes and this may explain their tendency to under report exacerbations to physicians. Some studies of patient with acute infective exacerbations of chronic bronchitis found that exacerbation frequency is an important determinant of health status in COPD and is thus one of the important outcome measures in COPD. Factors predictive of frequent exacerbations included daily cough, sputum and frequent exacerbations in the previous year [55]. Falls in peak flow and FEV_1 at exacerbation are generally small and not useful in predicting exacerbations, but larger falls in peak flow are associated with symptoms of dyspnoea, common colds and related to a longer recovery time from exacerbations. The combination of the symptoms of increased dyspnoea and the common cold at exacerbation with a prolonged exacerbation recovery suggests that viral infections may lead to more prolonged exacerbation.

The reasons for the incomplete recovery of symptoms and lung function are not clear, but may involve inadequate treatment or persistence of the causative agent. The incomplete physiologic recovery after an exacerbation could contribute to the decline in lung function with time. A study reveals that patients with a history of frequent exacerbations had a faster decline of FEV_1 compared to patients with a history of infrequent exacerbations [56].

Mechanisms underlying exacerbations

Although it has been assumed that exacerbations are associated with increased airway inflammation, there has been little information available on the nature of inflammatory markers, especially when studied closed to an exacerbation, as performing bronchial biopsies at exacerbation is difficult in patients with moderate to severe COPD.

Exacerbations may be triggered by a variety of factors, including viral or bacterial infection and air pollution. COPD exacerbations are frequently associated with upper respiratory tract infections and these are more common in the winter months, when

there are more respiratory viral infections in the community. Recent studies have shown that around half of COPD exacerbations were associated with viral infections and that the majority of these were due to rhinovirus, that is, viral exacerbations were associated with symptomatic colds and prolonged recovery [55, 57]. This finding is in agreement with the data that respiratory viruses produce longer and more severe exacerbations and have a major impact on healthcare utilization [58].

Airway bacterial colonizations have been found in approximately 30% of COPD patients, and these colonizations have been shown to be related to the degree of airflow obstruction and current cigarette smoking status. Bacterial colonization in COPD may be an important determinant of airway inflammation and therefore further long-term studies are required to determine whether bacterial colonization predisposes to decline in lung function, characteristic of COPD.

Pathophysiologic changes at COPD exacerbation

As seen before, the mechanical performance of respiratory muscles is reduced in patients with moderate to severe COPD. The airflow obstruction leads to hyperinflation, so that the respiratory muscles work far from the optimal extension and generate reduced inspiratory pressure. The load on the respiratory muscles is also increased in patients with airflow obstruction given by the presence of intrinsic positive end-expiratory pressure (iPEEP). With an exacerbation of COPD, the increase in airflow obstruction will further increase the load on the respiratory muscles and increase the work of breathing, precipitating respiratory failure in more severe cases. Minute ventilation may be normal, but the respiratory pattern will be irregular with increased frequency and decreased tidal volume. The resultant hypercapnia and acidosis will then reduce inspiratory muscles function, contributing to further deterioration of the respiratory failure.

Hypoxemia in COPD usually occurs due to a combination of ventilation-perfusion mismatch and hypoventilation, although arteriovenous shunting can also contribute to the acute setting. This causes an increase in pulmonary artery pressure, which can lead to salt and water retention and the development of edema. The degree of the ventilation-perfusion abnormalities increases during acute exacerbations and resolves over the following few weeks. Acidosis is an important prognostic factor in survival from respiratory failure during a COPD exacerbation and thus early correction of acidosis is an essential goal of therapy.

6   Diagnosis and assessment of COPD

COPD is a typical mid-to-late age pathology, especially observed in smokers or ex-smokers [59]. Chronic cough may precede the development of airway limitation by many years [13], even if some patients with limited airflow do not present any form of cough. The majority of patients seek medical attention after the development of dyspnea, which is usually persistent and becomes progressively more serious during everyday activities or at rest.

Airflow limitation is best measured by spirometry, as this is the most widely available, reproducible test of lung function. Although it does not fully capture the impact of COPD on a patient's health, it remains the gold standard for diagnosing the disease and monitoring its progression in specialized labs.

The main spirometric indices include the volume of air forcibly exhaled from the point of maximal inspiration (forced vital capacity, FVC), the volume of air exhaled during the first second of this maneuver (forced expiratory volume in one second, FEV_1), the peak of expiratory flow (PEF) and the FEV_1/FVC ratio. Spirometry measurements are evaluated by comparison with reference values based on age, height, sex and race.

Figure 1.6 shows a normal spirogram and a spirogram typical of patients with mild to moderate COPD. Patients with COPD typically show a decrease in both FEV_1 and FVC. The degree of spirometric abnormality generally reflects the severity of COPD. The presence of a postbronchodilator $FEV_1 < 80\%$ of the predicted value in combination with an $FEV_1/FVC < 70\%$ confirms the presence of airflow limitation that is not fully reversible [13]. The FEV_1/FVC on its own is a more sensitive measure of airflow limitation and an $FEV_1/FVC < 70\%$ is considered an early sign of airflow limitation in patients whose FEV_1 remains normal ($\geq 80\%$ predicted). This approach in defining airflow limitation is considered pragmatic in view of the fact that universally applicable reference values for FEV_1 and FVC are not available.

Assessment of COPD severity is based on the patient's level of symptoms, the severity of the spirometric abnormality, and the presence of complications such as respiratory failure, right heart failure, weight loss and arterial hypoxemia. The characteristic symptoms of COPD are cough, sputum production, and dyspnea upon exertion. COPD has a variable natural history and not all individuals follow the same course, however, it is generally a progressive disease, especially if a patient's exposure to noxious agents continues. A simple classification of disease severity into five stages is presented [60]:

- Stage 0: *At Risk* - Characterized by chronic cough and sputum production. Lung function, as measured by spirometry, is still normal.

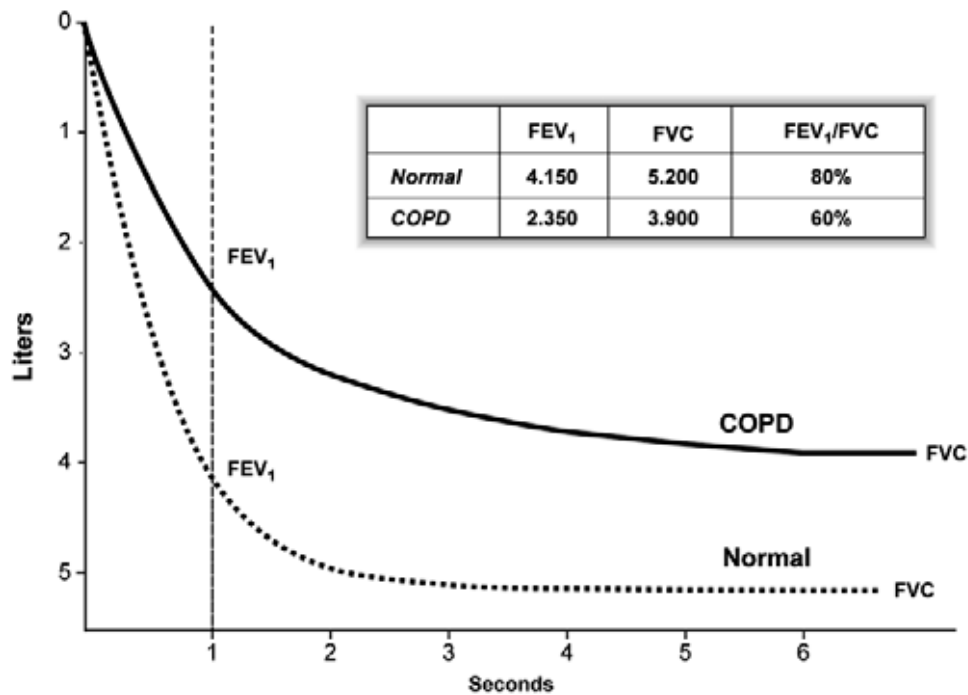|	FEV₁	FVC	FEV₁/FVC
Normal	4.150	5.200	80%
COPD	2.350	3.900	60%

Figure 1.6: Normal spirogram and spirogram typical of patients with moderate COPD.

- Stage I: *Mild COPD* - Characterized by mild airflow limitation ($FEV_1/FVC<70\%$ but $FEV_1>80\%$ predicted) and usually, but not always, by chronic cough and sputum production. At this stage, the individual may not even be aware that his or her lung function is abnormal.

- Stage II: *Moderate COPD* - Characterized by worsening airflow limitation ($50\%\leq FEV_1<80\%$ predicted) and usually progression of symptoms, with shortness of breath typically developing on exertion. This is the stage at which patients typically seek medical attention because of dyspnea or an exacerbation of their disease.

- Stage III: *Severe COPD* - Characterized by further worsening of airflow limitation ($30\%\leq FEV_1<50\%$ predicted), increased shortness of breath, and repeated exacerbations which have an impact on patients' quality of life.

- Stage IV: *Very Severe COPD* - Characterized by severe airflow limitation ($FEV_1<30\%$ predicted) or the presence of chronic respiratory failure.

The staging is based on airflow limitation as measured by spirometry. Specific FEV_1 cut-points (e.g., $<80\%$ predicted) are used for purposes of simplicity.

Respiratory failure may also lead to effects on the heart such as cor pulmonale (right heart failure). At this stage, quality of life is very appreciably impaired and exacerbations can threaten the life of the patient.

The diagnosis and staging of chronic respiratory diseases with spirometry have some limitations. As asthma and COPD are characterized by abnormal temporal variability of symptoms and fluctuations in the inflammatory state of the airways, the one-time punctual assessment of lung function in hospitals do not provide a comprehensive evaluation of the severity and progression. Second, all the spirometric indices are obtained using effort-dependent maneuvers that do require the supervision of clinicians or trained technician for the evaluation of their correctness and accuracy. However, also in this condition, some elderly patients and children may have difficulty completing spirometry with the desired quality standards [14, 15].

7 Assessment of temporal variability of airflow limitation and disease progression

Asthma and COPD are characterized by a high and abnormal variability of symptoms over time. An elegant mechanicistic definition of life considers an organism as an entity interacting with a complex and fluctuating external environment, continuously trying to maintain a state of internal dynamic equilibrium known as homeokinesis [10]. While fluctuations of physiological parameters are part of normal life, excess or lack of variability is normally associated to pathology or, eventually, to death. This is true also for asthma and COPD, where abnormal temporal variability of airway caliber and significant changes in functional and structural parameters are often observed [61, 62]. These observations suggest that a correct understanding of the disease and an accurate evaluation of its severity can not be based just on one-time assessment of clinical symptoms and on a punctual measurement of pulmonary function. A more comprehensive strategy, including statistical measures of objective parameters with time should provide a more comprehensive and clear picture of the pathology and its progression.

The clinical tool recommended in the most modern guidelines as a monitoring instrument for the measurement of such variability is the peak expiratory flow (PEF) meter [63]. PEF is the highest flow obtained during a forced expiration, starting immediately after a deep inspiration to the total lung capacity. However, PEF meters have some limitations that have prevented the development of accurate self-testing at home, a desirable goal in the chronic stabilization of patients suffering from airways obstruction [64, 65]. First, even if PEF meters provide data with good temporal resolution, many patients may have difficulty in executing correctly the forced maneuvers without the supervision of nurses, clinicians or trained technicians. Second, PEF is

effort dependent and it mainly reflects the caliber of large airways and therefore underestimates the degree of airflow limitation present in peripheral airways. Third, the association between PEF and airway hyperresponsiveness (AHR) is still unclear. Some studies found a correlation between PEF and bronchial responsiveness in adult asthmatics [66, 67] while others found weaker associations in patients treated with inhaled corticosteroids [68–71].

Despite these limitations, PEF is used as a measurement of temporal variability of airflow limitation and disease progression in both COPD and asthma [4–6]. In particular, Frey et al. [4] have recently shown that, in patients with asthma, it is possible to predict the likelihood of future exacerbations from the measurement of PEF fluctuations. Through detrended fluctuation analysis (DFA), these authors determined whether the statistical and correlation properties of the time series of PEF recordings could be used to predict the risk of subsequent exaggeration of airway instability.

This approach is very interesting and demonstrates that the temporal observation of respiratory parameters other than those related to the subjective self-report of symptoms may contribute to understand the effects of medications on the fluctuations of airway inflammation. However, even if the strategy of monitoring the temporal fluctuations of respiratory parameters seems to contribute in enhancing the comprehension of the mechanisms involved in the development and progression of chronic respiratory diseases, some limitations prevent the use of these results in clinical practice. First, the need to record a long time series of PEF separately for each subject in order to calculate the DFA exponent with a suitable accuracy. At least 400-500 daily measurements of PEF are required and this poses a problem of patients' compliance. Second, the long predictive window (30 days) in which the asthma attacks may occur with a high probability.

A speculation proposed is that the uncertainty in the prediction of extreme events may be partly explained by the fact that the authors analyzed the time series of a respiratory parameter, the peak of expiratory flow, which is effort dependent and that provides data with poor reproducibility when performed without supervision. Therefore, a more meaningful clinical outcome could be obtained by leaving the "spirometric approach" and concentrating on the analysis of the temporal fluctuations of reliable, accurate and non-effort-dependent parameters, as the respiratory mechanics indices measured by the forced oscillation technique (FOT) presented in the next Section.

Moreover, fluctuations of respiratory impedance measured by FOT can be used to monitor its temporal changes at different time scales. Gulotta et al. [72] have provided evidence that home monitoring of inspiratory resistance (Rinsp) may be a reliable tool to diagnose asthma and predict acute deterioration of airway function. The authors

used mean and standard deviation of Rinsp to build time series of its coefficients of variation (CV_{RINSP}) at 2, 4, 8, 16, and 32 consecutive days. At all time scales, CV_{RINSP} was significantly larger in subjects with asthma than in healthy subjects ($P<0.001$) (see Figure 1.7). A time window of only 4 days provided 90% sensitivity and specificity for CV_{RINSP} with respect to the diagnosis of the disease.
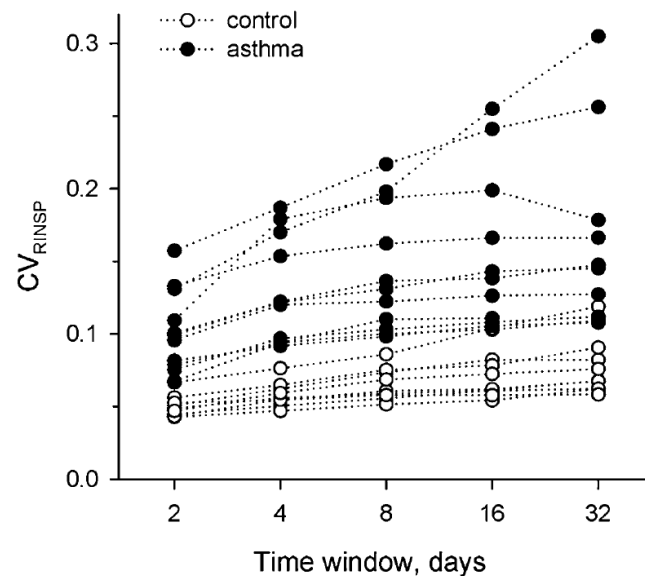

Figure 1.7: Long-term variability of Rinsp characterized by the coefficient of variation (CV_{RINSP}) in individual subjects with asthma (solid symbols) and healthy subjects (open symbols) plotted against time windows. *Adapted from: Gulotta et al. Monitoring the temporal changes of respiratory resistance: a novel test for the management of asthma. Am J Respir Crit Care Med. 2012 Jun 15;185(12):1330-1.*

As it will be investigated in Chapter 3, these insights can be applied to the case of COPD to advance knowledge about the disease.

8 The Forced Oscillation Technique (FOT)

The essence and unique advantage of FOT compared to other techniques commonly used in respiratory practice can be elucidated by contrasting its principle with that of the respiratory mechanical measurements that depend on respiratory maneuvers. For the FOT, external driving signals (i.e., forced oscillations) are used to estimate the mechanical properties of the respiratory system and many different forcing waveforms with frequency content different from that of the spontaneous breathing can be chosen.

FOT is completely non-invasive as it explores the mechanical properties of the lungs at the flows and volumes at which the patient actually breathes, whereas spirometry does it during a particular maximal expiratory maneuver rarely executed in the normal life activity. As a consequence, measurements obtained by FOT are likely to be more accurate and reproducible than spirometry.

FOT has also a consolidated theoretical background, as the hypothesis of linearity between the forcing signal and the measured output generally holds. Therefore the tools of linear system theory, such as impedance transfer function approximation, may be used [73]. However, the requirement of linearity implies the use of small amplitude oscillations which may leave undisclosed some nonlinear properties that manifest also during tidal breathing [74].

The most easily implementable FOT system (see Figure 1.8) measures the input impedance of the respiratory system (Zrs), namely the transfer function between flow (V') and pressure measured at the subject's mouth (Pao). The stimulus is generated by a loudspeaker and applied to the subject either by a mouthpiece or by a nasal mask. In the first case, the subject wears a nose clip and an operator firmly supports his/her cheeks in order to limit the artifactual role of the upper airways shunt. In order to measure V', the mouthpiece (or the nasal mask) is connected to a pneumotachograph. Pao is instead measured immediately after the mouthpiece or directly to the nasal mask. The pneumotachograph is in turn connected to a T-piece with one side connected to the loudspeaker, and the other one to a low-resistance and high-inertance tube that allows the subject to breathe without significant loss of the forcing signal. However, since this tube greatly increases the equipment dead space, the system is connected to a vacuum generator that extracting the expired air reduces the dead space of the equipment.

Zrs is a complex number that can be separated in its real part, the resistance (Rrs), and its imaginary part, the reactance (Xrs) and both appear as function of the frequency of the forcing stimulus (f). Rrs describes the dissipative mechanical properties of the respiratory system, whereas Xrs is related to the energy storage capacity and thus determined jointly by the elastic properties (the relationship between pressure and volume) dominant at low oscillation frequencies and the inertive properties (the relationship between pressure and volume acceleration), which become progressively more important with increasing f.

The conventional clinical application of FOT applies a medium frequency range, namely the imposed oscillations start from 2.4Hz (approximately one decade above the spontaneous breathing rate), and extend up to 20-30Hz. In this frequency range the healthy respiratory system exhibits a largely frequency-independent Rrs whose major
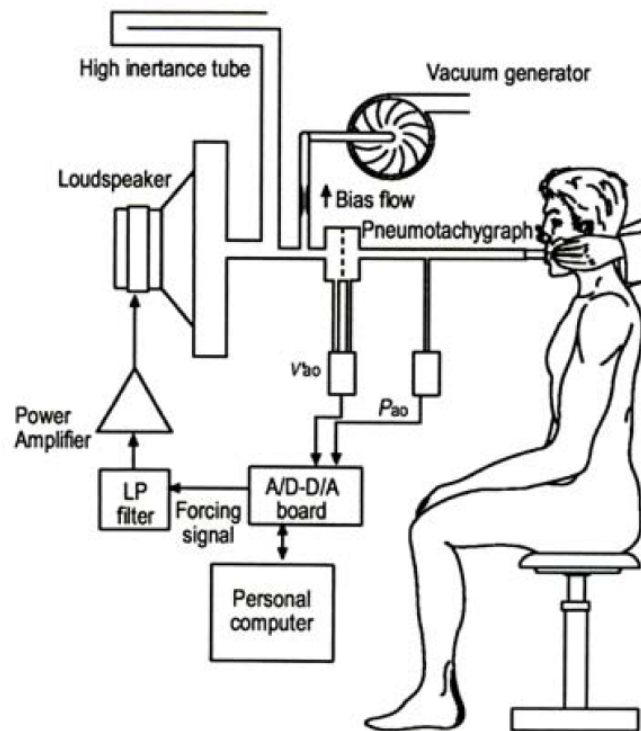
Figure 1.8: Instrumentation setup for the measurement with the forced oscillation technique.

component is airway resistance. Conversely, Xrs undergoes the transition from negative values (when the elastic reactance dominates) to positive values increasing with f (the dominance of inertial reactance). At the characteristic resonant frequency (f_{res}), where Xrs crosses zero, the elastic and inertial forces are equal in magnitude and opposite (see Figure 1.9).

In a typical clinical application composite forced oscillations, such as pseudorandom noise or recurrent impulses, are preferred and the output allows analyzing the frequency dependence of Zrs in health and disease. To prevent harmonic distortion and crosstalk arising from simple nonlinearities, the energy of the stimulus can be concentrated at frequencies that obey the non-sum-non-difference criteria [75], such that none of the frequencies are integer multiple of each other.

9   Conclusions

This chapter provided a suitable background for the understanding of the methods and of results described in the following chapters. Due to a chronical and degenerative
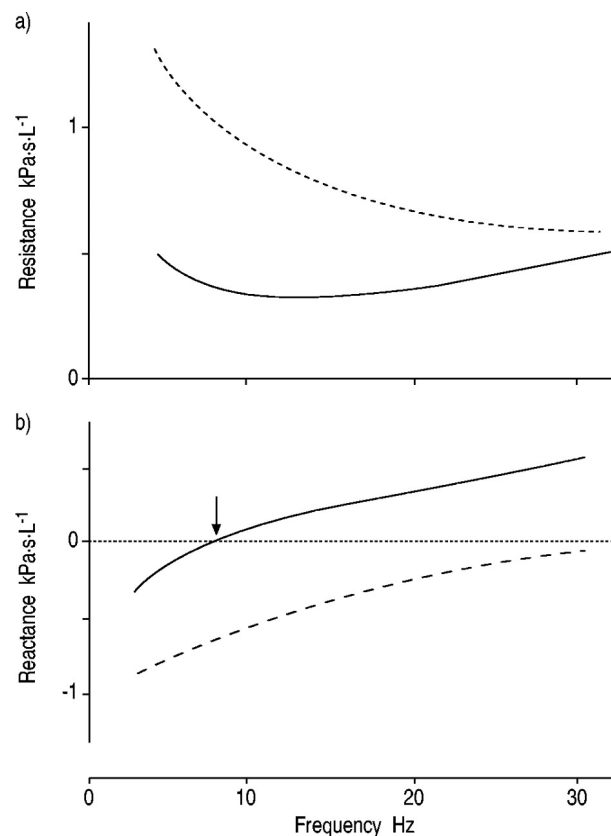
Figure 1.9: Frequency dependence of respiratory impedance of adults, in health (solid line) and disease (dashed line). Compared to the normal impedance data, the resistance is higher and negatively frequency-dependent, whereas reactance is lower. The arrow indicates the resonant frequency. *Source: Oostveen E. et al., The forced oscillation technique in clinical practice: methodology, recommendations and future developments, Eur.Respir J, 2003, vol. 22, no. 6, pp. 1026-1041.*

evolution which is characterized by continuative injury-and-repair processes, COPD implies a persistent structural remodeling of the airway wall with increasing collagen content and scar tissue formation that narrows the lumen and produces fixed airways obstruction. The peripheral airways become the major site of airways obstruction in COPD [29].

The dynamical evolution characterizing the COPD brings to abnormal rhythms and fluctuations of several respiratory parameters. A speculation proposed is that a long-term observation of objective parameters may help to characterize the progression of the disease over time and, possibly, early detect the onset of extreme future events.

Unfortunately, the diagnosis of this syndrome still relies on the measurements of spirometric indices that, given the effort-dependent maneuvers required, are not suitable for self-testing of airway obstruction. Therefore, a long-term and accurate observation

of respiratory parameters is still missing.

More meaningful clinical outcome could be obtained by leaving the spirometric approach and concentrating on the analysis of the temporal fluctuations of more reliable parameters that do not require the execution of forced respiratory maneuvers, as the respiratory mechanical indices measured by forced oscillation technique (FOT). As the measurement of Zrs does not require patient's collaboration, the results could be more accurate than spirometry and highly reproducible even if performed without supervision.

Chapter 2

Validation of a novel telemedicine system for chronic respiratory diseases' home monitoring

1 Introduction

The forced oscillation technique (FOT) is a promising method for providing a detailed analysis of respiratory mechanics during spontaneous breathing. Principal applications regard asthma and COPD, pathologies that have a tremendous impact on the economic and social infrastructure of the developed countries.

Recently, new commercial devices telemedicine devices employing this method have been developed. The capillary diffusion of such systems would permit a domiciliar monitoring of the lung function that, if effective, would represent a strategic solution for limiting costs.

In this Chapter a validation analysis of a commercial device developed and commercialized by a spin-off society of the Politecnico di Milano is presented. It is based on comparisons between measurements performed by the commercial device and measurements obtained by a reference laboratory FOT system taken as a gold standard. Results are supported by the employment of numerical tools of statistical significance.

2 Experimental setup

2.1 Commercial device

The novel telemedicine system object of validation (RESMON® PRO) is developed and commercialized by a spin-off society of the Politecnico di Milano (Restech Srl), in collaboration with the Laboratory of Biomedical Research (TBM Lab) of the Politecnico di Milano. The device detects flow and pressure signals at the airway openings and calculates the value of impedance of the tracheobronchial tree thanks to the application of the Forced Oscillation Technique (FOT, see Chapter 1 for an extensive explanation), besides assessing many other respiratory parameters and providing additional information about the test. The complete information set consists in:

- date and time of the test;
- frequency components of the stimulus (5Hz, 6Hz, 8Hz, 10Hz, 5-11-19Hz, pseudorandom noise);
- number of breaths recognized;
- percentage of breaths recognized;
- mean and standard deviation of breath-by-breath inspiratory resistance (for any of the frequency components, $Rinsp_{f[Hz]}$);
- mean and standard deviation of breath-by-breath expiratory resistance (for any of the frequency components, $Rexp_{f[Hz]}$);
- mean and standard deviation of breath-by-breath total breath resistance (for any of the frequency components, $Rtot_{f[Hz]}$);
- mean and standard deviation of breath-by-breath inspiratory reactance (for any of the frequency components, $Xinsp_{f[Hz]}$);
- mean and standard deviation of breath-by-breath expiratory reactance (for any of the frequency components, $Xexp_{f[Hz]}$);
- mean and standard deviation of breath-by-breath total breath reactance (for any of the frequency components, $Xtot_{f[Hz]}$);
- mean and standard deviation of breath-by-breath $Xinsp_{f[Hz]}$-$Xexp_{f[Hz]}$, with f[Hz] equal to the lowest frequency component of the stimulus (DeltaXrs);
- percentage of expiratory flow-limited breaths (EFL %);
- mean and standard deviation of breath-by-breath Rinsp5-Rinsp19 (only if multiple frequency is provided);
- mean and standard deviation of breath-by-breath inspiration time (Ti);
- mean and standard deviation of breath-by-breath expiration time (Te);

- mean and standard deviation of breath-by-breath Ti/Ttot;
- mean and standard deviation of breath-by-breath respiratory rate (RR);
- mean and standard deviation of breath-by-breath tidal volume (Vt);
- mean and standard deviation of breath-by-breath Vt/Ti;
- mean and standard deviation of breath-by-breath Vt/Te;
- mean and standard deviation of breath-by-breath ventilation (Ve);

Acquisitions were performed with a prototype (shown in Figure 2.1) made from the same hardware components as those commercialized items are equipped with, except for the external cover which is obtained from raw presswork. As regards the software installed, this is different from the clinical study's version in terms of main menu and data visualization, but the signal processing is standardized and common to all the devices produced. Such an approach makes the prototype's respiratory data outputs identical to those that would be obtained by certified devices. In addition, the software version for the clinical study administers a questionnaire meant to gain additional qualitative information on patient's health. The questionnaire is not important for the validation of the device but it might be correlated with respiratory parameters, therefore it will be introduced and evaluated in Chapter 3.
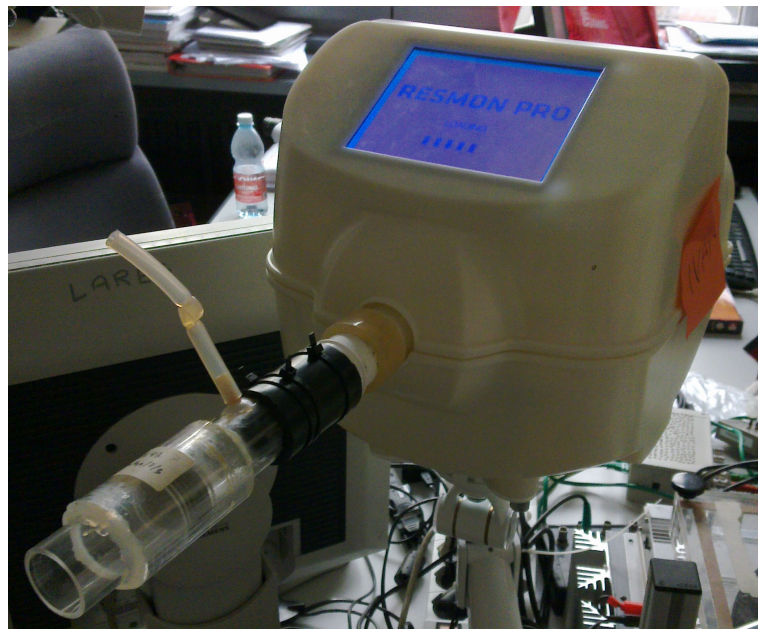

Figure 2.1: Prototype of the commercial device. The picture shows a test object applied to the inlet.

Flow and pressure sensors of the prototype were firstly calibrated by supplying known flow and pressure signals to the inlet. Due to the nonlinear characteristics of the mesh

placed inside the pneumotachograph, the relationship between the actual airflow and the signal provided by the transducer needs 7 coefficients to be correctly described. On the contrary, the pressure transducer measures the pressure drop between airway openings and atmosphere. Since the mesh is not involved and the sensor follows a linear law, this relationship can be effectively described by 2 coefficients. Periodical verifications where performed to check if the initial calibration was still valid. Calibration verification consists in applying a certified test object with known resistance and reactance characteristics, shown in Figure 2.2, to the device's inlet. An appropriate procedure implemented by the device's software assesses the test object impedance and compares it to that given manually by the user through a code (a checksum guarantees that the user inserts the right code). A pseudorandom noise, which consists in a composition of all the prime frequencies between 5 and 37Hz, is applied. The resulting couples of resistance and reactance are then compared to those obtained by evaluating the resistance and reactance characteristics of the test object (provided manually from screen through the above-mentioned code) at the same frequencies. According to the ERS guidelines for the application and interpretation of FOT [74], if any of the values given manually differs from the corresponding measured value by less than 10% the verification can be considered successful, otherwise the calibration must be executed again.

Figure 2.2: Test object used to verify the device's calibration.

In order to carry out offline processing and statistical analysis, data collected by the device were transferred to a desktop computer through ethernet collection and Secure Shell protocol. Whereas the device can select regularly performed breaths and calculate in real-time all the parameters listed at the beginning of the Subsection, it also saves raw data provided by analog-to-digital converters placed downstream of pressure and flow sensors (see Figure 2.3) and minor indexes. Offline processing consists in applying calibration coefficients to these raw signals to convert their units from 12-bit binary digits to proper pressure and flow units, and successively elaborate them so that respiratory parameters are obtained. The difference between the two approaches mainly concerns computational precision. For instance, real-time IIR digital filters as those implemented on the device introduce a delay between input and output signals which is not quantifiable in a deterministic way, whereas an offline processing permits to apply filters with the same frequency characteristics but with zero phase, such that no error is introduced. Also, all internal registers and data memory cells of the embedded digital signal processor are composed by a number of bit which is necessarily smaller than that of a general purpose desktop processor. As a consequence, approximations given by the latter result much finer.
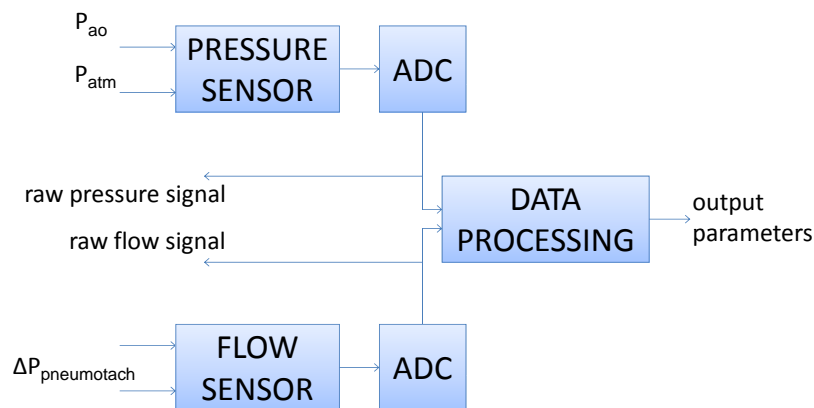

Figure 2.3: Simplified block diagram of the prototype's data acquisition and processing.

Another aspect that discriminates between online and offline processing is automatic breath selection. This consists in detecting zero-flow instants in order to discriminate inspiratory and expiratory phases of every breath, and successively discard breaths in which respiratory parameters exceed threshold values. For example, even for serious

COPD patients the tracheobronchial tree resistance is never supposed to break through $40 cmH_2OsL^{-1}$, hence values of this kind must be attributed to signal noise, glottis closure or disturbances of any other nature. The complete list of breath selection rules is:

- VtMin=0.1l;
- TtotMin=1s;
- TtotMax=10s;
- TinspMin=10s;
- TexpMin=10s;
- RtotMin=$0 cmH_2OsL^{-1}$;
- XtotMin=$-50 cmH_2OsL^{-1}$.

Offline processing has advantages in terms of both computational precision and correct automatic breath selection. Results provided by offline processing will be more reliable from a numerical point of view, and the set of breaths accepted will be optimized as this process is much more difficult to be performed and more prone to malfunctioning while data is being acquired rather than when the complete dataset is already saved and entirely visible. In Section 4 differences between the two approaches will be assessed and commented.

2.2 Reference FOT system

Validation of any kind of measuring instrument consists in comparing its output with the output of a second system or instrument (also called *gold standard*) that has a better accuracy. For instance, a thermometer can be put in a mixture of water and ice and in boiling water to validate its capability to correctly measure 0 and 100°C, or a traditional low quality balance scale can be validated by assuring that, putting the same weight on it and on a more accurate strain gauge scale, the two outputs are the same. In the case of this study, outputs of the prototype of the commercial devices are compared with those given by a reference FOT system properly assembled for validation aims (see Figure 2.4).

The reference FOT system has the same functionalities as those of the commercial device, but it is composed by the following discrete parts:

- large speaker;
- external amplifier able to provide sinusoidal power signals;
- high current power supply;

Figure 2.4: Reference FOT system.

