Aus dem Helmholtz Zentrum München

Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH)

Institut für Gesundheitsökonomie und Management in Gesundheitswesen

Institutsleiter: Prof. Dr. Reiner Leidl

A COMPARISON OF DISEASE–SPECIFIC HEALTH–RELATED QUALITY OF LIFE ASSESSMENT TOOLS AGAINST THE GENERIC EQ–5D–5L

LESSONS LEARNT FROM THREE CHRONIC LUNG DISEASES

Dissertation

zum Erwerb des Doktorgrades der Humanbiologie

an der Medizinischen Fakultät der

Ludwig–Maximilians–Universität zu München

vorgelegt von

Boglárka Lilla Szentes

aus Budapest

2021

Mit Genehmigung der medizinischen Fakultät

Der Universität München

Berichterstatter:	PD Dr. Larissa Schwarzkopf
Mitberichterstatter:	Prof. Dr. Erika von Mutius
	Prof. Dr. Nikolaus Kneidinger
Mitbetreuung durch den promovierten Mitarbeiter:	Prof. Dr. Rolf Holle
Dekan:	Prof. Dr. med. Thomas Gudermann
Tag der mündlichen Prüfung:	16.12. 2021

Dum spiro spero

(While I breathe, I hope)

/Cicero/

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass die Dissertation mit dem Thema

A COMPARISON OF DISEASE–SPECIFIC HEALTH–RELATED QUALITY OF LIFE ASSESSMENT TOOLS AGAINST THE GENERIC EQ–5D–5L
LESSONS LEARNT FROM THREE CHRONIC LUNG DISEASES

selbständig und ohne unerlaubte fremde Hilfe angefertigt und keine anderen, als die von mir angegebenen Schriften und Hilfsmittel benutzt wurden. Die den benutzten Werken wörtlich und inhaltlich entnommenen Stellen sind kenntlich gemacht.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 29.03.2022 Boglárka Lilla Szentes

Ort, Datum Unterschrift Doktorandin

Table of contents

List of abbreviations

AQLQ	Asthma Quality of Life Questionnaire
CAT	COPD Assessment Test
CCQ	Clinical COPD Questionnaire
COPD	Chronic Obstructive Pulmonary Disease
EQ–5D–5L	EuroQol 5 Dimensions
HRQL	Health–Related Quality of Life
ILD	Interstitial Lung Diseases
K–BILD	King’s Brief Interstitial Lung Disease Questionnaire
MESH	Medical Subject Headings
mMRC	Modified British Medical Research Council Dyspnea scale
QALY	Quality–Adjusted LIfe Years
SGRQ	St George's Respiratory Questionnaire

List of publications in this thesis

Szentes BL, Kreuter M, Bahmer T, Birring SS, Claussen M, Waelscher J, Leidl R, Schwarzkopf L. Quality of life assessment in interstitial lung diseases:a comparison of the disease–specific K–BILD with the generic EQ–5D–5L. Respir Res. 2018 May 25;19(1):101.
doi: 10.1186/s12931–018–0808–x.

Szentes BL, Schwarzkopf L, Kirsch F, Schramm A, Leidl R. Measuring quality of life in COPD patients: comparing disease–specific supplements to the EQ–5D–5L. Expert Rev Pharmacoecon Outcomes Res. 2020 Oct;20(5):523–529.
doi: 10.1080/14737167.2019.1662302.

Szentes BL, Schultz K, Nowak D, Schuler M, Schwarzkopf L. How does the EQ–5D–5L perform in asthma patients compared with an asthma–specific quality of life questionnaire? BMC Pulm Med. 2020 Jun 13;20(1):168.
doi: 10.1186/s12890–020–01205–8.

Böckmann D, Szentes BL, Schultz K, Nowak D, Schuler M, Schwarzkopf L. Cost-Effectiveness of Pulmonary Rehabilitation in Patients With Bronchial Asthma: An Analysis of the EPRA Randomized Controlled Trial. Value Health. 2021 Sep;24(9):1254-1262.
doi: 10.1016/j.jval.2021.01.017. Epub 2021 Apr 30. PMID: 34452704.

Summary

The lung is an essential part of the human body. There are numerous different chronic diseases that damage the lung and cause respiratory symptoms such as cough, wheezing, sputum and shortness of breath. The different chronic diseases have varying onset, intensity and severity, but they all impair the health–related quality of life (HRQL) of patients. The extent of the impairments varies and they have not yet been fully studied . It is important to understand not only how patients with the same disease differ from each other, but also how they differ from patients suffering from different lung diseases and from the general population. To assess these differences, two approaches are available. On the one hand, to measure differences between patients with the same disease, disease–specific questionnaires are available. These assess symptoms and impairments characteristic of the distinct disease. On the other hand, to draw conclusions across diseases and in comparison with the general population generic tools are available. Generic tools focus on aspects of everyday life and thus depict a different spectrum ofHRQL than disease–specific tools. The combined use of these tools could provide a global picture about the HRQL of patients.

The focus of this thesis is to evaluate disease–specific questionnaires in their disease area. In addition, it is investigated whether a well–known generic tool, the EuroQol 5 dimension (EQ–5D–5L) questionnaire can adapt to the problems of patients with lung–related diseases and thus depict their condition realistically. Because allocation decisions and treatment options partly depend on the results of HRQL questionnaires, the characterization and evaluation of the tools is of particular importance.

The first paper investigates the group of interstitial lung diseases (ILD), which is an umbrella term for more than 200 rare chronic lung diseases. At the time of the dissertation only one HRQL tool was available for this disease group, the King's Brief Interstitial Lung Disease (K–BILD) questionnaire. The K–BILD has not yet been validated in the German population and its use is limited. Therefore, we compared it with the EQ–5D–5L for evaluation purposes. Both tools are suitable to depict HRQL in patients and show the negative association between disease severity and HRQL impairments and are thus applicable in the German ILD population. Additionally, both tools react to the same main influencing factors;

however, the K–BILD reacts more strongly to clinical parameters, whereas the EQ–5D–5L is stronger on sociodemographic factors.

The second paper focuses on Chronic Obstructive Pulmonary Disease (COPD). As COPD has high prevalence with low morbidity, multiple HRQL tools exist e.g. the ’COPD Assessment Test’ (CAT), the ’St George's Respiratory Questionnaire’ (SGQR) or the ’Clinical COPD Questionnaire’ (CCQ). However, there is no consensus on the most suitable one in a clinical setting. This paper investigates the psychometric properties of three widely known disease–specific tools (CAT, SGRQ, CCQ) in comparison with and in interaction with the above–mentioned EQ–5D–5L. The chosen instrument, the CAT is short, easy to understand and administer, explains a large part of the variance in the EQ–5D–5L and is recommended for use in treatment decisions, thus providing physicians with a tool to support easy and fast medical decision–making.

The third paper in the thesis evaluates the implementation of the EQ–5D–5L in asthma patients to answer health economic questions. The evaluation is based on discriminatory power, reliability and sensitivity in comparison with a well–established disease–specific tool, the ’Asthma Quality of Life Questionnaire’ (AQLQ). Additionally the minimal important difference (MID) is calculated for the EQ–5D–5L in asthma patients, and the existing MID for the AQLQ is re–estimated in the study population. This evaluation is based on data from a randomized controlled trial with a waiting group design for pulmonary rehabilitation. This setting allowed the in–depth evaluation of these questionnaires between and within individuals. Although both tools showed good discriminatory power and reliability, the generic tool was lacking enough sensitivity to notice relevant HRQL changes in asthma patients. This result is also confirmed by the calculation of the MID. The detection of changes, however, would be very important in evaluating interventions. Therefore, the complementary use of disease–specific tools in health economic evaluations would be beneficial.

The fourth paper tests the knowledge gained about HRQL tools in a real–world setting following a common health economic evaluation approach. It investigates whether tools with different psychometric properties deliver different results in the cost–effectiveness analysis of the above–mentioned pulmonary rehabilitation programme in asthma patients.

Three endpoints were investigated, two disease–specific tools and one generic HRQL tool. The expectations based on paper three were fulfilled: the quality–adjusted life years derived from EQ–5D–5L show difficulties in depicting HRQL changes resulting from the rehabilitation programme, despite the fact that the disease–specific tool showed the beneficial and cost–effective outcome. Thus, the results underline the significance of the use of combined measures.

In conclusion, the thesis highlights the importance of the combination of generic and disease–specific questionnaires in patient care. Even though EQ–5D–5L performs fairly well in every disease area, the psychometric values of the disease–specific tools mostly outperform the EQ–5D–5L. As a consequence, patient–relevant aspects could be overlooked by choices based solely on EQ–5D–5L results, especially in asthma patients. The combined use of these tools would enable patient centered treatment decisions to be made and would allow the drawing of complex conclusions about HRQL.

Zusammenfassung

Die Lunge ist ein essentieller Teil unseres Körpers. Eine Vielzahl von chronischen Krankheiten kann die Lunge angreifen und respiratorische Symptome wie Husten, Keuchen, Auswurf oder Atemnot auslösen. Die verschiedenen chronischen Krankheiten unterscheiden sich hinsichtlich Einsetzen,Intensität und Schwere, welche die gesundheitsbezogene Lebensqualität (HRQL) der Betroffenen in unterschiedlicher Form beeinträchtigen. Die Ausmaße der Beeinträchtigungen sind bis jetzt nicht vollständig untersucht. Die HRQL–Unterschiede in den Patienten verschiedener Krankheitsgruppen und gegenüber der gesunden Population sind wichtige Aspekte um die HRQL der Betroffenen richtig zu verstehen. Um diese Beeinträchtigungen zu messen und quantifizieren, sind zwei Möglichkeiten verfügbar. Einerseits werden bei Patienten mit gleicher Erkrankung krankheitsspezifische Erhebungsinstrumente angewendet, die sich auf ihre spezifischen Symptome konzentrieren und dadurch die krankheitsspezifische Einschränkungen betonen. Zum anderen kommen generische Erhebungsinstrumente zur Anwendung, die eine Gegenüberstellung von Erkrankten mit der gesunden Population ermöglichen. Generische Instrumente fokussieren auf die Beeinträchtigungen in Alltagssituationen und liefern dadurch ein umfassenderes Bild über HRQL als krankheitsspezifische Instrumente. Eine kombinierte Nutzung dieser Instrumente könnte ein ausdifferenziertes Bild über die HRQL der Patienten bilden.

Diese Arbeit umfasst vier Manuskripte zur Evaluation verschiedener krankheitsspezifischer Erhebungsinstrumente in deren jeweiligen Krankheitsgebieten. Zusätzlich wird im Vergleich ein anerkannter generischer Fragebogen, der ’EuroQol 5 dimension questionnaire’ (EQ–5D–5L) untersucht, um zu ermitteln, ob dieser die Bedürfnisse der Patienten mit Lungenerkrankungen erkennen und deren Situationen realistisch abbilden kann. Allokationsentscheidungen und Behandlungsoptionen sind zum Teil von den Ergebnissen zur HRQL abhängig, deshalbist die Beschreibung und Beurteilung der Instrumente von hoher Bedeutung

Das erste Manuskript fokussiert sich auf Patienten mit interstitiellen Lungenerkrankungen (ILD). ILDs umfassen mehr als 200 chronische Lungenerkrankungen. Zu der Zeit der Analyse war für diese Patientengruppe nur der King’s Brief Interstitial Lung Disease (K–BILD)

Fragebogen verfügbar. Dieser war im deutschsprachigen Raum noch nicht validiert worden und kaum im Einsatz. Die Untersuchung im ersten Manuskript zieht einen Vergleich zwischen dem K–BILD und dem EQ–5D–5L. Die Ergebnisse zeigen, dass beide Fragebögen adäquat in der deutschen ILD Population einsetzbar sind und negative Assoziation zwischen Krankheitsschwere und Lebensqualität zeigen können. Beide Instrumente werden von den gleichen Einflussfaktoren beeinflusst, jedoch reagiert K–BILD ausgeprägter auf klinische Parameter wohingegen der EQ–5D–5L stärker auf soziodemographische Faktoren reagiert.

Das zweite Manuskript untersucht die Gruppe mit chronisch obstruktiver Lungenerkrankung (COPD), die einen beträchtlichen Teil der Weltbevölkerung ausmacht. Da COPD eine hohe Prävalenz und zugleich niedrige Mortalität hat, gibt es in dem Krankheitsfeld viele verschiedene Lebensqualitätsinstrumente, zum Beispiel den ’COPD Assessment Test’ (CAT), den St George's Respiratory Questionnaire (SGRQ) oder den ’Clinical COPD Questionnaire’ (CCQ). Allerdings gibt es keinen Konsensus über das im Klinikalltag am besten geeignete Instrument. Demzufolge werden im zweiten Manuskript die psychometrischen Eigenschaften von drei weit verbreiteten krankheitsspezifischen Instrumenten (CAT, CCQ, SGR) im Vergleich zu dem oben erwähnten EQ–5D–5L untersucht. Das gewählte Instrument; der CAT, ist kurz, gut verständlich, bildet Lebensqualität angemessen ab, erklärt ein großer Teil die Varianz der EQ–5D–5L und wird bei der Wahl von Behandlungsoptionen in der Routineversogung benutzt. Dadurch erhalten behandelnde Ärzte eine schnelle und einfache Unterstützung zur medizinischen Entscheidungsfindung.

Das dritte Manuskript untersucht, wie gut der EQ–5D–5L, für gesundheitsökonomische Fragestellungen bei Asthmapatienten einsetzbar ist. Die Untersuchung basiert auf Trennschärfe, Reliabilität und Sensitivität im Vergleich zum fest etablierten krankheitsspezifischen Fragebogen ’Asthma Quality of Life Questionnaire’ (AQLQ). Zusätzlich wird der ’minimal klinisch relevante Unterschied’ (MID) in Bezug auf den EQ–5D–5L in Asthmapatienten ermittelt und der bereits vorhandene MID für den AQLQ vergleichend neugeschätzt. Die Studienpopulation stammt aus einer randomisierten kontrollierten Studie über die Effektivität einer pneumologischen Rehabilitation mit Wartegruppendesign. Das Studiendesign erlaubt Vergleiche zwischen den Individuen sowie die Analyse intraindividueller Veränderungen über die Zeit. Obwohl beide Instrumente gute Trennschärfe und Reliabilität haben, ist die Sensitivität des EQ–5D–5L gegenüber

Lebensqualitätsänderungen in Asthmapatienten mittelmäßig. Dieses Ergebnis wird auch durch die MID-Bestimmung sichtbar. Demzufolge wäre die Nutzung eines zusätzlichen krankheitsspezifischen Instruments bei gesundheitsökonomischen Evaluationen empfehlenswert.