- double-piston air aspirator;
- external pressure and flow transducers;
- pneumotachograph with high resistance mesh;
- very high inertance circuit;
- National Instruments™ data acquisition board;
- laptop for offline processing.

The performances of most of the listed components are superior to those of the commercial device's parts. This difference is due to many project constraints, among which cost, weight, power consumption and dimensions play a leading role. For instance, the reference system can drive the speaker with pure sinusoidal stimuli whereas the commercial device must apply a signal which is more similar to a square wave in order to limit the power consumption. Also, the pneumotachograph used in the reference system opposes a big resistance to the airflow, so that the signal to noise ratio of the signals output by the sensors is higher than the prototype's one. The drawback of such a choice is that the system studied (i.e., the respiratory system of the patient) results much more "loaded" by the measurement instrument. Last, the possibility of uncoupling sensors from the remaining components shields them from negative effects given by electromagnetic interference and aspirator's vibrations. For these and other

minor reasons, the reference FOT system can be considered more accurate than the prototype, that is, it can be treated as the gold standard.

Although signals produced by the reference system are always transferred in real-time to the laptop to be memorized, no online processing was implemented. In fact, as it will be explained in the next Section, one of the targets of the validation is ascertaining that the front-end hardware correctly transduces physical signals into digital ones. This means that raw data from both systems were processed with the same offline algorithm, so that any difference in the final results could only be ascribed to a different accuracy in the two front-ends.

Finally, calibration of the pneumotachograph was performed before the first measurement and repeated periodically in order to guarantee results' reliability.

3   Experimental protocols

The process of validation carried out in this study does not only aim at stating whether the commercial device can measure the characteristics of the patient's respiratory system in a reliable way, but also tries to spot where the possible source of low accuracy is and proposes solutions and improvements. For this reason different experimental protocols were defined, each one treating a different aspect of the problem. They will be illustrated separately in the remaining part of the Section.

In order to assess whether the difference between data acquired with the two systems can be considered significantly different, graphical and numerical tools are used for all the experimental protocols. As regards graphical tools, *scatter plot* and *Bland-Altman plot* are considered.

A scatter plot (or scattergraph) is a type of mathematical diagram using cartesian coordinates to display values for two variables of a set of data (e.g., different methods of volume measurement applied to the same sample). As the left plot of Figure 2.5 shows, data are displayed as a collection of points each having the value of one variable determining the position on the horizontal axis and the value of the other variable determining the position on the vertical axis. If no dependent variable exists, as for the data considered in this Chapter, either can be plotted on either axis and the scatter plot will illustrate only the degree of correlation (not causation) between the two variables, which may be positive (rising), negative (falling) or null (uncorrelated). Moreover, through a scatter plot it is possible to visualize and quantify the level of correlation. If the points lie on a straight line the level of correlation is strong and a linear relation can well model the data, otherwise the correlation can be considered weak.

In this context the second graphical tool can be introduced, namely the Bland-Altman plot (right plot of Figure 2.5), that is a graph constructed by considering a set of n samples (e.g., objects of unknown volume). Two assays (e.g., again different methods of volume measurement) are performed on each sample, resulting in 2n data points. Each of the n samples is then represented on the graph by assigning the mean of the two measurements to the abscissa and the difference between them to the ordinate. Primary application of the Bland-Altman plot is the comparison between two clinical measurements that each provide some errors in their measure, or the comparison between a new measurement technique or method with a gold standard [76]. In Bland-Altman analysis it is also common to compute the limits of agreement, usually specified as average difference ±2 standard deviations of the difference. Due to specific characteristic of the analysis treated in this Chapter, different confidence limits will be considered. They will be introduced and given reasons for in Section 4.
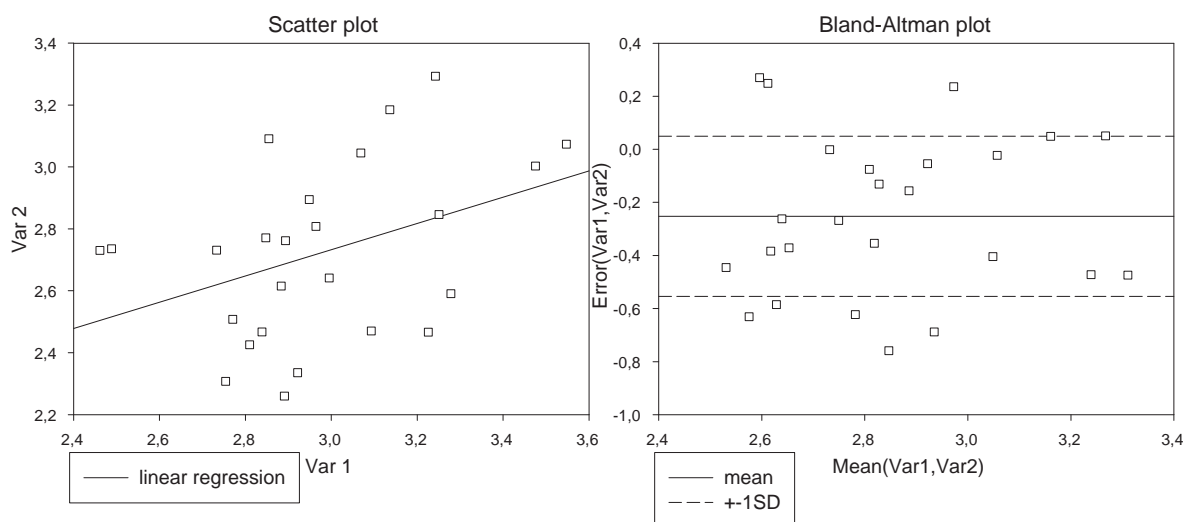

Figure 2.5: Graphical tools used in validation.

Complementarily to the above-mentioned plots, numerical tools of statistical significance are employed. Statistical hypothesis tests are methods for making decisions using data, whether from a controlled experiment or an observational study (not controlled). In statistics, a result is called statistically significant if it is unlikely to have occurred by chance alone, according to a pre-determined threshold probability called *significance level*. The phrase "test of significance" was coined by Ronald Fisher: "Critical tests of this kind may be called tests of significance, and when such tests are available we may discover whether a second sample is or is not significantly different from the first." [77].

In frequency probability, decisions about the data are almost always made using null-hypothesis tests. These are tests that answer the question "Assuming that the null

hypothesis is true, what is the probability of observing a value for the test statistic[1] that is at least as extreme as the value that was actually observed?" [78]. More formally, they represent answers to the question, posed before undertaking an experiment, of what outcomes of the experiment would lead to rejection of the null hypothesis for a pre-specified probability of an incorrect rejection. One use of hypothesis testing is deciding whether experimental results contain enough information to cast doubt on conventional wisdom. In this context, the *critical region* of the test is the set of all outcomes which cause the null hypothesis (H_0) to be rejected in favour of the alternative hypothesis (H_1).

Because of the characteristics of the data analyzed in this thesis, it will be made use of two-sample, paired tests. The attribute "two-sample" indicates that the test is based on a comparison of two samples, possibly of the same dimension, whereas the attribute "paired" means that the samples are dependent, that is, they consist in matched pairs of similar units or one group of units that has been tested twice. In particular, the *t*-test will be applied when datasets follow a gaussian distribution and the Wilcoxon signed rank test when only symmetry of the distribution can be guaranteed. When none of them can be applied, the two-sample Kolmogorov-Smirnov test will be used. Last, significance level α will be assigned a standard value of 5%.

3.1 Systems comparison

The most critical phase of the validation process consisted in assessing the accuracy of the commercial device's components that are directly involved in the transduction and the processing of input pressure and flow input signals. These can be summarized in the following two macro-blocks:

- front-end hardware;
- real-time processing algorithm.

To do so, specific testing protocols and processing approaches were defined. Two healthy young men and two healthy young women were tested for two weeks during working days with both systems (their physical characteristics are summarized in Table 2.1). The subjects were all university members and they did not have any prior clinical experience in pulmonary function tests. Each day, each subject made five acquisitions (at 5Hz, 6Hz, 8Hz, 10Hz, 5-11-19Hz) with one system and the same five acquisitions

---

[1] The statistic is value calculated from a sample, often to summarize the sample for comparison purposes.

with the other, making sure that at least 15 breaths were recorded ($\simeq$2 mins of spontaneous breathing). Altogether, these tests required about half an hour to be completed. Measurements were performed according to standard protocols [74], that is, subjects were seated with the head in a neutral or slightly extended position. Flexion of the head was avoided. During the measurements, subjects firmly supported their cheeks and the floor of the mouth using both hands and they wore a noseclip. The subjects were instructed to breathe quietly at FRC level. In addition, "free flow" mouthpieces and bacterial filters were interposed between patients' mouths and systems' inlets.

	Race	Sex	Age (years)	Weight (kg)	Height (m)
Subject #1	Caucasian	M	25	70	1.77
Subject #2	Caucasian	M	31	72	1.75
Subject #3	Caucasian	F	26	48	1.56
Subject #4	Caucasian	F	25	52	1.58

Table 2.1: Physical characteristics of the tested subjects. A large range of height was purposely chosen in order to obtain a wide span of values for the respiratory parameters (short subjects have greater resistance).

With regard to processing approaches three different datasets, listed below, were produced for any testing day:

- RESMONPROonline;
- RESMONPROoffline;
- REFERENCEoffline.

On the one hand RESMONPRO stands for "RESMON$^{\circledR}$ PRO, the commercial device" and REFERENCE indicates the "reference FOT system". On the other hand, the suffix online refers to results obtained from real-time processing and the suffix offline refers to results obtained from already acquired data. In more detail, in case of online processing most of the calculations are executed by the commercial device. These include calibration of raw signals, application of numerical filters, computation of lung mechanical impedance and breath recognition. Next, the desktop computer has the only duty to organize data in a way that permits a comparison with results obtained through offline processing. Offline processing is instead common to raw signals coming from both the commercial device and the reference system, and as already mentioned in Subsection 2.1 it differs from online processing by the fact that offline results are supposed to be more accurate. A small note that deserves to be mentioned is the fact that the pneumotachographs of the two systems follow two different laws, linear for

the reference system and non linear for the commercial one, therefore signal calibration must be adjusted coherently.

Comparison between the first two datasets stresses the effect of the online processing performed by the commercial device. Indeed, since the two algorithms are based on signals provided by the same flow and pressure sensors, any resulting difference must be ascribed to software aspects. On the contrary, comparison between the last two datasets serves the purpose of validating the front-end hardware. Provided that respiratory parameters do not change significatively throughout one single day [72], subsequent measurements that are executed by the two systems can be considered identical. As a consequence, different results might be ascribed to a different accuracy in "sensing" the mechanical characteristics of the tracheobronchial tree. Note that, as it will be clarified in the last part of this Subsection, this may not always be the case. Finally, comparison between the first and the last dataset puts together the foregoing concepts, so that an overall validation can be given.

Inter-day variability of respiratory parameters

The state of health of the respiratory system can be considered a variable with a time constant of some days, with some variability that depends on subjective peculiarities and most of all on the level of respiratory illness itself. Equivalently, one might say that the same conclusions can be drawn about respiratory parameters, as they are meant to describe the health status of the breathing apparatus. In reality other variables have to be taken into account, among which ventilation plays a leading role.

For reasons explained by fluid dynamics, the level of resistance offered by any kind of duct strongly depends on the speed of the fluid flowing inside it. Respiratory ducts make no exception, hence successive measurements could be considerably different if ventilation changes (see Section 4 for analytical description). For instance, a subject could get little fatigued when preparing for the test, so that ventilation slowly decreases during the completion of the test. However, much attention was given to the fact that patients were completely at rest just before starting the measurements. Also, a feeling common to all subjects is a sense of oppression when they first approach the device. This inevitably results in an increased ventilation that tends to go back to normal values after a sufficient number of tests. Last, a strong and systematic bias of ventilation can be given by a not efficient removal of the carbon dioxide residing in the high-inertance path after exhalation.

In order to increase the understanding of the dynamics governing inter-day variability of respiratory parameters, a comparison between REFERENCEoffline datasets

belonging to subsequent days is made. The reference system is preferred to the commercial device as it is considered more accurate, and only adjacent days are included in the comparison (e.g., Friday-Monday is rejected). Although a little change in the values of respiratory parameters can take place in 24 hours, the two datasets are expected to be practically identical because this fourth processing approach is defined in order to minimize the possibility of difference between them, and also because the subjects were in perfectly healthy throughout the observational time.

3.2 Test objects' measurement

An additional way to verify that the front-end hardware of both systems works properly consists in applying known impedances to the inlets and measuring them. In reality this was already assessed with the verification for the commercial device and with periodical recalibrations for the reference system, therefore another interesting test consists in comparing measurements of the same test object, possibly different from the one used in the calibration phase. Figure 2.6 shows two test objects that were acquired with both systems. The first is basically a pure resistance, as it is composed by 4 meshes placed in sequence and very close to each other, whereas the second is an impedance very similar to that used for calibration.

Figure 2.6: Test object #1 (left) and test object #2 (right).

Like in the testing protocols for systems comparison the objects were measured five times (at 5Hz, 6Hz, 8Hz, 10Hz, 5-11-19Hz) with both systems. Each acquisition lasted

about two minutes in order to provide a good amount of data. These pressure and flow signals are processed with the same offline algorithm, so that resistance and reactance traces are obtained. Moreover, mean values of resistance and reactance are calculated for both objects. Results of these tests will be illustrated in Section 4.

3.3 Breathing pattern comparison

The speed of the air flowing through the breathing airways has a strong impact on their mechanical properties, especially with regard to resistance. In particular, as it will be demonstrated both analytically and empirically in Section 4, the relationship between flow and resistance can be considered linear.

Although the commercial device's aspirator is dimensioned for expelling all the exhaled carbon dioxide, a comparison between the pulmonary ventilation required by the two aspirators is made. Table 2.2 lists their characteristics. The much wider space occupancy of the external aspirator, given by the bigger dimensions of the double-piston and the damping chambers, is compensated by a flow which is larger and more constant throughout the motor revolution, such that air is removed more efficiently and less noise is introduced in the hydraulic circuit.

System	Aspirator type	Input voltage (V)	Flow rate (L/min)	Mechanical RC filter efficiency
RESMON® PRO	single piston	~14,3	~12	low
REFERENCE	double piston	~14	~19	high

Table 2.2: Characteristics of the aspirators the systems are equipped with. Values of flow rate were preliminarily assessed by connecting a flow sensor to their air outlets.

The ventilation of four healthy subjects whose physical characteristics are reported in Table 2.3 was measured four times, during quiet breathing, making use of the prototype. In order to exclude the FOT, the motherboard was deactivated and external flow and pressure sensors were applied to the embedded pneumotachograph (see Figure 2.7(a)). Tests were single-blinded, in fact both aspirators were turned on but the high-inertance path was randomly connected to one of them (twice to the internal and twice to the external), and subjects were not revealed which aspirator was actually connected. Figure 2.7(b) offers a view of all the components, in which a paper cover hides the real

connection inside the prototype's case.

	Race	Sex	Age (years)	Weight (kg)	Height (m)
Subject #1	Caucasian	M	25	70	1.77
Subject #2	Caucasian	M	31	72	1.75
Subject #3	Caucasian	F	30	56	1.73
Subject #4	Caucasian	F	25	48	1.56

Table 2.3: Physical characteristics of the tested subjects.

4 Results

Results here presented are based on datasets purposely created to validate the front-end hardware of the commercial device. Moreover, results are often represented through scatter plots and Bland-Altman plots. All these tools are extensively illustrated in Section 3, therefore they will not be introduced again. The only detail that must be mentioned concerns the confidence limits of Bland-Altman plots. Whereas standard deviations give an idea of the dispersion of the points, confidence limits indicate the limits beyond which the difference between two measurements is not acceptable anymore. As defined in the ERS guidelines for the application and interpretation of FOT [74], a maximum error of 10% or $0,1 cmH_2OsL^{-1}$, whichever is greater, is allowed over the frequency range of interest. These boundaries give origin to the "ribbon-shaped" confidence area that can be observed for reactance Bland-Altman plots. Note that only the right side of the ribbon is visible for parameters that are far from zero, that is, all the parameters except for reactance.

Also, results are presented in a way that reflects the evolution of the validation process.

4.1 Calibration stability

Calibration was periodically verified throughout the time span in which measurements were performed. Also, a final check was made two months after the first calibration. Table 2.4 indicates that the front-end maintains the same characteristics over a span of at least two months. The first verification failed because a pipe connecting a sensor to the pneumotachograph was squeezed. Such episodes are due to not proper assembling of the components and can be solved easily. The stability of the results suggests that calibration should remain valid for much longer periods.
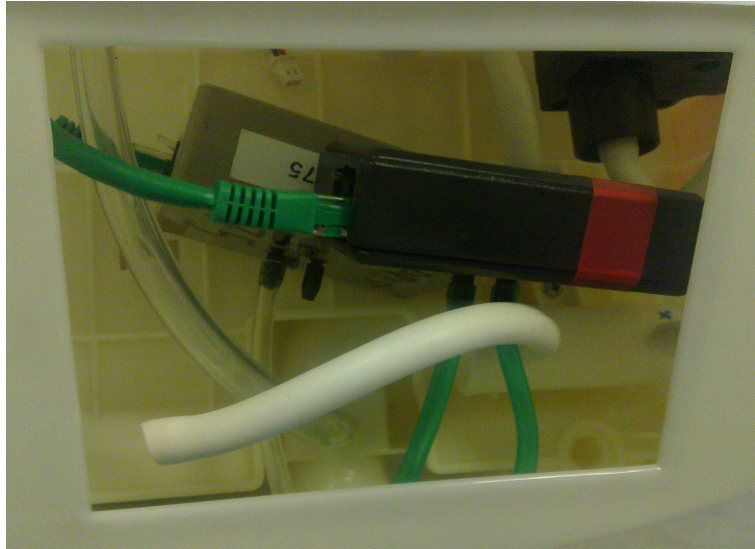
(a) External sensors and aspirators connection. The gray and the black boxes are, respectively, pressure and flow sensors. A situation in which the external aspirator (clear pipe) is connected to the high-inertance path and the internal aspirator (white pipe) collects ambient air is shown.
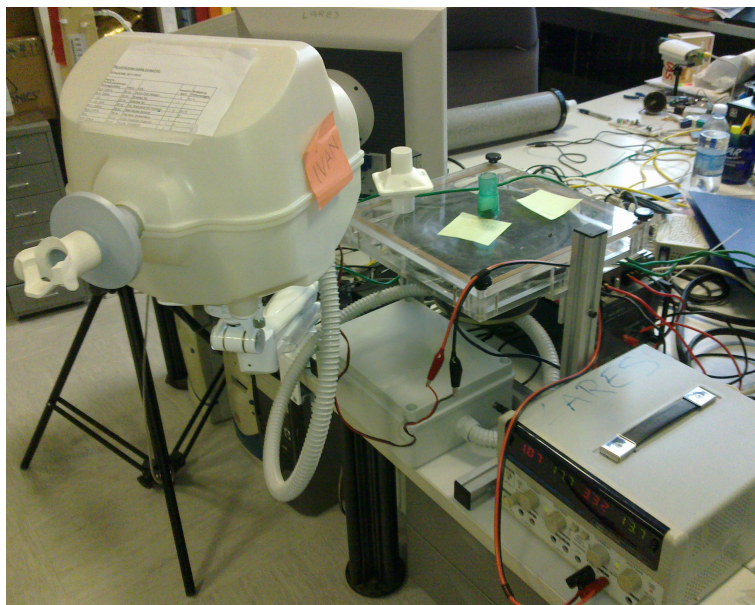
(b) Snapshot of the components. Prototype with paper cover, double piston external aspirator (grey box with wavy pipe) and 14V power supply are shown.

Figure 2.7: Experimental setup for breathing pattern comparison.

Date	Coefficients	Actual resistance	Measured resistance	Actual reactance	Measured reactance	Validation result
11-nov-11	M (cmH$_2$O*s/(L*Hz))	-0,02	-0,02	0,16	0,15	FAIL
	Q (cmH$_2$O*s/L)	2,67	2,36	0,19	0,07	
16-nov-11	M (cmH$_2$O*s/(L*Hz))	-0,02	-0,02	0,16	0,16	SUCCESS
	Q (cmH$_2$O*s/L)	2,67	2,61	0,19	0,24	
21-nov-11	M (cmH$_2$O*s/(L*Hz))	-0,02	-0,02	0,16	0,15	SUCCESS
	Q (cmH$_2$O*s/L)	2,67	2,58	0,19	0,37	
23-nov-11	M (cmH$_2$O*s/(L*Hz))	-0,02	-0,02	0,16	0,17	SUCCESS
	Q (cmH$_2$O*s/L)	2,67	2,63	0,19	0,12	
30-nov-11	M (cmH$_2$O*s/(L*Hz))	-0,02	-0,03	0,16	0,17	SUCCESS
	Q (cmH$_2$O*s/L)	2,67	2,73	0,19	0,21	
13-gen-12	M (cmH$_2$O*s/(L*Hz))	-0,02	-0,02	0,16	0,17	SUCCESS
	Q (cmH$_2$O*s/L)	2,67	2,55	0,19	0,25	

Table 2.4: List of the verifications carried out during the study. Coefficients indicate slope and intercept of the linear resistance-frequency and reactance-frequency relations characterizing the test object.

4.2   Validation of the front-end hardware

A comparison between inspiratory mean resistance and inspiratory mean reactance belonging to the datasets RESMONPROoffline and REFERENCEoffline is shown in Figure 2.8. Since the remaining respiratory parameters follow similar relationships, they are not visualized. Resistance has a regression coefficient very far from the ideal one ($m=0,27$cmH$_2$O^{-1}), whereas reactance shows a much better regression slope ($m=0,90$cmH$_2$O^{-1}) thanks to the fact that its frequency-dependence widens the span of values covered and because of is little ventilation-dependence, a concept that will be soon introduced. Nevertheless, points often fall outside the confidence area, making the quality of the results unsatisfactory.

Such an inconsistency between the two measurements can be ascribed to several causes. Besides a possible poor accuracy of the front-end hardware, the inter-test variability of the parameters in question could play a lead role. This intrinsic variability is surely increased by the different number of breaths recognized in each test, which inevitably affects the resulting parameters' mean values. Furthermore, a different level of ventilation can heavily affect the data, especially with regard to resistance.

For the mentioned reasons a comparison between successive acquisitions belonging to REFERENCEoffline is performed. Since the measurement instrument is the same and the compared acquisitions refer to subsequent days (e.g., Monday-Tuesday), points are expected to lie on the straight line $y = x$. Figure 2.9 demonstrates this is not the case, in fact the linear regression is still far from the ideal one ($m=0,74$cmH$_2$O^{-1}), and also a fair amount of points keeps falling outside confidence limits.

Results shown so far indicate that conclusions about the difference in accuracy between the front-end hardware of the two systems are strongly affected by inter-test variability. This phenomenon "masks" the actual disparity due to the different measurement systems, thus the information content of the data available is heavily weakened.

A different approach that aims at verifying how similarly the two front-ends measure hydraulic impedances is applying the same test objects to their inlets. Table 2.5 refers to the measurement of the objects depicted in Figure 2.6. Small differences imputable to a slightly wrong calibration of one or both front-ends are observable for Test obj #2, whereas significatively different values are present for Test obj #1. In this case only the reference system measures a reactance close to zero, which is the value expected. The commercial device provides a reliable evaluation at 5Hz, but tends to overestimate the reactance value for increasing frequencies.

In Figure 2.10 traces are shown visually for the case of multifrequency, which is
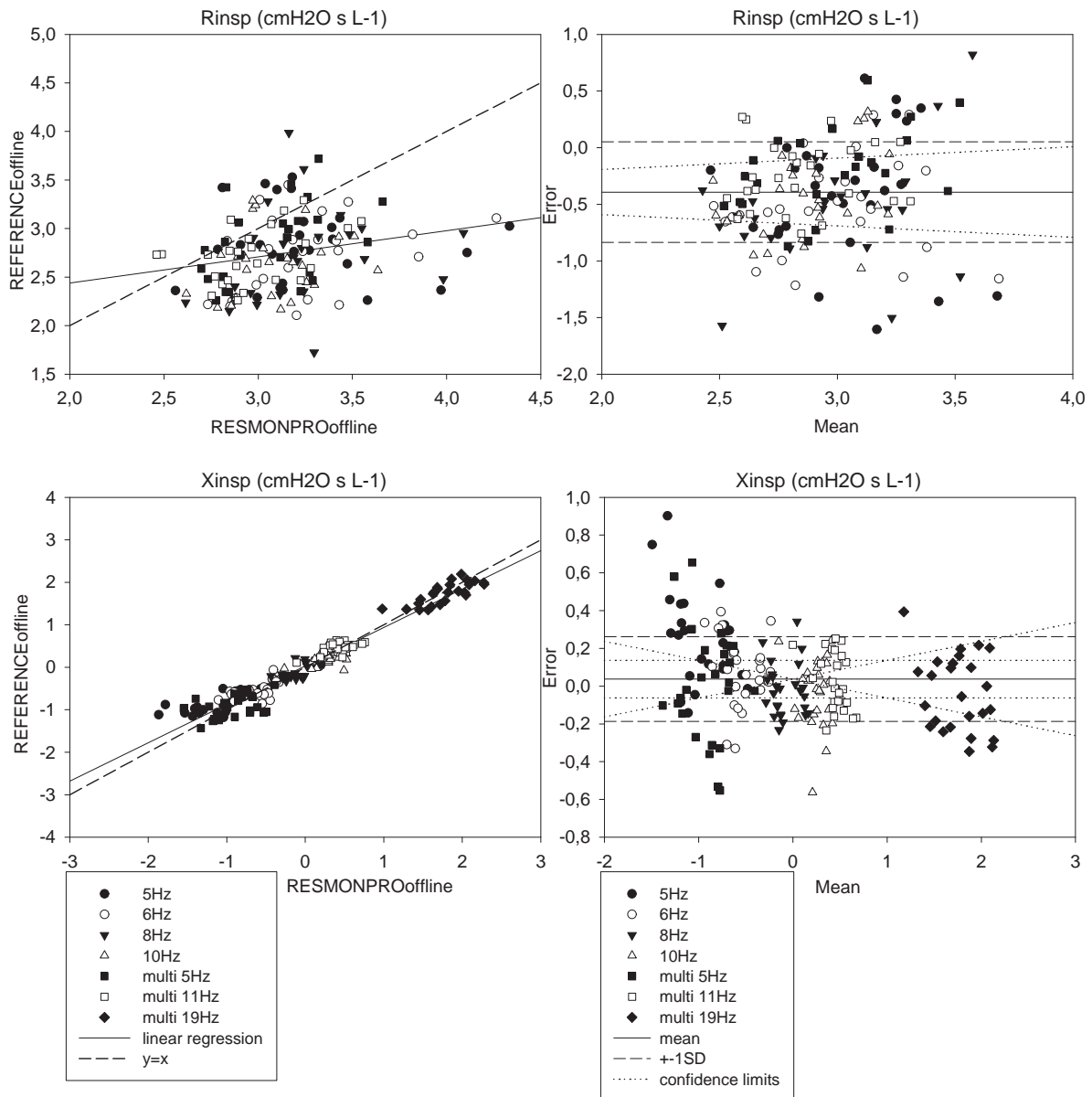
Figure 2.8: Comparison between inspiratory resistance (Rinsp) and reactance (Xinsp) acquired with both systems and processed with the same offline algorithm. Measures for each patient are included in order to increase the range of values covered (women have greater resistance).
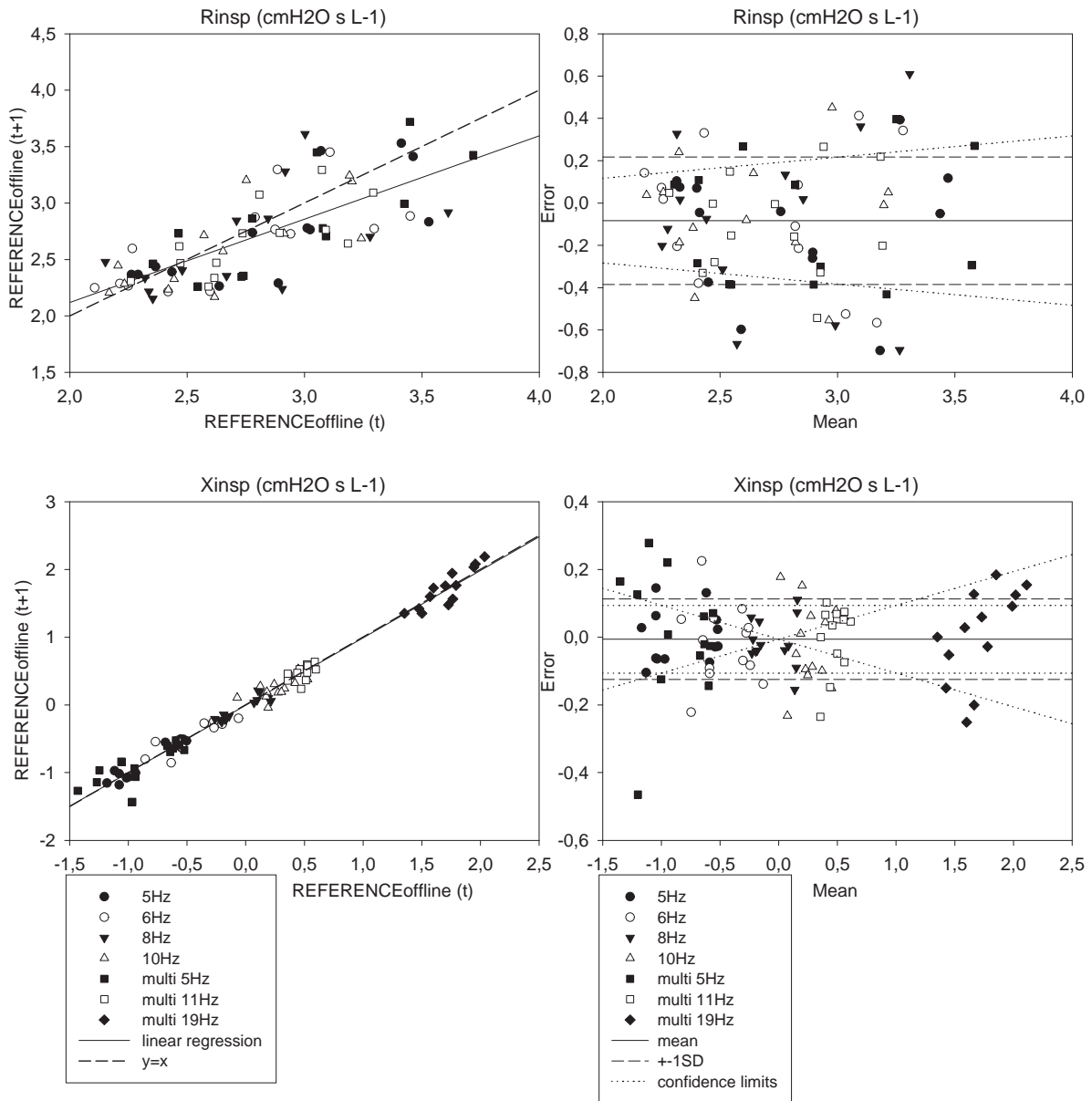
Figure 2.9: Comparison between inspiratory resistance (Rinsp) and reactance (Xinsp) acquired with the reference FOT systems in subsequent days and processed with the same offline algorithm. Measures for each patient are included in order to increase the range of values covered (women have greater resistance).

	Stimulus (Hz)	Rmean (cmH2O*s/L)			Xmean (cmH2O*s/L)		
		Commercial	Reference	Difference	Commercial	Reference	Difference
Test obj #1	5	7,098	7,347	-0,249	-0,039	0,054	-0,093
	6	7,624	7,315	0,309	0,496	0,087	0,409
	8	7,796	7,316	0,480	0,831	0,098	0,732
	10	7,856	7,301	0,554	1,036	0,161	0,875
		6,061	7,397	-1,336	1,740	0,078	1,662
	5-11-19	7,007	7,392	-0,385	1,315	0,180	1,135
		7,045	7,442	-0,398	1,467	0,271	1,196
Test obj #2	5	2,502	2,554	-0,052	0,868	0,885	-0,016
	6	2,524	2,589	-0,065	1,151	1,058	0,093
	8	2,514	2,597	-0,084	1,540	1,429	0,111
	10	2,597	2,629	-0,032	1,901	1,778	0,124
		2,258	2,554	-0,296	1,028	0,884	0,144
	5-11-19	2,321	2,626	-0,305	2,005	1,946	0,059
		2,109	2,763	-0,654	3,347	3,401	-0,054

Table 2.5: Mean resistance (Rmean) and reactance (Xmean) of a pure resistance (Test obj #1) and an impedance (Test obj #2) measured with both systems. Difference between the mean values are indicated.

way far the worst case because the overlapping of different pressure waves worsens the signal-to-noise ratio. The mesh placed in the prototype's pneumotachograph is much less resistive than that of the reference system, dimensioned for pediatric use. This results in a superimposed noise which, as expected, has a greater but acceptable magnitude for Test obj #2. Different is the situation for Test obj #1. In this case noise magnitude is rather marked and also the mean value of reactance at all frequencies is overestimated. Reasons for such a behaviour might be in part due to the fact that Test obj #2 was directly connected to the reference system's inlet, but it needed a couple of additional manifolds to adapt its diameter to that of the commercial device's inlet. As a consequence, the increased dead space inevitably results in an increased value of inertance of the resulting hydraulic load. However, this effect alone cannot motivate a bias and a variability of the reactance of such a magnitude. Other reasons could be represented by the scarce design quality of Test obj #1 and also on the bad seal given by its manifold.

The possible malfunctioning of the commercial device must be taken into account,

but results on the accuracy of the front-end hardware are encouraging for the only certified test object, namely Test obj #2.

4.3   Breathing pattern comparison

As anticipated in Section 3, ventilation is positively correlated with respiratory resistance. Figure 2.8 suggests that ventilation is greater when subjects breathe in the commercial device, as for a given test on the scatter plot the value of resistance along the abscissa is higher than the corresponding on the ordinate. Specific tests aimed at assessing the level of ventilation when the internal and the external ventilators are alternatively connected to the high-inertance path of the commercial device demonstrate that the internal aspirator is not sufficient to expel all the exhaled carbon dioxide (see Figure 2.11). In particular, subjects tend to keep the same tidal volume but increase the respiratory rate by 7% and ventilation by 9% when the internal aspirator is connected.

For completeness, a test of significance is applied to the breathing pattern parameters of the originally compared datasets, namely RESMONPROoffline and REFERENCEoffline. The consideration made at the beginning of this Subsection, that is, Figure 2.8 suggests that the commercial device causes a greater ventilation, finds further support in Table 2.6. Again, the commercial device increases ventilation by 22% but in this case the cause is tidal volume in place of respiratory rate. This might be due to the more resistive mesh of the reference system, which induces subjects to take deeper breaths. Moreover, a comparison between the two percentages concerning ventilation cannot be compared because of the following reasons:

- datasets for breathing pattern comparison are much smaller;
- subjects who participated in the two tests are not the same;
- tests were made with different experimental setups.

4.4   Validation of the real-time processing algorithm

Figure 2.12 offers a comparison between inspiratory mean resistance and inspiratory mean reactance belonging to the datasets RESMONPROonline and RESMONPROoffline. The remaining respiratory parameters are not visualized as they follow similar relationships. Both resistance and reactance generate a regression line very close to the bisector of the first quadrant ($m=1{,}05\text{cmH}_2\text{O}^{-1}$ and $m=1{,}01\text{cmH}_2\text{O}^{-1}$, respectively) and most of the points fall inside confidence limits. The small amount of outliers is given by the fact that the online algorithm implements numerical filters of lower order
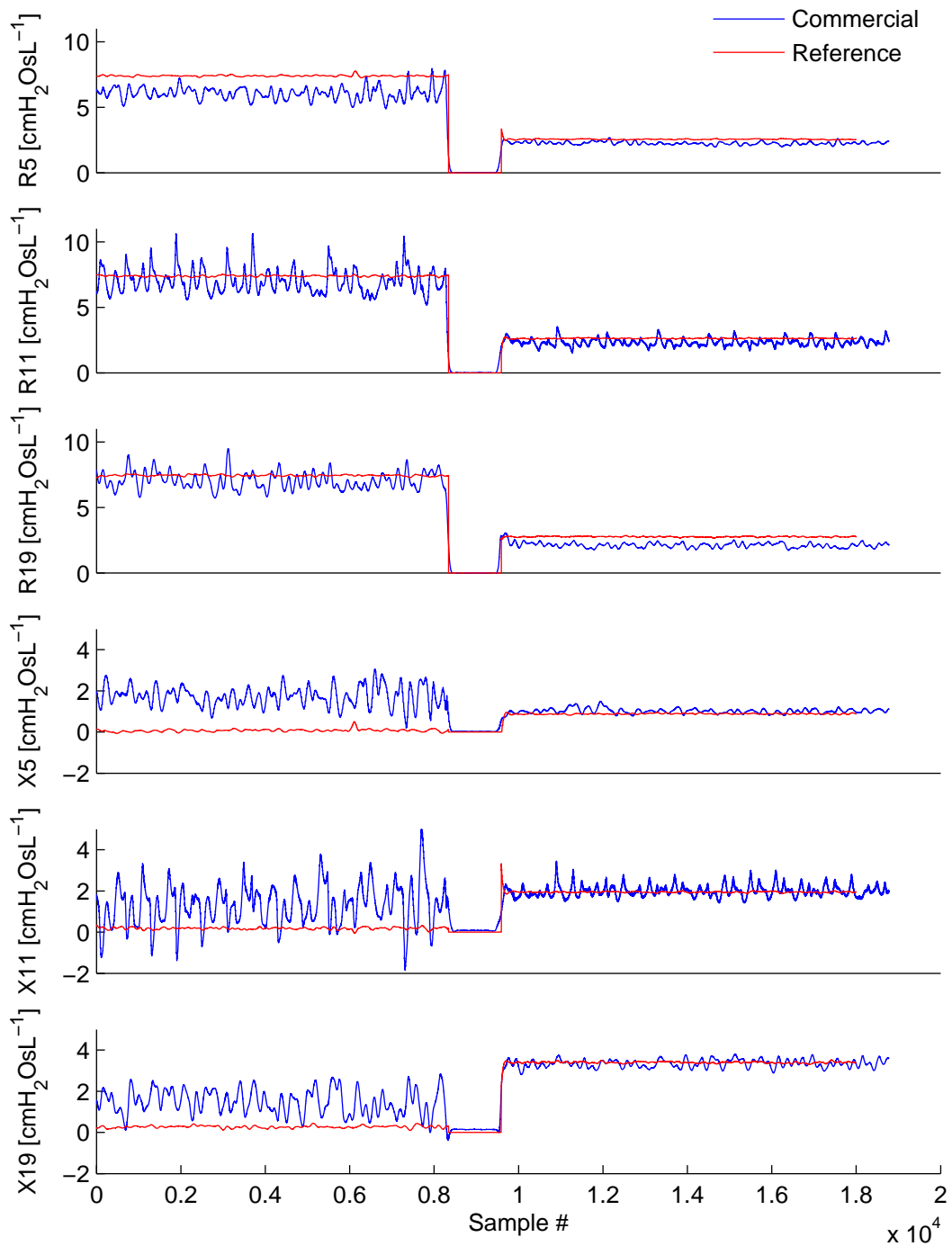
Figure 2.10: Traces of resistance and reactance for the multifrequency stimulus. Test obj #1 was measured first, then a zero resistance and reactance window indicates the moment in which it was removed in order to connect Test obj #2, which was measured last.
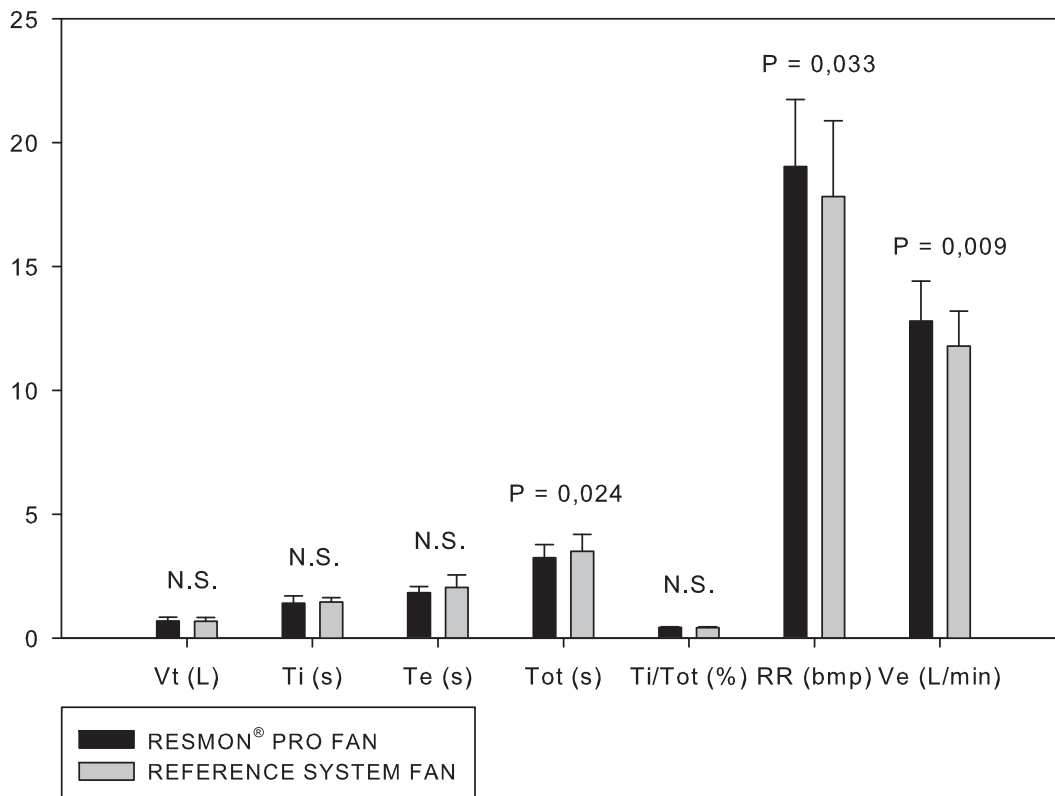
Figure 2.11: Error bar plot comparing the most important parameters of breathing pattern. Columns and error bars indicate, respectively, mean value and 1 standard deviation of the semi-dataset they refer to. Each dataset is composed by 16 values, 8 of them referring to the internal aspirator (RESMON® PRO FAN) and 8 to the external aspirator (REFERENCE SYSTEM FAN). Percentages indicate the P-value of a paired *t*-test applied to the two semi-datasets (α=5%). N.S. means "Not Significant".

and its automatic breath recognition has a bug, but the incidence of such wrong estimations can be considered negligible. Results indicate a successful validation of the real-time processing algorithm.

The bug in the automatic breath recognition, that has to be fixed, is illustrated in Figure 2.13. Automatic breath detection is based on the recognition of zero-crossing points in flow traces. In particular, when the transition is from negative to positive values a point of end expiration is identified and viceversa. Manifestations of the bug consist in points of end expiration mistaken for points of end inspiration and are caused by a local sequence of flow samples that occur with a very low incidence. Due to this, the following calculation of mean breath-by-breath respiratory parameters results affected by an error which is proportional to the number of mistaken breaths.
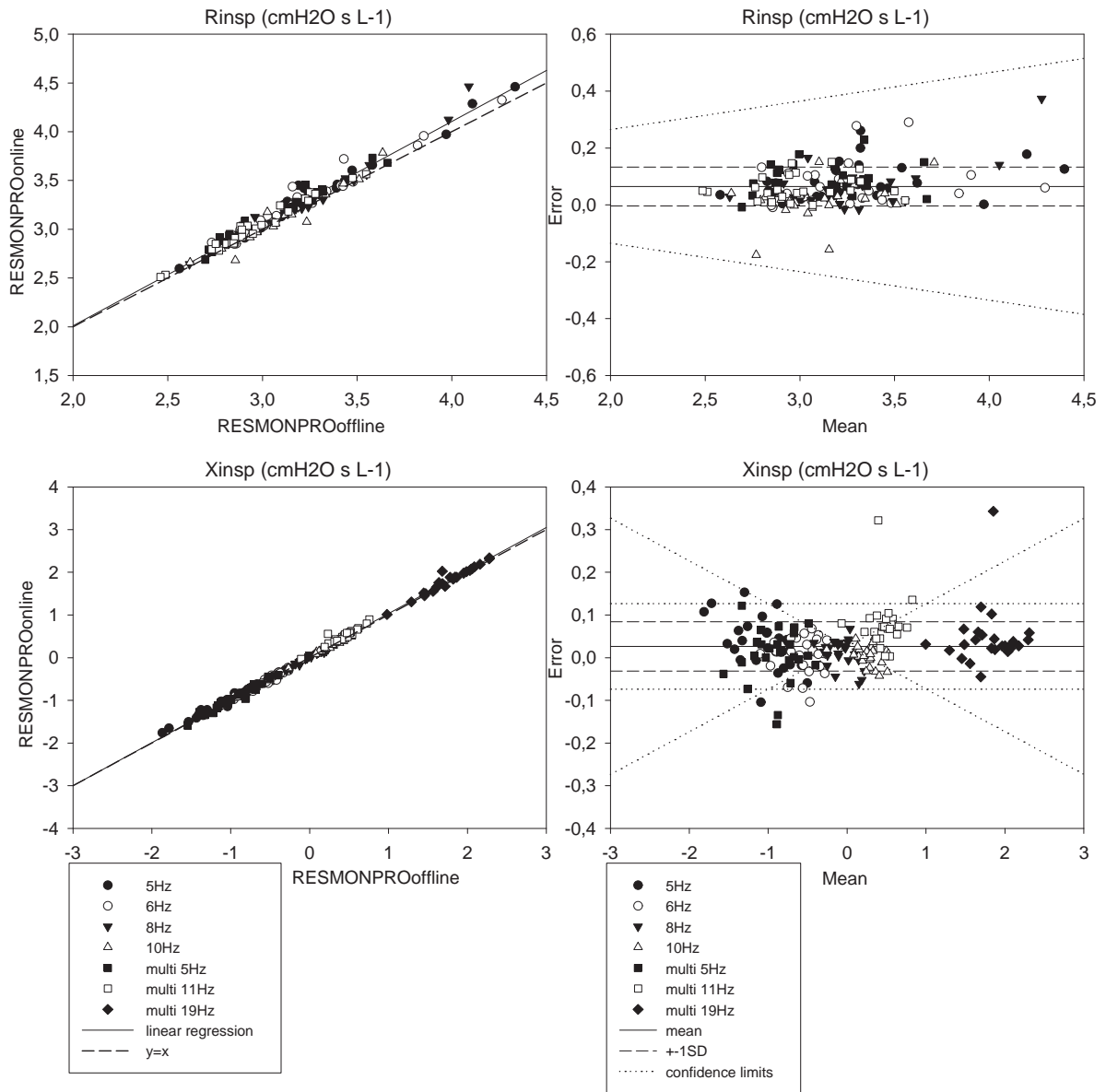
Figure 2.12: Comparison between inspiratory resistance (Rinsp) and reactance (Xinsp) acquired with the commercial device but processed with online and offline algorithms. Measures for each patient are included in order to increase the range of values covered (women have greater resistance).

| Parameter | Normality test | Significance test | Mean (Median) | | P-value |
			RESoffline[*]	REFoffline[**]	
Ve (L/min)	Passed (P = 0,874)	t-test	11,316	9,275	<0,001
Vt (L)	Passed (P = 0,124)	t-test	0,671	0,550	<0,001
RR (s)	Failed (P < 0,050)	signed rank test	16,625	16,949	0,441

Table 2.6: Significance tests applied to Ve, Vt and RR belonging to RESMONPROoffline(*) and REFERENCEoffline(**).
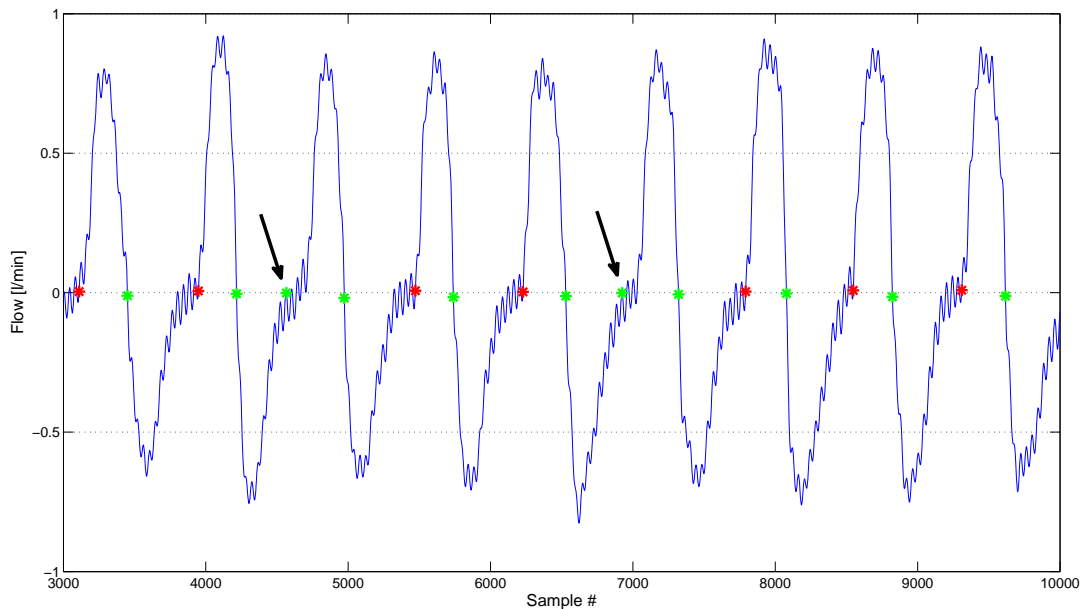

Figure 2.13: Automatic zero-crossing recognition for a flow trace. Green and red asterisks spot identified end-inspiratory and end-expiratory points, respectively. Black arrows indicate erroneous recognitions.

4.5 Overall validation

Considering the datasets defined in Subsection 3.1, of interest is the third and last comparison, namely RESMONPROonline vs. REFERENCEoffline. It gives a general and comprehensive validation of the commercial device, and expectedly will be a composition of the previously performed comparisons.

Results of the comparison are very similar to those of Figure 2.8, with regression coefficients little worsened by the small error introduced by the real-time processing

algorithm (m for Rinsp goes from 0,27 to 0,23cmH$_2$O^{-1}, whereas m for Xinsp goes from 0,90 to 0,89cmH$_2$O^{-1}). As previously demonstrated, at least part of this discrepancy is caused by the following "masking" phenomena:

- inter-test variability;

- positive correlation between ventilation and resistance.

Whereas the first point is not avoidable, the second can be canceled discovering the type of relation linking ventilation and resistance and normalizing data accordingly.

Pressure drop (or head loss), occurs in all piping systems because of the friction of the fluid's molecules among them and against the surface of the duct. Considering a stationary current of flow rate Q inside a straight duct of constant section, the general expression of head loss can be expressed by the Darcy-Weisbach equation:

$$\Delta p = \lambda \frac{L}{D} \frac{\rho V^2}{2} \tag{2.1}$$

where the pressure loss due to friction Δp (units: Pa or kg/ms^2) is a function of:

- the ratio of the length to diameter of the pipe, L/D;

- the density of the fluid, ρ (kg/m^3);

- the mean velocity of the flow, V (m/s);

- *Darcy friction factor*, a (dimensionless) coefficient of laminar, or turbulent flow, λ.

λ, also known as *flow coefficient*, is not a constant and depends on the parameters of the pipe and the velocity of the fluid flow, but it is known to high accuracy within certain flow regimes. It may be evaluated for given conditions by the use of various empirical or theoretical relations, or it may be obtained from published charts (Moody diagrams).

For laminar flows, it is a consequence of Poiseuille's law that $\lambda = A/Re$, where Re is the Reynolds number.

For turbulent flow friction forces are negligible with respect to turbulence effects, so that the friction factor results constant. Methods for finding it include using a diagram such as the Moody chart or solving equations such as the Colebrook-White equation, or the Swamee-Jain equation.

As a consequence it can be demonstrated that, in case of laminar flow, (2.1) becomes

$$\Delta p = K_1 V \tag{2.2}$$

whereas this dependence is quadratic for turbulent flow

$$\Delta p = K_2 V^2. \tag{2.3}$$

Deriving 2.2 and 2.3 with respect to V, the following equations for resistance are obtained:

$$\frac{\Delta p}{dV} = K_1 = R \tag{2.4}$$

$$\frac{\Delta p}{dV} = K_3 V = R \tag{2.5}$$

The tracheobronchial tree is a very complex hydraulic system therefore one global Reynolds number, or equivalently one resistance-flow relation, is not definable. Flow goes from turbulent in the bronchial tube to laminar towards alveoli, therefore the total Reynolds number is expected to be weighted by both regimes. Figure 2.14, showing the resistance-ventilation relation of a single subject, solves this in an empirical way. Note that only one subject must be taken into account otherwise the superimposition of different resistance-flow relations would alter the results. Though data are scarce in number they are enough to confirm a linear dependence, which means that turbulences present in the central airways are dominant.

The idea is applying this concept to the overall validation in order to cancel the masking effect given by the positive correlation between ventilation and resistance. Indeed, a normalization of the respiratory parameters with regard to ventilation is supposed to eliminate this dependance. Figure 2.15 depicts an important improvement for both resistance and reactance, as m moves from 0,23 to 1,05cmH$_2$O^{-1} and from 0,89 to 1,05cmH$_2$O^{-1} ,respectively. The situation is improved for Bland-Altman analysis as well, although a good amount of points keeps falling outside the confidence limits. Unfortunately, it is not possible to discriminate how much of the incongruence can be assignable to inter-test variability and how much is due to the error introduced by the commercial device.

However, the calculation of the coefficient of variation (CoV) of the time series of Rtot5 belonging to the datasets RESMONPROoffline and REFERENCEoffline shows how the commercial device provides a measure of the variability which is more similar to that of previous studies [79, 80]. Given any of the healthy subjects of Table 2.1 and the daily mean values of Rtot5 for the datasets RESMONPROoffline and REFER-ENCEoffline, the CoV is calculated as the ratio between the total standard deviation and the total mean. Considering all the subjects, the mean $\pm$ standard deviation of the CoV for RESMONPROoffline is 11.2%$\pm$4.3%, whereas it is 7.2%$\pm$2.7% for REF-ERENCEoffline. Results by Gimeno et al. [79] and by Dellaca et al. [80] indicate a CoV of 10.8%$\pm$3.2% and 15.4%, respectively.
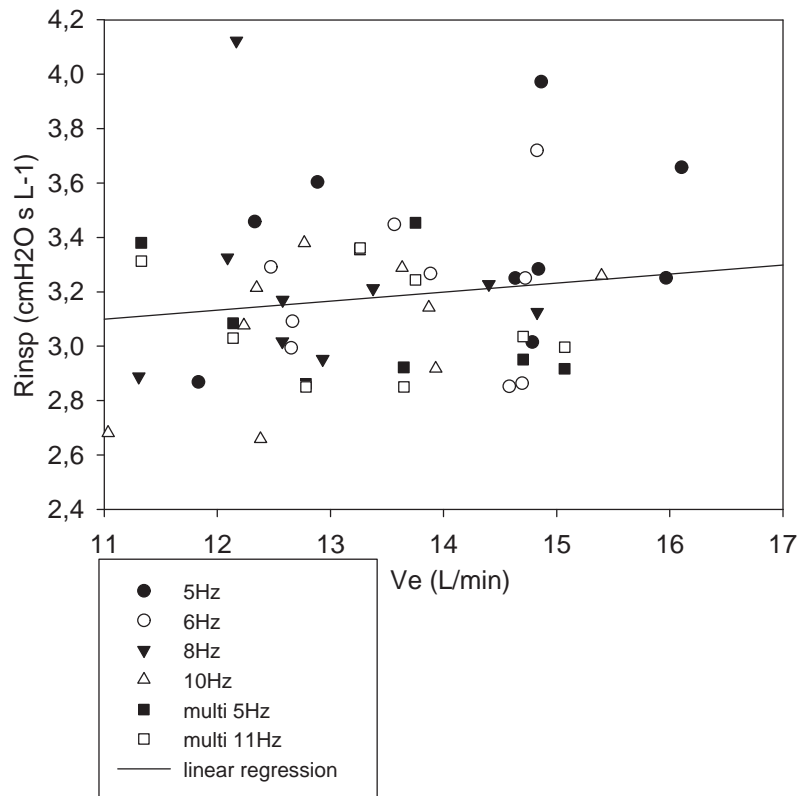
Figure 2.14: Resistance-ventilation relation of subject #1 (see Table 2.1).

5   Conclusions

In this Chapter a validation analysis of a commercial FOT device was performed. Results were obtained thanks to a comparison between measurement performed by the commercial device and measurements obtained by a reference FOT system taken as a gold standard. Through properly designed protocols the accuracy of the various macro-components of the device was investigated.

Results show that the signal processing software is fully validated. The error introduced by real-time IIR digital filters, which implies a non-deterministic delay between some of the processed signals, and that introduced by internal registers and data memory cells, which is caused by a non-optimal numerical approximation in the internal calculations (compared to those provided by a general purpose desktop computer used as a reference), affect the final output by a little bias. The magnitude of this bias is compliant with the ERS guidelines for the application and interpretation of FOT [74].

Results regarding the front-end hardware are more controversial. Direct validation of the front-end based on lung assessments is not implementable because of physiologic inter-test variability. Indeed, successive measurements performed by the commercial
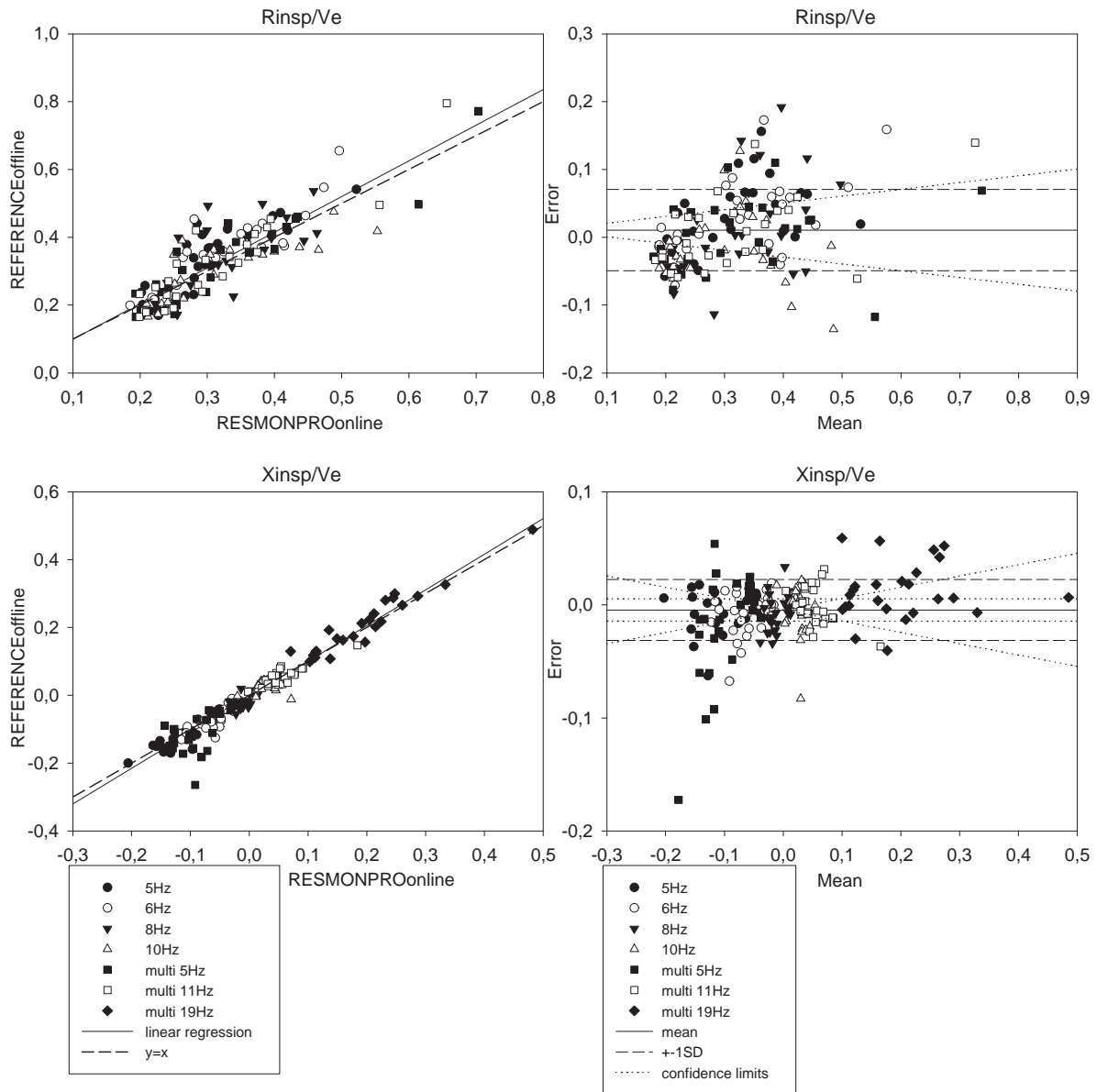
Figure 2.15: Comparison between inspiratory normalized resistance (Rinsp) and reactance (Xinsp) acquired with both systems and processed with the online algorithm in case commercial device and the offline algorithm in case of reference system. Normalization is performed with respect to ventilation. Measures for each patient are included in order to increase the range of values covered (women have greater resistance).

device and the reference system are systematically affected by a slightly different status of the lungs of the observed subjects. As a consequence, differences between values obtained by the two front-ends are composed by a summation of effects which is not divisible in its components. Several clues suggest that the actual difference in accuracy between the two systems is rather small, but this is not directly provable. Moreover, values for the CoV of time series obtained by the commercial device confirm results of previous studies.

Since the final aim of this work is observing how the characteristics of COPD evolve over time, a possible shift of the measured parameters would not strongly affect results. Also, time series of COPD patients display a marked variability. Given its magnitude, this variability is supposed to be larger than that erroneous introduced by the device by many orders.

Another comment regards the efficiency of the aspirator needed to expel the exhaled carbon dioxide. Results show that subjects tend to increase the respiratory rate because of the not complete extraction of the carbon dioxide present in the high-inertance path after exhalation. Therefore an alternative solution able to guarantee a greater flow rate should be taken into account.

Chapter 3

Home monitoring of COPD and data analysis

1 Introduction

Chronic obstructive pulmonary disease (COPD) is a chronic respiratory disorder characterized by the presence of airflow limitation and sudden exacerbations of the symptoms during which a severe inflammatory process occurs [13]. Such aggravations recur in a periodic fashion and often require hospitalization.

The dynamical evolution characterizing the COPD brings to abnormal rhythms and fluctuations of several respiratory parameters. A speculation proposed is that a long-term observation of objective parameters may help to characterize the progression of the disease over time and, possibly, early detect the onset of extreme future events.

In this Chapter results of analysis performed on clinical data are presented. The observational clinical study utilized as source of data moves from the above mentioned speculation. Its purpose is to measure daily variability of FOT data measured at home of a group of COPD patients in order to identify possible correlations between symptoms change, breathing pattern, lung mechanical impedance and occurrence of exacerbation. The estimated enrollment of 80 patients and the time frame consisting in daily assessments for 6 to 8 months make this study the largest and most detailed clinical trial ever defined for the observation of COPD.

In particular, the eligibility defines the enrollment of patients of age ranging from 40 to 80 years with characteristics of frequent exacerbators (>2 exacerbations in the past year).
Inclusion criteria are:

- COPD at stage 3 and 4 of GOLD classification (spirometric values after bron-

chodilator: $FEV_1/VC<$95th percentile of predicted and $FEV_1<$50% of predicted)

- patients who reported more than two exacerbations in the past year OR

- patients who required more than two hospital admission in the last year OR

- patients with ER admission in the last year due to acute respiratory failure

better if:

- depressive phenotype

- worsening of dyspnea during walk (measured by MRC-Medical Research Council score)

- malnutrition or obesity (BMI<19 or >25)

- patients with ER admission in the last year due to acute respiratory failure

- patients live alone

Exclusion criteria are:

- other respiratory diseases

- alpha-1 antitrypsin deficiency

- significant inflammatory diseases other than COPD

- organ or systemic diseases that may impair the ventilatory function (any restrictive pulmonary disease, cystic fibrosis and so on)

- prior lung surgery

- concomitant enrollment in other trials

- any major non-COPD disease or condition, such as uncontrolled malignancy, end-stage heart disease, liver or renal insufficiency (that requires current evaluation for liver or renal transplantation or dialysis), amyotrophic lateral sclerosis, or severe stroke, or other as deemed appropriate by investigator as determined by review of medical history and/or patient reported medical history

Whenever a new patient is enrolled, he/she is given an ID code that guarantees anonymity. It increases on the basis of the date of enrollment. Results will be shown for the first 8 patient enrolled: BC001, BC030, BC031, BC032, BC033, BC034, BC035 and BC036.

2 Methods

2.1 Data collection and processing

The network architecture adopted in the clinical study is derived from typical and well consolidated solutions for telemedicine systems and is adjusted to the specific features of the FOT device (RESMON® PRO, introduced in Chapter 2). A process running on the Linux-based operating system manages the data communication with a central data server [81]. During the transmission, data integrity and confidentiality are provided by end-to-end encryption via the standard TLS protocol, which also ensures mutual authentication between device and central server via public-key certificates. Public key cryptography is a fundamental and widely used technology which is employed by many cryptographic algorithms to securely transmit the data over the Internet. Its peculiar characteristic is the use of asymmetric key algorithms, which are used to create a mathematically-related key pair: a secret private key and a public known key. Whereas the public key is easily created from the private, the inverse transformation is almost impossible. No initial exchange of cryptographic keys is needed between the sender (the FOT device) and the recipient (the data server). The presence of a double-key mechanism allows protection of the confidentiality and integrity of a file, encrypting the message using the public key of the server, which can only be decrypted by the server itself using its own private key. Moreover, the private key of the FOT device is used to create a digital signature and thus allows protection of the authenticity of a file, which can then be easily verified by the server using the public key of the FOT device to decrypt the file.

The TCP/IP network connection to the data server is always established by using secure protocols (SSH) but over different physical layers, telephone landlines, mobile phone networks (GPRS), DSL or wireless broadband connections (Wi-Fi), whichever is available at patient's house (Figure 3.1). Through the data transmission, which is scheduled immediately after the test, the central server receives new (encrypted) files over a specific SSH socket, verifies the authenticity of the sender and the integrity of the data and then updates the internal database.

The data server is configured as a one-tier (single processor) hardware architecture and runs a bash script each time a FOT device creates a connection and wants to send its data. This script extracts the files from the archive (i.e., the binary file containing raw pressure and flow and an xml-file filled with respiratory parameters elaborated by the device, besides other supplementary files) and verifies that they are not corrupted before updating the internal database. The FOT device waits for the
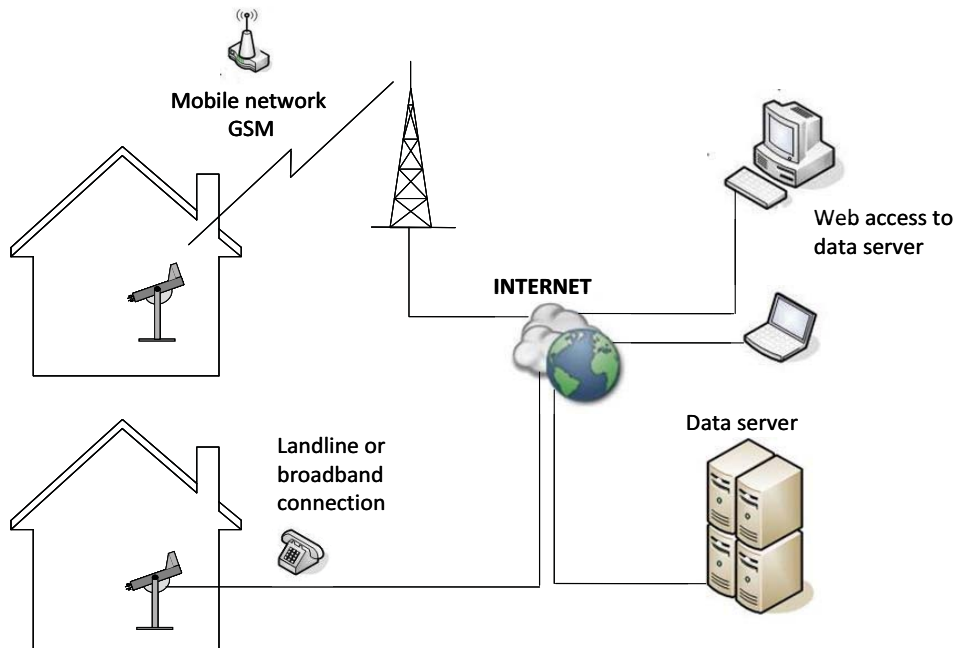
Figure 3.1: Home monitoring network architecture.

positive acknowledgment of this procedure and then closes the connection. However, if for any reason the transmission fails, the (uncorrupted) files are retransmitted at the next attempt of connection. In order to guarantee the last mentioned feature and increase the robustness of the data storage, each FOT device keeps a copy of each test saved in the local memory.