Das 4. Manuskript wendet die in Manuskript 3 gewonnenen Erkenntnisse und untersucht ob Instrumente mit unterschiedlichen psychometrischen Eigenschaften unterschiedliche Ergebnisse bei gesundheitsökonomischen Fragestellungen liefern. Es wird die Kosteneffektivität der oben genannten pneumologischen Rehabilitation von Asthmapatienten berechnet. Drei Endpunkte werden hierzu untersucht, darunter zwei krankheitsspezifische und ein generischer. Wie basierend auf Manuskript 3 zu erwarten, reagieren ’quality adjusted life years’ basierend auf EQ–5D–5L nicht ausreichend auf die positiven Einflüsse der Rehabilitation und Kosten–Effektivität ist klar zu verneinen. Demgegenüber deuten der krankheitsspezifische AQLQ sowie der Asthma–Kontroll–Test stark in Richtung Kosten–Effektivität der Rehabilitation. Die Ergebnisse betonen die Relevanz des kombinierten Einsatzes von verschiedenen Instrumenten.

Diese Arbeit hebt die Wichtigkeit der kombinierten Nutzung von generischen und krankheitsspezifischen Instrumenten hervor. Obwohl der EQ–5D–5L in allen untersuchten Indikationsgebieten prinzipiell geeignet ist, sind die psychometrische Eigenschaften der krankheitsspezifischen Tools tendenziell besser ausgeprägt. Dadurch könnten bei Entscheidungen, die nur auf EQ–5D–5L basiert sind, patientenrelevante Aspekte verloren gehen, vor allem im Krankheitsgebiet Asthma. Durch der kombinierten Nutzung der verschiedenen Tools könnte man zusammengesetzte Schlussfolgerungen über die Lebensqualität der Patienten ziehen und patientenrelevante Behandlungsentscheidungen besser treffen.

Chapter 1: General introduction

1.1 Health–related quality of life (HRQL) – a conceptual approach

Early medical and public health analyses focused mainly on the quantity of life (e.g. mortality, life expectancy or other objective clinical surrogate measures such as blood pressure), which are easier to assess and interpret than subjective parameters such as quality of life. However, the measurement of actual deaths or other quantifiable parameters does not provide significant insights into the patients' self–perception of how distinct diseases impair their daily lives. This problem occurs especially in chronic diseases with low mortality or in diseases characterized by phases of varying symptom intensity. Furthermore, through developments in medical industry, life expectancy in general increases constantly through time. This makes the evaluation of how a lifetime is subjectively judged by the individuals concerned even more important than a bare quantification of lifetime gained. Therefore, the concept of quality of life has gained significant importance in the last few decades (1). Quality of life was included in the medical subject headings (MESH) terms in 1977 for the first time (2) and was defined as a generic measure. MESH terms are a controlled vocabulary list to categorize publications based on their contents (3). The inclusion of quality of life showed its increasing use and the acknowledgement of its importance. The concept of quality of life is meant to measure the influence of not only the physical condition, but also capture judgements on the social, political and physical environment. Social sciences were able to work with this concept; however, the medical field needed different guiding principles, and the concept of health–related quality of life (HRQL) was introduced.

The Center for Disease Control and Prevention defined HRQL as an "individual's or a group's perceived physical and mental health over time" (4) leaving for example the political influence unmeasured. The concept has also been developed through the increasing interest in HRQL. Additional to the generic HRQL tools, that focus on comparison between groups with different diseases, disease–specific measures were established step by step for several indications. The increase in these various tools has led to the development and frequent use of common psychometric and measurement properties to evaluate the variety of tools.

1.1.1 Generic HRQL

Generic questionnaires address differences between diseased groups and the healthy population or between groups with different diseases. Generic tools focus on common aspects of life (e.g. anxiety, mobility, mental or general health), not on the aspects common to the disease (e.g. severe head pain by migraine). Generic tools are often used to help with resource allocation problems and therefore there are established benchmarks for their application in medical and health economic decision–making. Here, a combination of lifetime (quantitative measure) and HRQL often takes place. One of the most important and well–known methods in quantifying the gain of the generic HRQL is the calculation of quality–adjusted life years (QALY) (5). QALYs are a measure of the life years gained after a treatment or intervention weighted by the achieved HRQL (5). One well–known tool for this purpose is the short form 36 health (SF–36) generic questionnaire (6). The SF–36 contains 36 questions about health divided into three categories: functional status, wellbeing, and overall evaluation of health (7). One other common way to calculate QALYs is with the use of the EuroQol 5 Dimension questionnaire (EQ–5D–5L) and valuing the resulting health states with country–specific utilities (8). The EQ–5D–5L contains five questions each with five possible answers about mobility, self–care, usual activities, pain/discomfort and anxiety/depression and a visual analogue scale with values between zero and 100 (9). With these measures, a cost–utility calculation is possible, where costs and health improvements (measured as QALYs) are weighted against each other (10). This enables the calculation of thresholds above which a treatment is not considered cost–effective anymore. As QALYs and costs are generic measures, these thresholds can be used across various diseases and treatments. Furthermore, it can help to allocate resources to healthcare sectors, where most QALYs can be gained through an intervention. The National Institute for Health and Clinical Excellence in the UK obliged health care professionals to use this measure in health technology assessments (5), which underlines its importance in the health economics and outcomes field. In Germany however, there is a debate about the correctness of the utility calculations. Critique includes ethical and methodological concerns (11) as for example the lack of inclusion of uncertainty by the creation of value sets (12). Therefore, the ’Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen’ declined the use of cost–utility analyses (13). Nevertheless, generic measures are still very important to be able to evaluate

the burden of different diseases in various populations and to provide help for healthcare professionals, especially with ageing populations and increasing healthcare costs.

1.1.2 Disease–specific HRQL

To complement generic tools, disease–specific questionnaires focus on comparison between patients with the same disease to better understand the impact of disease severity on specific aspects of life. The diversity of diseases implies the large variety and number of disease–specific measures. Their quantification and assessment varies from disease to disease. However, all tools focus on the physical and/or mental impact of the disease on the patients' HRQL. These tools were meant to assess whether and how disease severity influences HRQL in the evaluated population, and where and why differences between patients may occur. Furthermore, treatment benefits can be evaluated with the use of these tools. The results might help to decide on treatment decisions, especially in incurable diseases, where providing a high HRQL is of utmost importance as the potential to extend life expectancy is limited. A detailed description of some chosen HRQL tools is listed with a short description of the examined disease in Chapter 1.2.

1.1.3 Validation through measurement properties

In contrast to clinical measures, HRQL values and changes are much more difficult to quantify and understand because of their theoretical nature. With the increasing focus on HRQL, the available variety of tHRQL tools is also increasing. Tools have to be able to provide answers to various research questions, e.g. which patient group is most affected, what is the impact of the disease development, or which part of life is most influenced by the disease. Furthermore, it should be easily administered in various situations, in routine care, in a hospital or sent by post. To ensure all of these properties and make different tools comparable, various methods were implemented. Some of the most important measurements in patient–centred health–care management are listed below.

One important aspect is validity, i.e. whether the tool measures what it should measure. Categories of validity provide answers to different research questions e.g concurrent validity draws comparison with gold standards, whereas construct validity measures the overall intended construct. Known–group validity and discriminative validity measure the tools'

abilities to distinguish between patients with different severity grades or between diseased and healthy people. Furthermore, predictive validity evaluates the sensitivity (identification of people with disease) and specificity (identification of people without disease) of the tools (14).

Additionally, the tools have to provide consistent results when repeating the tests in the same population without changing the parameters, i.e. they have to be reliable. They have to be reliable through time (test–retest reliability), through people (inter–rater reliability) and internally (internal consistency) (15). Reliability ensures that tools do not provide results by chance but by scientific value.

To evaluate interventions the assessment of meaningful changes is of great help. One of the most commonly used methods is the calculation of the minimally important difference (MID). The MID reflects according to Jaeschke et al. 'the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management.' (16). There are several ways to calculate the MID, anchor– or distribution–based methods (17) or with the help of the Global Ratings of Change (GROC) Scale (18). Distribution–based methods use the statistical distribution of the score to calculate MID: whereas anchor– based methods have an external anchor to link the changes to (19). The GROC measures the self–reported perceived change in HRQL of the patients through a defined time period (18). This can be combined with the measured changes in the HRQL tools in the same period and a MID can be calculated. The MID can be used as an indicator of treatment success or to calculate the amount of benefit a patient gains through an intervention, which in turn can be used in resource allocation decisions. However, MIDs are meant to measure intra individual changes and should not be used across populations in a cross–sectional setting.

Time needed to fill out the questionnaire provides important information about feasibility in different settings. Furthermore, the percentage of fully answered questions helps to determine how easily understandable the questionnaire is. Additionally, it detects problems if patients interrupt before finishing because of the length of the questionnaire or because of other factors (e.g. difficult or offensive wording).These two measures can also be an

indicator as to whether outside help is needed for application of the tool. The response rate is also a crucial indicator of feasibility.

Depending on the research question, different measurement properties can be given greater importance. Research on the different tools facilitates the decision about the implementation of the different tools. Uniform evaluation of the tools in a real–world setting is the first step in taking the perception of the patients into account in treatment decisions.

1.2 Chronic lung diseases

One of the most important organs of the human body is the lung. Several diseases can affect its health, from a regular cold to lung cancer. This thesis focuses on three chronic lung diseases, interstitial lung diseases (ILD), chronic obstructive pulmonary disease (COPD) and asthma. A short description and explanation of the relevance of these diseases is given below together with their HRQL implications.

1.2.1 Interstitial Lung Diseases

Short description and HRQL implications

ILD is an umbrella term for more than 200 rare heterogeneous lung diseases (20). Therefore, severity, mortality and other clinical outcomes vary substantially in this group. One common feature is the scarring of the lung, resulting in breathing difficulties, shortness of breath, cough and reduced life expectancy (21). Identification and correct classification of the specific diseases is difficult because of the great variety and the unspecific symptoms (22). Thus the prevalence of ILDs is probably underestimated and because of the great variation in the disease, not yet well researched (20).

Not only the direct influence of the disease on the lung, but the unknown causes of the disease, its chronic character and the length of time until diagnosis are all detrimental to the HRQL of patients. Even though there are numerous diseases in this group, most of the studies focus on bigger subgroups, e.g. sarcoidosis (23) or idiopathic pulmonary fibrosis (24,25), leaving several subgroups unevaluated. Even studies focusing on more prevalent

diseases (26) lack the use of disease–specific quality of life tools. Some studies used COPD–specific tools but comparing COPD with the variety of ILDs has proved difficult (26–28). As a result, Patel et al. developed a disease–specific questionnaire, the Kings Brief Interstitial Lung Disease questionnaire (K–BILD) (29). This is a questionnaire with three domains (breathlessness and activities, psychological impact, and chest symptoms) and an overall score to depict HRQL without excluding any subtypes of ILD. This questionnaire was translated by Kreuter et al. (30), but a German validation –beyond the semantic validation– was not yet available at the time of this thesis. Furthermore, the practicability of the questionnaire was not yet fully understood in comparison with a generic questionnaire.

1.2.2 Chronic Obstructive Pulmonary Disease

Short description and HRQL implications

COPD is a chronic disease affecting the lung and causing various symptoms such as breathlessness, cough and wheezing (31). Currently, there is no cure for the disease, merely the symptoms can be treated. Owing to its chronic character and high prevalence COPD is of great importance. According to the World Health Organization there were more than 250 million cases worldwide in 2016 (32). Furthermore, it is estimated to be the fourth leading cause of death by 2030 (33).

The progression of the disease is continuous, but relatively slow. Nonetheless, COPD patients have a shortened life expectancy compared with the healthy population, especially for patients with high disease severity (34). One of the most common disease severity classifications of COPD is based on lung function values, e.g. on the Forced Expiratory Pressure in 1 Second percent predicted (31). But it is not only clinical values that influence disease severity and thus treatment and medication decisions. According to the Global Initiative for Chronic Obstructive Lung Disease, there are two important HRQL questionnaires that have to be considered in treatment decisions (31); the questionnaire from the Modified British Medical Research Council (mMRC) and the COPD Assessment Test (CAT). In addition, because of the increasing interest in COPD numerous different HRQL questionnaires have been developed over the years (35), e.g. the St. George's Respiratory Questionnaire (SGRQ) (36) or the Clinical COPD Questionnaire (CCQ) (37). The tools differ

greatly in length starting from a single–question tool (mMRC) to tools with as many as 50 questions (SGRQ). The single–item tool covers only dyspnoea whereas longer questionnaires include intensity of coughing, sleep problems or social aspects. Studies using these tools conclude that the disease causes serious impairments in patients' on HRQL (38,39) and also causes high caregiver burden (40,41). Owing to the diversity of the tools and the research questions behind them, there is not yet a gold standard to measure HRQL in COPD patients in a clinical setting. This would be beneficial to enable efficient patient–centered treatment in a routine setting.

1.2.3 Asthma

Short description and HRQL implications

Asthma is a non–communicable respiratory disease characterized by chronic inflammation of the airways and recurrent attacks of breathlessness, wheezing or coughing, so–called asthma attacks (42). Asthma currently has no cure, but the symptoms can be controlled through proper treatment. The morbidity of the disease varies with socioeconomic status and other factors (43), but the analysis by the Global Burden of Disease study puts its worldwide prevalence at over 300 million people in 2016 (44). Thus, its public health relevance is not to be underestimated.

Several studies have shown the negative impact of asthma on HRQL (45, 46), which results from the various symptoms with sudden occurrence. There are, aside from medication, several methods that can help reduce the symptoms, e.g. respiratory physiotherapy (47), patient education (48) or special exercise training (49,50), and thus can increase HRQL. To measure HRQL and the impact of these methods, there are several tools available (51–53). One of the most common tools is the 'Asthma Quality of Life Questionnaire' (AQLQ) developed by Juniper et al. (52,54). AQLQ is a questionnaire with 32 questions in four domains (symptoms, activity limitations, emotional function, and environmental exposure) (52,54). It is a reliable and responsive tool that is available in many languages and has good use in clinical trials (55). Furthermore, AQLQ can complement health economic analyses through calculating the AQL–5D, a preference–based measure and thus calculating QALYs (56). Other studies used the SGRQ (57) or the Asthma Control Questionnaire (58). However,

there is no benchmark questionnaire yet, and the research question influences the tool (55). Furthermore, the research on generic tools is limited (59) and needs to be followed up.

1.3 Datasets used

All analyses are based on primary datasets, i.e. directly won through questionnaires and/or medical assessment. Using primary data is an ideal way to find answers to specific questions, as one can adapt the questionnaires and measurements taken. Furthermore, there is no data modification (e.g. summarizing data) throughout the process, unlike in some secondary data.

For manuscript one the questionnaires were distributed by medical staff in the corresponding hospitals. The established and used questionnaires were selected by the research team and were entered manually by medical personnel. The questionnaires used in publication two were selected by the group of researchers responsible for the study, including the author of this dissertation. The questionnaires were distributed by post, and the results were entered manually. Additionally, for some patient characteristics claims data were utilized. For the third and fourth publications, the data were generated through a randomized controlled trial. Patients receiving treatment had the questionnaires administered in the clinic, whereas patients in the waiting group and for the longitudinal analyses received the questionnaires by post. Data entry was done by medical personnel. Data quality in all of the analyses was high and partial double data entry was conducted to ensure this. However, it has to be kept in mind that HRQL scores are self–reported values; therefore, differences in the nuance of the answers may occur.