The content of the xml file is eventually organized in a relational database designed in the MySQL environment, which runs on the data server. This allows a web-based access to the database which is used, in this work, to download additional information necessary for data processing. However an exhaustive description of the database is not necessary because results are based on an offline processing of the raw files. Details will be illustrated hereafter.

Data processing software

The data processing here presented consists in the application of an algorithm to obtain statistical properties of respiratory parameters and additional indexes for each patient on a daily basis. The strategy chosen is assigning to each patient a data structure containing an amount of information sufficient to calculate the statistical properties of the mechanical parameters for any test (mean and inter-breath variability in the simplest case), but optimizing the memory occupancy. In particular, given the data

structure of a patient, it is possible to obtain the value of respiratory parameters with a resolution equal to the inspiratory and expiratory phases of each breath of the test. Provided that the pressure signal produced in multifrequency (i.e., a combination of waves at 5,11 and 19 Hz), the complete set of 13 parameters is:

- inspiratory resistance at 5Hz (Rinsp5);
- inspiratory reactance at 5Hz (Xinsp5);
- inspiratory resistance at 11Hz (Rinsp11);
- inspiratory resistance at 19Hz (Rinsp19);
- expiratory resistance at 5Hz (Rexp5);
- expiratory reactance at 5Hz (Xexp5);
- whole breath resistance at 5Hz (Rtot5);
- whole breath reactance at 5Hz (Xtot5);
- Xinsp5-Xexp5 (DeltaXrs);
- inspiration time (Ti);
- inspiration time normalized by the total breath time (Ti/Ttot);
- respiratory rate (RR);
- tidal volume (Vt);

Note that the first 9 parameters are referred to as parameters of *respiratory mechanics*, whereas the last 4 are called parameters of *breathing pattern*.

When defining the concept of the data cruncher the first issue faced was deciding what data to use, in fact raw data is recorded for every test but the device is also able to process raw signals and give final results immediately. Although this real-time process proved to be rather reliable (see Chapter 2), an offline approach gives much more flexibility and computational precision. For instance, real-time IIR digital filters as those implemented on the device introduce a delay which is not quantifiable in a deterministic way. The stochastic component of the delay is due to the not-linear phase of IIR filters and depends on the local frequency characteristics of the input signal. An offline processing permits to apply zero phase filters with the same frequency characteristics, therefore output and input traces result perfectly synchronized and no error is introduced.

However, the main advantage of an offline processing is the possibility to create custom data structures and to tune the algorithm in a flexible manner, as it resides in a single central host rather than on several devices spread all over the territory covered by the clinical study.

Figure 3.2 illustrates a simplified structure of the MATLAB® data cruncher. First of all, raw data is transferred from the central server to the data processing host through a multithread FTP client (gFTP). The program compares the two folders that have to be synchronized and copies only new data in the target folder. Secondly baselines corresponding to every test are obtained through a web access to the database. As explained in Chapter 2 the baseline is the constant term of the calibration function for the flow sensor, and it is measured any time a test is performed since this assessment requires a very simple procedure. The preliminary phase of the algorithm is completed with the uncompression of new raw data files and some checks about the success of the previous actions.
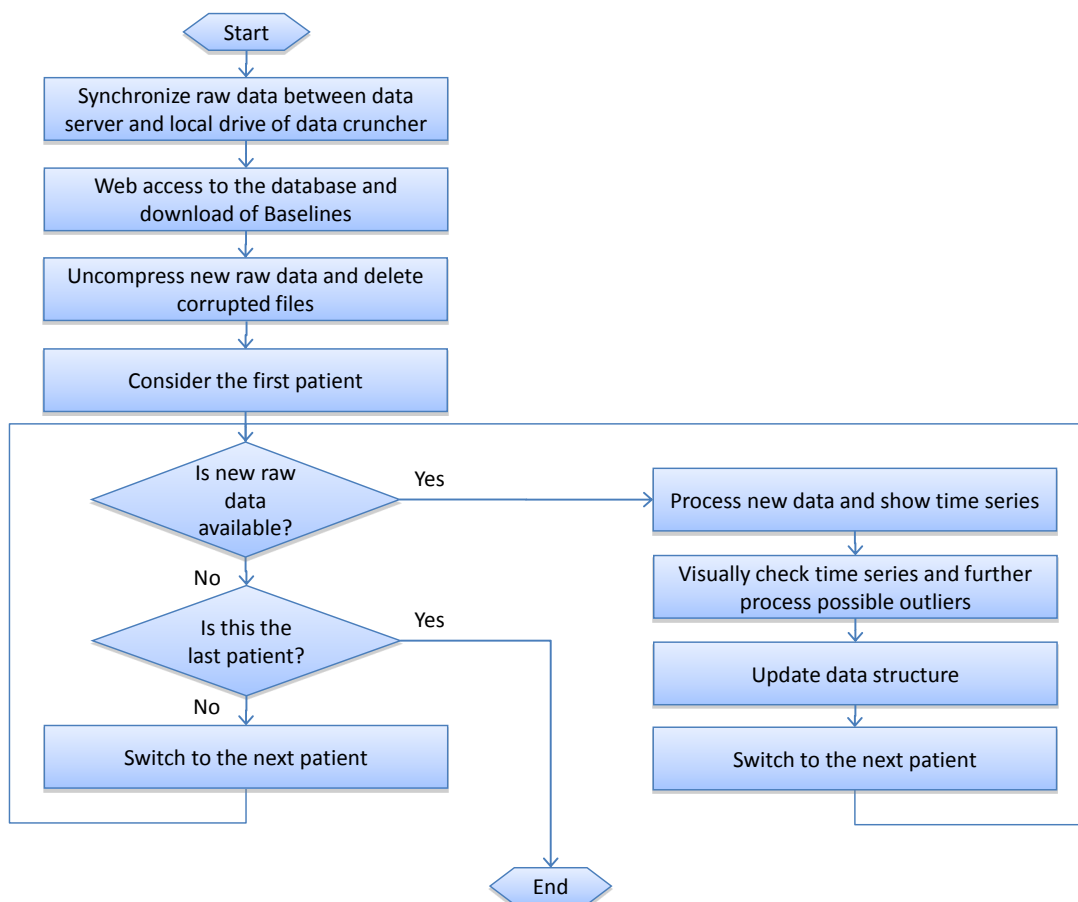
Figure 3.2: Simplified flow chart of the data processing algorithm.

The core of the data cruncher is represented by a cycle which scans any path present in the main raw data folder and looks for patients with new data available. Whenever

new data are found, they are calibrated using the baseline and the remaining parameters that are contained in a text file located in the patients' folders. Note that, as proved in Chapter 2, the calibration is very stable therefore calibration files are not supposed to be updated throughout the progress of the clinical study. However, the clinical protocol indicates that calibration must be verified by measuring test objects every two months in order to make sure acquisitions are reliable. Successively, calibrated data are processed in a way that allows the update of the data structure they belong to, but before doing so time series are obtained from the data structure and are shown on screen. The aim of this is to guarantee a visual check of the parameters before saving new data in order to treat possible outliers. To make it clearer, Figure 3.3(a) represents how the user can interact with the software and look for outliers. Through the buttons in the bottom left it is possible to switch between pages that show different respiratory parameters and answers to the queries of the CAT-based questionnaire. Note that before the proper test some questions are administered to the patient in order to gain additional qualitative information on their health (see Subsection 2.5 for more information). Moreover, the user can select acquisitions that are suspected to be outliers. This process is supported by an automatic outlier detection that shows mean and confidence intervals equal to 2 standard deviations for every time series and assigns a pink colour to acquisitions in which at least one parameter lies outside the confidence limits. Red points indicate the remaining new acquisitions, grey points identify training acquisitions (the first ten), black ones estimate missing values whereas blue and light blue ones stand for old acquisitions which have been previously accepted without any modification or with manual correction of the breaths selected, respectively. This last point will be now better clarified: whenever the user selects one or more possible outliers, the corresponding points are surrounded by black squares as those visible in the last portion of the time series in Figure 3.3(a). Subsequently all the selected tests are visualized one by one in a different page (represented in Figure 3.3(b)), through which the user can visualize several traces of the test at issue (default ones are Chest Wall Volume, R at 5Hz and X at 5Hz). Apposite coloured markers identify automatically recognized breaths, and the user can modify them if he/she is not satisfied with the default selection. If a new set of breaths is defined, the acquisition is considered "manually corrected" and coloured light blue at the following simulation, otherwise it is tagged "correct" and coloured blue. Finally, if the user does not consider the acquisition good enough, an additional procedure tags the entire test as "discarded" so that it is not visualized anymore from that time on.

Black points require a more detailed explanation that will be supported by a thorough analysis in Subsection 2.3. Briefly, patients are required to execute one test per

day, possibly at the same time, but statistically they tend to miss some tests. In these situations missing values are estimated by a custom model of Locally Weighted Regression, so that the time series results complete. One should also note that this method represents a solution for estimating discarded acquisitions, which however have never been obtained to date.

Last, the algorithm provides for saving the survey's responses in an additional data structure. Since this information could be effectively related to the aggravation of the disease, in Subsection 2.5 the correlation between qualitative data obtained from the CAT-based questionnaire and quantitative data recorded during the tests will be assessed.

2.2   Automatic outlier detection

Time series data are often subject to outliers or discordant observations. Their detection has become an important part of time series analysis because outliers can lead to wrong model specification, biased parameter estimate and poor forecasts. The most widely used way to deal with outliers is the diagnostic approach, in which diagnostic methods are applied to the residuals of a model of the observed phenomenon to identify possible outliers. Once they are identified a model that incorporates them is proposed, so that the outliers' effect can be assessed and, as a consequence, a robust parameter estimate is obtained [82].

It is clear that an approach of this kind cannot prescind from the definition of a model of the data. Tsay [83], Ljung [84] and Kaya [85] studied the effect of outliers through ARMA models. Tsay [86] extended his results to ARIMA models, that are suitable solution for processes formed by the summation of stationary processes (they are also known as *integrated processes*). Kaiser and Maravall [87] proposed a similar solution for seasonal outliers. It must be specified that state-space models are often used as well, but for univariate time series a Kalman filter will always have an ARIMA representation. The transition and and observation matrices determine the autoregressive part of the ARIMA equivalent, whereas its moving-average part depends on the Kalman gain matrix K [82].

As it will be extensively analyzed in the next Subsection, parametric approaches as those mentioned above need some strong underlying hypotheses that cannot be guaranteed for the data of this study. Thus a different and more general approach is needed. For the data at disposal, automatic outlier detection is very hard to be implemented in a reliable way since outliers can be due to several and unpredictable causes. The most typical is a bad breath recognition, that might lead to mean intra-test
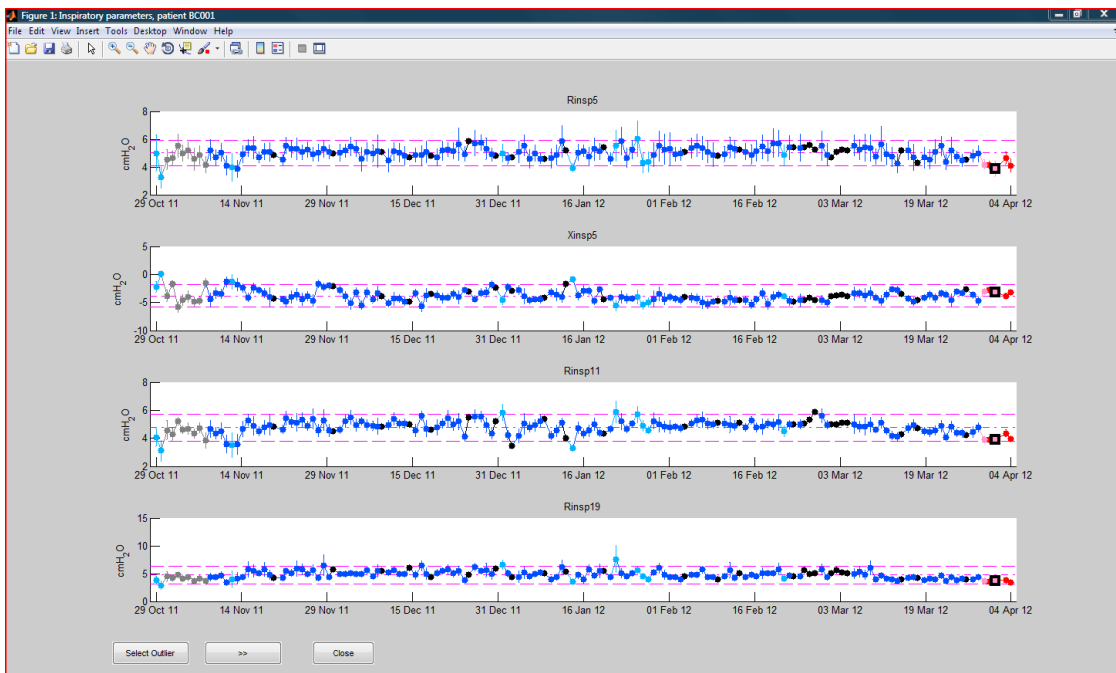
(a) Time series visualization. Grey points are training acquisitions required by the patients to be at ease with the device; blue points are old correct acquisitions; light blue points are old manually corrected acquisitions; red points are new acquisitions inside the confidence limits; pink points are new acquisitions automatically detected as outliers; black points are missing data estimated by a custom model of Locally Weighted Regression. Error bars indicate inter-breath standard deviation.
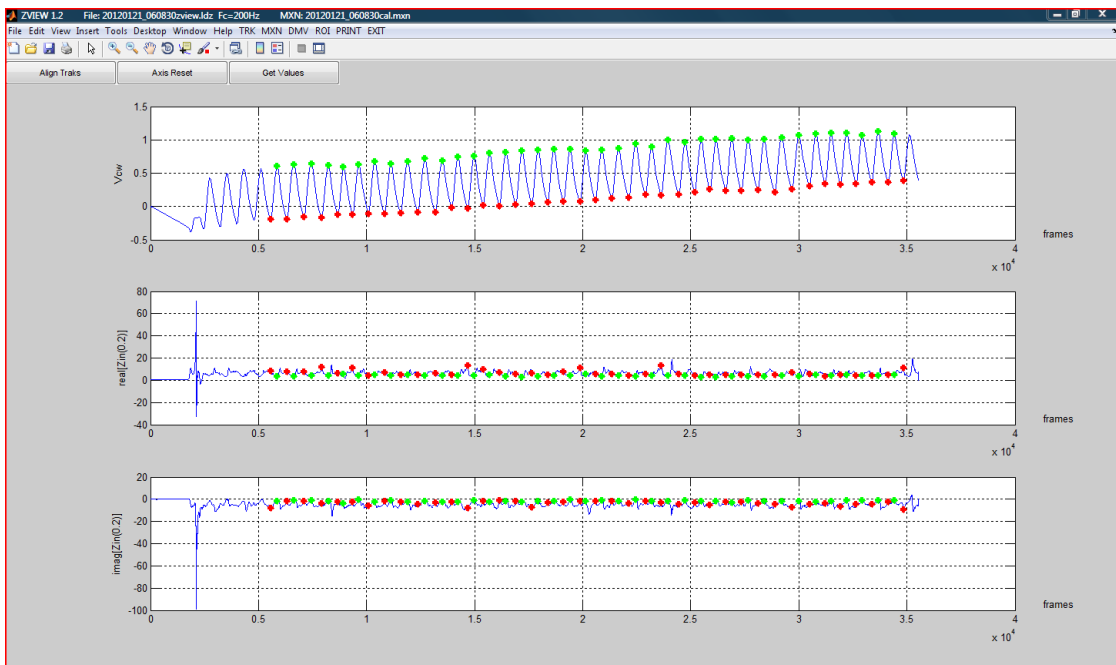
(b) Treatment of outliers. Default traces are Chest Wall Volume, R5 and X5, but the user can visualize many other parameters. Red and green crosses spot end of expiration and end of inspiration. The first five breaths are purposely discarded as patients need some time to reach steady breathing.

Figure 3.3: Graphical user interface of the data cruncher.

respiratory parameters (points in Figure 3.3(a)) far from typical values and intra-test standard deviations (error bars in Figure 3.3(a)) larger than usual. For these reasons time series must be checked visually and automatic recognition can only support the user during this process.

As many natural processes, respiratory parameters tend to distribute normally around the mean value. Given a gaussian density function, a commonly used method for detecting outliers consists in excluding values that fall outside confidence limits of a certain percent, say 95% (it approximately corresponds to 2 standard deviations). Probability densities during pathological conditions make no exception, besides when an aggravation of the disease occurs. For instance, COPD is characterized by periodical exacerbations during which respiratory parameters assume "extreme" values (see Chapter 1). During an exacerbation respiratory resistance tends to increase and, on the contrary, respiratory reactance assumes more negative values. Since exacerbations occur rarely these atypical values are much less frequent than normal ones, therefore the pathology's effect on probability distributions consists in a loss of symmetry. In particular, the longer tail will be the right one for resistance and the left one for reactance.

On the light of COPD dynamics one could argue that a logarithmic transformation might solve or at least limit the issue of skewness, so that the method proposed is still applicable. This seems to fit well with both resistance and reactance as, after calculating their absolute values, either distribution will spread more towards positive values. In this way the logarithmic function will compress large values (in modulus) of resistance and reactance and stretch small ones, leading to distributions closer to the gaussian (see Figure 3.4).
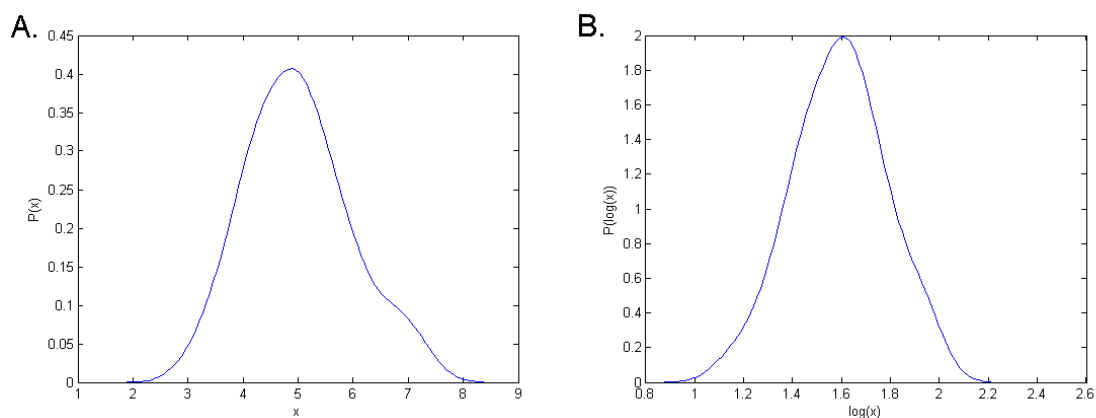

Figure 3.4: Positively skewed normal distribution (A.) and same function after logarithmic transformation (B.).

Logarithmic transformation is a widely used and effective method but it has some drawbacks. First of all probability densities should assume either positive or negative values and not cover both. In the latter case the absolute value of the original dataset will lead to a distribution which is far from the original one (see Figure 3.5(a)). On the contrary, Figures 3.5(b) and 3.5(c) show that positive distributions remain the same and negative distributions are simply mirrored with respect to the y-axis. Since reactance measurements often lie around $0cmH_2OsL^{-1}$, this method is applicable only to the other parameters. In addition, a second quite obvious disadvantage is that logarithmic transformation increases the level of asymmetry if for any reason the skew is negative. As already mentioned, probability densities of COPD patients are supposed to have a skewness far from zero but this is not always the case. During the observational time a given patient could not encounter any exacerbation or, more in general, exacerbations could not have a strong effect on the respiratory parameters' distributions. Anyway, this last option can be considered very unlikely.

For the just mentioned reasons a third outlier detection method is proposed and tested. If the probability density is much different from the gaussian, mean and standard deviation are not good indexes to describe it. On the contrary percentiles prescind from the shape of the distribution they are referred to, therefore they provide correct information for any probability distribution function. In order to keep a level of coherence with the solutions already proposed, which is a confidence of 95%, this last method considers median, 2.5 and 97.5 percentiles in place of mean and 2 standard deviations. In order to decide what method to use, these will be evaluated on patient's data in Section 3.

2.3   Methods for missing data estimate

In statistics, missing data occur when no data value is stored for a variable in the current observation. Missing data are a common occurrence and can have a significant effect on the conclusions that can be drawn from the data, therefore an unbiased estimate of them is necessary to guarantee the best inference about the underlying process.

In the present study lacking values can basically occur for two reasons. The most typical cause is nonresponse, which means that the patient misses a test for any reason. For example he/she can forget to make the test or the device can fail to turn on, so that for a given day no data are available. Another less frequent cause is low quality acquisitions. When the user checks the time series, he/she looks for potential outliers and controls single parameter plots. In case the acquisition is considered badly performed or too noisy, the user tags it as discarded (refer to Subsection 2.1 for an
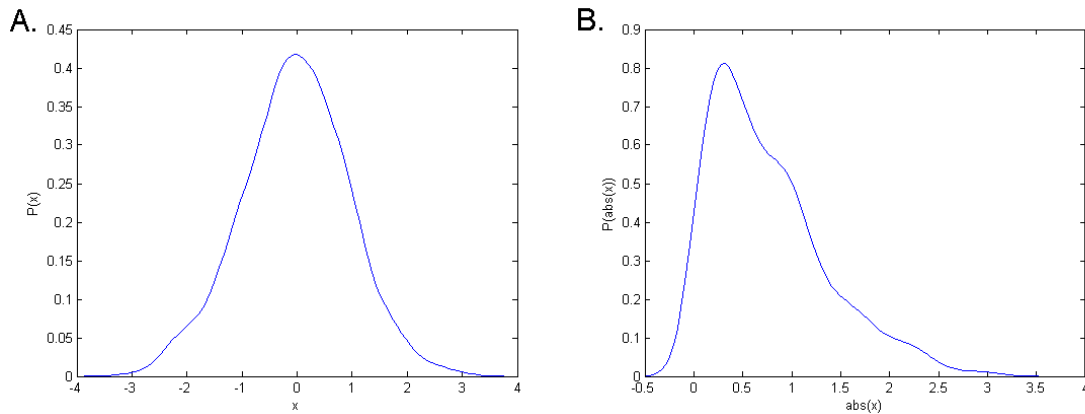
(a) If the Probability Density Function (PDF) assumes both positive and negative values, the PDF of the modulus of the dataset tends to concentrate near zero. Note that the small negative portion of the curve, that should not be present, is due to an error introduced by the interpolation method used to obtain the trace.
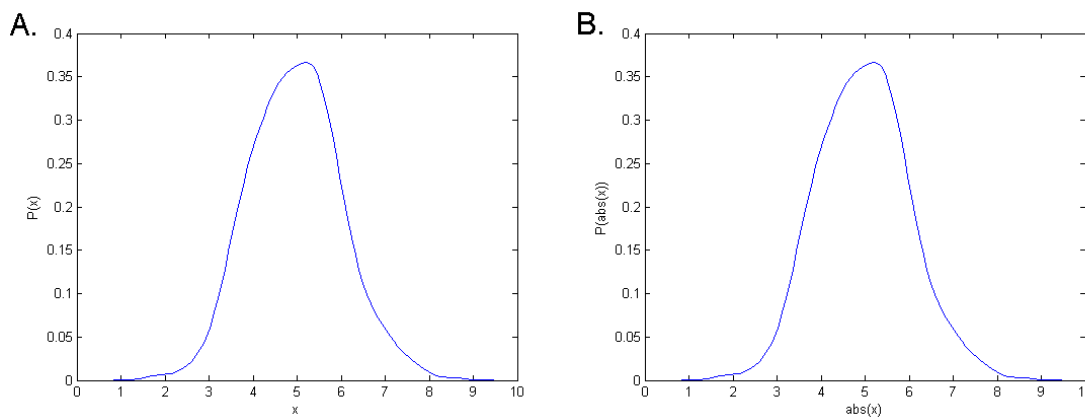
(b) If the PDF is completely positive, the operation of absolute value does not affect the function.
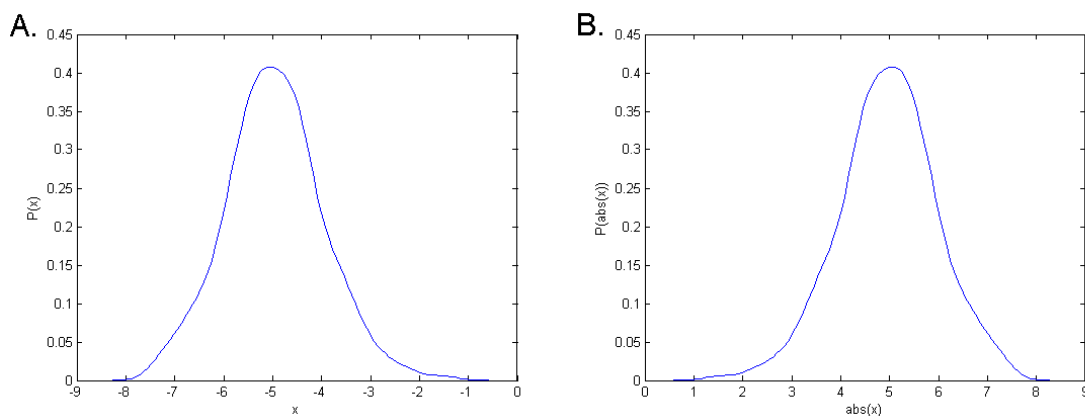
(c) If the PDF is completely negative, the modulus of the dataset has a distribution which is reversed with respect to the y-axis.

Figure 3.5: Quasi-normal probability distributions before (A.) and after (B.) calculating their absolute values.

accurate explanation). Doing so, information about the test is still present in the data structure but it is transparent to the processes analyzing the data structure itself.

Regardless the cause, the consequence is information missingness that has to be filled in. After an extended literature review [82, 88, 89], suitable methods for missing data estimate were identified and evaluated. They will be treated one at a time in the following passages.

Imputation

Standard statistical methods have been developed to analyze rectangular data sets. Traditionally, the rows of the data matrix represent units (also called cases, observations or subjects depending on the context) whereas the columns represent variables measured for each unit. In real situations it often happens that some of the entries are not present, therefore such a data matrix is not rectangular (see Figure 3.6). For example, respondents in a household survey may refuse to report income. In an industrial experiments some variables' values can be missing because of mechanical breakdowns unrelated to the experimental process. In an opinion survey some individuals may be unable to express a preference for one candidate over another.
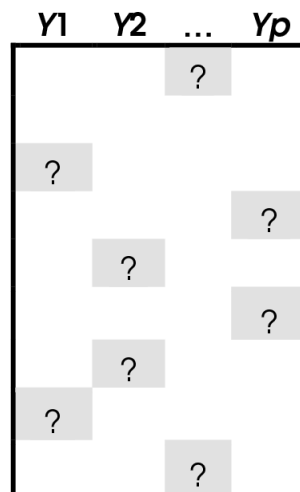

Figure 3.6: Example of nonresponse. Rows correspond to observations, columns to variables.

Unlike rectangular data sets, incomplete ones require a statistical analysis which is not straightforward anymore. Typical strategy, termed *complete-case analysis*, excludes units that have missing value codes for any of the variables involved in the analysis. Such an approach is generally inappropriate as the investigator is usually interested in making inferences about the entire target population, rather than the portion of the

target population that would provide responses in all relevant variables of the analysis. Imputation-based procedures address this issue, in fact through them missing data are filled in and the resultant completed data can be analyzed with standard methods. Commonly used methods include *hot deck* imputation, where individual missing values of an incomplete unit are drawn from "similar" responding units; *mean* imputation, where missing variables are substituted with means from responding units of the same variable; and *regression* imputation, where the missing variables for a unit are estimated by predicted values from the regression on the known variables for that unit.

These estimate methods (also called models for nonresponse) can be applied to impute one value for each missing item (single imputation) or to impute more than one value (resampling and multiple imputation). The last two allow appropriate assessment of imputation uncertainty as they provide a probabilistic estimate of the missing value rather than a deterministic estimate by applying the chosen estimate method several times. Briefly, in the case of resampling one can implement *bootstrap*, in which a set of estimates for the missing value is given by applying the estimate method to sets obtained by simple random sampling with replacement of the original entire sample; or *jackknife*, which follows the same scheme but uses a resampling where estimates are based on dropping a single observation from the sample. The case of multiple imputation yet reflects the same procedure, but it is more flexible as combined inferences can be obtained by applying to the entire sample the same model for nonresponse (e.g., hot deck or any model that imputes draws) or different models for nonresponse (e.g., to display the sensitivity of inference to those models). Once the set of estimates is obtained, its mean value represents the final missing value estimate and the uncertainty of the estimate is measured by the standard deviation of the set if the distribution is normal, or by other indexes in case the distribution is more general.

Imputation methods are widely used in statistics but best performances are obtained with non-rectangular data sets. This is not the case of the data of this study, in fact missing acquisitions correspond to missing values for any variable. In other words the data matrix has entire missing rows but still results rectangular. In cases like this basic imputation methods are yet applicable (e.g., mean imputation), but the results obtained are rather naive. In addition basic methods can be ascribed to other approaches that will be illustrated further on.

For the listed reasons imputation is not considered a satisfactory way to treat missing values.

Linear prediction

The main objectives of time series modeling and analysis are understanding the dynamic or time-dependent structure of the observations of a single series (*univariate* time series analysis) and ascertaining the leading, lagging and feedback relationships among several series (*multivariate* time series analysis). Knowledge of the dynamic structure helps to produce accurate forecasts of future observations and design optimal control schemes. Equivalently, models can be used to estimate missing observations on the basis of previous data as if they were future unknown realizations of the underlying process.

Within the boundary of univariate time series analysis, interest often centers on modeling the temporal inertia of an observable time series for forecasting future observations. Since the 1970s, this serial dependence of the data has been represented with accuracy through a class of mixed autoregressive moving-average models (ARMA) [90]. Basic modeling specification techniques for applying these models to real data can be listed as follows:

- tentative specification or identification of the model;
- efficient estimate of model parameters;
- diagnostic checking of fitted model for further improvement.

The first point employs statistics that can be readily calculated from the data and allow the user to tentatively select a model, that is, determine the number of coefficients for the autoregressive and the moving-average parts of the model (p, q). Without going deep into this point, it is sufficient to say that through the calculation and the convergence study of sample autocorrelation function, sample partial autocorrelation function and sample extended autocorrelation function it is possible to identify the model.

The second point, which is efficient in the statistical sense of making best use of the information contained in the data, is based on assumptions about the distributional properties of the data. In concrete terms it uses standard statistical inference procedures: *least squares* for AR models and *maximum likelihood* for ARMA models. The practical results of using either of these two procedures are similar and effectively lead to the following scheme: apply the model to predict successive values of the recorded time series data; choose the parameters that minimize the sum of squares of the resulting one-step ahead prediction errors. The question of whether k-step ahead errors would be better used in the sum of squares criterion has been well studied. The answer might be yes if the model does not perfectly describe the data and if multistep prediction is the object of the modeling. However, assuming that the data do arise from

the chosen model, one-step ahead errors are best used even if k-step prediction is the object of the modeling.

The last point is needed to discern if the model provides a good fit to the data. The model may not fit the data because it was not well chosen or because it was poorly estimated, even though it is capable of a good fit to the data. To guarantee a positive checking four targets must be met. First of all the residuals, which are estimates of the innovations, must set up a white noise. This simply requires a sample statistical analysis as the one applied to the original series, and its result must show no evidence of autocorrelation. Secondly, the spectral features of the fitted model must match those of the original data. Since the residual spectrum is the ratio of the sample spectrum of the series to the fitted model spectrum, the more they depart from each other the more information remains in the residual rather than in the output of the model. A third aspect of model checking is validation by forecasting out-of-sample values of the series. A proportion of data at the end of the series is withheld and various forecasts of these data are produced using the model fitted to the earlier part. The objective is consistency between the forecast limits and the data. This in itself does not demonstrate a good fit, but lack of consistency is clear evidence of model inadequacy. Finally, a model that fits well with the foregoing criteria should also be tested by fitting an extended model, in which the autoregressive or moving average order is increased. Formal tests can then be used to check that no significant improvement in fit can be achieved by such an extension.

Although linear modeling is a powerful tool to identify the serial temporal dependence of a time series, the overview given demonstrates how such an approach requires time consuming procedures to properly tune the model. Applying these concepts to the study would mean collecting some identification/fitting data for each patient, designing the model and periodically check its reliability. If the time series are not extended enough, identification and fitting could result poor. Furthermore a missing value estimate of this kind would provide guess on the basis of data recorded before the missing day, whereas future data is usually available and should be taken into account in the estimate. Even more importantly, ARMA modeling of time series requires that the underlying process is stationary, namely it has constant expected value and variance for all t and the covariance depends only on the separation lag k and not on t. COPD time series are much expected to be nonstationary, in fact the disease is poorly reversible, it gets progressively worse over time and is characterized by periodical exacerbations [91]. Since the "drifting" behaviour of the time series cannot be known and subtracted in advance, ARMA identification and fitting would result rough.

The last issue can be addressed by allowing some of the zeros of the autoregres-

sive polynomial to equal one. The model obtained, which is known as *autoregressive integrated moving-average model* (ARIMA), can fit a nonstationary series as long as it derives from an *integrated process*[1]. Since the coefficients of the integrated part of the model are by definition set to one, identification and fitting of the remaining parts follow the already introduced scheme. For completeness, it must be also said that an ARIMA model will also have a Kalman filter representation therefore the results obtained using one approach or the other will exactly be the same.

Although ARIMA modeling provides a solution for nonstationarity, the hypothesis that the phenomenon studied produces an integrated process is rather strong for the data of the study. In addition other limitations remain, such as the fact that the model is time-invariant. For these reasons, linear prediction can be considered a practicable way to treat missing values but it is not free from criticisms.

Adaptive prediction

In problems of filtering, prediction and control the features of the system can often be subjected to variation hard to predict. To cope with this situations it is suggested to make use of adaptive approaches, in which the dynamics of the filter, predictor or controller is progressively updated on the light of the actual behaviour of the system. Figure 3.7 represents the case of prediction, in which the system that executes the adjustment is called *tuner*. Typically the tuning is carried out automatically through a feedback that monitors the current performance.

In every adaptive system a primary role is played by techniques of adaptive identification. Since traditional methods of parameter identification aim at providing constant values for the parameters, it is not surprising that they are not suitable when identifying systems whose dynamic characteristics are subjected to variations, especially when these are due to accidental or hardly predictable factors. The idea is to render the already introduced fitting procedures (i.e., *least squares* and *maximum likelihood*) recursive and to add an exponential oblivion factor μ $(0 < \mu < 1)$. Recursion is needed when the data have to be processed in real-time, as in the case of real-time fitting, whereas μ has the task of differently weighing recent and old data. The last available point is weighted by the first power of μ, the one before by its second power and so on, giving rise to an exponential weighting function with a time constant of $1/(1 - \mu)$ which is commonly called *window* of the algorithm. For instance, if $\mu = 0.99$ the window is of 100 data points, if $\mu = 0.98$ the window is of 50 data points, if $\mu = 0.96$

---

[1]Integrated processes are processes formed by the summation of a stationary process, therefore their first difference is white noise.
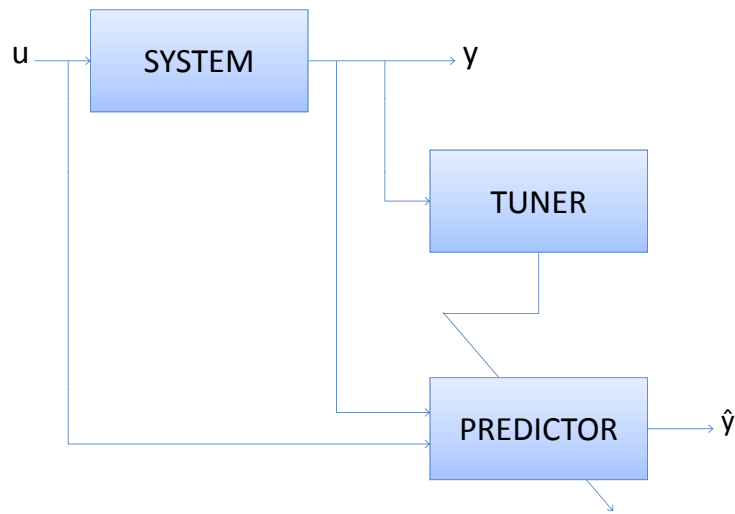
Figure 3.7: Self-tuning adaptive predictor.

the window is of 25 data points and so on.

The oblivion versions of the fitting procedures are fundamental for the estimate of time-variant parameters, a situation that occurs very frequently. By tuning the value of the oblivion coefficient it is possible to modulate the sensitivity of the algorithm to most recent data. The smaller the value of μ, the more reactive the update of the parameters, in fact new data participate more strongly in the parameter estimate.

Although adaptive prediction is a suitable solution for treating missing data, it is not free from drawbacks. As for all parametric methods, some assumptions on statistical characteristics of the data are required, and as already mentioned these characteristics cannot be guaranteed a priori. With respect to the foregoing procedures, adaptive prediction has the advantage of providing a good estimate of time-variant parameters, but the nature of the data does not rule out the possibility that the family of model itself, besides the parameters, can change when the disease aggravates.

Nonparametric approach

Nonparametric statistics covers techniques that do not assume a fixed model structure of the data. The term *nonparametric* is not meant to imply there is a complete lack of parameters, but that the number and nature of the parameters are flexible and not fixed in advance. In these techniques individual variables are assumed to belong to parametric distributions, and assumptions about the types of connections among

variables are also made. These techniques include, among others:

- *nonparametric regression*, which refers to modeling where the structure of the relationship between variables is treated nonparametrically, but where nevertheless there may be parametric assumptions about the distribution of model residuals;

- *nonparametric hierarchical Bayesian models*, which allow the number of latent variables to grow as necessary to fit the data, but where individual variables still follow parametric distributions and even the process controlling the rate of growth of latent variables follows a parametric distribution.

As nonparametric methods make fewer assumptions, their applicability is much wider than that of the corresponding parametric methods. In particular they may be applied in situations where less is known about the application in question. Data speak for themselves, thereby avoiding subjectivity in selecting a specific parametric model. Also, due to the reliance on fewer assumptions, nonparametric methods are more robust and often easier to use. Due to both this simplicity and their greater robustness, nonparametric approaches are seen by some statisticians as leaving less room for improper use and misunderstanding.

The wider applicability and increased robustness of nonparametric tests come at some costs. First, in cases where a parametric test would be appropriate, nonparametric tests have less power. In other words, a larger sample size can be required to draw conclusions with the same degree of confidence. Second, a higher complexity in the mathematical argumentation is involved.

Because of their flexibility, nonparametric techniques may serve as a first step in the process of finding an adequate parametric model. If no model can be found that describes the underlying structure adequately, as in the case of this study, then the results of nonparametric estimate may be used directly for forecasting, describing the characteristics of the time series or estimating missing values. This is why nonparametric approach is chosen as the most suitable for the data of the study.

2.4   Ad-hoc Locally Weighted Regression method

The rationale used to develop the customized method for missing data consists in the fact that values right before and right after the missing point contain information about it. This a reasonable assumption from a pathophysiologic point of view, as previous studies [72] showed how the today's health condition of a patient influences the condition of the next 4-5 days. In particular, today's condition influences tomorrow's condition very much, whereas the effect on the following days gradually decreases. The

method for missing data estimate must therefore meet these concepts and at the same time it must show robustness against noise that may mask the information content of the data. *Locally Weighted Regression* (LWR) addresses both the issues of taking into account a "local" portion of the data and "trusting" the values of this portion differently on the basis of how much they are close to the missing point, as it relies on the fact that a window of the data is weighted according to a given kernel, also referred to as weighting function, in order to obtain an estimate of the central point of the window. With regard to robustness of the estimate, a tentative estimate based on kernels of different width and an assessment of these attempts through the calculation of a cost function consent to make the algorithm adaptive to the local variability of the data. These functionalities will be illustrated in detail after an introduction to LWR.

LWR, also referred to as LOWESS or LOESS[2], is a regression modeling method that combines multiple regression models in a k-nearest-neighbour-based model.

LWR thus builds on "classical" methods, such as linear and nonlinear least squares regression. It addresses situations in which the classical procedures do not perform well or cannot be effectively applied without undue labor. LWR combines much of the simplicity of linear least squares regression with the flexibility of nonlinear regression. It does this by fitting simple models to localized subsets of the data to build up a function that describes the deterministic part of the variation in the data, point by point. In fact, one of the chief attractions of this method is that the data analyst is not required to specify a global function of any form to fit a model to the data, but only to fit segments of the data.

The trade-off for these features is increased computation. Because it is so computationally intensive, LWR would have been practically impossible to use in the era when least squares regression was being developed. Most other modern methods for process modeling are similar to LWR in this respect. These methods have been consciously designed to use the current computational ability to the fullest possible advantage of achieving goals not easily achieved by traditional approaches.

Plotting a smooth curve through a set of data points using this statistical technique is referred to as a LOESS curve, particularly when each smoothed value is given by a weighted quadratic least squares regression over the span of values in its neighbourhood. When each smoothed value is given by a weighted linear least squares regression over the span, this is known as a LOWESS curve. However, authors often treat LOWESS and LOESS as synonyms.

LOESS, originally proposed by Cleveland [93] and further developed by Cleveland

---

[2]LOESS is a later generalization of LOWESS; although it is not a true initialism, it may be understood as standing for "LOcal regrESSion" [92]

and Devlin [94], specifically denotes a method that is also known as *locally weighted polynomial regression*. At each point in the data set a low-degree polynomial is fitted to a subset of the data, with explanatory variable values near the point whose response is being estimated. The polynomial is fitted using weighted least squares, giving more weight to points near the point whose response is being estimated and less weight to points further away. The value of the regression function for the point is then obtained by evaluating the local polynomial using the explanatory variable values for that data point. The LOESS fit is complete after regression function values have been computed for each of the n data points.

Many of the details of this method are flexible. First, the subsets of data used for each weighted least squares fit are determined by a nearest neighbour algorithm. A user-specified input to the procedure called the *bandwidth* or *smoothing parameter* (α) determines how much of the data is used to fit each local polynomial. The value of α is the proportion of data used in each fit, therefore the subset of data used in each weighted least squares fit comprises the $n\alpha$ (rounded to the next larger integer) points closest to the point at which the response is being estimated. α is called smoothing parameter because it controls the flexibility of the LOESS regression function. Large values of α produce the smoothest functions that wiggle the least in response to fluctuations in the data. Small values of α, on the contrary, cause the regression function to conform to the data. Using too small values for the smoothing parameter is not desirable, however, since the regression function will eventually start to capture the random error in the data. Useful values of the smoothing parameter typically lie in the range 0.25 to 0.5 for most LOESS applications. Figure 3.8 shows a case in which α is set to the optimal value.

Second, the local polynomial fits to each subset of the data are almost always of first or second degree, that is, either locally linear (in the straight line sense) or locally quadratic. Using a zero degree polynomial turns LOESS into a weighted moving average, but such a simple local model might not always approximate the underlying function well enough. Higher-degree polynomials would work in theory, but yield models that are not really in the logic of LOESS. LOESS is based on the idea that any function can be well approximated in a small neighbourhood by a low-order polynomial and that simple models can be fit to data easily. High-degree polynomials would tend to overfit the data in each subset and are numerically unstable, making accurate computations difficult.

Third, as already mentioned the weight function has the task to give more influence on the local model parameter estimate to points near the the one whose response is being estimated. The traditional weight function used for LOESS is the tricube
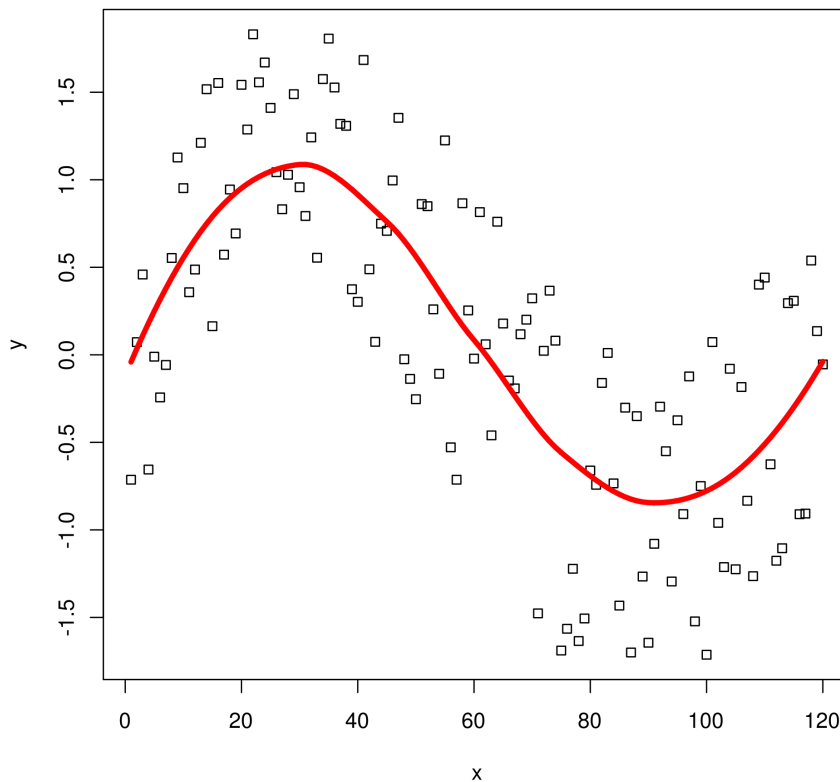
Figure 3.8: LOESS curve fitted to a population sampled from a sine wave with uniform noise added. Thanks to the optimal tuning of the smoothing parameter α, the LOESS curve approximates the original sine wave.

weight function, however any other weight function that satisfies the properties listed by Cleveland [93] could be conveniently used for best treating the data at disposal. Some popular kernel functions are exhibited in Figure 3.9. The gaussian kernel has infinite support, whereas the remaining kernels weigh a specific point on the basis of the distance between that point and the point of estimate, after scaling the distance so that the maximum absolute gap over the points in the localized subset of data is exactly one.

Standard LOESS implementations are meant to smooth noisy datasets. This means that the entire time series is given as input to the regression function, and the output results in the same x-axis points to which smoothed y-axis values correspond. In case of missing data a slightly different approach is needed, since the output time series must have the same data points as the input plus estimated values for the missing acquisitions. An increase in the number of data points from the input to the output can
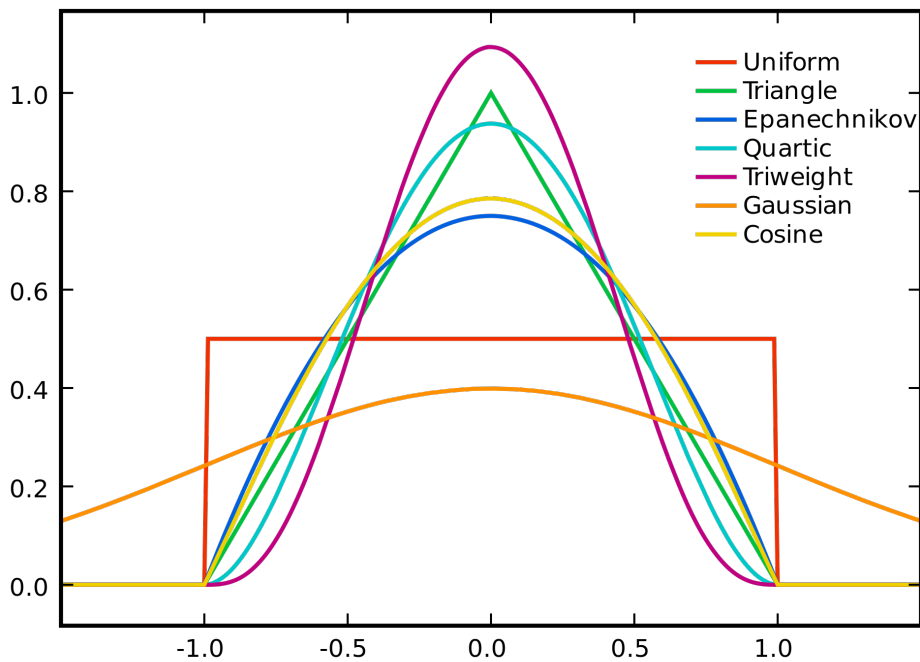
Figure 3.9: Commonly used kernel functions.

be achieved by leaving the central position of the weighting function empty. As Figure 3.10(b) illustrates, this approach consent to obtain an output value in correspondence of which no input value is present. In addition, since the objective is not smoothing the original time series, the method must be applied only when acquisitions are missing, therefore the computational power needed to perform traditional LOESS methods is drastically reduced.

The choice of the weighting function is strongly related to the purpose of filling in missing values. In normal conditions the greatest weight is given to the value falling in the central position of the kernel, whereas in case of missing value the two points adjacent to it are weighted the most. In order to give them a large relative weight with respect to the remaining points, a kernel with a wide plateau is preferable. The tricube, besides being one of the traditional weight functions used for LOESS, is the one that best meets this necessity (see Figure 3.10). Since standard LOESS methods often use different kernels and do not allow the user to choose one rather than another, also this issue requires a self implementation of the LWR method. Moreover, the k-nearest neighbour approach used by standard LOESS guarantees that a certain number of data points participate in the regression regardless their distance along the abscissa. This is important to guarantee an optimal amount of regression information, but could mean that data very far from the point of interest are taken into account. An assumption

of this kind is not acceptable because of the pathophysiologic time inertia of COPD, therefore an application in which the kernel function cover a certain amount of days is preferred. The main disadvantage is represented by the fact that the method could undergo a loss of robustness due to the small amount of data covered, that is, when missing data are very close to one another. This last issue, which is strongly related to the optimization of the kernel width, will be hereafter discussed and solved.

Traditional LOESS methods allow the designer to adjust a limited number of features, among which the bandwidth α. As already mentioned, it usually consists in the percentage of data that is supposed to participate in the regression. Depending on the objective of the data analysis the designer can manually set it. In particular, the greater its value, the smoother the output time series. In case of missing data estimate the best kernel width cannot be evaluated manually as the designer has not any reference to use as comparison for the estimate. Again, the larger the kernel the more conservative the estimate. Furthermore, assuming that such a task is viable, the designer would be required to study every missing value separately, which is an option that must be avoided. As a consequence, an automatic method capable of evaluating the optimal kernel width on the basis of the local characteristics of the data is required.

Smoothing parameter setting

There is a number of ways one can set the smoothing parameter α [95]. This thesis is not meant to discuss all possible available statistics but rather give a panoramic view of suitable techniques. The method used by Cleveland and Devlin [94] is to set α such that the reference point being predicted has a predetermined amount of support, that is, α is set so that the total sum of weights n is close to some target value. This has the disadvantage of requiring assumptions about the noise and smoothness of the function being learned. Another technique, used by Schaal and Atkeson [96], sets α to minimize the cross validation error on the training set. In LWR this is done by evaluating and minimizing the mean squared cross validation error $\mathrm{MSE_{cross}}$ with respect to α. A third method, still described by Schaal and Atkeson [96], is based on the prediction intervals $\mathrm{I_q}$, which are expected bounds of the prediction error at a query point x_q (see [96] for a thorough mathematical description of both methods). Besides using the intervals to assess the confidence in the fit at a certain point, they provide another optimization measure when they are minimized with respect to α.

The rationale behind the last two methods can be better explained graphically. Figure 3.11 shows the case of linear LWR applied to a time series with locally linear behaviour. The estimator variance, which can be to some extent estimated by either
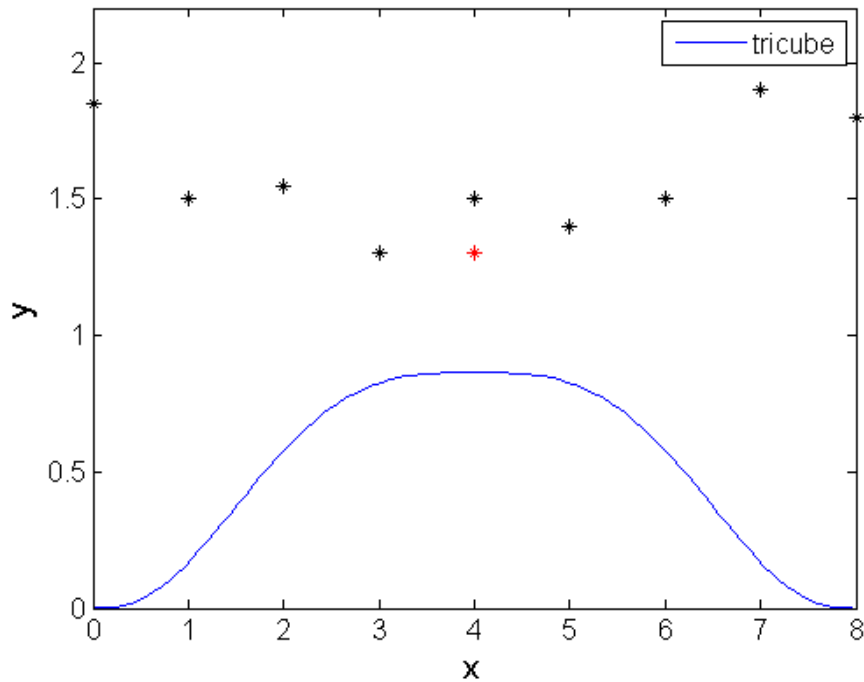
(a) Traditional kernel-based smoothing. Quadratic LWR is applied to the black points such that the one falling in the central position of the window is substituted by the new red value.
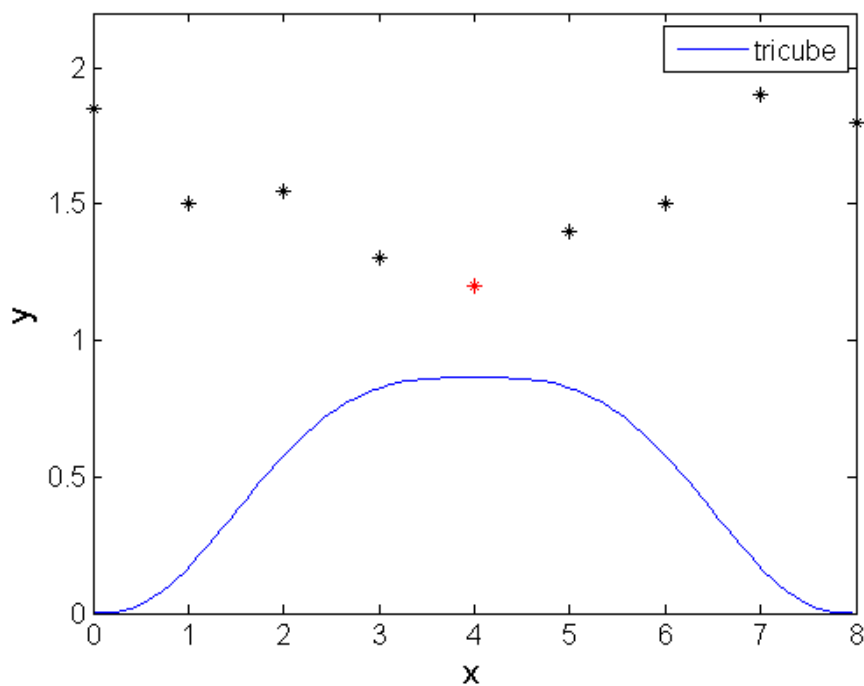
(b) Application of quadratic kernel regression for estimating one missing value. The wide plateau of the kernel ensures that a large weight is given to the points closest to the missing value.

Figure 3.10: Quadratic LWR based on tricube weighting function.

method, is minimized when the kernel includes as many training points as can be accommodated by the model. Too large a kernel includes points that degrade the fit; too small a kernel neglects points that would increase confidence in the fit. Since both methods seem suitable for optimizing the missing data estimate, they will be evaluated on patient's data in Section 3.
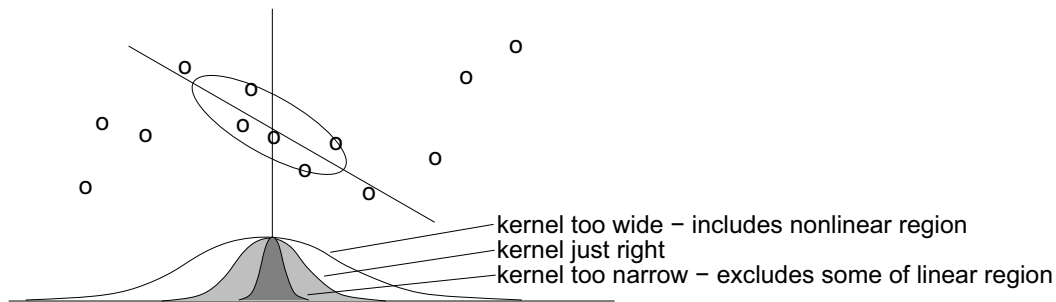

Figure 3.11: Smoothing parameter optimization in linear LWR.

A correct setting of α is strongly related to the choice of the fitting polynomial. Linear interpolation is preferable if the noise level is high as it tends to give a conservative estimate of the query point. On the contrary, when the locally interpolated data follow a nonstationary evolution (e.g., data points in Figure 3.10) a linear fitting is not able to represent trends of greater degree, that is, it underfits. Quadratic interpolation works very well in the latter case but it could output extreme values if the data is very noisy and the number of points participating in the regression is small. Choosing a wide kernel when the data are little reliable (i.e., when the data is noisy) is a reasonable approach both from the point of view of compensating the low information level of the data and the point of view of making a quadratic fitting more similar to the linear alternative. Since a smart tuning of the smoothing parameter emphasizes pros and avoids cons of quadratic regression, this is selected as the winning approach.

Estimate of imputation uncertainty

A key problem with most approaches for missing data estimate is that inferences based on the filled-in data do not account for imputation uncertainty. Thus standard errors computed from the filled-in data are systematically underestimated, P-values of tests are too small and confidence intervals are too narrow. In other words, if the added variance due to nonresponse is not incorporated into the model the resulting estimate would always be too conservative.

Multiple imputation, first proposed by Rubin [97], is a typical approach for adding imputation uncertainty to the estimate but it cannot be adapted to LWR. The most

attractive alternative would be to quantify the variance due to nonresponse on the basis of the statistical properties of the neighbourhood of each missing point, and add this variance to the "deterministic" estimate. Doing so the final estimate would result affected by little stochastic disturbance when the time series is stationary, whereas it would be affected by heavy stochastic disturbance in portions of the time series showing high variability. As previously mentioned the goodness of the prediction can be evaluated through prediction intervals I_q. For each missing value, I_q could be used to set the level of additive noise that guarantees the minimal modification to the statistical properties of the time series before and after the missing data estimate.

Two methods that follow the just mentioned guidelines are here proposed and will be evaluated in section Section 3. The first method contaminates the LWR estimate by adding a random number drawn from a Student's t-distribution with a number of degrees of freedom n which is provided by the LWR method. By their nature prediction intervals represent confidence limits (at a given percentage $1 - \alpha$) of a Student's t-distribution, therefore this kind of approach can be considered the most straightforward. The second method uses the same prediction intervals to limit a uniform distribution from which, as in the first case, an additive random number is drawn. In this way the risk of drawing numbers very far from the mean value is avoided, so as to improve the robustness of the estimate.

2.5   Analysis of the CAT-based questionnaire

A major goal in the treatment of COPD is to ensure that the patient's health is optimized. However, despite the availability of clinical guidelines to manage COPD, there is continued evidence to suggest that a substantial proportion of patients are not achieving the level of treatment success that may be possible [49, 98]. In addition to routine clinical evaluations, a critical step in management is to obtain, from the patient, reliable and valid information on the impact of COPD on their health status. This would include information on daily symptoms, activity limitation and other manifestations of the disease. A standardized patient-centered assessment tool, covering key attributes of COPD health, should facilitate information gathering and improve communication between patient and clinician. In addition to an overall score, an ideal tool should be able to identify specific areas of greater severity to serve as a focal point for targeted management or the evaluation of management goals, thereby improving both the process and the outcome of care.

Available disease-specific health status measures, obtained by questionnaires modules, are widely used in clinical trials or in clinical practice. Among them, the COPD

Assessment Test (CAT) provides the best trade-off between measurement consistency and intricacy of the questions [99]. Indeed, length and complexity of the scoring algorithms are too complex for routine use in clinical practice.

A criticism made about questionnaire-based measures consists in the fact that the information gathered is by its nature subjective. This might correspond to strongly biased total scores and to answers extremely sensitive to little mindful evaluations of the symptoms. Such sources of misrepresentation are evaluated by administering a version of the CAT purposely adapted to be shown on the screen of the FOT device at the beginning of each test and comparing the obtained qualitative information with quantitative data recorded during the following FOT measurement. However, results of these investigation should be cautiously used to make conclusions about the original CAT as it is presented in a slightly different format, necessary to guarantee intelligibility on the device's screen.

For any patient with sufficient amount of tests, correlation between the CAT-based questionnaire and quantitative data is evaluated by the following statistical tools:

- Simple Linear Regression (SLR) between the questionnaire's total score and the respiratory parameters taken singularly;
- Multiple Linear Regression (MLR) between the questionnaire's total score and the respiratory parameters taken all together;
- MLR between the questionnaire's total score and the most correlated parameters resulting from global MLR.

The total score is calculated by summing every numerical single score (see Figure 3.12). Since some answers are ordinal categorical, they are substituted with increasing natural numbers. Following is the complete list of questions:

1. How often have you coughed today? (Min: Never; Max: All the time; Answers: 0|1|2|3|4|5)
2. How much phlegm (mucus) do you feel in your chest? (Min: Not at all; Max: Full; Answers: 0|1|2|3|4|5)
3. Do you feel your chest tight? (Min: Not at all; Max: Very; Answers: 0|1|2|3|4|5)
4. When you walk up a hill or one flight of stairs, how much are you breathless? (Min: Not at all; Max: Very; Answers: 0|1|2|3|4|5)
5. How much do you feel limited doing any activities at home? (Min: Not at all; Max: Very; Answers: 0|1|2|3|4|5)
6. How much do you feel confident leaving your home despite your lung condition? (Min: Very; Max: Not at all; Answers: 0|1|2|3|4|5)

7. How have you slept last night? (Min: Soundly; Max: Bad; Answers: 0|1|2|3|4|5)
8. How much energy do you have? (Min: A lot; Max: Not at all; Answers: 0|1|2|3|4|5)
9. How much sputum have you coughed up (in tablespoons)? (Answers: 0|1|2|3|4+)
10. Sputum color: (Answers: Clear(0)|White(1)|Yellow(2)|Brown(3)|Green(4))
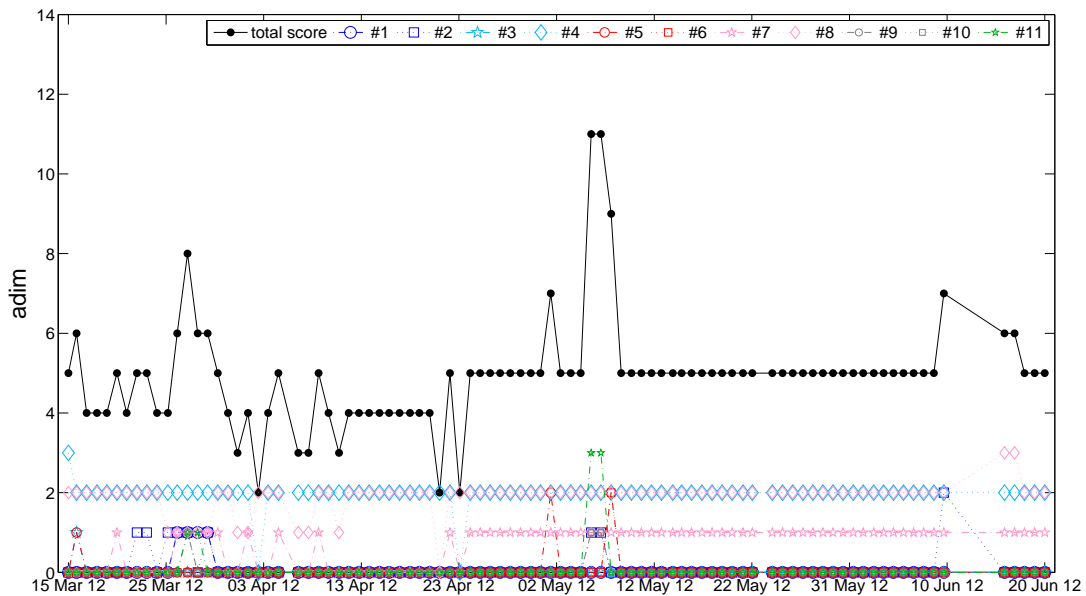11. Sputum consistency: (Answers: None(0)|Watery(1)|Thin(2)|Thick(3))

Figure 3.12: CAT-based questionnaire's responses for patient BC031.

The goal of linear regression is ascertaining whether the CAT-based questionnaire is systematically correlated with any of the respiratory parameters taken singularly, or with any set drawn from them. In particular, putting together all the possible combination of single and multiple parameters, 8192 different groups are evaluated for each patient.

In SLR, the 13 respiratory parameters are sorted from the most to the least correlated and only the first 3 among the significantly correlated ones (P<0.05) are considered. Patient without any correlated parameter are also indicated. Patients in the training phase are not considered due to the small dimension of their datasets.

Equivalently, in MLR groups of parameters are sorted from the most to the least correlated according to their adjusted coefficient of determination R^2_{adj}, often described as the proportion of variance "accounted for", "explained" or "described" by regression. If the regression is perfect R^2_{adj} equals 1; if the regression is a total failure, R^2_{adj} equals 0.

Given the huge amount of groups, MLR is repeated twice. First, all the combinations are processed and only the most correlated group is considered for each patient. Second, MLR is performed again considering all the combinations of only the parameters correlated with every patient's total score in the first regression. Doing so, it is possible to discriminate whether the total score can systematically explain a certain group of parameters.

2.6   Principal component analysis

One of the difficulties inherent in multivariate statistics is the problem of visualizing data that have many variables. Fortunately, in these cases groups of variables often move together, in fact more than one variable might be measuring the same driving principle governing the behavior of the system. In many systems there are only a few of such driving forces, but an abundance of instrumentation enables the measurement of dozens of system variables. When this happens, a winning solution is to simplify the problem by replacing a group of variables with a single new variable [100].

The principal component analysis (PCA) is a quantitatively rigorous method for achieving this simplification. The method generates a new set of variables, called principal components, that are linear combinations of the original variables. As Figure 3.13 illustrates all the principal components are orthogonal to each other, therefore redundant information is avoided. The principal components as a whole form an orthogonal basis for the space of the data, in which components are increasingly important as the variance they own increases. Note that the variance "explained" by a variable indicates the amount of information that variable provides about the underlying phenomenon.

The full set of principal components is as large as the original set of variables, but it is commonplace to consider an amount of components sufficient to exceed 80% of the total variance of the original data. By examining plots of these few new variables, researchers often develop a deeper understanding of the driving forces that generated the original data.