1.4 Objectives and contents of this dissertation

The main objective of this dissertation is to evaluate one existing generic (EQ–5D–5L) and various disease–specific HRQL tools in patient groups with different chronic lung diseases. On the one hand, it is important to analyse and acknowledge the differences between

generic and disease–specific tools, to understand their working mechanism and the results derived. On the other hand, the different diseases are in different stages regarding the availability and routine application of disease–specific HRQL tools. To be able to provide patient–relevant care, to support resource allocation problems and to assist the healthcare section in decision making, the description of the status quo is important. Additionally, it is necessary to provide resources to define gold standard questionnaires. This should help to create comparable results throughout different study populations.

The three diseases investigated have different stages regarding HRQL assessment and face different problems, which was addressed within the scope of this thesis. With ILDs, because of their rare and heterogeneous nature, ILD–specific questionnaires barely exist, and their psychometric properties are not yet fully understood. COPD faces an opposite problem. As COPD is a relatively well–understood chronic disease with high prevalence, it is the object of many studies. The high interest in the disease has resulted in several HRQL questionnaires, but a lack of consensus about the gold standard. Asthma is placed somewhere in between the other two disease groups. It has a relatively high prevalence but because the disease varies so much throughout its course, capturing HRQL is problematic and results are volatile. Therefore, with the use of observational primary data, this thesis aims to address the specific problems of the disease areas.

The first paper investigates the use of one of the first ILD–specific tools, the K–BILD in the German ILD population. The aim of the evaluation was to understand whether the K–BILD is able to depict HRQL in the population and to find the main factors influencing HRQL. As ILDs are non–curable, detecting influencing factors, e.g. comorbidities is of the utmost importance. Furthermore, the paper compared the EQ–5D–5L with the K–BILD and concluded that the generic tool can also suitably measure HRQL in ILD patients. Both tools react to the same influencing factors: lung function measures and comorbidity sum score. Thus, they provide an important insight into the lives of ILD patients.

Paper two aims to find a suitable COPD questionnaire in a clinical setting. As HRQL results can influence treatment options, a reliable and easily usable tool is needed. Three preselected established disease–specific tools, the CAT, the CCQ and the SGRQ, were investigated alongside the EQ–5D–5L. Criteria such as response rate, floor and ceiling effects,

and explanatory power were taken into consideration. The CAT performed best out of the selection, especially in combination with the EQ–5D–5L and is thus suggested for clinical use. It is reliable, short, easy to fill out, explains the highest variance in the EQ–5D–5L and is recommended in international guidelines as a tool for treatment decisions.

Paper three uses data from a randomized controlled trial, which investigates the effectiveness of pulmonary rehabilitation in asthma patients. As rehabilitation programmes are often evaluated with the use of HRQL endpoints, it is important to evaluate the different HRQL tools in this scenario. With the use of a well–established disease–specific questionnaire this paper looks into the measurement properties of the EQ–5D–5L, especially focusing on its ability to detect changes. Reliability and discriminatory power are similar in the tools; however the disease–specific questionnaire outperforms the EQ–5D–5L in detecting changes. Furthermore, this paper provides MID measures for all investigated tools to facilitate decision–making. Concluding from the results, the complementary use of disease–specific tools with economic evaluations would be beneficial.

Paper four focuses on the implications of paper three and tests whether tools with different psychometric properties deliver different results in health economic analyses. The use of EQ–5D–5L was tested in a cost–effectiveness analysis within the scope of the randomized controlled trial in the population in paper three. The analysis provides a comparison of the cost–effectiveness results with a disease–specific tool and with the EQ–5D–5L. Even though the disease–specific tool clearly states the cost–effectiveness of the rehabilitation programme, the EQ–5D–5L was not able to depict this result. As laid out in paper three, the generic tool cannot adequately detect changes in the HRQL of asthma patients. Thus the use of an additional tool should be encouraged in health economic analyses to provide a more realistic picture.

In conclusion, in all three disease areas disease–specific tools provide more insight about the lung–related HRQL than the EQ–5D–5L. Thus, decisions based solely on EQ–5D–5L could miss important aspects of HRQL. Nevertheless, the EQ–5D–5L is an important tool for comparison across diseases and for health economic analyses. Therefore, this dissertation recommends the combined use of generic and disease–specific tools to provide a global picture about the HRQL of patients and foster patient–centred treatment.

1.5 Individual contribution from the author

The author of the thesis contributed substantially to the concept of all included articles. She conducted data preparation and plausibility checks for all four publications.

Furthermore, she conducted all the analyses, drafted the manuscripts and served as corresponding author throughout the publication process for articles 1, 2 and 3.

Regarding article 4, she co–supervised the underlying master thesis, did preliminary analyses and supported the entire publication process.

References

1. Pennacchini M, Bertolaso M, Elvira MM, De Marinis MG. A brief history of the Quality of Life: its use in medicine and in philosophy. Clin Ter. 2011;162(3):e99–e103.
2. Medicine USNLo. MeSH Descriptor Data 2020 [cited 18 August 2020] Available from: https://meshb.nlm.nih.gov/record/ui?ui=D011788.
3 National Library of Medicine 2020 [cited 2. March 2021] Available from: https://www.nlm.nih.gov/mesh/intro_preface.html.
4. Center for Disease Control and Prevention . Health–related quality of life; 2018 October 31 [cited 2020 August 30]. Available from: https://www.cdc.gov/hrqol/index.htm.
5. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull. 2010;96:5–21.
6. Brazier J, Roberts J, Deverill M. The estimation of a preference–based measure of health from the SF–36. J Health Econ. 2002 Mar;21(2):271–92. doi: 10.1016/s0167–6296(01)00130–8. PMID: 11939242.
7. Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF–36 health survey questionnaire: new outcome measure for primary care. BMJ. 1992 Jul 18;305(6846):160–4. doi: 10.1136/bmj.305.6846.160. PMID: 1285753; PMCID: PMC1883187.
8. Brazier JE, Rowen D, Lloyd A, Karimi M. Future Directions in Valuing Benefits for Estimating QALYs: Is Time Up for the EQ–5D? Value Health. 2019 Jan;22(1):62–8. doi: 10.1016/j.jval.2018.12.001. PMID: 30661635.
9. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five–level version of EQ–5D (EQ–5D–5L). Qual Life Res. 2011 Dec;20(10):1727–36. doi: 10.1007/s11136–011–9903–x. Epub 2011 Apr 9. PMID: 21479777; PMCID: PMC3220807.
10. Drummond MF, Sculpher M, Claxton K, Stoddart GL Torrance GW. Methods for the Economic Evaluation of Health Care Programmes (4th edition). Oxford: Oxford University Press, 2015.
11. Pettitt, D. A., Raza, S., Naughton, B., Roscoe, A., Ramakrishnan, A., Ali, A., Davies, B., Dopson, S., Hollander, G., Smith, J. A., & Brindley, D. A. (2016). The limitations of QALY: a literature review. Journal of Stem Cell Research and Therapy, 6(4), Article: 1000334.
12. Kharroubi SA, Beyh Y. The importance of accounting for the uncertainty around the preference–based health–related quality–of–life measures value sets: a systematic review. J Med Econ. 2019 Jul;22(7):671–683. doi: 10.1080/13696998.2019.1592178. Epub 2019 Mar 27. PMID: 30841768.
13. Ivandic, V. Requirements for benefit assessment in Germany and England – overview and comparison. Health Econ Rev 4, 12 (2014). https://doi.org/10.1186/s13561–014–0012–8.
14. Asunta P, Viholainen H, Ahonen T, Rintala P. Psychometric properties of observational tools for identifying motor difficulties – a systematic review. BMC Pediatr. 2019;19(1):322.
15. Del Greco L, Walop W, McCarthy RH. Questionnaire development: 2. Validity and reliability. *CMAJ*. 1987;136(7):699–700.
16. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10(4): 407–15.
17. Johnston BC, Ebrahim S, Carrasco–Labra A, et al. Minimally important difference estimates and methods: a protocol. *BMJ Open*. 2015;5(10):e007953. Published 2015 Oct 1. doi: 10.1136/bmjopen–2015–007953.
18. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther. 2009;17(3):163–70.
19. Ousmen, A., Touraine, C., Deliu, N. et al. Distribution– and anchor–based methods to determine the minimally important difference on patient–reported outcome questionnaires in

oncology: a structured review. Health Qual Life Outcomes 16, 228 (2018). https://doi.org/10.1186/s12955–018–1055–z.
20. Demedts M, Wells AU, Antó JM, Costabel U, Hubbard R, Cullinan P, Slabbynck H, Rizzato G, Poletti V, Verbeken EK, Thomeer MJ, Kokkarinen J, Dalphin JC, Taylor AN. Interstitial lung diseases: an epidemiological overview. Eur Respir J Suppl. 2001 Sep;32:2s–16s. PMID: 11816822.
21. Naghavi M, Wang H, Lozano R, Davis A, Liang X, Zhou M, et al. Global, regional, and national age–sex specific all–cause and cause–specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015;385:117–171. doi: 10.1016/S0140–6736(14)61682–2.
22. Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence–based guidelines for diagnosis and management. Am J Respir Crit Care Med. 2011;183(6):788–824. doi:10.1164/rccm.2009–040GL.
23. De Vries J, Drent M. Quality of life and health status in sarcoidosis: a review of the literature. Clin Chest Med. 2008;29(3):525–ix. doi:10.1016/j.ccm.2008.03.003.
24. Yu X, Li X, Wang L, et al. Pulmonary Rehabilitation for Exercise Tolerance and Quality of Life in IPF Patients: A Systematic Review and Meta–Analysis. Biomed Res Int. 2019;2019:8498603. Published 2019 Mar 21. doi:10.1155/2019/8498603.
25. Rozenberg D, Sitzer N, Porter S, et al. Idiopathic Pulmonary Fibrosis: A Review of Disease, Pharmacological, and Nonpharmacological Strategies With a Focus on Symptoms, Function, and Health–Related Quality of Life. J Pain Symptom Manage. 2020;59(6):1362–1378. doi:10.1016/j.jpainsymman.2019.12.364.
26. Kreuter M, Swigris J, Pittrow D, Geier S, Klotsche J, Prasse A, et al. Health related quality of life in patients with idiopathic pulmonary fibrosis in clinical practice: insights–IPF registry. Respir Res. 2017;18(1):139.
27. Someya F, Nakagawa T, Mugii N. The COPD Assessment Test as a Prognostic Marker in Interstitial Lung Disease. Clin Med Insights Circ Respir Pulm Med. 2016;10:27–31.
28. Berry CE, Drummond MB, Han MK, Li D, Fuller C, Limper AH, et al. Relationship between lung function impairment and health–related quality of life in COPD and interstitial lung disease. Chest. 2012;142(3):704–11.
29. Patel AS, Siegert RJ, Brignall K, Gordon P, Steer S, Desai SR, et al. The development and validation of the King's Brief Interstitial Lung Disease (K–BILD) health status questionnaire. Thorax. 2012;67(9):804–10.
30. Kreuter M, Birring SS, Wijsenbeek M, Wapenaar M, Oltmanns U, Costabel U, et al. German Validation of the "King's Brief Interstitial Lung Disease (K–Bild) Health Status Questionnaire". Pneumologie. 2016;70(11):742–6.
31. Singh D, Agusti A, Anzueto A, Barnes PJ, Bourbeau J, Celli BR, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease: the GOLD science committee report 2019. Eur Respir J. 2019;53(5).
32. World Health Organization. Chronic Obstructive Pulmonary Disease Key facts; 2017 Dec 1 [cited Aug 25 2020]. Available from: https://www.who.int/news–room/fact–sheets/detail/chronic–obstructive–pulmonary–disease–(copd).
33. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006 Nov;3(11):e442. doi: 10.1371/journal.pmed.0030442.
34. Shavelle RM, Paculdo DR, Kush SJ, Mannino DM, Strauss DJ. Life expectancy and years of life lost in chronic obstructive pulmonary disease: findings from the NHANES III Follow–up Study. Int J Chron Obstruct Pulmon Dis. 2009;4:137–48. doi:10.2147/copd.s5237.
35. Weldam SW, Schuurmans MJ, Liu R, Lammers JW. Evaluation of Quality of Life instruments for use in COPD care and research: a systematic review. Int J Nurs Stud. 2013;50(5):688–707.
36. Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. Respir Med. 1991 Sep;85 Suppl B:25–31; discussion 33–7. doi: 10.1016/s0954–6111(06)80166–6. PMID: 1759018

37. Zhou Z, Zhou A, Zhao Y, Chen P. Evaluating the Clinical COPD Questionnaire: A systematic review. Respirology. 2017 Feb;22(2):251–62. doi: 10.1111/resp.12970. PMID: 28102972.
38. Wacker ME, Jorres RA, Karch A, Koch A, Heinrich J, Karrasch S, et al. Relative impact of COPD and comorbidities on generic health–related quality of life: a pooled analysis of the COSYCONET patient cohort and control subjects from the KORA and SHIP studies. Respir Res. 2016;17(1):81.
39. Wacker ME, Jorres RA, Karch A, Wilke S, Heinrich J, Karrasch S, et al. Assessing health–related quality of life in COPD: comparing generic and disease–specific instruments with focus on comorbidities. BMC Pulm Med. 2016;16(1):70.
40. Kar S, Zengin N. The relation between self–efficacy in patients with chronic obstructive pulmonary disease and caregiver burden. Scand J Caring Sci. 2020 Sep;34(3):754–61. doi: 10.1111/scs.12780. Epub 2019 Oct 27. PMID: 31657060.
41. Pinto RA, Holanda MA, Medeiros MM, Mota RM, Pereira ED. Assessment of the burden of caregiving for patients with chronic obstructive pulmonary disease. Respir Med. 2007 Nov;101(11):2402–8. doi: 10.1016/j.rmed.2007.06.001. Epub 2007 Jul 12. PMID: 17624751.
42. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention 2018 [cited 2020 Sept 10]. Available from: www.ginasthma.org.
43. Cruz AA, Bateman ED, Bousquet J. The social determinants of asthma. Eur Respir J. 2010 Feb;35(2):239–42. doi: 10.1183/09031936.00070309. PMID: 20123842.
44. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017; 390: 1211–59.
45. Williams SA, Wagner S, Kannan H, Bolge SC. The association between asthma control and health care utilization, work productivity loss and health–related quality of life. J Occup Environ Med. 2009;51(7):780–5.
46. Hernandez G, Dima AL, Pont A, Garin O, Marti–Pastor M, Alonso J, et al. Impact of asthma on women and men: comparison with the general population using the EQ–5D–5L questionnaire. PLoS One. 2018;13(8):e0202624.
47. Bruton A, Lee A, Yardley L, Raftery J, Arden–Close E, Kirby S, et al. Physiotherapy breathing retraining for asthma: a randomised controlled trial. Lancet Respir Med. 2018;6(1):19–28.
48. Gibson PG, Powell H, Coughlan J, Wilson AJ, Abramson M, Haywood P, et al. Self–management education and regular practitioner review for adults with asthma. Cochrane Database Syst Rev. 2003;1:CD001117.
49. Mendes FA, Goncalves RC, Nunes MP, Saraiva–Romanholo BM, Cukier A, Stelmach R, et al. Effects of aerobic training on psychosocial morbidity and symptoms in patients with asthma: a randomized clinical trial. Chest. 2010;138(2):331–7.
50. Turner S, Eastwood P, Cook A, Jenkins S. Improvements in symptoms and quality of life following exercise training in older adults with moderate/severe persistent asthma. Respiration. 2011;81(4):302–10.
51. Juniper EF, Guyatt GH, Epstein RS, Ferrie PJ, Jaeschke R, Hiller TK. Evaluation of impairment of health related quality of life in asthma: development of a questionnaire for use in clinical trials. Thorax. 1992 Feb;47(2):76–83. doi: 10.1136/thx.47.2.76. PMID: 1549827; PMCID: PMC463574.
52. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. Eur Respir J. 1999 Oct;14(4):902–7. doi: 10.1034/j.1399–3003.1999.14d29.x. PMID: 10573240.
53. Apfelbacher CJ, Jones CJ, Frew A, Smith H. Validity of three asthma–specific quality of life questionnaires: the patients' perspective. BMJ Open. 2016 Dec 22;6(12):e011793. doi: 10.1136/bmjopen–2016–011793. PMID: 28007706; PMCID: PMC5223641.
54. Juniper EF, Buist AS, Cox FM, Ferrie PJ, King DR. Validation of a standardized version of the asthma quality of life questionnaire. Chest. 1999;115(5):1265–70.
55. Wilson SR, Rand CS, Cabana MD, et al. Asthma outcomes: quality of life. J Allergy Clin Immunol. 2012;129(3 Suppl):S88–S123. doi:10.1016/j.jaci.2011.12.988.