In this study the PCA is applied to the complete set of parameters, which is formed by the 13 respiratory parameters obtained from the FOT measurement and the questionnaire's total score. For simplicity and visualization reasons, only the first two components are taken into account. Indeed, as it will be shown in Section 3, they suffice to make conclusions about the information content of the parameters available.
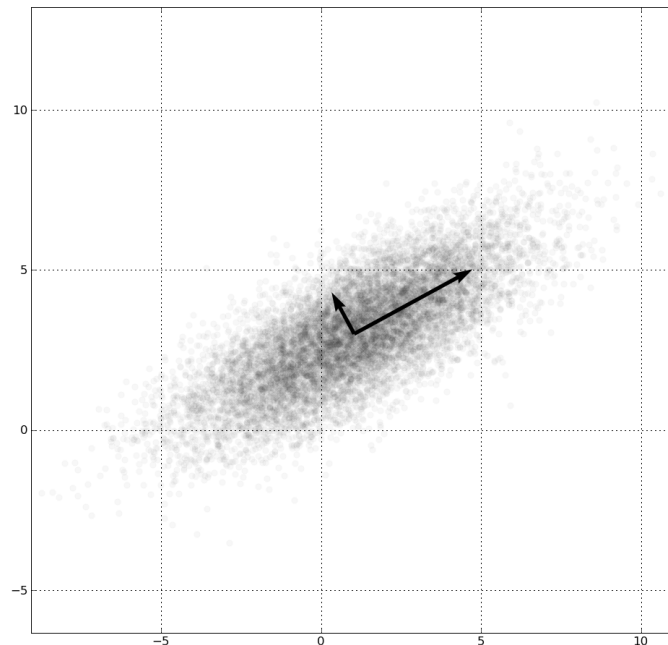
Figure 3.13: PCA of a bivariate gaussian distribution centered at (1,3) with a standard deviation of 3 in roughly the (0.878, 0.478) direction and of 1 in the orthogonal direction. The vectors shown are the eigenvectors principal components.

2.7 Variability analysis at different time scales

Homeokinesis is the ability of an organism interacting with a variable external environment to maintain an organized internal environment, which fluctuates within acceptable limits by dissipating energy in a state far from thermodynamic equilibrium [10]. This definition implies that, while continuous fluctuations are part of normal life, alterations in the fluctuations always signify abnormal physiology. The disruption of such rhythms often leads to collapse or even the death of the organism [11, 12].

If homeokinesis is healthy and pathology is lack or excess of fluctuations in some observed parameters, it is clear that variability itself contains an encoded message that needs to be deciphered, quantified, analyzed and understood. Understanding this message should shed light into what is considered consider health and which are the mechanisms that bring to illness.

Due to the complexity of natural and biological feedbacks and the presence of unknown environmental factors, it is becoming clear that variability can not be analyzed using a classic reductionist approach. This is noticeably true also for the altered vari-

ability observed in the principal chronic respiratory pathologies, bronchial asthma and COPD.

A previous study has proposed a variability analysis at different time scales for 10 asthmatic patients [72]. Another study has measured one COPD patient to obtain values for the daily and the overall coefficient of variation (CoV) of the resistance at 5Hz [80]. This work aims at completing the mentioned analyses by assessing the variability at different time scales of the parameters identified by the PCA as the most representative of the disease, namely the impedance at 5Hz and the tidal volume. For the 8 COPD patients available, standard deviations of Rinsp5, Rexp5, Rtot5, Xinsp5, Xexp5, Xtot5 and Vt are evaluated at the following time scales:

- 1s (within-breath variability, non definable for Vt);
- 1min (inter-breath variability);
- 2d (inter-2-days variability);
- 4d (inter-4-days variability);
- 8d (inter-8-days variability);
- 16d (inter-16-days variability);
- 32d (inter-32-days variability, calculable only for some patients).

Results are compared with those of 10 normal subjects and 10 mild to moderate asthmatic patients from a previously recorded dataset[3]. These followed a protocol very similar to that defined in this work, with the only difference that test were repeated in the morning and in the evening. Therefore, for these datasets the within-day variability (e.g., at 1 day) is computable. Furthermore, comparisons between the three groups are evaluated by the analysis of variance (ANOVA). In its simplest form, ANOVA provides a statistical test of whether or not the means of several groups are all equal, and therefore generalizes the t-test to more than two groups. Doing multiple two-sample t-tests would result in an increased chance of committing a *type I error* (i.e., when the null hypothesis is true, but is rejected). For this reason, ANOVA methods are useful in comparing three or more means.

3 Results

3.1 Outlier automatic detection

The need to automatically recognize outliers arises from the fact that every new acquisition has to be manually checked by the user in order to make sure it was performed

---

[3]Gobbi A. Home Monitoring of Respiratory Mechanics in Patients with Chronic Obstructive Lung Diseases.

correctly. Although the task is rather simple to be carried out, with a consistent number of patient it might become very time consuming and in turn the quality of the check could at some point worsen. For this reason three method of automatic outlier detection are designed and evaluated:

- for each parameter, values that fall outside $\mu \pm 2\sigma$ of the logarithmic transformation of the parameter are considered potential outliers;
- for each parameter, values that fall outside $\mu \pm 2\sigma$ of the parameter are considered potential outliers;
- for each parameter, values that fall outside 2.5 and 97.5 percentiles of the parameter are considered potential outliers.

The first method is based on the fact that exacerbations tend to generate a skewness in the probability distributions. Results are presented for one patient (patient code: BC001), but these can be extended to the others as well. Figure 3.14(a) shows some parameters of BC001 with the relative confidence limits. As explained in Subsection 2.2 logarithmic transformation cannot be applied to reactance time series, therefore confidence limits are shown only in the first and third plot. A watchful eye can easily recognize that this method is not working well as outliers lie only below the lower edge. Indeed, if the logarithmic PDFs were actually symmetric the probability of having lower and higher outliers would be exactly the same. Figure 3.14(b) further clarify this point, in fact it demonstrates how the PDF of Rexp5 is originally quasi-gaussian and it becomes more asymmetric after the logarithmic transformation. As a result, the standard deviation is not significative of the real distribution of the parameter anymore.

At the current stage logarithmic transformation cannot be considered satisfactory. Parameters tend to follow a quasi-gaussian distribution or, in some cases, they present a longer left tail. Both features are therefore not suitable for logarithmic transformation as the result would be a distribution with a larger skewness. Moreover, logarithmic transformation cannot be applied to parameters that assume both negative and positive values. However, such an approach could be taken into account for patients who have had at least one exacerbation during the observational time.

The second method starts from the assumption that every distribution is gaussian, so it does not apply any transformation to the time series. A comparison between Figure 3.15 and Figure 3.14(a) demonstrates that mean and standard deviation calculated from the non-transformed time series are more factual in highlighting potential outliers, besides being applicable to reactance time series as well. Despite their simplicity, linear confidence limits are better than logarithmic ones when the patient does not suffer from aggravations of the disease.
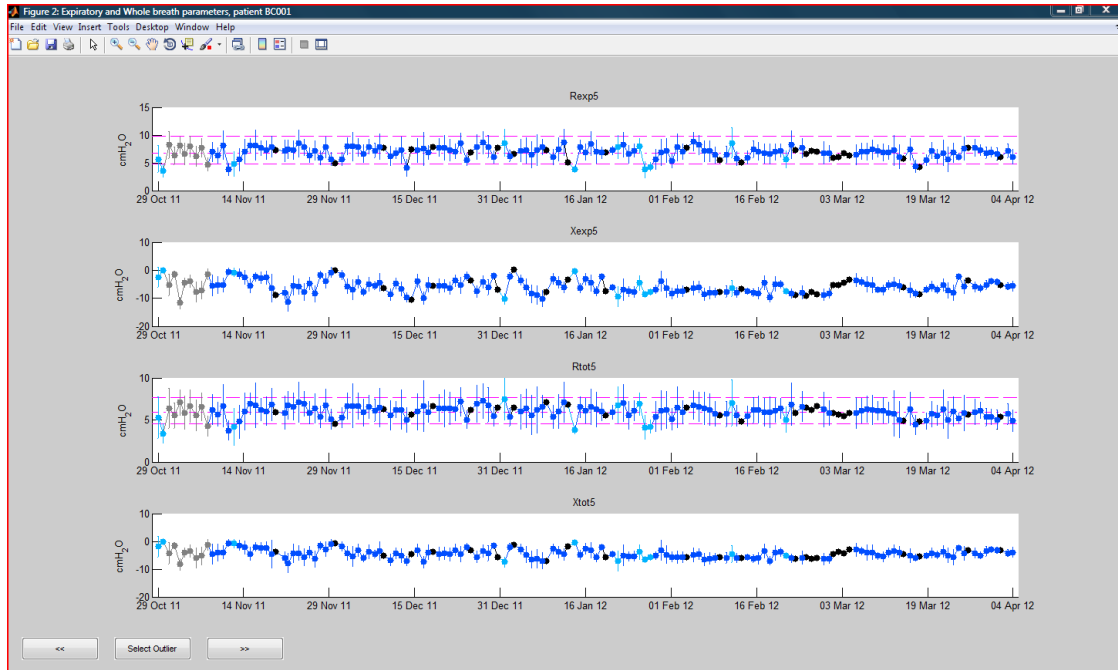
(a) Time series visualization. Confidence limits are calculated from a logarithmic transformation of the time series and are reconverted to a linear scale in order to be visualized in the plots. This is why the distance between the two limits and the mean value is slightly asymmetrical.
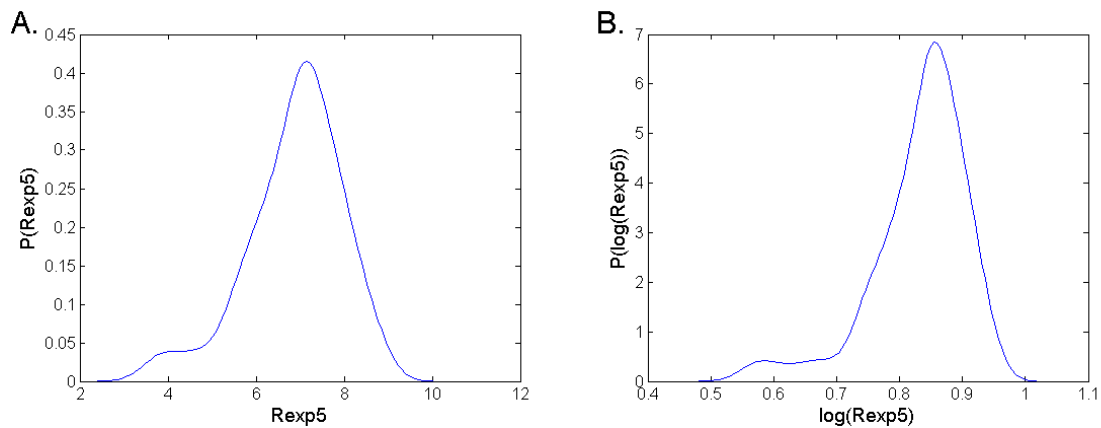
(b) PDF of Rexp5 in before (A.) and after (B.) logarithmic transformation.

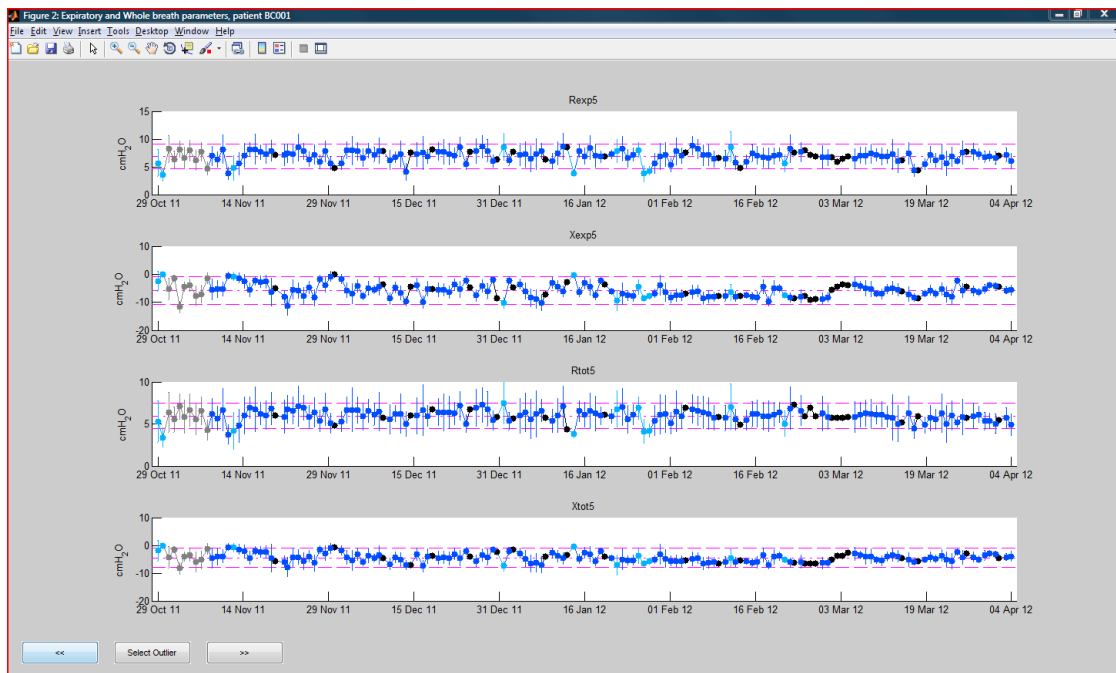Figure 3.14: Logarithmic confidence limits for patient BC001.

Figure 3.15: Linear confidence limits for patient BC001.

The last method counts on the independence from any kind of distribution as it calculates and draws median, 2.5 and 97.5 percentiles. Percentiles, indeed, are the values of the variable outside which a certain percent of the observations falls. Thus, taking say the 2.5 and 97.5 percentiles would be exactly the same as taking $\mu \pm 2\sigma$ only if the original distribution is gaussian. Applying this strategy can be very truthful when the distribution is well detailed regardless the specific function followed, but it proved to be rather raw when the data points are a few. Figure 3.16 illustrates the same parameters as those used to make conclusions about the first two methods, but it considers patient BC033 at the initial stage. Since data points are a few, estimating median and percentiles can suggest a distribution which is really far from the underlying one, as for parameter Xtot5.

In conclusion, patient's data show that linear confidence limits are the ones that, on average, work best. Logarithmic confidence limits might be chosen for patients experiencing at least one exacerbation during the observational time, whereas percentiles can be effectively used after about one month of observations. However, it must be also said that automatic outlier detection is not mostly crucial as the manual support given by the user remains unavoidable. The drawing of horizontal bars giving an idea of the distribution of the values is by itself helpful, regardless the method used to calculate those limits.
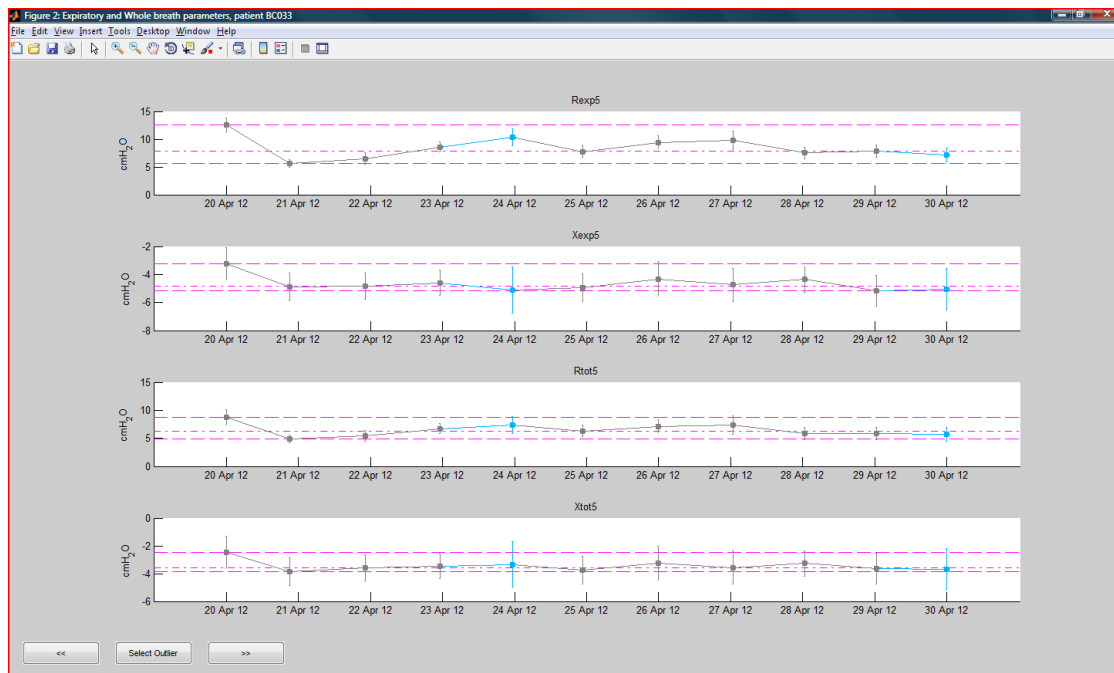
Figure 3.16: Median, 2.5 and 97.5 percentiles for patient BC033.

3.2   Missing data estimate

The ad-hoc LWR is designed to attempt estimating the value of all parameters for every missing acquisition. This is done by considering a number of values in the neighbourhood of the missing point and applying LWR equipped with a tricube weighting function. Since the estimate strongly depends on the width of this window, some attempts are executed and evaluated with the objective of minimizing a cost function. Typical applications span from a small ensemble of points to the totality amount of data and choose the best solution, but this could go against the mechanisms of the real system modeled. Indeed, of most importance and often hard is to conciliate mathematical constructions and the underlying pathophysiologic characteristics. Windows larger than nine days are avoidable, since a previous study [72] underlined that the time constant of obstructive pathologies is most likely four days. This means that the today's value of whatever parameter is affected by its evolution until 4 days ago and, in turn, today's value will affect the next 4 values, resulting in a total amount of 9 days. On the other hand, since LWR needs at least 4 points to be carried out, larger windows have the advantage of gathering a sufficient amount of points even when the missingness of the data is sustained. Furthermore, too many attempts could result in a slowdown of the algorithm. The trade-off chosen for the designed LWR model is to consider four attempts with windows of 4,6,8 or 10 surrounding points. Note that, by saying surrounding points, it is not meant the k nearest neighbours but the k nearest

days regardless the fact that in those days the acquisitions were missing or not. For example, if the window encompasses four days before and four days after the reference missing point and in this time lapse four data points are missing, the regression will be performed on the remaining four points. If instead five or more points are missing, the only chance to estimate the reference point will be broaden the window to the maximum width. In case there are not enough points even with the largest window, the estimate fails due to the too strong missingness of the local data.

Regarding the cost function to be applied to each attempt, two formulas are compared: mean squared cross validation error MSE_{cross} and prediction intervals I_q (see [96] for a thorough mathematical description of both methods). To a first approximation they both try to minimize the estimated variance, which can be expressed as the ratio between the conditional variance and the total sum of weights. Either component increases as the window becomes wider, but they do this with a different slope. As a consequence, at some window width the ratio is minimum.

Despite both formulas work in a similar way, they present a different bias towards larger of narrower windows. MSE_{cross} favours the most wide window almost ever, whereas I_q sometimes trusts smaller ones. Figure 3.17 illustrates this result by plotting the time series of Xexp5 for patient BC001 in the two cases. Grey points are training acquisitions, blue points are correct values, light blue ones are manually corrected tests and black ones are estimated missing values. Estimated points do not show intra-test variability (indicated by the vertical error bars) as this is not important to be estimated, but they are associated with horizontal bars indicating the width of the optimal tricube kernel. Optimization based on MSE_{cross} is often biased towards large windows even though local data is little noisy, whereas optimization based on I_q chooses narrow windows when data show a parabolic trend. This behaviour reflects pathophysiologic dynamics better, in fact when the phenotype of the disease is stable a small number of surrounding recorded values suffice to make conclusions on today's health condition.

For the mentioned reasons I_q is the cost function chosen for kernel width optimization. However, simulations show that the application of either cost function results in a similar estimate of missing values.

The objective of LWR is exploiting the information contained in the data surrounding any missing value the most. Nevertheless, as mentioned in Subsection 2.4, this process cannot account for imputation uncertainty. In other words imputation models can not incorporated the added variance due to nonresponse by themselves, resulting in estimates that are too conservative. Four cases are presented and compared in order to quantify this phenomenon and propose solutions for it:

- original time series with missing data;
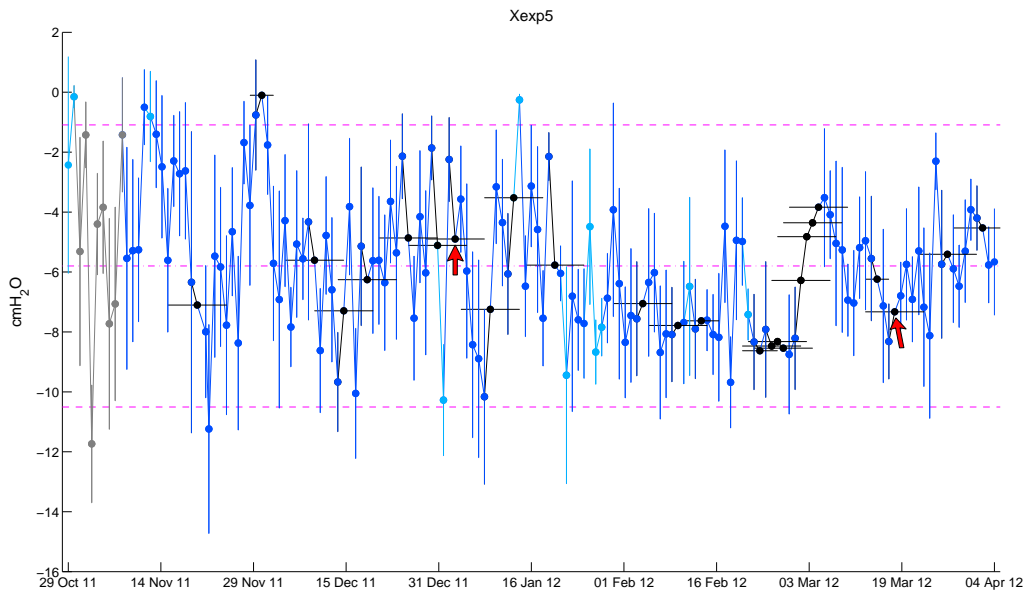
(a) MSE$_{cross}$ minimization. The window width, indicated by horizontal error bars, is often preferred to be as wide as possible. Even when the local surrounding points seem to accomodate the quadratic model the largest window is chosen (red arrows).
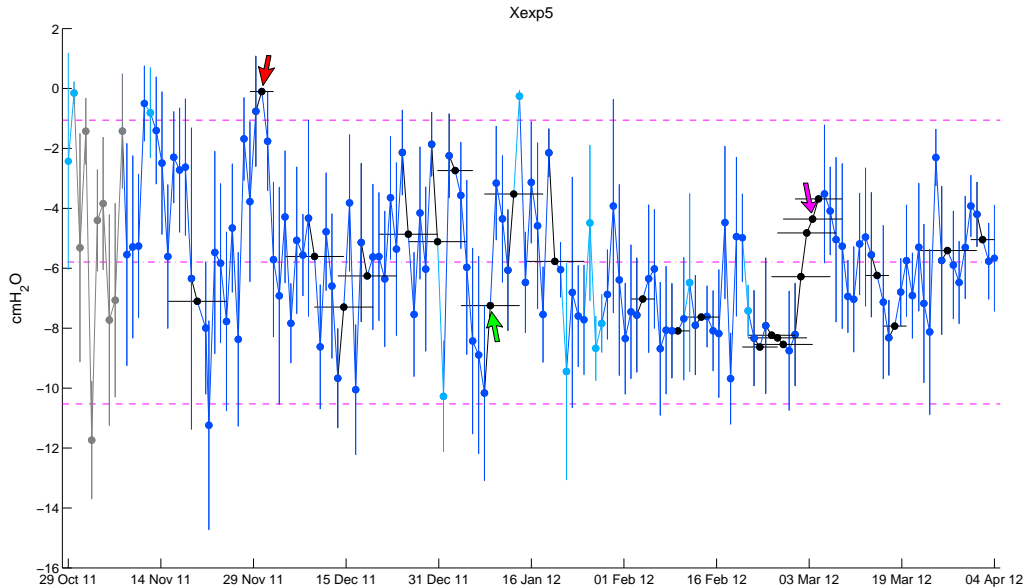
(b) I$_q$ minimization. Horizontal error bars show that small windows are preferred when the surrounding points can be accommodated by the quadratic model (red arrow). On the contrary, when the data are not trustworthy (i.e., noisy, green arrow) large windows are chosen. Magenta arrow indicates the mandatory choice of wide windows when missingness is strong.

Figure 3.17: Kernel width optimization in LWR. Parameter Xexp5, patient BC001.

- time series with missing data filled in by LWR;
- time series with missing data filled in by LWR with a properly dimensioned additive uniform random noise;
- time series with missing data filled in by LWR with a properly dimensioned additive Student's t-distributed random noise.

The comparison between the first two wants to demonstrate how not taking into account imputation uncertainty causes the probability distribution of a given parameter to become narrower and more concentrated around the modal value. The comparisons between the first and the third or the first and the last have the goal to preserve as much as possible the initial distribution even after filling in missing values. Results are presented for the dataset which is most extended and has the greatest missingness, namely that of patient BC001. The big amount of missing data was due to a temporary communication breakdown during the observational period. Since this dataset can be considered worst-case senario in terms of number of data points to estimate, results are very likely to remain valid for the other patients as well.

Besides providing an optimization measure for missing data estimate, prediction intervals are functional tools for assessing the confidence in the fit at a reference point. Thus they can be taken as measure of the imputation uncertainty whenever a parameter value is missing, as shown in Figure 3.18. The power of such a solution consists in the fact that highly variable portions of the data are likely to have highly variable unobserved values, whereas more stationary portions are not supposed to be filled in with values far away from the local trend. Before showing the results, a brief introduction to Student's t-distribution and its relationship with prediction intervals will be given.

In probability and statistics, Student's t-distribution (or simply the t-distribution) is a family of continuous probability distributions that arises when estimating the mean of a normally distributed population in situations where the sample size is small and population standard deviation is unknown. In addition the t-distribution is suitable to describe data affected by outliers or by measurement noise as in the case of this study. It plays a role in a number of widely used statistical analyses, including the Student's t-test for assessing the statistical significance of the difference between two sample means, the construction of confidence intervals for the difference between two population means and in linear regression analysis. The overall shape of the probability density function of the t-distribution resembles the bell shape of a normally distributed variable with mean 0 and variance 1, except that it is a bit lower and wider. This means that it is more prone to produce values that fall far from its mean. However there are many different t-distributions depending on the number of degrees of freedom
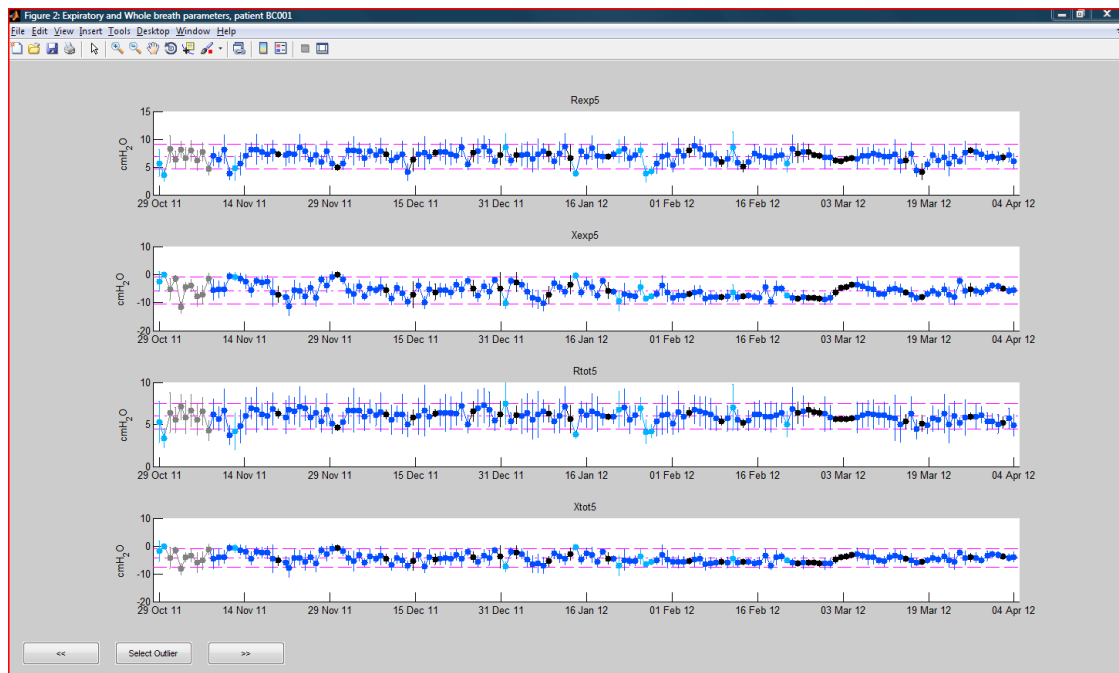
Figure 3.18: Missing data estimate based on I_q minimization. Black points are values estimated by LWR and black error bars are I_q corresponding to the window widths that minimize them (level of confidence for I_q is 80%). When surrounding points are stationary I_q are small; when surrounding points are noisy I_q increase.

ν[4], and as ν grows the t-distribution approaches the normal distribution with mean 0 and variance 1 (see Figure 3.19).

Prediction intervals, in the framework of LWR, stand for confidence limits of the local subset of data used for carrying out the quadratic regression. Starting from the assumption that the population (i.e., the observed pulmonary properties) follows a normally or quasi-normally distributed PDF, the draw of a sample from this population will follow a t-distribution with the following degrees of freedom:

$$\nu = n' - p'$$

where n' and p',

$$n' = \sum_{i=1}^{n} w_i^2$$
$$p' = \frac{n'}{n}p$$

are slight modifications of the number of data in regression n and the number of regression parameters p commonly used in unweighted linear regression. Note that w_i

---

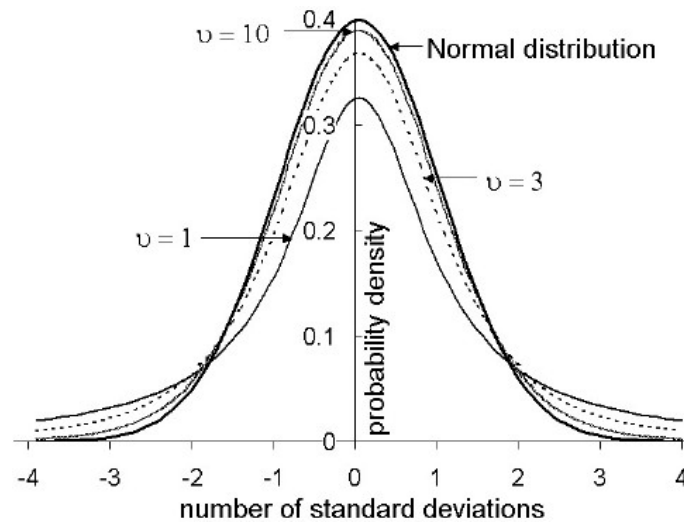[4]The number of degrees of freedom refers to the number of independent observations in a set of data.

Figure 3.19: A comparison of the normal distribution and t-distributions with 1, 3 and 10 degrees of freedom ν. When ν equals 20 or more, the t-distribution and the normal distribution may be considered equivalent for the purposes of calculating uncertainties.

are the weights for the surrounding data, and they are normalized with respect to the biggest weight (i.e., the biggest weight equals 1). Since n can range from 4 to 10 and p is equal to 3, the number of degrees of freedom is always small. As a consequence, the resulting t-distribution is low and wide. Figure 3.20 shows the effect of such a PDF on the estimate of missing data. The algorithm starts from the "deterministic" estimate given by LWR (shown in Figure 3.17(b)) and, for each estimated value, it adds to the estimate a random number drawn from the t-distribution defined by ν (or equivalently I_q). The flatness caused by small values of ν gives rise to very extreme draws, therefore this approach cannot be accepted.

In order to increase the robustness of the estimate a random noise with finite support is chosen. Uniform random noise can be easily dimensioned using properly scaled prediction intervals I_q. By "properly scaled", it is meant that the magnitude of the additive random noise must cause the filled-in data to have the same statistical properties as the original data. In particular, the parameter through which it is possible to optimize the estimate is the level of confidence $1 - \alpha$ of the prediction intervals. the greater $1 - \alpha$, the greater I_q, and in turn the larger the interval $b - a$ in Figure 3.21.

The strategy used to optimize $1 - \alpha$ is based on the following indexes:

- median value;
- interquartile range;
- skewness;
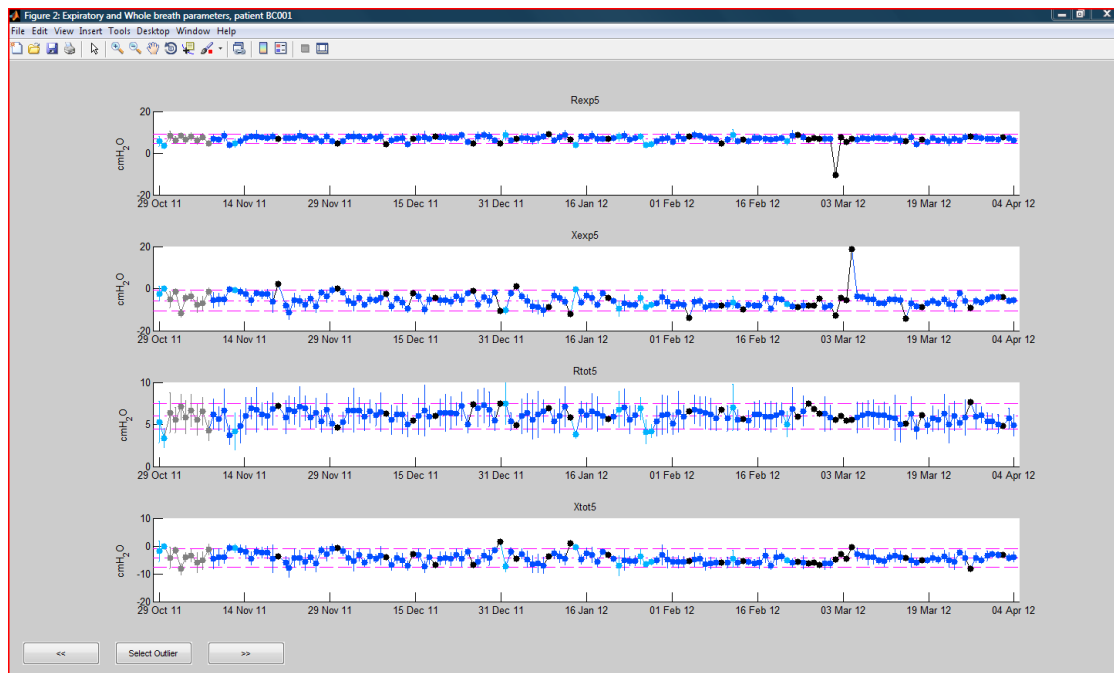- two-sample Kolmogorov-Smirnov test.