56. Young TA, Yang Y, Brazier JE, Tsuchiya A. The use of rasch analysis in reducing a large condition–specific instrument for preference valuation: the case of moving from AQLQ to AQL–5D. Med Decis Making. 2011;31(1):195–210.
57. Ortega HG, Liu MC, Pavord ID, Brusselle GG, FitzGerald JM, Chetta A, Humbert M, Katz LE, Keene ON, Yancey SW, Chanez P; MENSA Investigators. Mepolizumab treatment in patients with severe eosinophilic asthma. N Engl J Med. 2014 Sep 25;371(13):1198–207. doi: 10.1056/NEJMoa1403290. Epub 2014 Sep 8. Erratum in: N Engl J Med. 2015 Apr 30;372(18):1777. PMID: 25199059.
58. Jia CE, Zhang HP, Lv Y, Liang R, Jiang YQ, Powell H, Fu JJ, Wang L, Gibson PG, Wang G. The Asthma Control Test and Asthma Control Questionnaire for assessing asthma control: Systematic review and meta–analysis. J Allergy Clin Immunol. 2013 Mar;131(3):695–703. doi: 10.1016/j.jaci.2012.08.023. Epub 2012 Oct 8. PMID: 23058645.
59. Ilmarinen, P., Juboori, H., Tuomisto, L.E. et al. Effect of asthma control on general health–related quality of life in patients diagnosed with adult–onset asthma. Sci Rep 9, 16107 (2019). https://doi.org/10.1038/s41598–019–52361–9.

Chapter 2 – Article 1

Quality of life assessment in interstitial lung diseases: a comparison of the disease–specific K–BILD with the generic EQ–5D–5L

Published in:

Szentes BL, Kreuter M, Bahmer T, Birring SS, Claussen M, Waelscher J, Leidl R, Schwarzkopf L. Quality of life assessment in interstitial lung diseases: a comparison of the disease–specific K–BILD with the generic EQ–5D–5L. Respir Res. 2018 May 25;19(1):101. doi: 10.1186/s12931–018–0808–x.

Szentes et al. Respiratory Research (2018) 19:101
https://doi.org/10.1186/s12931-018-0808-x

Respiratory Research

RESEARCH Open Access

Quality of life assessment in interstitial lung diseases:a comparison of the disease-specific K-BILD with the generic EQ-5D-5L

Boglárka Lilla Szentes[1*], Michael Kreuter[2], Thomas Bahmer[3], Surinder S. Birring[4,5], Martin Claussen[3], Julia Waelscher[2], Reiner Leidl[1,6] and Larissa Schwarzkopf[1]

Abstract

Background: Patients with interstitial lung diseases (ILD) have impaired health-related quality of life (HRQL). Little is known about the applicability of the disease-specific King's Brief Interstitial Lung Disease questionnaire (K-BILD) and the generic EQ-5D-5L in a German setting.

Methods: We assessed disease-specific (K-BILD) and generic HRQL (EQ-5D experience based value set (EBVS) and Visual Analog Scale (VAS)) in 229 patients with different ILD subtypes in a longitudinal observational study (HILDA). Additionally, we assessed the correlation of the HRQL measures with lung function and comorbidities. In a linear regression model, we investigated predictors (including age, sex, ILD subtype, FVC percentage of predicted value (FVC%pred), DLCO percentage of predicted value, and comorbidities).

Results: Among the 229 patients mean age was 63.2 (Standard deviation (SD): 12.9), 67.3% male, 24.0% had idiopathic pulmonary fibrosis, and 22.3% sarcoidosis. Means scores were as follows for EQ-5D EBVS 0.66(SD 0.17), VAS 61.4 (SD 19.1) and K-BILD Total 53.6 (SD 13.8). K-BILD had good construct validity (high correlation with EQ-5D EBVS (0.71)) and good internal consistency (Cronbach's alpha 0.89). Moreover, all HRQL measures were highly accepted by patients including low missing items and there were no ceiling or floor effects. A higher FVC % pred was associated with higher HRQL in all measures meanwhile comorbidities had a negative influence on HRQL.

Conclusions: K-BILD and EQ-5D had similar HRQL trends and were associated similarly to the same disease-related factors in Germany. Our data supports the use of K-BILD in clinical practice in Germany, since it captures disease specific effects of ILD. Additionally, the use of the EQ-5D-5L could provide comparison to different disease areas and give an overview about the position of ILD patients in comparison to general population.

Keywords: ILD, Health-related quality of life, K-BILD, EQ-5D-5L, Comorbidities

Background

Interstitial lung diseases (ILDs) comprise more than 200 rare diseases, which are characterized by varying degrees of inflammation and fibrosis of the lung, and are associated with serious quality of life impairments in affected people [1–5]. There were attempts to quantify the health-related quality of life (HRQL), however most previous analyses focused on the most prevalent forms of ILDs (i.e. Idiopathic Pulmonary Fibrosis (IPF), Sarcoidosis) e.g. Kreuter et al. provide data about the German IPF population [6] or did not apply ILD-specific assessment tools [6–13]. Instead, among others, questionnaires originally designed for patients with Chronic Obstructive Pulmonary Disease (COPD) were tested in ILD-populations: e.g. the COPD Assessment Test [7] and the St George's Respiratory Questionnaire (SGRQ) [5, 8]. The suitability of these questionnaires to reflect ILD-specific aspects of HRQL remains up to discussion. Moreover, given the heterogeneous clinical course of ILDs, a transferability

* Correspondence: boglarka.szentes@helmholtz-muenchen.de
[1]Helmholtz Zentrum München – German Research Center for Environmental Health (GmbH), Institute of Health Economics and Health Care Management, Comprehensive Pneumology Center Munich (CPC-M), Member of the German Center for Lung Research (DZL), Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany
Full list of author information is available at the end of the article

of HRQL findings among patients with IPF or sarcoidosis to other ILD subtypes is a highly sensitive issue.

Keeping these drawbacks in mind, all studies cited suggest impaired HRQL in ILD patients but comprehensive analyses of HRQL in ILDs accounting for many different subtypes and focusing on disease-specific questionnaires are sparse.

Recently the King's Brief Interstitial Lung Disease Questionnaire (K-BILD) [9] has been proposed as the first and so far only ILD-specific HRQL assessment tool. The K-BILD is a validated [10] and clinically oriented HRQL tool [11]. Evidence shows that K-BILD is a suitable HRQL measure in different countries; e.g. in UK [9] and in Italy, France, Sweden and the Netherlands as shown by Wapenaar et al. [10] However, until now there is no study using K-BILD in a German setting beyond Kreuter et al. that have translated and validated the questionnaire [12] in 2016.

To compare the disease burden of ILD patients with the general population or with patients suffering from different diseases, the use of a generic HRQL instrument is recommended, since generic questionnaires measure overall HRQL and not just disease-specific primarily symptom-driven aspects, which would not apply for every group [14]. The EuroQol group developed the generic EuroQol five dimensional 5-Level (EQ-5D-5L) questionnaire, which is the improved version of the well-known and well-established 3-level version; EQ-5D-3L [13]. Thus, we assume that the 5L version would provide a good insight in the generic HRQL in the ILD patients and allows the comparison with disease-specific measures. The use of the EQ-5D-5L in lung disease patients is spare so far [10, 15, 16], and there is no validation in the ILD disease-area yet.

Therefore, in the first step we aimed to investigate the suitability of the K-BILD in Germany to measure ILD-specific HRQL. In the second step we aimed to measure psychometric values of the EQ-5D-5L compared to the K-BILD and thereby contribute to a validation of the generic HRQL measure in the disease group ILD. Furthermore, we want to give first insights into HRQL of ILD patients and its predictors in a German tertiary care setting.

Methods

Study population and data collection

Data is derived from the ongoing HILDA (**H**ealth Care in **ILD** **A**mbulance Visitors) study. This observational study addresses outpatients diagnosed with any ILD subtype who presented to the outpatient practices of two large German tertiary care centers for ILD in Germany (Thoraxklinik Heidelberg, LungenClinic Großhansdorf). Heidelberg is a city in south-west Germany whereas Grosshansdorf is in the Northern part. Participants were recruited sequentially over a period of six months starting in November 2016. The local Ethics Committees of Heidelberg and Luebeck approved the study (reference number S-200/2013, and AZ: 16-192, respectively). Participants provided written informed consent.

Individuals who were 18 years of age and older, with ILD confirmed by the ILD boards of the respective centers, with an expected survival time of more than 12 months and with sufficient knowledge of the German language were eligible for the HILDA-study. Participants were grouped to one of the following ILD subtypes: 'IPF', sarcoidosis, Hypersensitivity Pneumonitis (HP), other Idiopathic Interstitial Pneumonias (than IPF) ('other IIP'), and other ILDs based on the differential diagnosis of the treating clinician. 'Other IIP' accounts for idiopathic interstitial pneumonia, non-specific interstitial pneumonia, desquamative interstitial pneumonia, cryptogenic organizing pneumonia, lymphocytic interstitial pneumonia, respiratory bronchiolitis-associated interstitial lung disese, pleuropulmonary fibroelastosis, and acute interstitial pneumonia, while the 'other' group includes every other subtype not listed above.

HRQL assessment

The patients' self-reported HRQL was assessed at time of inclusion (at baseline) into the HILDA registry as part of their regular ambulance visits using EQ-5D-5L (generic HRQL) [17] and K-BILD [9] (disease-specific HRQL).

a) EQ-5D-5L

The generic EQ-5D-5L consists of two parts, the descriptive system and a Visual Analogue Scale. The descriptive system addresses five different dimensions ('mobility', 'self-care', 'usual activities', 'pain/discomfort', and 'anxiety/depression'), each with a five point Likert-Scale. The answering pattern can be transferred to a utility between 0 and 1 (the higher the better) by distinct (nation-specific) scoring algorithms [18–21]. We chose the Germany-specific experience-based value set (EQ-5D EBVS) from Leidl et al. [18] for calculation of values. The Visual Analog Scale (VAS) allows valuing current health on a thermometer scale between 0 and 100, with higher values indicating better health.

b) K-BILD

K-BILD measures health impairments induced by ILD. The questionnaire covers 15 questions spread out in three domains ('breathlessness and activity', 'chest symptoms' and 'psychological impact') via a seven point Likert Scale. The total (cross-domain) score and domain-specific subscores range from zero to 100 with higher values indicating better health. Scores can be calculated through a

predefined, not a patient-reported scoring algorithm, which is provided by the authors upon request [9].

Assessment of covariables

To reflect potential impact factors on HRQL we accounted for comorbidity burden, clinical aspects and the patients' sociodemographic background.

Comorbid conditions were derived from patients' history and medical records. We considered the following comorbid conditions based on previous evidence on their ILD-relevance: pulmonary hypertension, arterial hypertension, coronary heart disease, congestive heart failure, other cardiovascular disease, diabetes mellitus, emphysema/COPD, lung cancer, depression, gastroesophageal reflux disease, renal failure, obstructive sleep apnea, thromboembolism, and malignant tumors other than lung cancer [22–24]. In addition, physicians were allowed to list up to three relevant comorbid conditions not included in the pre-selection. Comorbidity burden was operationalized as sum of all documented conditions, therefore ranging between zero and 17.

As clinical routine, we measured forced vital capacity percent predicted (FVC % pred), and diffusing capacity of the lungs for carbon monoxide percent predicted (DLCO % pred) as functional parameters.

We assessed further basic characteristics by questionnaire-based self-reports of the patients.

Statistical analysis

For our analyses, only patients with complete information on HRQL and on all covariables relevant for regression analyses were included (complete case analysis). The data from the HRQL questionnaire were considered complete if the total score could be calculated. To avoid selection bias, we compared patient characteristics of those with incomplete and complete questionnaires before finally excluding any patients from further analyses. Subsequently, we assessed floor and ceiling effects; defined as > 15% of the participants achieving the best/worst HRQL score [25]. Besides, correlations between the HRQL measures, lung function parameters and the comorbidity sum score were quantified by Spearman's rank coefficient to examine associations. We considered correlations < 0.3 as weak, those ≥ 0.3 and < 0.7 as moderate and those ≥ 0.7 as strong [26]. Furthermore, internal consistency was assessed for the K-BILD domains and total score with Cronbach's alpha.

Influencing factors on HRQL were investigated via separate linear regression analyses using EQ-5D EBVS, VAS, K-BILD and the K-BILD domains as the respective outcome variables and sex, age (in years), education (basic ≤ 9 years, secondary 10-11 years, higher ≥ 12 years of schooling), employment status (full-time, part-time, unemployed), clinic location (to control also for climate differences), smoking status (current, former, never smoker), lung function parameters, disease subtype and comorbidity sum score as the independent variables. Reference categories were male, higher education, retired, study center Heidelberg, smoker and 'other ILD' respectively. Since the reference category for ILD subtype is more arbitrary than for the other covariables, we conducted least squares mean comparisons to detect further differences among ILD subtypes.

Given the extended recruitment period (November – April) we also investigated the potential impact of seasonal fluctuation of respiratory symptoms by including time of enrolment (winter yes/no) into our regression models. Since this more complex approach did not have a substantial additional explanatory effect, we consciously disregarded seasonal aspects within the analyses to support a straightforward interpretation.

Within a secondary analysis, we included all 14 pre-selected comorbidities to examine the influence of the distinct conditions on HRQL. Furthermore, in a sensitivity analysis we imputed the missing values except for our outcome variables (EQ-5D EBVS, VAS and K-BILD, $n = 9$). For missing categorical values we used the median of the observation (education $n = 7$, smoking status $n = 2$), and for missing continuous values the mean of the observation (DLCO % pred $n = 16$, FVC % pred $n = 4$). For the variable employment we imputed 'full-time' under 65 years of age and 'retired' above; according to the German retirement policies ($n = 2$) [27]. Additionally, we conducted the secondary analysis with imputing the lowest DLCO % pred values, in case patients with missing DLCO values were not able to take the test and thus assuming low DLCO values.

Statistical analyses were performed with SAS software (SAS Institute Inc., Cary, NC, USA, version 9.4), and p-values of 0.05 or less were considered statistically significant.

Results

Patient characteristics

Out of the 268 patients we included 229 into final analyses after excluding 39 (14.6%) with incomplete data. The excluded patients were similar to the finally included study population except for their FVC % pred (included: 70.4 vs excluded: 53.1 $p < 0.0001$) and HRQL (EQ-5D EBVS: 0.66 vs 0.59 $p = 0.032$; VAS: 61.4 vs. 49.1 $p = 0.0005$; K-BILD 53.6 vs 48.2 $p = 0.0166$).

The majority of the included patients was male (67.3%) with mean age of 63.2 (standard deviation: 12.9) and around half of the patients were retired. IPF was present in 24.0% of the patients, 22.3% presented with sarcoidosis, 11.4% HP, 9.2% 'other IIP', and 33.2% other ILDs (Table 1). Descriptive results stratified by center are shown in the online supplement; patients in

Table 1 Baseline characteristics of the participants

Characteristic		Total sample	Excluded		p-value
N = 268		Mean/%	Mean/%	Missing	
		229(85.5)	39 (14.5)		
Sex *n(%)*	Male	154 (67.3)	22 (56.4)	0	0.1876
Age *Mean (SD)*		63.2 (12.9)	62.0 (13.6)	0	0.677
Education	Basic	99 (47.1)	15 (50)	9	0.6694
n(%)	Secondary	59 (28.1)	8 (26.7)		
	Higher	52 (24.8)	7 (23.3)		
Employment	Full-time	57 (24.9)	11 (29.7)	2	0.1379
n(%)	Part-time	24 (10.5)	1 (2.7)		
	Unemployed	30 (13.1)	9 (24.3)		
	Retired	118 (51.5)	16 (43.2)		
Smoking status	Current smoker	9 (3.9)	3 (8.1)	2	0.3797
n(%)	Former smoker	139 (60.7)	19 (51.4)		
	Never smoker	81 (35.4)	15 (40.5)		
ILD subtypes	IPF	55 (24.0)	6 (15.4)	0	0.1145
n(%)	Sarcoidosis	51 (22.3)	7 (18.0)		
	Hypersensitivity pneumonitis	26 (11.4)	1 (2.6)		
	Other IIPs[a]	21 (9.2)	5 (12.8)		
	Others	76 (33.2)	20 (51.3)		
DLCO% predicted *Mean (SD)*		44.2 (17.2)	44.2 (17.2)	17	0.3187
FVC % predicted *Mean (SD)*		77.4 (18.9)	53.1 (17.5)	4	<.0001
Mean number of comorbidities *Mean (SD)*		2.7 (1.8)	2.8 (2.0)	0	0.8936
EQ-5D-5L *Mean (SD)*	EBVS	0.66 (0.17)	0.6 (0.2)	6	0.0320
	VAS	61.4 (19.1)	49.1 (17.9)	5	0.0005
K-BILD	Total score	53.6 (11.7)	48.2 (10.7)	2	0.0166
Mean(SD)	Breathlessness and activity	41.1 (20.6)	31.6 (23.3)	0	0.0052
	Chest symptoms	64.4 (22.2)	57.4 (22.4)	1	0.0920
	Psychological impact	52.2 (13.8)	47.1 (12.1)	2	0.1014

Percentages in the excluded group show the percent of valid answers. *Abbreviations*: *SD* Standard deviation, *EQ-5D* EBVS-EQ-5D experience based value set, VAS-Visual Analog Scale, IPF-idiopathic pulmonary fibrosis, IIP-idiopathic interstitial pneumonia, EBVS-experience based value. [a]inlcuding non-specific interstitial pneumonia, desquamative interstitial pneumonia, cryptogenetic organizing pneumonia, lymphocytic interstitial pneumonia, respiratory bronchiolitis-associated interstitial lung disease, pleuropulmonary fibroelastosis, and acute interstitial pneumonia

Heidelberg were older, have more frequently basic education, were retired more often, showed higher comorbidity score and lower DLCO % pred values but showed no difference in the outcome variables (Additional file 1). The most frequent comorbidity was arterial hypertension (41.3%), followed by coronary heart disease (19.5%) and diabetes mellitus (15.9%). All other comorbidities were present in less than 10% of the study population.

ILD-specific and generic HRQL

K-BILD showed the least missing values, followed by VAS and EQ-5D EBVS with two (0.7%), five (1.87%) and six (2.24%) missing values respectively.

There was no indication for ceiling or floor effects in any outcome parameter. Regarding generic HRQL, 29 (12.7%) patients had the maximum possible score for EQ-5D EBVS, three the maximum VAS score, but no one zero. There was only one patient each within the best and worst categories for the K-BILD. K-BILD domains showed also no ceiling or floor effects ('breathlessness and activity' worst 5.2%, best 4.4%, 'chest symptoms' 0.4% vs 14.4%, 'psychological impact' 0.4% vs 0.4%).

Altogether, ILD-specific HRQL had lower values relative to their scale than generic HRQL (EQ-5D EBVS: 0.66 and VAS: 61.4 vs K-BILD: 53.6) (Fig. 1) with the highest impairments occurring in the 'breathlessness and activity' domain (unadjusted mean

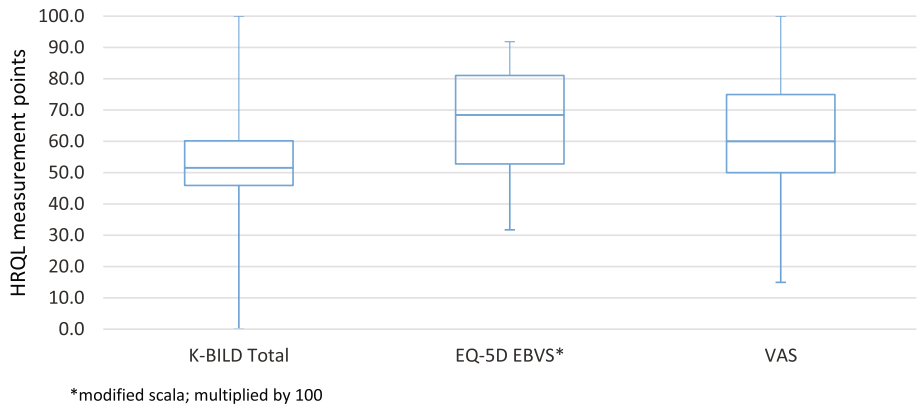

Fig. 1 Unadjusted health-related quality of life results. Abbreviations: EQ-5D EBVS-EQ-5D experience based value set, VAS-Visual Analog Scale

scores: 41.1 vs. 52.2 'psychological impact' and 64.4 'chest symptoms').

Correlation of HRQL assessment tools and internal consistency

K-BILD total score correlated strongly with the EQ-5D EBVS (0.71), but only moderately with the VAS (0.55). All instruments had weak or moderate correlations with lung function parameters and co-morbidity burden (Table 2). The K-BILD domains showed stronger correlations to the EQ-5D EBVS than to the VAS. Looking at correlations between the K-BILD and the EQ-5D dimensions, the strongest correlation was found for the 'breathlessness and activity' (K-BILD) with 'usual activities' (EQ-5D) (− 0.69, $p < 0.00001$) and 'mobility' (EQ-5D) (− 0.65, $p < 0.0001$). However, further correlations were moderate at best (Table 3).

The K-BILD total score showed the highest internal consistency with Cronbach's alpha of 0.89, followed by the 'breathlessness and activity' domain, 'chest symptoms' and psychological impact with values 0.87, 0.74 and 0.73 respectively.

Impact factors on ILD-specific and generic HRQL

In the primary analysis, the strongest influencing factor for all of the HRQL measures and their domains was FVC % pred (Table 4). Older age, higher education and

Table 2 The correlation between health status and lung function

	EQ-5D EBVS	VAS	K-BILD Total	K-BILD Breath	K-BILD Chest	K-BILD Psych	FVC % predicted	DLCO % predicted	Comorbidity sum score
EQ-5D EBVS	1								
VAS	0.58 (<.0001)	1							
K-BILD Total	0.71 (<.0001)	0.55 (<.0001)	1						
K-BILD Breath	0.71 (<.0001)	0.58 (<.0001)	0.86 (<.0001)	1					
K-BILD Chest	0.60 (<.0001)	0.47 (<.0001)	0.78 (<.0001)	0.61 (<.0001)	1				
K-BILD Psych	0.60 (<.0001)	0.49 (<.0001)	0.93 (<.0001)	0.67 (<.0001)	0.69 (<.0001)	1			
FVC % predicted	0.30 (<.0001)	0.21 (0.0013)	0.29 (<.0001)	0.36 (<.0001)	0.22 (0.0006)	0.24 (0.0002)	1		
DLCO % predicted	0.17 (0.0106)	0.14 (0.0409)	0.27 (<.0001)	0.35 (<.0001)	0.12 (0.0821)	0.22 (0.001)	0.47 (<.0001)	1	
Comorbidity sum score	−0.26 (<.0001)	− 0.25 (0.0002)	− 0.21 (0.0012)	− 0.28 (<.0001)	−0.16 (0.0173)	− 0.16 (0.0174)	− 0.09 (0.1569)	− 0.26 (<.0001)	1

Abbreviations; K-BILD Breath- K-BILD Breathlessness and activity, K-BILD Chest- K-BILD Chest symptoms, K-BILD Psych- K-BILD Psychological Impact, FVC % pred –Forced vital capacity % predicted, DLCO % predicted- Carbon monoxide diffusing capacity % predicted. In brackets we reported *p*-values. We considered correlations < 0.3 as weak, ≥0.3 and < 0.7 as moderate and ≥ 0.7 as strong

Table 3 Relationship between the different HRQL domains

	EQ-5D Mobility	EQ-5D Self-care	EQ-5D Usual activities	EQ-5D Pain/ discomfort	EQ-5D Anxiety/ depression
EQ-5D Mobility	1				
EQ-5D Self-care	0.51 (<.0001)	1			
EQ-5D Usual activities	0.65 (<.0001)	0.54 (<.0001)	1		
EQ-5D Pain/ discomfort	0.40 (<.0001)	0.29 (<.0001)	0.49 (<.0001)	1	
EQ-5D Anxiety/ depression	0.39 (<.0001)	0.26 (<.0001)	0.44 (<.0001)	0.24 (0003)	1
K-BILD Breath	−0.65 (<.0001)	− 0.48 (<.0001)	− 0.69 (<.0001)	− 0.45 (<.0001)	− 0.38 (<.0001)
K-BILD Chest	− 0.49 (<.0001)	− 0.31 (<.0001)	−0.52 (<.0001)	− 0.45 (<.0001)	−0.35 (<.0001)
K-BILD Psych.	−0.47 (<.0001)	−0.33 (<.0001)	− 0.53 (<.0001)	−0.39 (<.0001)	− 0.51 (<.0001)

Abbreviations; K-BILD Breath- K-BILD Breathlessness and activity, K-BILD Chest- K-BILD Chest symptoms, K-BILD Psych- K-BILD Psychological Impact. In brackets, we reported *p*-values. We considered correlations < 0.3 as weak, ≥0.3 and < 0.7 as moderate and ≥ 0.7 as strong

working full time were associated with higher EQ-5D EBVS but did not significantly influence K-BILD or VAS. Patients classified as other ILDs had worse HRQL compared to IPF patients (measured with EQ-5D EBVS, VAS and 'breathlessness and activity' domain) or compared to 'other IIP' patients (measured with EQ-5D EBVS and with 'psychological impact' domain) (Table 4). Least square mean comparisons between the remaining groups showed in two cases a difference in the primary analysis. Regarding EQ-5D EBVS sarcoidosis patients had lower values than patients did in the 'other IIP' group (− 0.098 p = 0.0278). Moreover, they were significantly more impaired in the 'breathlessness and activity' domain of

Table 4 Results of regression analyses for the primary analysis (with comorbidity score)

				K-BILD			
		EQ-5D EBVS	VAS	Total	Breathlessness and activity	Chest symptoms	Psychological impact
	Parameter	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate
Sex *(Ref = Male)*	Female	−0.016	2.63	−0.99	−1.39	− 0.90	− 1.49
Age		0.003**	−0.02	0.12	0.06	0.22	0.13
Education	Basic	−0.05*	−5.44	−1.85	−4.47	− 1.66	− 0.84
(Ref = higher)	Secondary	0.03	−0.65	0.24	0.16	2.86	0.14
Employment	Full-time	0.11**	6.05	3.78	5.85	7.19	2.31
(Ref = retired)	Part-time	0.07	7.9	3.39	7.26	2.13	3.27
	Not employed	−0.02	−3.38	−0.23	−1.03	−3.49	−0.19
Clinic *(Ref = GH)*	Heidelberg	0.03	1.86	2.28	3.61	4.25	3.05
Smoking status *(Ref = smoker)*	Never smoker	0.03	0.75	6.00	3.01	6.45	9.53*
	Former Smoker	−0.02	0.45	−2.69	−4.52	−4.71	−2.59
FVC % pred		0.002***	0.19*	0.15**	0.28***	0.27**	0.16**
DLCO % pred		−0.0004	−0.06	0.07	0.20*	−0.004	0.06
Disease Subtype *(Ref = other)*	IPF	0.07*	8.39*	3.51	7.85*	3.62	2.71
	Sarcoidosis	−0.001	−0.49	− 0.90	−4.34	− 5.71	0.67
	HP	0.02	0.73	3.33	3.29	4.89	4.43
	Other IIPs[1]	0.09*	8.78	4.70	5.48	1.16	7.29*
Comorbidity sum score		−0.03***	−2.72**	−1.51**	−3.06***	− 2.68**	− 1.44*

Values depicted are the beta estimates of regression coefficients. *Abbreviations*: IPF-idiopathic pulmonary fibrosis, HP-Hypersensitivity pnemonitis, GH-Großhansdorf.
[1]including: non-specific interstitial pneumonia, desquamative interstitial pneumonia, cryptogenetic organizing pneumonia, lymphocytic interstitial pneumonia,respiratory bronchiolitis-associated interstitial lung disease, pleuropulmonary fibroelastosis, and acute interstitial pneumonia. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$

K-BILD compared to IPF patients (− 12.19 $p = 0.0089$) (Additional file 2).

The secondary analysis revealed two comorbidities with a significant influence; arterial hypertension was associated with a lower EQ-5D EBVS (− 0.05 $p = 0.0441$) and with a lower score in the 'breathlessness and activity' domain (− 6.85 $p = 0.0173$) (Table 5). Additionally, depression had a strong negative association with the 'chest symptoms' domain (− 17.04 $P = 0.0029$).

Within the sensitivity analyses ($n = 257$) results changed only marginally in terms of effect sizes as well as in terms of significant levels. Altogether, the results of the main analyses were mirrored almost perfectly without any noteworthy exceptions.