Figure 3.20: Missing data estimate with additive Student's t-distributed random noise. Due to the flatness of the t-distribution, estimates can fall far away from the surrounding values.

The value of $1 - \alpha$ chosen is the one that leaves these quantities as unchanged as possible before and after missing data estimate. Moreover, in order to better perform this optimization, PDFs of all 13 parameters are calculated and compared before and after missing data filling. Comparison consists in counting the number of parameters for which the post-processed PDFs have a greater maximum than that of the pre-processed PDFs or, viceversa, counting the number of parameters for which the post-processed PDFs have a smaller maximum than that of the pre-processed PDFs (see Figure 3.22 for better explanation). When this number equals 6 and 7, $1 - \alpha$ can be considered optimized. Simulations indicate that the best value for $1 - \alpha$ is 80%.

Final results are obtained as follows. First of all, three time series are obtained for each parameter: original time series with unfilled missing data, time series with missing data filled in by plain LWR and time series with missing data filled in by LWR with additive uniform random noise. Secondly, median, interquartile range and skewness are calculated from the three datasets, such that one value is obtained for each parameter. Taking together all the parameters, three new datasets composed by 13 elements are obtained for each index (they will be referred to as *original*, *no noise* and *uniform noise* from now on). The objective is demonstrating that *original* and *uniform noise* are samples coming from the same distribution, whereas no target is required for the
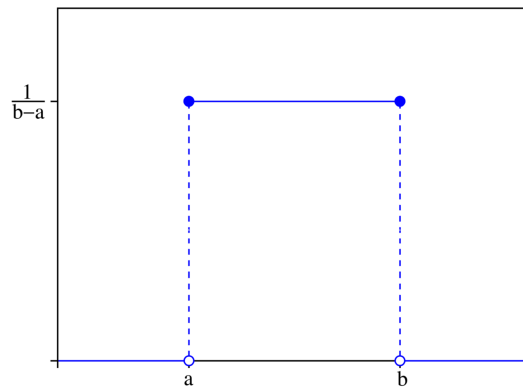
Figure 3.21: Continuous uniform distribution from which random numbers are drawn. The adaptation of this distribution to missing data estimate consists in defining a=-I$_q$ and b=I$_q$.
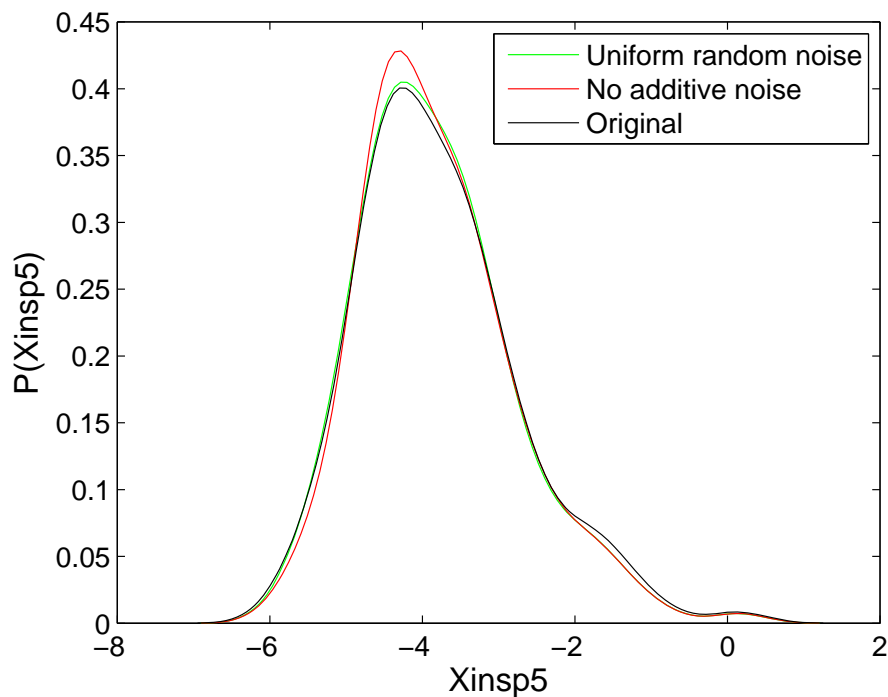
Figure 3.22: PDF of Xinsp5, patient BC001. Whereas plain LWR narrows the probability distribution (red line), LWR "disturbed" by additive uniform random noise leaves the probability distribution almost unaltered (green line). Since the post-processed PDF maximum (green line) is greater than the pre-processed one (black line), parameter Xinsp5 increases the counter of post-processed narrower parameters. Comparison are performed for all the 13 parameters in order to quantify this counter. The objective is tuning $1 - \alpha$ so that the counter equals 6 or 7 ($\simeq 13/2$).

comparison between *original* and *no noise*. Since it is not possible to guarantee that these made-up datasets follow either a gaussian or a symmetric distribution, a paired *sign test* for zero median is performed. The output P is the probability of observing the hypothesized result (median difference is not zero), by chance if the null hypothesis (median difference between the pairs is zero) is true. Small values of P cast doubt on the validity of the null hypothesis. Finally, time series for each parameter are directly compared by the *two-sample Kolmogorov-Smirnov test*. Unlike the sign test, this statistic can be applied to the initial datasets, therefore one P-value per parameter are obtainable. Its objective is to determine if two independent random samples are drawn from the same underlying continuous population (null hypothesis), without making any assumption about the nature of the distribution. Comparisons are once again made between the original time series and those filled in by plain LWR and between the original time series and those filled in by LWR with additive uniform random noise.

Table 3.1 summarizes the results. For each index, the first P-value refers to the sign test applied to the first two columns, which are values for the original time series with unfilled missing data and values for the time series with missing data filled in by plain LWR. The second P-value still refers to the sign test but the input columns are the first and the third, namely values for the original time series with unfilled missing data and values for the time series with missing data filled in by LWR with additive uniform random noise. Considering a standard significance level of 5%, a conclusion that can be drawn about the median value is that both approaches do not alter this index. On the contrary, there is a statistically significant difference for the interquartile range if no noise is added to the LWR estimate (Figure 3.22 confirms this result graphically). Estimating the added variance due to nonresponse through a properly dimensioned additive uniform random noise substantially increases the P-value, allowing one to considerate the interquartile range preserved even though new estimated data is added to the original dataset. Conclusions about the skewness are exactly the same as those about median, as both P-values are greater than 0.05. Finally, the two-sample Kolmogorov-Smirnov test consents to draw the conclusion that both approaches do not generate any distance between the empirical distribution functions of the two samples. Results are positive but they should not be considered as definitive as the foregoing because of the low statistical power of this statistic. This is due to the fact that it requires continuous distributions without any kind of restriction about their shape and it considers two independent samples as input, whereas the samples at disposal can be considered paired.

For completeness, also a *Wilcoxon signed rank test* for zero median is performed. This method is a more statistically powerful variant of the sign test, but it assumes a

	median value				interquartile range				skewness				2-sample ks test	
	original	no noise	P-value	u. noise	P-value	original	no noise	P-value	u. noise	P-value	original	no noise	P-value	u. noise	P-value	P-value	P-value
Rinsp5	5,034	5,030		5,034		0,581	0,578		0,590		-0,682	-0,673		-0,611		**1,000**	**1,000**
Xinsp5	-3,966	-3,970		-3,970		1,271	1,191		1,223		0,839	0,876		0,854		**1,000**	**1,000**
Rinsp11	4,836	4,853		4,826		0,532	0,527		0,534		-0,655	-0,638		-0,665		**1,000**	**1,000**
Rinsp19	4,706	4,706		4,618		0,974	0,933		1,009		0,532	0,528		0,519		**1,000**	**1,000**
Rexp5	7,112	7,063		7,009		1,369	1,314		1,384		-0,925	-0,917		-0,761		**1,000**	**0,995**
Xexp5	-5,745	-5,771		-5,745		3,364	3,272		3,309		0,212	0,296		0,250		**1,000**	**1,000**
Rtot5	6,073	6,065	**0,146**	6,025	**0,109**	0,991	0,870	**0,003**	0,932	**1,000**	-0,872	-0,841	**0,267**	-0,759	**0,267**	**1,000**	**0,981**
Xtot5	-4,497	-4,535		-4,549		2,237	2,249		2,321		0,296	0,376		0,365		**1,000**	**1,000**
DeltaXrs	2,040	2,083		2,077		2,149	2,102		2,264		0,181	0,041		0,068		**1,000**	**1,000**
Ti	1,463	1,454		1,459		0,134	0,133		0,135		-0,211	-0,137		-0,164		**1,000**	**1,000**
Ti/Ttot	0,400	0,400		0,400		0,014	0,012		0,013		-0,266	-0,289		-0,250		**1,000**	**1,000**
RR	16,388	16,537		16,510		1,865	1,737		1,722		0,415	0,362		0,340		**1,000**	**1,000**
Vt	0,724	0,723		0,724		0,101	0,098		0,097		-1,623	-1,507		-1,498		**1,000**	**1,000**

Table 3.1: Results for missing data estimate. Median, interquartile range and skewness are evaluated for each parameter, and the values obtained are compared by the *sign test* to discriminate whether missing data estimate alters the statistical properties of the respiratory parameters (null hypothesis is "median difference between the pairs is zero"). The first P-value is referred to the datasets contained in the first two columns whereas the second P-value refers to the first and the third column. P-values in the last two columns result from direct comparisons performed by the *two-sample Kolmogorov-Smirnov test*. Again, the first P-value is referred to original time series vs. time series treated by plain LWR and the second P-value is referred to original time series vs. time series treated by LWR with additive uniform random noise.

symmetric distribution for the differences between the two variables. Results are not shown as they confirm those of the sign test.

3.3 Analysis of the CAT-based questionnaire

Any time a patient makes a test, a CAT-based questionnaire is administered before the proper FOT assessment. A correlation study between the questionnaire's total score and the respiratory parameters is performed. The goal is assessing whether the questionnaire can provide information on the impact of COPD on patient's health status in a systematical and reliable way.

A visual comparison between Figure 3.23 and Figure 3.24 gives a first qualitative picture of measurement consistency given by the CAT-based questionnaire. According to the total score, patients can be divided in the following groups:

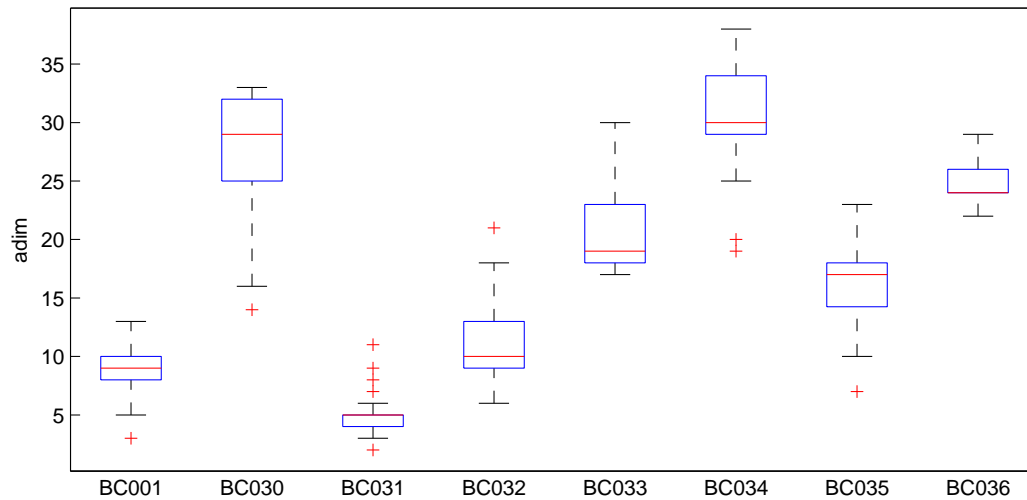- Patients BC030, BC034, BC036: COPD has a strong impact on their lives.

Figure 3.23: Box plot of the CAT-based questionnaire's total score for each patient available. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers ($w = 1.5$ corresponds to approximately $\pm 2\sigma$ and 99.3% coverage if the data are normally distributed) and outliers are plotted individually.

- Patients BC033, BC035: COPD has a moderate impact on their lives.
- Patients BC031, BC031, BC032: COPD has a weak impact on their lives.

Given the just mentioned rankings, one would expect to appreciate extreme values of resistance and reactance for patients belonging to the first group, medium values for patients of the second group and values not far from normal ones for patients of the third group. Objective values depart from this expectation, in fact only patients BC030 and BC031 have a clear correspondence between the "perceived" impact of the disease and the real status of it. Although the phenotype of the pathology could not always clearly correspond to the level of obstruction of the airways, the magnitude of this pattern excludes any possible misinterpretation of the data.

Quantitative results confirm a very weak correlation between respiratory parameters and questionnaire's total score. Figure 3.25 represents, for any patient, a plot of simple linear regression (SLR) between the total score and the most correlated parameter. Only 5 patients out of 8 meet the threshold of significance, and among these there is not any parameter that systematically correlates with the total score.

Results of SLR can be interpreted as follows. The CAT-based questionnaire can not systematically guarantee a good measure of the impact of COPD on a person's life, and how this changes over time. For instance, as shown in Figure 3.12, it might happen
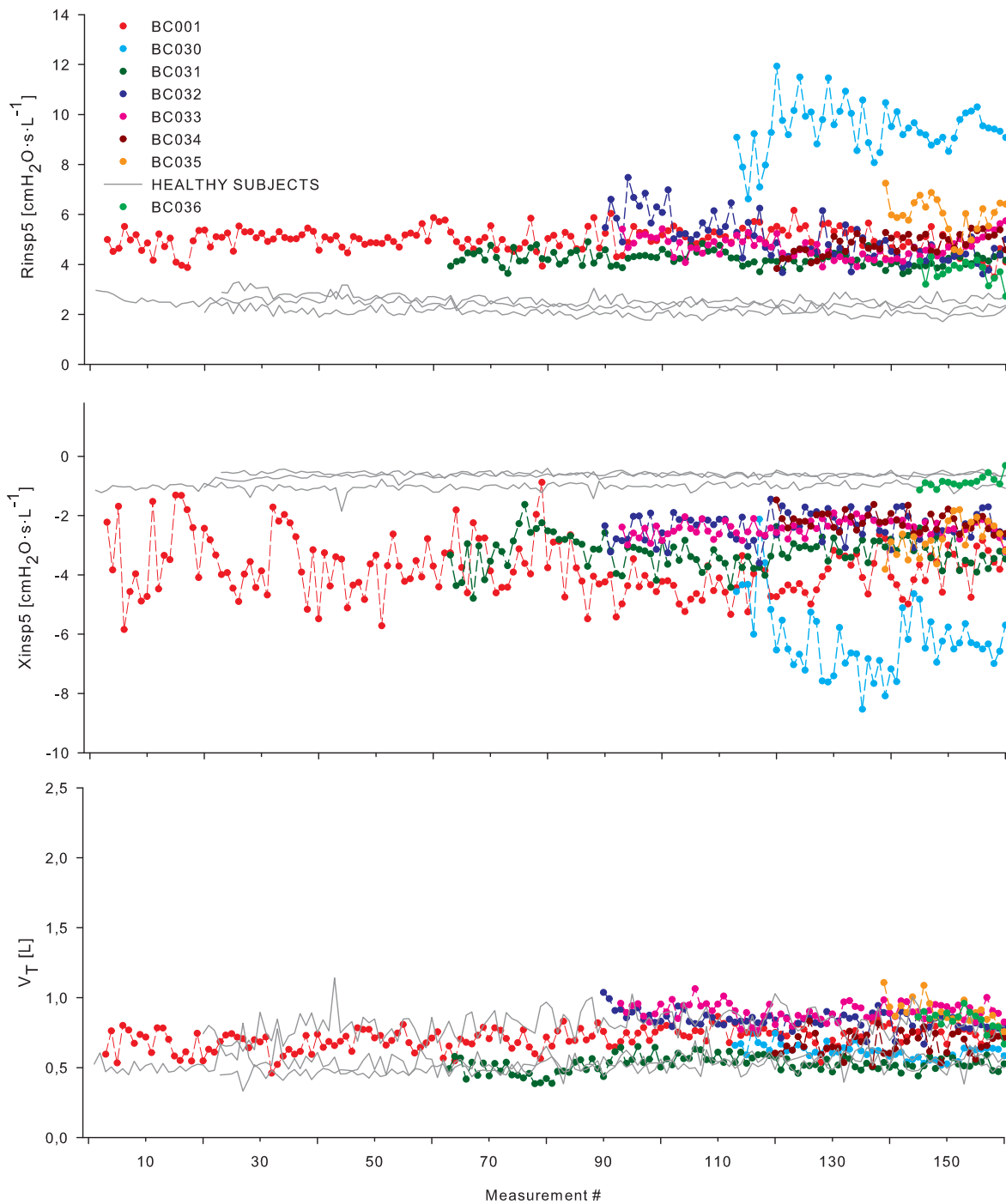
Figure 3.24: Time series of inspiratory resistance at 5Hz (Rinsp5), inspiratory reactance at 5Hz (Xinsp5) and tidal volume (Vt) for 8 COPD patients participating in the clinical study and 3 healthy subjects. The disease strongly affects the lung mechanical impedance, whereas the tidal volume remains comparable. *Data on healthy subjects taken from a previuosly recorded dataset: Gobbi A. Home Monitoring of Respiratory Mechanics in Patients with Chronic Obstructive Lung Diseases.*

Figure 3.25: Simple linear regression plots. For each patient, up to 3 parameters significantly correlated with the CAT-based questionnaire's total score are listed (P<0.05). In case no significantly correlated parameter is present, the most correlated is shown.
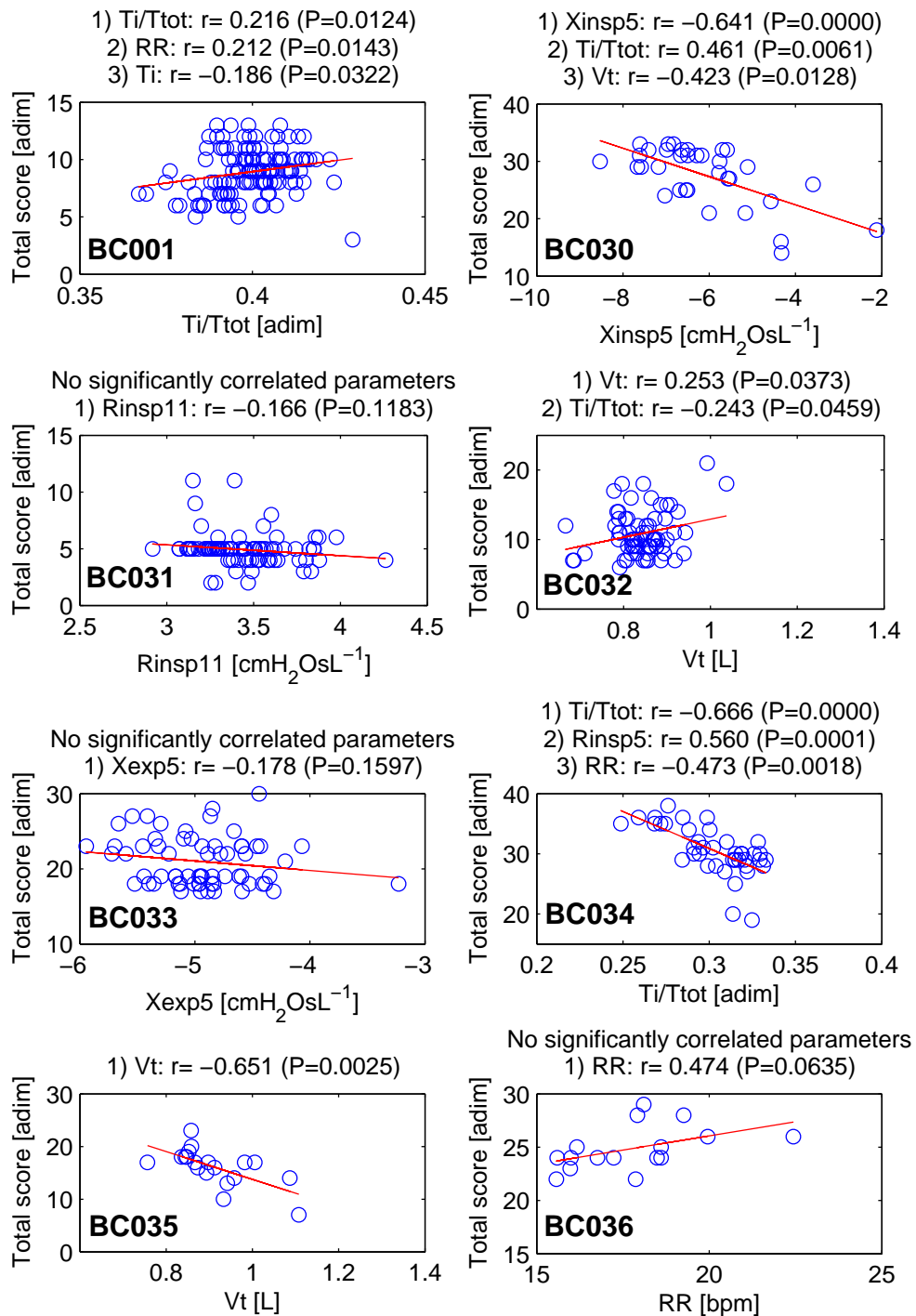
that a person is not able to evaluate the change of the symptoms day by day. As a consequence, he/she would tend to repeat the same answers for several days, so that the informative content of the questionnaire would result drastically reduced. Furthermore, it might happen that the questionnaire can measure only the change of symptoms and not the general gravity of the pathology even though the person responds critically to all the questions every day. This is confirmed by comparing patient BC030 and patient BC034. They both perceive the disease as extremely debilitating (see Figure 3.23) and have a good critical ability (see the high correlation coefficients and low P-values in Figure 3.25), but the level of obstruction, both central and peripheral, is much more pronounced in patient BC030.

On the basis of the data available, the following rules about the validity of the CAT-based questionnaire can be defined:

- if the answers given to the questionnaire are not perfectly mindful, it is likely to result useless;
- if the answers given to the questionnaire are mindful, it is likely to consistently measure how the disease evolves but it might not capture the real magnitude of the disease's impact on the person's life.

Even though the total score does not correlate with any parameter taken singularly, it could provide systematic correlation with groups of parameters. For this reason an analysis of multiple linear correlation (MLR) is performed between the total score and all the 8178 possible combinations of the remaining 13 parameters. Results of this preliminary analysis are omitted because the second round of MLR, which is applied to all the combinations of parameters successfully selected by the preliminary analysis (i.e. Rinsp5, Rexp5 and Xexp5), provides results much more meaningful. Table 3.2 indicates that, among {Rinsp5,Rexp5}, {Rinsp5,Xexp5}, {Rexp5,Xexp5} and {Rinsp5,Rexp5,Xexp5}, only groups of two parameters best correlate with the total score. However results should be commented only for patient BC034, who is the only one that largely meets the threshold of significance. Patients BC030 and BC035 are very close to the threshold, whereas the remaining ones show a very poor correlation.

Results of MLR further strengthen the hypothesis that the CAT-based questionnaire cannot be used as a tool for the punctual assessment of the degree of lung obstruction in COPD.

3.4 Principal component analysis

Time series at disposal clearly show drifting patterns common to many parameters. This behaviour suggests that the underlying phenomenon can be effectively represented by

	Parameters of the most correlated group			R^2adj	P-value
BC001	Rinsp5 (P = 0,0198)	Rexp5 (P = 0,1375)	—	0,0265	0,064733
BC030	Rinsp5 (P = 0,0208)	Rexp5 (P = 0,2679)	—	0,1319	0,042382
BC031	Rinsp5 (P = 0,1685)	Rexp5 (P = 0,1114)	—	0,0191	0,160753
BC032	Rinsp5 (P = 0,2282)	Xexp5 (P = 0,1256)	—	0,0090	0,278090
BC033	Rexp5 (P = 0,6025)	Xexp5 (P = 0,1396)	—	0,0044	0,327162
BC034	Rinsp5 (P < 0,001)	Rexp5 (P = 0,0107)	—	0,3932	0,000028
BC035	Rinsp5 (P = 0,0177)	Xexp5 (P = 0,0422)	—	0,2279	0,049242
BC036	Rexp5 (P = 0,3619)	Xexp5 (P = 0,3156)	—	-0,0633	0,587926

Table 3.2: Multiple linear regression results. For each patient, the group of parameters chosen among Rinsp5, Rexp5 and Xexp5 that best correlates with the CAT-based questionnaire's total score are listed. Additional information consists in the P-values of every parameter taken singularly and on the global evaluation of the regression (R^2_{adj} and global P-value).

a much lower number of variables. In this context, the application of PCA moves from the need of understanding which parameters convey most of the information on the underlying phenomenon.

Figure 3.26 shows that two variables, obtained by a linear combination of the original ones, suffice to represent from the 55.9% to the 80.7% of the total variance represented by the entire set of parameters. Moreover, the arrangement of the original parameters in the new 2-D cartesian space consents to assert that:

- the CAT-based questionnaire's total score lies very close to the origin for most of the patient;
- on average, parameters of respiratory mechanics[5] distribute along the first principal component;

---

[5]As stated in Section 2, parameters describing the lung mechanical impedance are referred to as parameters of *respiratory mechanics*

- on average, parameters of breathing pattern[6] distribute along the second principal component.

The first point may be considered as an integration of the previous results. Besides being little correlated with respiratory parameters, the total score of the CAT-based questionnaire explains an amount of information about the pathology that is very small on average. Even for the patients that complete the questionnaire in a mindful way (i.e., BC030, BC034 and BC035), the modest length and the oblique orientation of the segment named "Tscore" indicates that the total score is quite stationary and superfluous.

The second point, interestingly, indicates that the parameters of respiratory mechanics are the most sensitive to any deterioration or improvement of the health status. Moreover, the fact that resistances and reactances have opposite proportionality coefficients (i.e., the two components of the vector representing them) confirms how they both tend to increase in modulus when the disease aggravates.

The third point shows a rather surprising characteristic of the data, that is, breathing pattern parameters tend to be aligned with the second principal component. This means that they can represent little information due to their small variability, but this information is different from the one explained by the first principal component. The perpendicularity between the two groups of parameters implicates that, if one group undergoes a variation of any kind, this variation not necessarily will correspond to a variation on the other group. In other words, parameters of respiratory mechanics and parameters of breathing pattern might be used to represent different aspect of the pathology.

The plots of Figure 3.26 also suggest that a plot representing the average of the single PCAs could eliminate the inter-subject variability and, at the same time, reduce the level of noise embedded in the original data. The averaging of the single PCAs can be done by separately calculating the absolute value of the components of the 14 vectors representing the total set of parameters, and plotting the new sets of vectors in the principal component 2-D space. Encouraging results are depicted in Figure 3.27. As expected, inter-subject variability is strongly reduced and hence parameters of respiratory mechanics and breathing pattern result strongly aligned with the two principal components. In particular, any of the resistances and respiratory rate (or equivalently Ti) seem to be the variables that best represent the status of the pathology.

Despite the modest amount of data available, PCA is able to strongly emphasize

---

[6]As stated in Section 2, parameters describing respiratory volumes and respiratory frequencies are referred to as parameters of *breathing pattern*
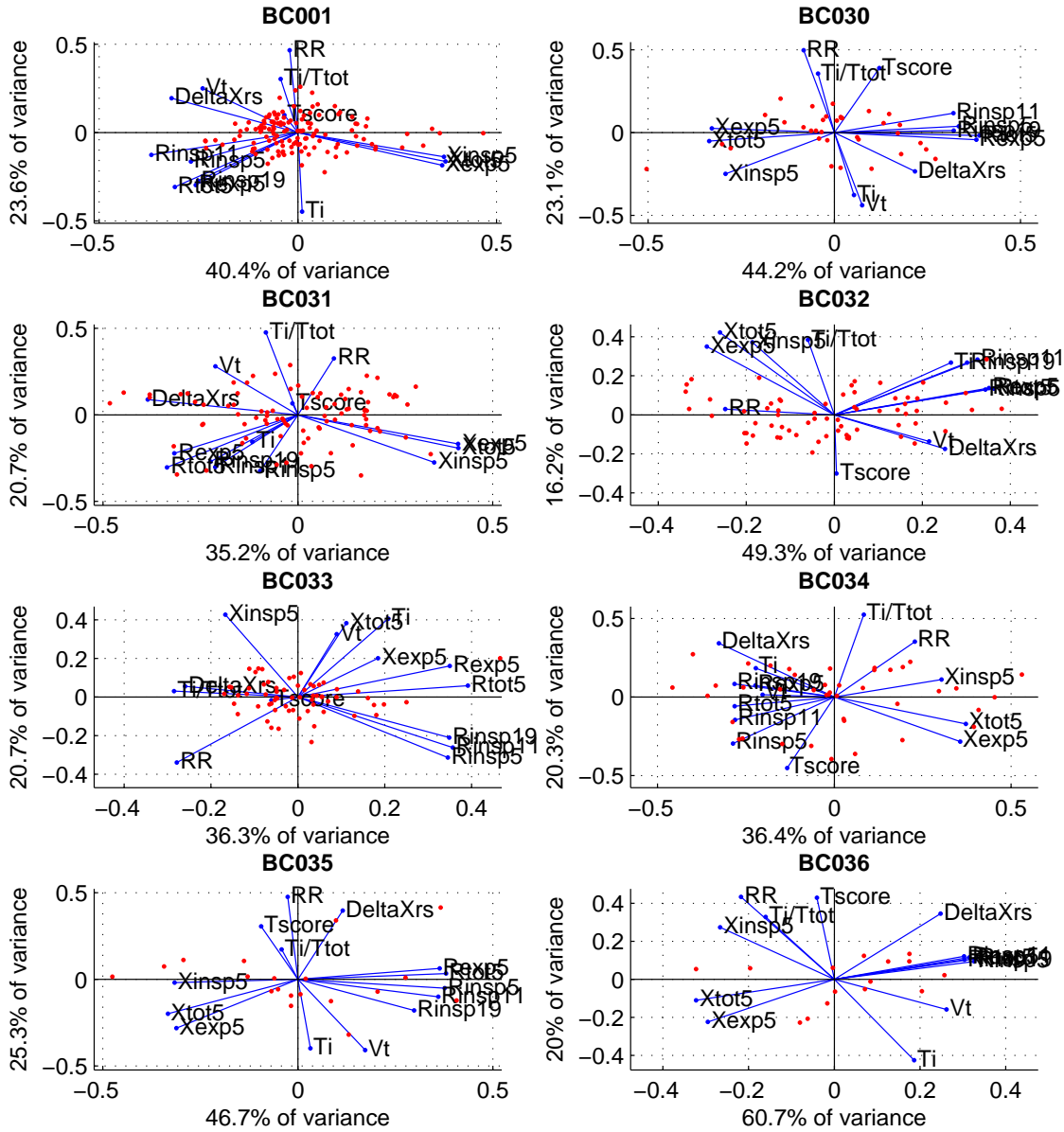
Figure 3.26: Biplot representing the distribution of the measurements (red points) in the new cartesian space produced by the principal component analysis. Blue segments indicate the level of proportionality between the original parameters and the first 2 principal components. The percentage of total variance explained by the two principal components is indicated as well.
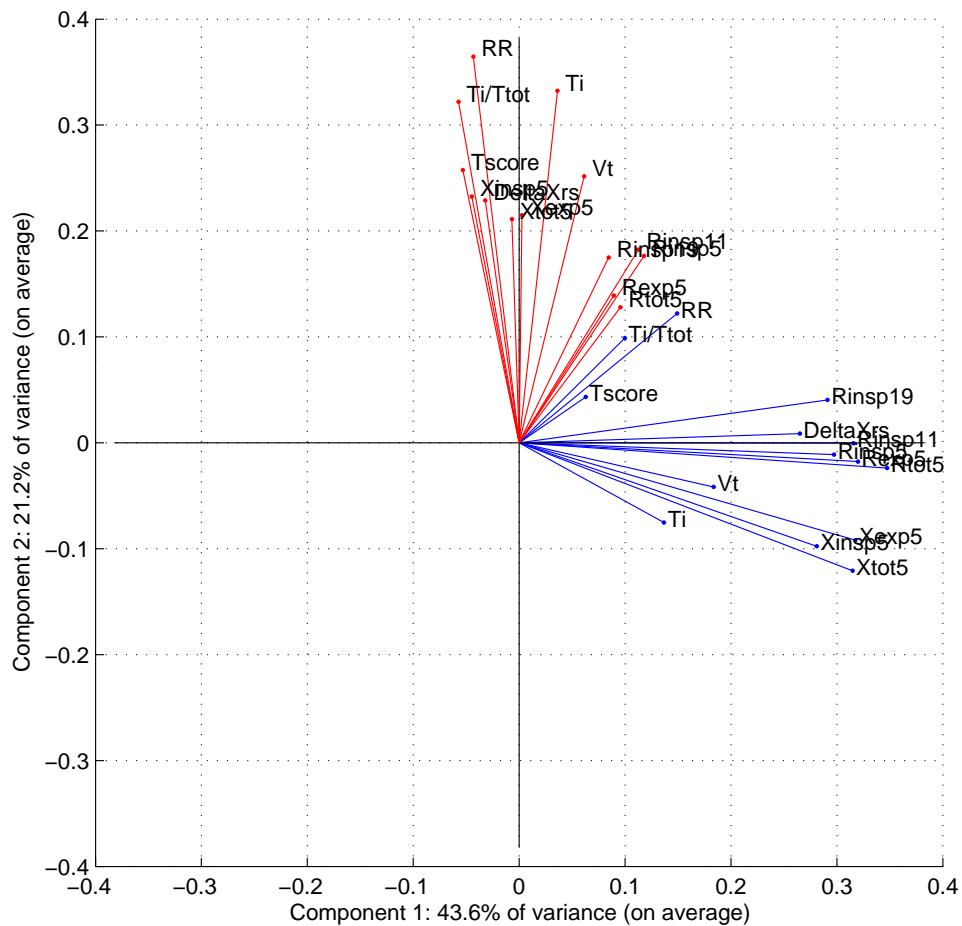
Figure 3.27: Average of the PCAs performed singularly for each patient. All the parameters occupy only the positive portions of the principal components because of the absolute value operation, but this does not affect the conclusion drawn whatsoever. Patient BC033 is excluded from the averaging operation because of its peculiar PCA results. Given the small amount of patients, its inclusion in the average would distort the results too strongly.

which variables might be used to capture the pathophysiologic principles governing the COPD. Furthermore, results are expected to substantially improve as the number of patients participating in the clinical study increases.