Discussion

Here, we provide first comprehensive data on HRQL in real life settings in Germany of a large ILD cohort and compared a ILD-specific HRQL questionnaire (K-BILD) with the generic EQ-5D-5L in order to examine its suitability to measure HRQL of ILD patients in a German setting.

In summary our results show, that K-BILD is well accepted among patients (low number of missing values) and its results in Germany are comparable to those of other studies [9, 10, 28], thus supporting the use of K-BILD in Germany. Additionally, further analysis showed the EQ-5D-5L to have properties similar to the K-BILD and hence allowing its use in the ILD disease group, and open up comparability of ILD disease burden in terms of HRQL to that of other diseases as well as to HRQL in the general population.

Both instruments lack floor and ceiling effects, indicating that they should be able to detect changes in the patients HRQL over time, which is important for further clinical research. Accordingly, the implementation of these tools could promote better understanding of the impairments in ILD in different countries, among different study populations and the HRQL development throughout time.

Our study was the first applying the K-BILD in a German observational study and comparisons to international evidence need to be interpreted keeping different healthcare environment and patient preferences in mind. Three studies from different European contries reported comparable mean K-BILD scores as our study [9, 10, 28] Additionally, in line with our findings, the ILD patients of the Wapenaar study had the greatest impairment in the 'breathlessness and activity' domain, followed by 'psychological impact' and 'chest symptoms'. This strongly supports the international transferability and applicability of K-BILD. Despite the high concordance of K-BILD scores cross-nationally and the lack of any other ILD-specific questionnaire, the tool is sparsely used. Our results suggest that a more frequent use would be beneficial.

Table 5 Results of regression analyses for the comorbidities in the secondary analysis

			K-BILD			
	EQ-5D EBVS	VAS	Total	Breathlessness	Chest symptoms	Psychological impact
	Est.	Est.	Est.	Est.	Est.	Est.
Comorbidities						
Pulmonary hypertension	−0.02	2.99	2.43	1.74	5.71	3.33
Arterial hypertension	−0.05*	−2.96	−2.75	−6.85*	−1.04	− 2.08
Coronary heart disease	−0.02	−4.70	−2.06	− 2.90	−5.97	− 2.51
Congestive heart failure	0.05	−4.57	2.12	3.22	3.81	1.81
Other CVD	0.02	10.96	4.07	0.98	16.60	5.34
Diabetes mellitus	−0.03	−0.96	−1.89	0.35	−3.94	−1.91
Emphysema/COPD	−0.06	− 3.02	− 3.18	−6.03	−9.11	−2.38
Lung cancer	0	0	0	0	0	0
Depression	−0.08	−5.27	−5.36	−5.69	−17.04**	−5.87
GERD	0.04	−6.87	−0.86	−1.72	−4.39	−2.62
Renal failure	0.08	2.83	3.79	8.87	6.28	1.82
OSAS	−0.05	−6.11	−1.71	−3.66	−8.22	−1.44
Thromboembolism	−0.11	−2.82	−2.27	0.31	−9.70	−4.93
Malignant tumor	−0.01	−13.66	−0.68	3.60	−13.06	−1.28

Values depicted are the beta estimates of regression coefficients,all adjusted for age, sex, education, employment,clinic location, smoking status, FVC % pred, DLCO % pred, disease suptype. *Abbreviations* Est-estimates, *IPF* idiopathic pulmonary fibrosis, HP: hypersensitivity pnemonitis [1]including: non-specific interstitial pneumonia, desquamative interstitial pneumonia, cryptogenetic organizing pneumonia, lymphocytic interstitial pneumonia,respiratory bronchiolitis associated interstitial lung disease, pleuropulmonary fibroelastosis, and acute interstitial pneumonia. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$

K-BILD showed strong correlation with the EQ-5D EBVS, suggesting that it measures similar aspects. At the same time, K-BILD showed stronger correlations with lung function parameters than EQ-5D-5L. This emphasizes the assumption that K-BILD may be more suitable to detect the impairments originating from ILD. Since EQ-5D-5L might not be sensitive enough in case of disease-specific conditions [29], it is especially important to find a valid instrument measuring disease-specific burden to foster a more targeted patient-centered ILD management.

The trend of lower correlations between the HRQL domains highlights once more the difference between generic and disease-specific questionnaires, but the high correlation of the K-BILD and EQ-5D EBVS allows us to still assume a good overall picture about the HRQL. Our results show in almost all cases slightly lower but still comparable correlation coefficients (EQ-5D, VAS vs K-BILD and its domains) than from the language validation from Wapenaar et al. [10]. Furthermore, reliability measured with Chronbach's alpha showed comparable results (good or moderate) as in Patel et al. [9] and in Wapenaar et al. [10], proving the consistency of K-BILD.

As expected EQ-5D EBVS reacted more sensitive to sociodemographic factors (age, sex) and socioeconomic status (employment) than K-BILD, since generic instruments are known to implicitly cover generic health aspects more comprehensively than disease-specific ones. The unexpected results of the association of higher age with higher HRQL could occur because older people have lower expectations regarding HRQL or maybe because of further undetected covariables, for which are not accounted in this setting. The disease-specific K-BILD showed in the domains 'breathlessness and activity' and 'chest symptoms' a high sensitivity for the lung function value FVC % pred. These findings are in contradiction with Coelho et al., who did not find any association by applying Medical Outcomes Study Short Form 36 -item questionnaire and SGRQ [8]. This could be due to the lower number of patients or due to use of a different HRQL assessment tool in their study. Worth mentioning, HRQL assessment tools are meant to quantify important aspects of the patients' subjective well-being than objective clinical outcomes, therefore low correlations between lung function values and HRQL seem to be tolerable.

Results from King et al. as well as from Kreuter et al. suggest that increasing comorbidity burden is associated with increased morbidity and mortality [30, 31]. Apart from these hard outcomes, the number of comorbid conditions has apparently also a detrimental effect on self-rated HRQL of ILD patients. Therefore, improved comorbidity management in ILD patients might not only reduce the mortality risk itself but also contribute to improve HRQL. In this regard, the generic EQ-5D EBVS and VAS react stronger to comorbidity than K-BILD. The reason could be the high severity of the ILD and thus overpowering other comorbidities in the psychological aspects. Even though our study population differed regarding their HRQL from the excluded patients (showed significantly lower values), the sensitivity analysis confirmed our primary results.

The secondary analysis revealed that, despite the significant association between comorbidity burden and HRQL, only a few distinct comorbidities seem decisive for HRQL. Depression was negatively associated with HRQL in the 'chest symptoms' domain. This could reflect the high burden of the underlying ILD or could be due to the general association between chest pain and depression independent of ILD [32]. The lack of association between HRQL and other comorbidities could be due to the low number of patients for the distinct comorbidities. Furthermore, these results suggest the additivity of the effects of the comorbidities on HRQL.

There is no clinical evidence for the measured higher impairment of sarcoidosis patients in comparison with other subtypes yet; further research is needed in this regard. A possible explanation could be that compared to diseases restricted to the lungs, e.g. IPF or HP, sarcoidosis is a systemic disease with systemic consequences.

Our findings have to be interpreted under some caveats. As with any observational cross-sectional study, our results show associations but causality cannot be tested. Controlling for confounders was the best strategy to address this issue but some important confounder might have been overlooked. Furthermore, this study was voluntary, therefore selection bias cannot be ruled out. Given the high accordance of our findings to international evidence on K-BILD, we consider selection bias to be of minor importance. Additionally, the results may not be generalizable to populations outside Germany.

The crucial strength of our study is that we applied K-BILD for the first time in a German setting in a comparatively large patient cohort consisting of individuals with various subtypes located on quite distant geographic location. This enables general conclusions on the suitability of K-BILD as the disease-specific HRQL measurement of choice.

Conclusion

In conclusion, K-BILD and EQ-5D revealed similar HRQL trends and were sensitive to the same disease-related factors. K-BILD reacted more sensitively to ILD-specific aspects of HRQL rendering it a valuable complementary measure to the generic EQ-5D-5L. Therefore, we propose that K-BILD should be implemented as standard tool in clinical practice.

Additional files

Additional file 1: Baseline characteristics stratified by clinic. (DOCX 17 kb)

Additional file 2: Least squares means comparison among disease subtypes for the primary analysis. (DOCX 17 kb)

Abbreviations

COPD: Chronic obstructive pulmonary disease; DLCO % pred: Diffusing capacity of the lungs for carbon monoxide percent predicted; EQ-5D EBVS: EQ-5D experience based value set; EQ-5D-5L: EuroQol 5 dimensions; FVC % pred: Forced vital capacity percent predicted; HP: Hypersensitivity pneumonitis; HRQL: Health-related quality of life; IIP: Idiopathic interstitial pneumonia; ILD: Interstitial lung diseases; IPF: Idiopathic pulmonary fibrosis; K-BILD: King's Brief Interstitial Lung Disease Questionnaire; SGRQ: St George's Respiratory Questionnaire; VAS: Visual Analog Scale

Acknowledgments

We thank our study nurses, Britta Markworth and Beate Schaufler for their patience and support throughout data collection.

Availability of data and materials

The datasets generated and analyzed during the current study are not publicly available due to them containing information that could compromise research participant privacy, but are available from the corresponding author on reasonable request.

Authors' contributions

All authors were involved in the conception of the research. MK, MC and LS initiated the project and decided on research questions and study design. BS designed analyses, programmed the statistical models and drafted the manuscript in close coordination with MK and TB regarding clinical and medical aspects and with SB regarding K-BILD interpretation. JW developed the classification of ILD subtypes and contributed to the interpretation of the results. All co-authors proofread the manuscript critically and approved its final version.

Ethics approval and consent to participate

The local Ethics Committees of Heidelberg and Luebeck approved the study (reference number S-200/2013, and AZ: 16-192). Participants provided written informed consent.

Competing interests

SB developed the K-BILD. All other authors declare no competing interests.

Publisher's Note

Author details

[1]Helmholtz Zentrum München – German Research Center for Environmental Health (GmbH), Institute of Health Economics and Health Care Management, Comprehensive Pneumology Center Munich (CPC-M), Member of the German Center for Lung Research (DZL), Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany. [2]Center for Interstitial and Rare Lung Diseases Pneumology and Respiratory Critical Care Medicine, Thoraxklinik University of Heidelberg, Translational Lung Research Center Heidelberg (TLRC), Member of the German Center for Lung Research (DZL), Roentgenstrasse, 169126 Heidelberg, Germany. [3]LungenClinic Grosshansdorf GmbH Pneumology, Member of the Airway Research Center North (ARCN), German Center for Lung Research (DZL), Wöhrendamm, 80 22927 Großhansdorf, Germany. [4]Centre of Human & Aerospace Physiological Sciences, School of Basic & Medical Biosciences Faculty of Life Sciences & Medicine, King's College London, London WC2R 2LS, UK. [5]Department of Respiratory Medicine, King's College Hospital, Denmark Hill, Brixton, London SE5 9RS, UK. [6]Munich Center of Health Sciences, Ludwig-Maximilians-Universität München, Ludwigstr. 28, 80539 Munich, RG, Germany.

Received: 27 February 2018 Accepted: 8 May 2018
Published online: 25 May 2018

References

1. Chang JA, Curtis JR, Patrick DL, Raghu G. Assessment of health-related quality of life in patients with interstitial lung disease. Chest. 1999;116:1175–82.
2. Swigris JJ, Kuschner WG, Jacobs SS, Wilson SR, Gould MK. Health-related quality of life in patients with idiopathic pulmonary fibrosis: a systematic review. Thorax. 2005;60:588–94.
3. Belkin A, Swigris JJ. Health-related quality of life in idiopathic pulmonary fibrosis: where are we now? Curr Opin Pulm Med. 2013;19:474–9.
4. De Vries J, Drent M. Quality of life and health status in interstitial lung diseases. Curr Opin Pulm Med. 2006;12:354–8.
5. Berry CE, Drummond MB, Han MK, Li D, Fuller C, Limper AH, Martinez FJ, Schwarz MI, Sciurba FC, Wise RA. Relationship between lung function impairment and health-related quality of life in COPD and interstitial lung disease. Chest. 2012;142:704–11.
6. Kreuter M, Swigris J, Pittrow D, Geier S, Klotsche J, Prasse A, Wirtz H, Koschel D, Andreas S, Claussen M, et al. Health related quality of life in patients with idiopathic pulmonary fibrosis in clinical practice: insights-IPF registry. Respir Res. 2017;18:139.
7. Someya F, Nakagawa T, Mugii N. The COPD assessment test as a prognostic marker in interstitial lung disease. Clin Med Insights Circ Respir Pulm Med. 2016;10:27–31.
8. Coelho AC, Knorst MM, Gazzana MB, Barreto SS. Predictors of physical and mental health-related quality of life in patients with interstitial lung disease: a multifactorial analysis. J Bras Pneumol. 2010;36:562–70.
9. Patel AS, Siegert RJ, Brignall K, Gordon P, Steer S, Desai SR, Maher TM, Renzoni EA, Wells AU, Higginson IJ, Birring SS. The development and validation of the King's brief interstitial lung disease (K-BILD) health status questionnaire. Thorax. 2012;67:804–10.
10. Wapenaar M, Patel AS, Birring SS, Domburg RTV, Bakker EW, Vindigni V, Skold CM, Cottin V, Vancheri C, Wijsenbeek MS. Translation and validation of the King's brief interstitial lung disease (K-BILD) questionnaire in French, Italian, Swedish, and Dutch. Chron Respir Dis. 2017;14:140–50.
11. Patel AS, Siegert RJ, Keir GJ, Bajwah S, Barker RD, Maher TM, Renzoni EA, Wells AU, Higginson IJ, Birring SS. The minimal important difference of the King's brief interstitial lung disease questionnaire (K-BILD) and forced vital capacity in interstitial lung disease. Respir Med. 2013;107:1438–43.
12. Kreuter M, Birring SS, Wijsenbeek M, Wapenaar M, Oltmanns U, Costabel U, Bonella F. German validation of the "King's brief interstitial lung disease (K-Bild) health status questionnaire". Pneumologie. 2016;70:742–6.
13. Rabin R, Gudex C, Selai C, Herdman M. From translation to version management: a history and review of methods for the cultural adaptation of the EuroQol five-dimensional questionnaire. Value Health. 2014;17:70–6.
14. Aktekin LA, Eser F, Malhan S, Oksuz E, Keskin D, Bodur H. A comparison of four different HRQoL generic questionnaire in five different patient groups. Rheumatol Int. 2009;30:63–7.
15. Lin FJ, Pickard AS, Krishnan JA, Joo MJ, Au DH, Carson SS, Gillespie S, Henderson AG, Lindenauer PK, MA MB, et al. Measuring health-related quality of life in chronic obstructive pulmonary disease: properties of the EQ-5D-5L and PROMIS-43 short form. BMC Med Res Methodol. 2014;14:78.
16. Khan I, Morris S, Pashayan N, Matata B, Bashir Z, Maguirre J. Comparing the mapping between EQ-5D-5L, EQ-5D-3L and the EORTC-QLQ-C30 in non-small cell lung cancer patients. Health Qual Life Outcomes. 2016;14:60.
17. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20:1727–36.
18. Leidl R, Reitmeir P. An experience-based value set for the EQ-5D-5L in Germany. Value Health. 2017;20(8):1150–6.
19. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, Lloyd A, Scalone L, Kind P, Pickard AS. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health. 2012;15:708–15.
20. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. Health Econ. 2018;27(1):7-22
21. M MV KMV, MAAE S, de Wit GA, Prenger R, AS E. Dutch tariff for the five-level version of EQ-5D. Value Health. 2016;19:343–52.
22. Margaritopoulos GA, Antoniou KM, Wells AU. Comorbidities in interstitial lung diseases. Eur Respir Rev. 2017;26:160027.