3.5   Variability analysis at different time scales

Figure 3.28 reports the average variability observed at different time scales for healthy, asthmatic and COPD subjects. Two-way ANOVA for repeated measurements is used to test for significant differences between COPD and asthmatic patients. Results for the remaining two differences are not displayed as they are superfluous in this work.

Short-term fluctuations in the parameters of respiratory mechanics of COPD patients are significantly greater than those of healthy subjects and asthmatic patients. As the latter suffer from mild to moderate asthma, short-term variability might constitute a measure of the severity of the disease and, at the same time, represent the effects of pathophysiologic mechanisms on both the central (resistance) and peripheral (reactance) airways. Differently, the long-term variability confirms that asthma and COPD are characterized by similar pathological changes in the central airways, whereas inflammatory processes in the peripheral airways and in lung parenchyma are much more pronounced in COPD patients. This is suggested by the fact that fluctuations for temporal scales greater than one day are comparable in asthma and COPD only for parameters of resistance. Parameters of reactance, on the contrary, are characterized by much larger fluctuations in COPD patients.

Moreover, considering the plots in the left column, it is possible to notice how in COPD patients the separation of the inspiration from the expiration part of the resistance helps increasing the sensitivity to changes in airway obstruction [101, 102], as the expiratory part of each breath may be influenced by nonlinearities due to the presence of expiratory flow limitation (EFL), especially in severe and very severe degrees of airway obstruction. Further studies on severe and very severe COPD subjects might include analysis of the differences between expiratory and inspiratory indexes, as it has been demonstrated that they have high sensitivity and specificity when assessing the degree of EFL in these patients [103]. Note that the knees present for variability at 16 and 32 days are artifacts given by the limited number of COPD patients with an amount of data sufficient to calculate those standard deviations, therefore they should not be considered as reliable as the remaining points.

Interestingly, considering the plots in the first three rows, it can be appreciated that the variability has the same magnitude for inspiratory, expiratory and whole breath parameters of respiratory mechanics, regardless the fact that they are resistances and
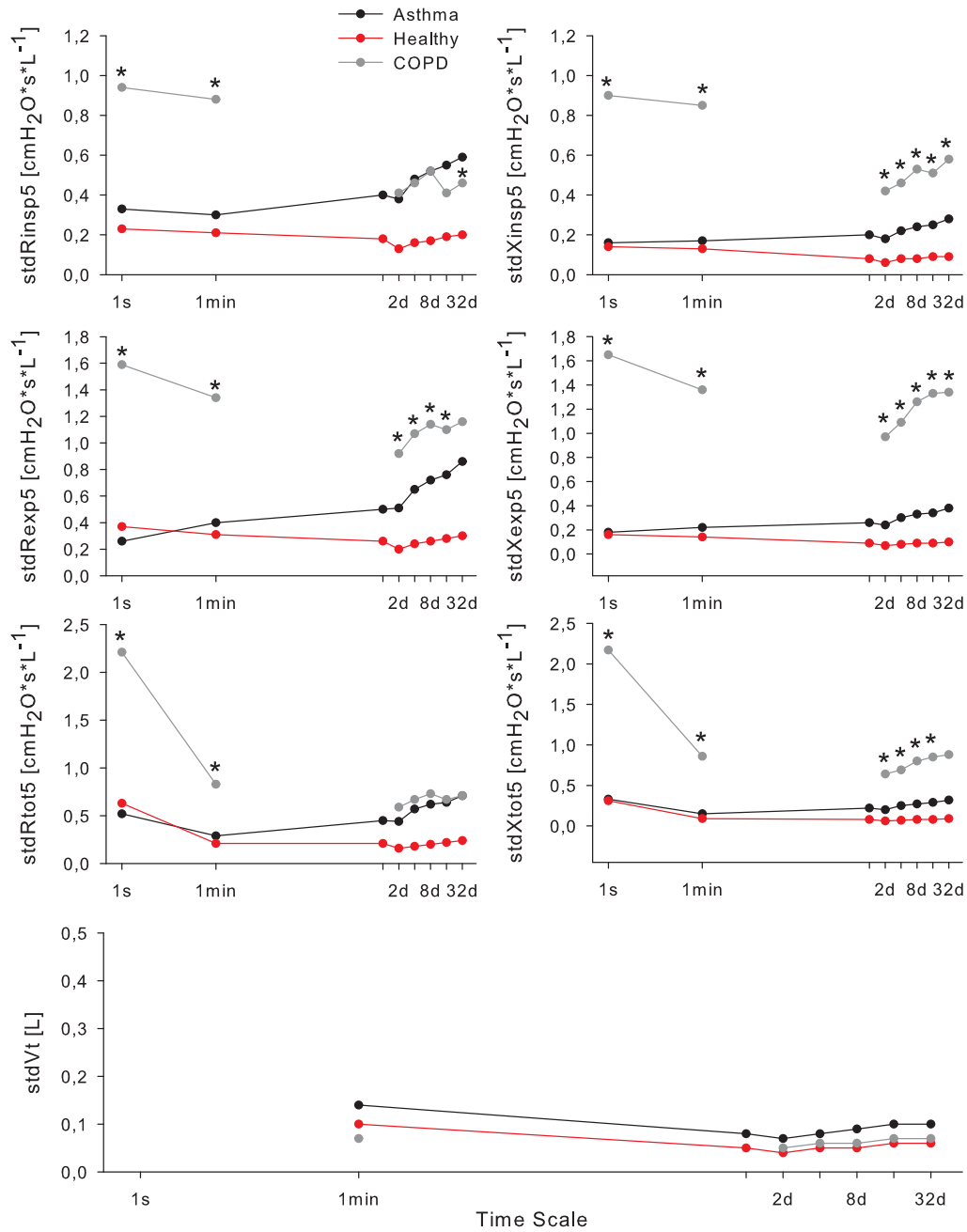
Figure 3.28: Variability at all time scales (semilog plot) of Rinsp5, Rexp5, Rtot5, Xinsp5, Xexp5, Xtot5 and Vt. Red plots are the average variability for the control group while the black and the grey plots are the average variability of the asthmatics and the COPD. The variability at 1 second corresponds to the within-breath variability, at 1 minute to the inter-breath variability, at 1 day to the within-day variability and from 2 days on at the inter-day variability. $*$ = significant difference between COPD and asthmatic patients at a given time scale. P-value at least <0.05 for all. Note that the knees noticeable for variability at 16 and 32 days are artifacts given by the limited number of COPD patients with an amount of data sufficient to calculate those standard deviations.

reactances. This confirms the observations made in PCA, that is, resistance and reactance are inversely proportional but they contain a very similar level of the total variance explained by the complete set of parameters.

As indicated in Section 2, traces for asthmatic and control groups also contain the inter-day variability, as for these subjects one test in the morning and one test in the evening were performed. The characteristic knee present in all the red and black plots of Figure 3.28 indicates that the variability at 1 day is always greater than that at 2 days. This behaviour is caused by an additional amount of variability associated to the diurnal variations of airway caliber and obstruction, especially in asthmatic patients. In order to increase the comprehension of the peculiarities distinguishing COPD and asthma, in future studies it would be interesting to assess whether within-day COPD recordings follow the same pattern as well.

Lastly, the observation of the level of variability of Vt reinforces previous findings of this thesis. The absence of significant difference between the three groups confirms what it can be inferred from Figure 3.24 and the PCA. Indeed, tidal volume is only slightly correlated with parameters of respiratory mechanics, and its content of total variance is quite limited at all time scales.

4   Conclusions

In this Chapter an analysis of respiratory parameters belonging to COPD patients at stage 3 and 4 of GOLD classification is performed. Methods applied to the data encompass automatic outlier detection, missing data estimate, analysis of the CAT-based questionnaire administered by the commercial device deployed in the clinical study (RESMON® PRO), principal component analysis (PCA) and variability analysis at different time scales.

A completely automatic outlier detection is not definable since outliers can be due to several and unpredictable causes. Thus time series must be checked visually and automatic recognition can only support the user during this process. At the current stage, the approach that demonstrates to be most flexible is represented by linear confidence limits. However, logarithmic confidence limits might be chosen for patients experiencing at least one exacerbation during the observational time, whereas percentiles can be effectively used after about one month of observations.

Anyway, a possible future improvement is identified. Although the drawing of horizontal bars giving an idea of the distribution of the values is by itself helpful, regardless the method used to calculate those limits, the knowhow acquired during the

operative visual check of the time series suggests that the standard deviation of daily acquisitions might represent an interesting measure of the quality of the acquisitions themselves. In this way it could be possible to combine both the check on mean values and that on intra-test variability. The definition of an effective threshold for the standard deviation, however, could be rather difficult to be defined. Indeed, more severe pathological conditions are usually accompanied with larger variability of the respiratory parameters. For this reason, the use of the coefficient of variation (CoV) could represent a strategic and effective solution.

The issue of missing data is addressed through nonparametric approach. Flexibility of nonparametric methods, indeed, consent to well estimate missing observation despite the small amount of hypotheses made about the underlying process. The application of an appositely defined method of locally weighted regression (LWR) to the data at disposal show extremely positive results. The method is very powerful and, to the best knowledge of the authors, it is the first application of LWR to missing data estimate.

Results of the actual analysis of respiratory parameters comprehend a study of the significance of the CAT-based questionnaire administered by the RESMON® PRO, PCA and monitoring of temporal changes at different time scales. Despite the modest amount of data currently available, interesting conclusions can be drawn.

First, results firmly demonstrate that the CAT-based questionnaire cannot be used as a tool for the punctual assessment of the degree of lung obstruction in COPD.

Second, PCA indicates that two variables, opportunely defined as a linear combination of the original set of 14 parameters, suffice to explain a fair percentage of the total information content represented by the complete set of parameters. Furthermore, the first principal component is very much related to the parameters of respiratory mechanics whereas the second principal components is somewhat aligned with breathing pattern parameters. This means that the parameters of respiratory mechanics are the most sensitive to any deterioration or improvement of the health status and that parameters of respiratory mechanics and parameters of breathing pattern might be used to represent different aspect of the pathology.

Third, the monitoring of temporal changes at different temporal scales illustrates how short-term fluctuations of parameters of respiratory mechanics might constitute a measure of the severity of the disease and, at the same time, represent the effects of pathophysiologic mechanisms on both the central and peripheral airways. Differently, the long-term variability confirms that asthma and COPD are characterized by similar pathological changes in the central airways, whereas inflammatory processes in the peripheral airways and in lung parenchyma are much more pronounced in COPD patients. In future developments it would be interesting to employ the CoV in place

of standard deviation given its greater robustness against large variations in the mean value, and also the assessment of within-day variability may provide insights into the diurnal variations of airway caliber and obstruction in COPD.

General Conclusions

The purpose of this work can be split in two macro-areas. Firstly, the validation of a new strategy for the home monitoring of chronic respiratory diseases, namely a novel commercial telemedicine system based on the forced oscillation technique (FOT). Secondly, the analysis of clinical data obtained through the employment of this device in the largest and most detailed clinical trial ever defined for the observation of COPD.

It has been suggested that respiratory disorders result when the organism, or part of it, becomes too far from a condition of dynamic equilibrium known as homeokinesis. In chronic respiratory diseases, the characteristic abnormal fluctuations of the respiratory parameters contain a message that might be associated with the mechanisms leading to a status of severe airway obstruction and, eventually, to collapse or death. Up to now, a long-term and reliable daily assessment of lung function has not been feasible with standardized clinical tools, even though its potential usefulness has been proposed since several years as a valid alternative to the punctual assessment of airway obstruction.

Despite the huge progresses in medical devices and in the information and communication technologies, the long-temporal assessment of lung function still relied on the use of written diaries, questionnaires and unreliable peak-flow spirometric indices, which provide data with good temporal resolution but poor accuracy. To overcome this limitation, a home monitoring network for respiratory diseases based on the novel FOT device is introduced. Though not still entered in the common clinical practice, FOT has the advantage of being a simple and reliable instrument to estimate non-invasively the mechanical properties of the respiratory system.

The validation of the device was performed by measuring the respiratory parameters of a group of normal subjects with both the commercial device and a reference FOT system. Results provide generally positive feedback, however a direct and univocal assessment of the difference in accuracy between the two systems is not provable. A more comprehensive evaluation requires additional specific tests that were not performed in this work. In particular, the parallel measurement of the same subject obtainable by connecting the sensors of both systems to the front-end hardware of the commercial device would rule out the "masking" phenomena that affected previous measurements.

Moreover, results indicate that the internal aspirator needed to expel the the exhaled carbon dioxide causes an increase in the respiratory rate, hence a new solution able to guarantee a greater flow rate should be taken into account.

The analysis of respiratory parameters was carried out on 8 COPD patients (at stage 3 and 4 of GOLD classification) who participated in the clinical study this thesis is based on. Methods applied to the data encompass automatic outlier detection, missing data estimate, analysis of the CAT-based questionnaire administered by the commercial device, principal component analysis (PCA) and monitoring of temporal changes of respiratory parameters.

Preliminary results confirm the usability and the effectiveness of the software infrastructure designed to manage the large amount of clinical data acquired. Nevertheless, the knowhow acquired during the operative phase of the work consented to identify aspects to be improved in automatic outlier detection and automatic breath selection.

Outcomes of linear regression and PCA firmly demonstrate that the CAT-based questionnaire cannot be used as a tool for the punctual assessment of the degree of lung obstruction in COPD. In particular, two among the 14 parameters at disposal are enough to explain a fair amount of the total variance. These can be identified in one of the parameters of resistance and in one between respiratory rate and inspiratory time. However, given the modest number of patients available, a weaker assumption should be made, that is, parameters of respiratory mechanics[7] represent most of the information of the underlying phenomenon whereas parameters of breathing pattern[8] explain a smaller amount.

Finally, the monitoring of temporal changes of respiratory parameters illustrates how short-term fluctuations of parameters of respiratory mechanics might constitute a measure of the severity of the disease and, at the same time, represent the effects of pathophysiologic mechanisms on both the central and peripheral airways. Differently, the long-term variability confirms that asthma and COPD are characterized by similar pathological changes in the central airways, whereas inflammatory processes in the peripheral airways and in lung parenchyma are much more pronounced in COPD patients. A criticism that can be made about this variability analysis is that results might be slightly overestimated due to the assessment of the standard deviation at different time scales. Indeed, the use of the coefficient of variation in place of the standard deviation would normalize the magnitude of the fluctuations with respect to the mean value,

---

[7]Parameters describing the lung mechanical impedance are referred to as parameters of *respiratory mechanics*

[8]Parameters describing respiratory volumes and respiratory frequencies are referred to as parameters of *breathing pattern*

giving a more factual quantification of the variability.

In conclusion, this work shows that a home daily monitoring of respiratory mechanics may contribute to advance the understanding of COPD and to improve the clinical management of patients suffering from this disease.

Bibliography

[1] Lopez AD, Shibuya K, Rao C, Mathers CD, Hansell AL, Held LS, et al. Chronic obstructive pulmonary disease: current burden and future projections. Eur Respir J. 2006;27(2):397-412.

[2] Murray CJL, Lopez AD. The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. Cambridge, MA: Harvard University Press; 1996.

[3] Murray CJL, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. Lancet 1997;349(9064):1498-504.

[4] Frey U, Brodbeck T, Majumdar A, Taylor DR, Town GI, Silverman M, Suki B. Risk of severe asthma episodes predicted from fluctuation analysis of airway function. Nature 2005;438(7068):667-670.

[5] Casale R, Pasqualetti P. Cosinor analysis of circadian peak expiratory flow variability in normal subjects, passive smokers, heavy smokers, patients with chronic obstructive pulmonary disease and patients with interstitial lung disease. Respiration. 1997;64(4):251-6.

[6] Thiadens HA, De Bock GH, Dekker FW, Huysman JA, Van Houwelingen JC, Springer MP, Postma DS. Value of measuring diurnal peak flow variability in the recognition of asthma: a study in general practice. Eur Respir J. 1998 Oct;12(4):842-7.

[7] Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N. Development and first validation of the COPD Assessment Test. Eur Respir J. 2009 Sep;34(3):648-54.

[8] Glass L. Synchronization and rhythmic processes in physiology. Nature 2001;410(6825):277-284.

[9] Peng CK, Havlin S, Stanley HE, Goldberger, AL. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series, Chaos. 1995;5(1):82-87.

[10] Que CL, Kenyon CM, Olivenstein R, Macklem PT, Maksym GN. Homeokinesis and short-term variability of human airway caliber. J Appl Physiol. 2001 Sep;91(3):1131-41.

[11] Dougherty CM, Burr RL. Comparison of heart rate variability in survivors and nonsurvivors of sudden cardiac arrest. Am J Cardiol. 1992 Aug 15;70(4):441-8.

[12] Papaioannou VE, Maglaveras N, Houvarda I, Antoniadou E, Vretzakis G. Investigation of altered heart rate variability, nonlinear properties of heart rate signals, and organ dysfunction longitudinally over time in intensive care unit patients. J Crit Care. 2006 Mar;21(1):95-103; discussion 103-4.

[13] Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. Am J Respir Crit Care Med 2001;163(5):1256-1276.

[14] Allen SC. Spirometry in old age. Age Ageing 2003;32(1):4-5.

[15] Brouwer AF, Roorda RJ, Brand PL. Home spirometry and asthma severity in children. Eur Respir J. 2006;28(6):1131-1137.

[16] Rigau J, Farré R, Roca J, Marco S, Herms A, Navajas D. A portable forced oscillation device for respiratory home monitoring. Eur Respir J. 2002 Jan;19(1):146-50.

[17] Rigau J, Burgos F, Hernández C, Roca J, Navajas D, Farré R. Unsupervised self-testing of airway obstruction by forced oscillation at the patient's home. Eur Respir J. 2003 Oct;22(4):668-71.

[18] American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 1995;152:S77-121.

[19] Barnes P, Shapiro S, Pauwels R. Chronic obstructive pulmonary disease: molecular and cellular mechanisms. Eur Respir J. 2003;22:672-688.

[20] Samet JM. Definitions and methodology in COPD research. New York: Marcel Dekker; 1989. p. 1-22.

[21] Birring SS, Brightling CE, Bradding P, Entwisle JJ, Vara DD, Grigg J, et al. Clinical, radiologic, and induced sputum features of chronic obstructive pulmonary disease in nonsmokers: a descriptive study. Am J Respir Crit Care Med 2002;166(8):1078-83.

[22] Barnes PJ, Shapiro SD, Pauwels RA. Chronic obstructive pulmonary disease: molecular and cellular mechanisms. Eur Respir J 2003;22(4):672-88.

[23] Barnes PJ. Mediators of chronic obstructive pulmonary disease. Pharmacol Rev 2004;56(4):515-48.

[24] Rahman I. Oxidative stress in pathogenesis of chronic obstructive pulmonary disease: cellular and molecular mechanisms. Cell Biochem Biophys 2005;43(1):167-88.

[25] Ito K, Ito M, Elliott WM, Cosio B, Caramori G, Kon OM, et al. Decreased histone deacetylase activity in chronic obstructive pulmonary disease. N Engl J Med 2005;352(19):1967-76

[26] Hogg JC. Pathophysiology of airflow limitation in chronic obstructive pulmonary disease. Lancet 2004;364(9435):709-21.

[27] Mullen JB, Wright JL, Wiggs BR, Pare PD, Hogg JC. Reassessment of inflammation of airways in chronic bronchitis. BMJ (Clin Res Ed) 1985; 291:1235-9.

[28] Rennard SI. Inflammation and repair processes in chronic obstructive pulmonary disease. Respir Crit Care Med 1999; 160: S12-6.

[29] Hogg JC, Macklem PT, Thurlbeck WM. Site and nature of airway obstruction in chronic obstructive lung disease. N Engl J Med 1968;278:1355-60.

[30] Leopold JG, Goeff J. Centrilobular form of hypertrophic emphysema and its relation to chronic bronchitis. Thorax 1957; 12:219-35.

[31] Hogg JC, Chu F, Utokaparch S, Woods R, Elliott WM, Buzatu L, et al. The nature of small-airway obstruction in chronic obstructive pulmonary disease. N Engl J Med 2004;350(26):2645-53.

[32] O'Donnell DE, Revill SM, Webb KA. Dynamic hyperinflation and exercise intolerance in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2001;164(5):770-7.

[33] Burgel PR, Nadel JA. Roles of epidermal growth factor receptor activation in epithelial cell repair and mucin production in airway epithelium. Thorax 2004;59(11):992-6.

[34] Barbera JA, Peinado VI, Santos S. Pulmonary hypertension in chronic obstructive pulmonary disease. Eur Respir J 2003;21(5):892-905.

[35] Wouters EF, Creutzberg EC, Schols AM. Systemic effects in COPD. Chest 2002;121(5 Suppl):127S-30S.

[36] Agusti AG, Noguera A, Sauleda J, Sala E, Pons J, Busquets X. Systemic effects of chronic obstructive pulmonary disease. Eur Respir J 2003;21(2):347-60.

[37] Tirimanna PR, van Schayck CP, den Otter JJ, van Weel C, van Herwaarden CL, van den Boom G, et al. Prevalence of asthma and COPD in general practice in 1992: has it changed since 1977? Br J Gen Pract 1996;46(406):277-81.

[38] Halbert RJ, Natoli JL, Gano A, Badamgarav E, Buist AS, Mannino DM. Global burden of COPD: systematic review and meta-analysis. Eur Respir J 2006.

[39] Halbert RJ, Isonaka S, George D, Iqbal A. Interpreting COPD prevalence estimates: what is the true burden of disease? Chest 2003;123(5):1684-92.

[40] Fukuchi Y, Nishimura M, Ichinose M, Adachi M, Nagai A, Kuriyama T, et al. COPD in Japan: the Nippon COPD Epidemiology study. Respirology 2004;9(4):458-65.

[41] Menezes AM, Perez-Padilla R, Jardim JR, Muino A, Lopez MV, Valdivia G, et al. Chronic obstructive pulmonary disease in five Latin American cities (the PLATINO study): a prevalence study. Lancet 2005;366(9500):1875-81.

[42] COPD Prevalence in 12 Asia-Pacific Countries and regions: Projections based on the COPD prevalence estimation model. Regional COPD Working Group. Respirology 2003;8:192-8.

[43] National Heart, Lung, and Blood Institute. Morbidity & Mortality: Chartbook on Cardiovascular, Lung, and Blood Diseases. Bethesda, MD: US Department. of Health and Human Services, Public Health Service, National Institutes of Health; 1998.

[44] Soriano JR, Maier WC, Egger P, Visick G, Thakrar B, Sykes J, et al. Recent trends in physician diagnosed COPD in women and men in the UK. Thorax 2000;55:789-94.

[45] Chapman KR. Chronic obstructive pulmonary disease: are women more susceptible than men? Clin Chest Med 2004;25(2):331-41.

[46] Schellevis FG, Van de Lisdonk EH, Van der Velden J, Hoogbergen SH, Van Eijk JT, Van Weel C. Consultation rates and incidence of intercurrent morbidity among patients with chronic disease in general practice. Br J Gen Pract 1994;44(383):259-62.

[47] Jemal A, Ward E, Hao Y, Thun M. Trends in the leading causes of death in the United States, 1970-2002. JAMA 2005;294(10):1255-9.

[48] Chapman KR, Mannino DM, Soriano JB, Vermeire PA, Buist AS, Thun MJ, et al. Epidemiology and costs of chronic obstructive pulmonary disease. Eur Respir J 2006;27(1):188-207.

[49] Mannino DM, Homa DM, Akinbami LJ, Ford ES, Redd SC. Chronic obstructive pulmonary disease surveillance - United States, 1971-2000. MMWR Surveill Summ 2002;51(6):1-16.

[50] Buist AS, Vollmer WM. Smoking and other risk factors. Textbook of respiratory medicine. Philadelphia: WB Saunders; 1994. p. 1259-87.

[51] Seemungal TAR, Donaldson GC, Bhowmik A, Jeffries DJ, Wedzicha JA. Time course and recovery of exacerbations in patients with chronic obstructive pulmonary disease. Respir Crit Care Med 2000; 161: 1608-1613.

[52] Burge S, Wedzicha JA. COPD exacerbations: definitions and classifications. Eur Respir J 2003; 21: Suppl 41, 46s-53s.

[53] Bhowmik A, Seemungal TAR, Sapsford Rj. Relation of sputum inflammatory markers to symptoms and physiological changes at COPD exacerbations. Thorax 2000; 55:114-200.

[54] Anthonisen NR, Manfreda J, Warren CPW, Hershfield ES, Harding GK, Nelson NA. Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease. Ann Intern Med 1987; 106: 196-204.

[55] Hill AT, Bayley D, Stockley RA. The interrelationship of sputum inflammatory markers in patients with chronic bronchitis. Respir Crit Care Med 1999; 160: 893-8.

[56] Kelly CA, Gibson GJ. Relation between FEV1 and peak expiratory flow in patients with chronic obstructive pulmonary disease. Thorax 1988; 43: 335-6.

[57] Ball P, Harris JM, Lowson D, Tillotson G, Wilson R Acute infective exacerbations of chronic bronchitis. Q J Med 1995, 88:61-8.

[58] Seemungal Tar, Harper-Owen R, Bhowmik A, et al Respiratory viruses, symptoms and inflammatory markers in acute exacerbations and stable chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2001; 164: 1618-23.

[59] Mannino DM, Gagnon RC, Petty TL, Lydick E. Obstructive lung disease and low lung function in adults in the United States: data from the National Health and Nutrition Examination Survey, 1988-1994. Arch Intern Med. 2000 Jun 12;160(11):1683-9.

[60] Fletcher C, Peto R. The natural history of chronic airflow obstruction. BMJ 1977; 1:1645-8.

[61] Bellia V, Cibella F, Coppola P, Greco V, Insalaco G, Milone F, Oddo S, Peralta G. Variability of peak expiratory flow rate as a prognostic index in asymptomatic asthma. Respiration. 1984;46(3):328-33.

[62] Hetzel MR, Clark TJ, Branthwaite MA. Asthma: analysis of sudden deaths and ventilatory arrests in hospital. Br Med J. 1977 Mar 26;1(6064):808-11.

[63] Quanjer PH, Lebowitz MD, Gregg I, Miller MR, Pedersen OF. Peak expiratory flow: conclusions and recommendations of a Working Party of the European Respiratory Society. Eur Respir J Suppl. 1997 Feb;24:2S-8S.

[64] Carvalhaes-Neto N, Lorino H, Gallinari C, Escolano S, Mallet A, Zerah F, Harf A, Macquin-Mavier I. Cognitive function and assessment of lung function in the elderly. Am J Respir Crit Care Med. 1995 Nov;152(5 Pt 1):1611-5.

[65] Gayrard P, Orehek J, Grimaud C, CHarpin J. Bronchoconstrictor effects of a deep inspiration in patients with asthma. Am Rev Respir Dis. 1975 Apr;111(4):433-9.

[66] Higgins BG, Britton JR, Chinn S, Cooper S, Burney PG, Tattersfield AE. Comparison of bronchial reactivity and peak expiratory flow variability measurements for epidemiologic studies. Am Rev Respir Dis. 1992 Mar;145(3):588-93.

[67] Ryan G, Latimer KM, Dolovich J, Hargreave FE. Bronchial responsiveness to histamine: relationship to diurnal variation of peak flow rate, improvement after bronchodilator, and airway calibre. Thorax. 1982 Jun;37(6):423-9.

[68] Brand PL, Duiverman EJ, Postma DS, Waalkens HJ, Kerrebijn KF, van Essen-Zandvliet EE. Peak flow variation in childhood asthma: relationship to symptoms, atopy, airways obstruction and hyperresponsiveness. Dutch CNSLD Study Group. Eur Respir J. 1997 Jun;10(6):1242-7.

[69] Brand PL, Duiverman EJ, Waalkens HJ, van Essen-Zandvliet EE, Kerrebijn KF. Peak flow variation in childhood asthma: correlation with symptoms, airways obstruction, and hyperresponsiveness during long-term treatment with inhaled corticosteroids. Dutch CNSLD Study Group. Thorax. 1999 Feb;54(2):103-7.

[70] Simons FE. A comparison of beclomethasone, salmeterol, and placebo in children with asthma. Canadian Beclomethasone Dipropionate-Salmeterol Xinafoate Study Group. N Engl J Med. 1997 Dec 4;337(23):1659-65.

[71] Verberne AA, Frost C, Roorda RJ, van der Laag H, Kerrebijn KF. One year treatment with salmeterol compared with beclomethasone in children with asthma. The Dutch Paediatric Asthma Study Group. Am J Respir Crit Care Med. 1997 Sep;156(3 Pt 1):688-95.

[72] Gulotta C, Suki B, Brusasco V, Pellegrino R, Gobbi A, Pedotti A, Dellacà RL. Monitoring the temporal changes of respiratory resistance: a novel test for the management of asthma. Am J Respir Crit Care Med. 2012 Jun 15;185(12):1330-1.

[73] Kaczka DW, Ingenito EP, Lutchen KR. Technique to determine inspiratory impedance during mechanical ventilation: implications for flow limited patients. Ann Biomed Eng. 1999 May-Jun;27(3):340-55.

[74] Oostveen E, MacLeod D, Lorino H, Farré R, Hantos Z, Desager K, Marchal F; ERS Task Force on Respiratory Impedance Measurements. The forced oscillation technique in clinical practice: methodology, recommendations and future developments. Eur Respir J. 2003 Dec;22(6):1026-41.

[75] Michaelson ED, Grassman ED, Peters WR. Pulmonary mechanics by spectral analysis of forced random noise. J Clin Invest. 1975 Nov;56(5):1210-30.

[76] Hanneman SK. Design, analysis, and interpretation of method-comparison studies. AACN Adv Crit Care. 2008 Apr-Jun;19(2):223-34.

[77] Fisher RA. Statistical Methods for Research Workers. Edinburgh: Oliver and Boyd 1925, p.43.

[78] Cramer D, Howitt D. The SAGE Dictionary of Statistics: A Practical Resource for Students in the Social Sciences. 2004, p.76.

[79] Gimeno F, van der Weele LT, Koëter GH, de Monchy JG, van Altena R. Variability of forced oscillation (Siemens Siregnost FD 5) measurements of total respiratory resistance in patients and healthy subjects. Ann Allergy. 1993 Jul;71(1):56-60.

[80] Dellaca RL, Gobbi A, Pastena M, Pedotti A, Celli B. Home monitoring of within-breath respiratory mechanics by a simple and automatic forced oscillation technique device. Physiol Meas. 2010 Apr;31(4):N11-24. Epub 2010 Feb 24.

[81] Gobbi A, Milesi I, Govoni L, Pedotti A, Dellaca RL. A New Telemedicine System for the Home Monitoring of Lung Function in Patients with Obstructive Respiratory Diseases. International Conference on eHealth, Telemedicine, and Social Medicine 2009;117-122.

[82] Pena D, Tiao GC, Tsay RS. A Course in Time Series Analysis. John Wiley, New York, 2001.

[83] Tsay RS. Time Series Model Specification in the Presence of Outliers. Journal of the American Statistical Association 1986 Mar;81(393):132-141.

[84] Ljung GM. On Outlier Detection in Time Series. Journal of the Royal Statistical Society 1993;55(1993):559-567.

[85] Kaya A. Statistical Modelling for Outlier Factors. Ozean Journal of Applied Sciences 2010;3(1).

[86] Tsay RS. Outliers, Level Shifts, and Variance Changes in Time Series. Journal of Forecasting 1988;7:1-20.

[87] Kaiser R, Maravall A. Seasonal Outliers in Time Series. 2001.

[88] Little RJA, Rubin DB. Statistical Analysis with Missing Data, 2nd Edition. John Wiley, New York, 2002.

[89] Bittanti S. Identificazione dei Modelli e Sistemi Adattativi. Pitagora Editrice, Bologna, 2002.

[90] Box GEP, Jenkins G. Time Series Analysis: Forecasting and Control. Holden-Day, 1976.

[91] Simpson CR, Hippisley-Cox J, Sheikh A. Trends in the epidemiology of chronic obstructive pulmonary disease in England: a national study of 51,804 patients. Br J Gen Pract. 2010 July 1;60(576):e277-e284.

[92] Fox J. Nonparametric Regression: Appendix to An R and S-PLUS Companion to Applied Regression, January 2002.

[93] Cleveland WS. Robust Locally Weighted Regression and Smoothing Scatterplots. Journal of the American Statistical Association 1979;74:829-836.

[94] Cleveland WS, Devlin SJ. Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. Journal of the American Statistical Association 1988;83:596-610.

[95] Cohn DA, Ghahramani Z, Jordan MI. Active Learning with Statistical Models. Journal of Artificial Intelligence Research 4 1996;129-145.

[96] Schaal S, Atkeson C. Robot Juggling: An Implementation of Memory-based Learning. Control Systems 1994;14(1):57-71.

[97] Rubin DB. Multiple Imputations in Sample Surveys. Proc. Survey Res. Meth. Sec., Am. Statist. Assoc. 1978;20-34.

[98] Mathers CD, Stein C, Ma Fat D, et al. Global Burden of Disease 2000: Version 2 Methods and Results. Geneva, World Health Organization, 2000;1-108.

[99] Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N. Development and first validation of the COPD Assessment Test. Eur Respir J. 2009 Sep;34(3):648-54.

[100] Jolliffe IT. Principal Component Analysis, 2nd edition, Springer, 2002.

[101] Kubota M, Shirai G, Nakamori T, Kokubo K, Masuda N, Kobayashi H. Low frequency oscillometry parameters in COPD patients are less variable during inspiration than during expiration. Respir Physiol Neurobiol. 2009 Apr 30;166(2):73-9. Epub 2009 Feb 6.

[102] Dellaca RL, Pompilio PP, Walker PP, Duffy N, Pedotti A, Calverley PM. Effect of bronchodilation on expiratory flow limitation and resting lung mechanics in COPD. Eur Respir J. 2009 Jun;33(6):1329-37. Epub 2009 Jan 22.

[103] Dellaca RL, Santus P, Aliverti A, Stevenson N, Centanni S, Macklem PT, Pedotti A, Calverley PM. Detection of expiratory flow limitation in COPD using the forced oscillation technique. Eur Respir J. 2004 Feb;23(2):232-40.