23. Raghu G, Amatto VC, Behr J, Stowasser S. Comorbidities in idiopathic pulmonary fibrosis patients: a systematic literature review. Eur Respir J. 2015;46:1113–30.
24. Fulton BG, Ryerson CJ. Managing comorbidities in idiopathic pulmonary fibrosis. Int J Gen Med. 2015;8:309–18.
25. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res. 1995;4:293–307.
26. Palfreyman S, Mulhern B. The psychometric performance of generic preference-based measures for patients with pressure ulcers. Health Qual Life Outcomes. 2015;13:117.
27. Hofacker D, Naumann E. The emerging trend of work beyond retirement age in Germany. Increasing social inequality? Z Gerontol Geriatr. 2015;48:473–9.
28. Ferrara G, Carlson L, Palm A, Einarsson J, Olivesten C, Skold M. Idiopathic pulmonary fibrosis in Sweden: report from the first year of activity of the Swedish IPF-registry. Eur Clin Respir J. 2016;3:31090.
29. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull. 2010;96:5–21.
30. King C, Nathan SD. Identification and treatment of comorbidities in idiopathic pulmonary fibrosis and other fibrotic lung diseases. Curr Opin Pulm Med. 2013;19:466–73.
31. Kreuter M, Ehlers-Tenenbaum S, Palmowski K, Bruhwyler J, Oltmanns U, Muley T, Heussel CP, Warth A, Kolb M, Herth FJ. Impact of comorbidities on mortality in patients with idiopathic pulmonary fibrosis. PLoS One. 2016;11:e0151425.
32. Kim Y, Soffler M, Paradise S, Jelani QU, Dziura J, Sinha R, Safdar B. Depression is associated with recurrent chest pain with or without coronary artery disease: a prospective cohort study in the emergency department. Am Heart J. 2017;191:47–54.

Chapter 3 – Article 2

Measuring quality of life in COPD patients: comparing disease–specific supplements to the EQ–5D–5L

Published in:

Szentes BL, Schwarzkopf L, Kirsch F, Schramm A, Leidl R. Measuring quality of life in COPD patients: comparing disease–specific supplements to the EQ–5D–5L. Expert Rev Pharmacoecon Outcomes Res. 2020 Oct;20(5):523–529. doi: 10.1080/14737167.2019.1662302.

Chapter 4 – Article 3

How does the EQ–5D–5L perform in asthma patients compared with an asthma–specific quality of life questionnaire?

Published in:

Szentes BL, Schultz K, Nowak D, Schuler M, Schwarzkopf L. How does the EQ–5D–5L perform in asthma patients compared with an asthma–specific quality of life questionnaire? BMC Pulm Med. 2020 Jun 13;20(1):168. doi: 10.1186/s12890–020–01205–8.

BMC Pulmonary Medicine

RESEARCH ARTICLE Open Access

How does the EQ-5D-5L perform in asthma patients compared with an asthma-specific quality of life questionnaire?

Boglárka L. Szentes[1*], Konrad Schultz[2], Dennis Nowak[3], Michael Schuler[4] and Larissa Schwarzkopf[1,5]

Abstract

Background: Asthma patients experience impairments in health-related quality of life (HRQL). Interventions are available to improve HRQL. EQ-5D-5L is a common generic tool used to evaluate health interventions. However, there is debate over whether the use of this measure is adequate in asthma patients.

Methods: We used data from 371 asthma patients participating in a pulmonary rehabilitation (PR) program from the EPRA randomized controlled trial. We used four time points: T0 randomization, T1 start PR, T2 end PR, T3 3 months follow-up. We calculated floor and ceiling effects, intra-class correlation (ICC), Cohen's d, and regression analysis to measure the sensitivity to changes of EQ-5D-5 L (EQ-5D index and Visual Analog Scale (VAS)) and the disease-specific Asthma Quality of Life Questionnaire (AQLQ). Furthermore, we estimated the minimally important difference (MID). Based on the Asthma Control Test (ACT) scores, we defined three groups: 1. ACT-A (ACT> 19) controlled asthma, 2. ACT-B (14 < ACT≤19) not well-controlled asthma, and 3. ACT-C (ACT≤14) very poorly controlled asthma.

Results: Only the EQ-5D index showed ceiling effects at T2 and T3 (32%). ICC (between T0 and T1) was moderate or good for all measures. Cohen's d at T2 and T3 was better at differentiating between ACT-A and ACT-B than between ACT-B and ACT-C. The EQ-5D index showed moderate effect sizes (0.63–0.75), while AQLQ showed large effect sizes (0.74–1,48). VAS was responsive to pronounced positive and negative ACT changes in every period, and AQLQ mostly to the positive changes, whereas the EQ-5D index was less responsive. We estimated a MID of 0.08 for the EQ-5D index, 12.3 for VAS, and 0.65 for AQLQ.

Conclusion: All presented HRQL tools had good discriminatory power and good reliability. However, EQ-5D-5 L did not react very sensitively to small changes in asthma control. Therefore, we would suggest using supplementary measures in addition to EQ-5D-5 L to evaluate asthma-specific interventions more comprehensively.

Trial registration: German Clinical Trial Register, DRKS00007740 (date of registration: 05/15/2015), https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00007740. The registration took place prospectively.

Keywords: EQ-5D-5 L, AQLQ, ACT, Asthma, Health-related quality of life, Responsiveness, Reliability, MID

* Correspondence: szentesboglarka@gmail.com
[1]Institute of Health Economics and Health Care Management, Helmholtz Zentrum München (GmbH) – German Research Center for Environmental Health, Comprehensive Pneumology Center Munich (CPC-M), Member of the German Center for Lung Research (DZL), Ingolstädter Landstraße 1, 85764 Neuherberg, Germany
Full list of author information is available at the end of the article

Background

Asthma is a respiratory disease characterized by chronic inflammation of the airways. Asthma patients experience cough, wheeze, and shortness of breath in varying intensity and frequency [1]. This symptom profile is associated with impairments in health-related quality of life (HRQL) [2–4]. These symptoms can be reduced by adequate drug therapy [1] and through several supplementary management strategies (e.g., patient education [5], respiratory physiotherapy [6], and exercise training [7, 8]), which would increase asthma control and thus presumably HRQL as well.

Two groups of HRQL assessment tools exist, disease-specific and generic ones. Disease-specific assessment tools are developed for specific diseases. They mainly focus on the impact of disease symptoms and the related consequences, but might also cover aspects of disease-associated impairments in social participation or emotional and general wellbeing. They enable comparisons between patients at different stages of the same disease and help to monitor disease development. In contrast, generic assessment tools can be applied across different diseases because they focus on impairments in general health-related aspects of life. Thus, comparisons between different disease areas or with the general population become possible. However, they might not always fully capture HRQL impairments in the context of disease-specific symptoms, especially in the early stages of a disease [9].

One of the most commonly used generic assessment tools is the EQ-5D-5L from the EuroQol group [10], which is a multi-attribute utility instrument (MAUI) for health economic evaluation. It allows the calculation of quality adjusted life years (QALY) [11], an important measure applied in cost–utility studies. Cost–utility studies are approaches, which evaluate and compare health interventions by assessing the costs of an intervention (for example, a pulmonary rehabilitation (PR)) in relation to its health effects. Based on this so-called incremental cost-effectiveness ratio and on additional information, a decision about implementation can be made. Another important aspect to facilitate this decision is the concept of minimally important difference (MID). According to Jaeschke et al. [12], the MID reflects "the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management." QALYs and MIDs reflect strategies that take into account different points of view to support decision making in the health care sector, and both approaches have their own reasons for being. Different countries set different priorities regarding the use of one or the other strategy. Furthermore, different stake holders (policy decision makers, clinicians, payers) and different research questions might favor one or the other parameter.

There is debate over whether the use of the generic EQ-5D is adequate in asthma patients. Whalley et al. The three-level version has already raised some concerns, e.g., its inefficient ability to differentiate between different levels of asthma control [13] or that it might miss clinically important changes in asthma control, which is closely associated with higher HRQL [14] . To overcome this issue, a five-level version of the EQ-5D, the EQ-5D-5 L, was developed, which allows more flexibility regarding the description of health states. Thus, a higher sensitivity to change was expected. However, based on a qualitative study in asthma patients, Whalley et al. [15] argued that, even after refinement of the levels, the dimensions per se are lacking in some asthma-relevant aspects. Furthermore, Hyland et al. [16] criticized the low correlation of EQ-5D-5L with lung function values. Hernandez et al. evaluated the metric properties of the EQ-5D-5 L in a cross-sectional setting to confirm the previous results [17]. They found good construct validity and good discriminative ability between health-related groups. Nevertheless, they did not assess responsiveness to changes and did not compare the EQ-5D-5L with a disease-specific assessment tool.

Therefore, our aim is to investigate whether the EQ-5D-5L is suited to measure HRQL in asthma patients in a longitudinal setting, whether it is reliable, and if it is responsive to changes in asthma control, compared with the established disease-specific Asthma Quality of Life Questionnaire (AQLQ). Furthermore, we aim to provide a MID value for the five-level version for asthma patients, which has not to our knowledge been provided in previous studies.

Methods

We used data from the EPRA study, a randomized controlled trial (RCT) using a wait-list control group assessing the effectiveness of PR among asthma patients (Registered in Deutschen Register Klinischer Studien No. DRKS00007740, the ethics committee of Bayerischen Landesärztekammer approved the study No. 15017). After approval for rehabilitation (T0), patients were randomized to the intervention group (IG) or control group (CG). The IG started the 3-week PR 4 weeks after randomization (T1: start of PR; T2: end of PR), whereas the CG started PR 5 months after randomization (T3). Further details of the study have been published elsewhere [18]. We assessed HRQL and asthma control at T0, T1, T2, and T3 in both groups. For the subsequent analyses, we only included patients with no missing values in the HRQL measures at any time point until T3 to avoid bias through imputation.

Furthermore, we pooled the data from both groups. Figure 1 shows the timeline and the time point of the statistical tests described in the statistical analysis section.

We assessed disease severity and HRQL using the following measures:

Asthma control test (ACT)

The ACT is a self-administered questionnaire to evaluate asthma control [19]. It contains five questions with five possible answers addressing asthma symptoms in the previous 4 weeks. The sum score ranges between 5 and 25; values > 19 represent controlled asthma, and values < 20 are regarded as uncontrolled not well-controlled asthma, as defined by the GINA guidelines [20]. A change of three points is regarded as a MID [21]. For parts of our analyses, we grouped patients into three categories according to their achieved ACT score: ACT-A as well-controlled asthma (ACT score > 19), ACT-B as not well-controlled asthma (16–19), and ACT-C as very poorly controlled asthma (5–15).

Asthma quality of life questionnaire (AQLQ)

The standardized version of the AQLQ is an asthma-specific HRQL assessment tool containing 32 questions in four domains (symptoms, activity limitations, emotional function, and environmental exposure) [22, 23]. The questions cover the last 2 weeks prior to the survey. Each question has to be answered on a 7-point Likert scale. The overall score ranges between 1 and 7, with the latter indicating the best HRQL. A change of 0.5 points is regarded as a MID [24].

EQ-5D-5L

The EQ-5D-5L is a generic HRQL measure from the EuroQol group [25], which evaluates the current health state of the patients. It consists of two parts: The first part is the EQ-5D descriptive system with five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression); each represented by five different levels (from experiencing no problems to extreme problems). Combining the dimension-specific levels across the five dimensions yields distinct health states, which form the basis for a preference-based valuation (utility). Country-specific tariffs exist for this valuation. We used the German Tariff from Ludwig et al. [26], which ranges between – 0.661 and 1; the higher the value, the better the HRQL. The second part of the EQ-5D-5L is the visual analog scale (VAS). The VAS is a vertical thermometer assessing self-rated health with values from 1 to 100, with 100 indicating the best HRQL.

Global rating of change scale (GROC)

The GROC is a rating scale with 15 categories assessing the self-reported change in global health. Patients with improvement and deterioration are symmetrically distributed around zero [12, 27], with negative values representing deterioration and positive values representing improvement. We grouped patients according to their perceived changes into four groups following Juniper et al. [24]: "no change" (GROC [– 1; 1], "small change" (GROC [– 3; – 2] and [2; 3]), "moderate change" (GROC [– 5; – 4] and [4; 5]), and "large change" (GROC [– 7; – 6, 6; 7]). Additionally, we split those groups according to

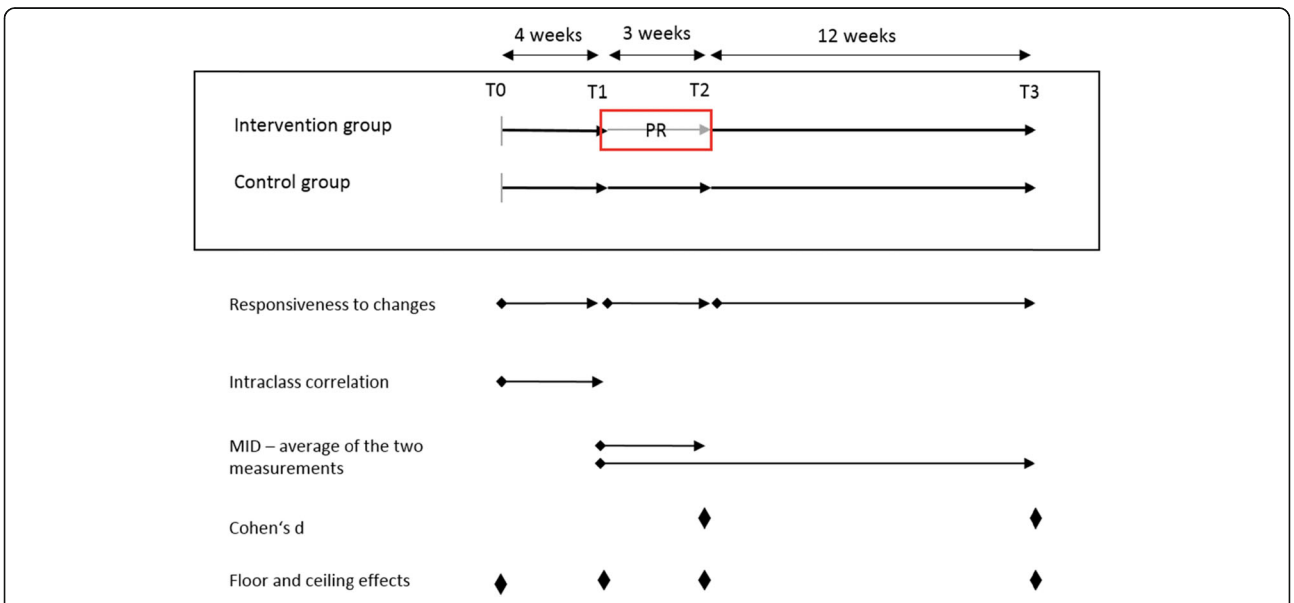

Fig. 1 Study design of the RCT and time points of the conducted pooled statistical analyses. Abbreviations: PR: pulmonary rehabilitation, T0: randomization, T1: start PR, T2: end PR, T3: 12 weeks follow-up

the direction of change to calculate a MID for deterioration and for improvement. We assessed the GROC at T2 and T3 (reference to change was the health state at T1 in both cases).

Statistical analysis and assessing measurement properties

All analyses were performed with SAS (SAS Institute Inc., Cary, NC, USA, version 9.4), and p-values of 0.05 or less were considered statistically significant. We looked at floor and ceiling effects at every time point, defined as > 15% of the patients reaching the best/worst HRQL score [28]. Furthermore, we calculated known-group validity, intra-class correlation (ICC), responsiveness to ACT changes, and the MID.

Known-group validity

Known-group validity (Cohen's d) is used to evaluate the ability of the HRQL tools to differentiate between disease severity groups. Cohen's d was assessed as the mean adjusted differences in HRQL scales between the ACT groups, divided by their pooled standard deviation at T2 or T3. We adjusted for group (IG/CG), age, sex, smoking status, body mass index (BMI), and employment status before PR (yes/no) to compensate for changes not originating from a change in ACT. Cohen's d was considered small between 0.2 and 0.5, moderate from 0.5 to 0.8, and large above 0.8 [29].

Intra-class correlation

To estimate the reliability of the HRQL questionnaires, we evaluated ICC (two-way random effects, absolute agreement, single rater) [30] between T0 and T1 for patients who were stable according to their ACT. We considered patients as stable if their ACT score changed by less than the MID. ICC > 0.9 was regarded as high, 0.75–0.9 as good, 0.5–0.75 as moderate, and < 0.5 as poor [31].

Responsiveness to ACT change

To estimate the responsiveness of HRQL scales associated with a change in ACT, we conducted different regression analyses for each HRQL scale. The dependent variable was the HRQL change score (ΔHRQL) in three periods (period 1: T1−T0, period 2: T2−T1, and period 3: T3−T2). The independent variables were ACT change (ΔACT) in five categories (ΔACT ≥MID, 0 < ΔACT< MID, ΔACT = 0, 0 > ΔACT≥MID, ΔACT≤MID) in the respective period, group (IG/CG), age, sex, BMI, smoking status, employed before PR (yes/no), and previous HRQL at T0, T1, or T2 respectively. ΔACT = 0 was the reference group. The ACT categories are based on the approach of Sullivan et al. [14], who analyzed the responsiveness of the EQ-5D and an asthma-specific questionnaire to changes in asthma control. As a sensitivity analysis, we calculated a quantile regression model for the quantiles 0.5 and for the extremes 0.1 and 0.9, which enables us to portray varying reactions to a continuous ACT change. As there is no hard evidence for the relationship to be linear, considering reactions at different starting points might give deeper insights. This analysis included the same adjustment variables.

Minimal important difference (MID)

We measured the GROC at T2 and T3 and considered a small GROC change as the minimal important change. We calculated the MID separately for improvement and deterioration, as well as combined using the absolute value of the changes. In analogy to Juniper et al. [24], who analyzed MIDs for the AQLQ, the mean of the two measurements (T2 and T3) was considered as the MID. This analysis strategy creates comparability between the disease-specific and generic HRQL tools and enables a cross-validation of our results with existing MIDs for AQLQ.

Results

Study population

The study sample included 371 patients: 199 (53.6%) were in the CG and 172 (46.4%) in the IG. The mean age was 51.4 years (SD: 5.6), and 58.5% of the population was male. Around 50% of the patients were current or previous smokers, and more than 80% were employed before the PR. Baseline HRQL did not differ in the groups, HRQL gains of the IG exceeded that of the CG regarding every measure. The whole development of the HRQL stratified by groups can be seen in Table 1, along with further characteristics.

Properties of the HRQL questionnaires

Floor and ceiling effects

None of the questionnaires used showed floor effects at any time point. Only the EQ-5D index showed ceiling effects at T2 and T3 with 55 (32%) patients each (Additional file 1).

Reliability

AQLQ and the EQ-5D index showed a good ICC (0.82, 95% confidence interval (CI) [0.78; 0.886] and 0.78 CI [0.72; 0.83]); VAS showed moderate ICC (0.62 CI [0.53: 0.70]).

Known-group validity

At T2, there were 185 (49.9%) patients in ACT-A, 72 (19.4%) in ACT-B, and 114 (30.7%) in ACT-C. At T3, there were 164 (44.2%) patients in ACT-A, 94 (25.3%) in ACT-B, and 113 (30.5%) in ACT-C. Adjusted mean scores for the ACT groups at T2 and T3 can be found in Fig. 2. Cohen's d was similar for the EQ-5D index at

Table 1 Characteristics of the study population stratified by group

			All	Control group	Intervention group
N (%)	T0		**371**	**199 (53.6)**	**172 (46.4)**
Male *N (%)*	T0		217 (58.5)	112 (56.3)	105 (61.1)
Age *Mean (SD)*	T0		51.4 (8.6)	51.4 (8.6)	51.4 (8.6)
BMI *Mean (SD)*	T0		29.8 (5.9)	30.3 (5.8)	29.1 (6.1)
Smoking status	T0	Current smoker	50 (13.5)	26 (13.1)	24 (14.0)
		Ex-smoker	142 (38.3)	69 (34.7)	73 (42.4)
		Never smoker	179 (48.3)	104 (52.3)	75 (43.6)
Employed (yes) *N (%)*	T0		322 (86.8)	176 (88.4)	146 (84.9)
ACT	T0		13.2 (3.7)	13.3 (3.8)	13.1 (3.5)
Mean (SD)	T1		15.1 (4.1)	14.8 (4.0)	15.5 (4.1)
	T2		18.1 (5.0)	15.3 (4.6)	21.4 (3.2)
	T3		18.0 (4.8)	15.8 (4.2)	20.6 (4.0)
AQLQ	T0		3.97 (0.93)	3.92 (0.91)	4.03 (0.95)
Mean (SD)	T1		4.21 (1.00)	4.09 (1.02)	4.36 (0.97)
	T2		4.87 (1.22)	4.19 (1.00)	5.66 (0.96)
	T3		4.90 (1.18)	4.41 (1.02)	5.47 (1.09)
EQ-5D index	T0		0.77 (0.20)	0.77 (0.19)	0.77 (0.21)
Mean (SD)	T1		0.80 (0.19)	0.78 (0.19)	0.82 (0.19)
	T2		0.84 (0.18)	0.79 (0.20)	0.90 (0.15)
	T3		0.84 (0.20)	0.80 (0.20)	0.88 (0.18)
VAS	T0		57.2 (16.9)	57.0 (17.6)	57.5 (16.2)
Mean (SD)	T1		60.3 (17.4)	59.6 (18.5)	61.2 (16.2)
	T2		68.0 (19.4)	58.6 (18.5)	78.9 (14.1)
	T3		67.1 (19.1)	59.2 (17.6)	76.2 (16.6)

Abbreviations: *BMI* body mass index, *ACT* Asthma Control Test, *AQLQ* Asthma Quality of Life Questionnaire, *VAS* Visual Analog Scale

every measuring point, whereas VAS was able to discriminate better between well-controlled asthma and not well-controlled asthma than between more severe cases. A similar pattern emerged for AQLQ, but with mostly higher values. Further details on Cohen's d are presented in Table 2.

Responsiveness

The overall responsiveness of a change in asthma control (measured in categories) of the HRQL tools was moderate. In most cases, AQLQ and VAS could differentiate between patients staying stable vs. patients reaching the |MID| on the ACT scale. The EQ-5D index was responsive to changes in only one period (period 3, detecting high negative changes) (Table 3). However, the confidence intervals between adjacent groups frequently overlapped, providing less reliable results for all HRQL measures (Table 3). The sensitivity analysis showed that every HRQL tool reacts positively to an increase in ACT (Table 4); however, the EQ-5D index and AQLQ were not significant in quantile 0.1. Furthermore, there was a gradient change of HRQL in AQLQ and the EQ-5D index through the quantiles, but VAS turned out to be more volatile.

MID

According to GROC at two time points, we identified (combining deterioration and improvement) mean MIDs in the pooled analysis of 0.67 [0.61; 0.74] for AQLQ, 12.28 [10.94; 13.61] for VAS, and 0.09 [0.07; 0.1] for the EQ-5D index (Table 5). Except for the EQ-5D index, we examined a gradient change in HRQL with increasing magnitude of the GROC change. In the analyses stratified for direction of change, the gradient changes appeared in all HRQL measures with regard to improvement. In case of deterioration, a large negative change was associated with positive values in the first measurement, except for the VAS. At the second measurement (T1–T3), the gradient change was detectable for every tool for deterioration and improvement.

Discussion

Our study contributed to the discussion about the suitability of EQ-5D-5 L in measuring asthma severity and

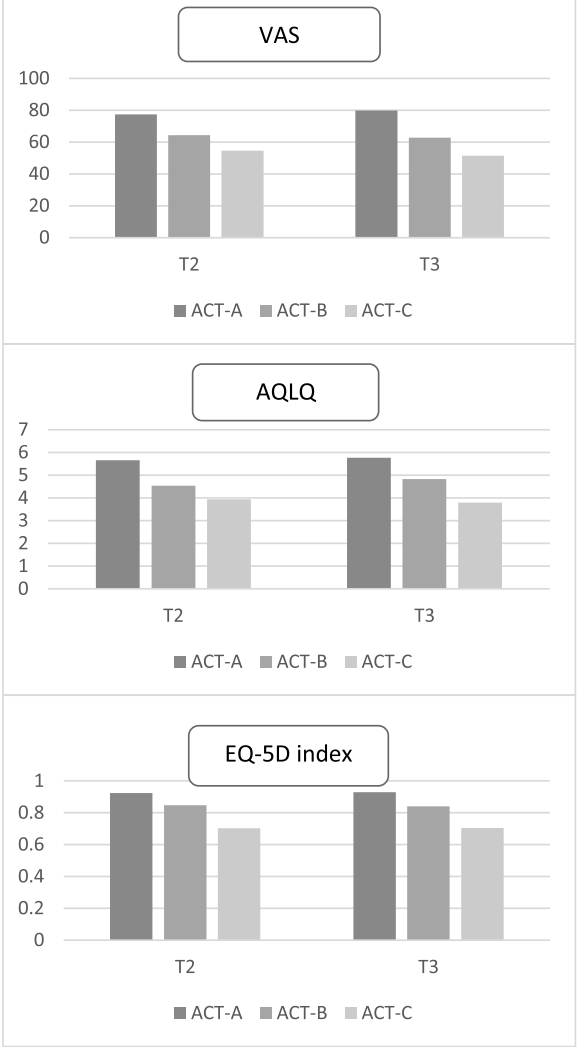

Fig. 2 Adjusted mean scores for the ACT groups at T2 and T3. All differences between the groups were significant at the 0.05 level. Abbreviations: ACT: Asthma Control Test, ACT-A: well-controlled asthma (ACT score > 19), ACT-B: not well-controlled asthma (16–19), and ACT-C: very poorly controlled asthma (5–15)

Table 2 Known-group validity at T2 and T3

Cohen's d		ACT A vs. ACT-B	ACT B vs. ACT-C
T2	AQLQ	1.48	0.74
	EQ-5D index	0.72	0.73
	VAS	0.97	0.57
T3	AQLQ	1.25	1.35
	EQ-5D index	0.75	0.63
	VAS	1.33	0.70

Abbreviations: *AQLQ* Asthma Quality of Life Questionnaire, *VAS* Visual Analog Scale, *ACT* Asthma Control Test, *ACT-A* well-controlled asthma (ACT score > 19), *ACT-B* not well-controlled asthma (16–19), and *ACT-C* very poorly controlled asthma (5–15)

asthma development over time. We assessed its reliability, its ability to differentiate between disease severity, and its responsiveness to changes. As a comparator, we used an established disease-specific questionnaire, the AQLQ. Furthermore, we calculated estimates for the MIDs to facilitate the evaluation of interventions in the disease area asthma.

In a cross-sectional setting, AQLQ showed the best discriminatory power between the asthma severity states, although it showed variation across time points. In contrast, Cohen's d for the EQ-5D index was stable across time points (T2 vs. T3) and different severity levels (ACT-A|ACT-B vs. ACT-B|ACT-C), but lower. Furthermore, AQLQ and VAS had a higher ability to differentiate between patients with asthma control or notand without asthma control (ACT-A vs. ACT-B) compared with differentiating between not well-controlled and very poorly controlled asthma (ACT-B vs. ACT-C). As the goal is to reach asthma control for most of the interventions, the differentiation between different degrees of uncontrolled not controlled asthma might be considered of secondary value. The results suggest that AQLQ, the EQ-5D index as well as VAS are all suited to detect patient groups with low HRQL and greater need for disease control, e.g., patients eligible for PR. Hernandez et al. [17] conducted similar analyses in their study, although using different distinguishing factors, e.g., the number of chronic conditions, asthma control and inhaler use [17]. This makes a comparison of the results difficult. When comparing groups with different asthma control, Hernandez et al. found a better ability of the EQ-5D index to differentiate between the groups compared with VAS [17], which we cannot confirm. Furthermore, the ceiling effect shown in their work is smaller than that we observed (26.5% vs. 32% for the EQ-5D index). The study samples differed in age, female/male ratio, disease severity, and the tariffs used [3, 17]. Additionally, our study sample also included patients with a lower level of asthma control. This might explain the slightly different results.

An important aspect in health economics is the evaluation of health interventions. Therefore, HRQL tools should be reliable and responsive to changes to enable evidence-based recommendations regarding health care interventions. In a longitudinal approach, we assessed reliability (ICC) between T0 and T1, where none of the patients had yet received PR and their ACT score stayed stable. Reliability was moderate for VAS but good for the EQ-5D index and AQLQ. Without interventionA, asthma-related components of HRQL without intervention tend to be more stable than generic health, which might explain the observed higher reliability of the AQLQ. Additionally, AQLQ reflects a time period of 4 weeks, whereas EQ-5D-5 L asks for current health only,

Table 3 Responsiveness of the different HRQL measures to changes in ACT—results of the regression analyses

		N (%)	β-coefficient	*p*-value	95% confidence intervals	
AQLQ change	3 ≤ ΔACT	185 (49.9)	**0.63**	< 0.0001	0.417	0.850
(T2–T1)	0 < ΔACT< 3	*66 (17.8)*	**0.03**	0.824	−0.199	0.250
	ΔACT = 0	*41 (11.1)*	**Ref. cat.**	.	.	.
	−3 < ΔACT< 0	*50 (13.5)*	**−0.14**	0.226	−0.374	0.089
	ΔACT≤−3	*29 (7.8)*	**−0.23**	0.095	−0.494	0.040
AQLQ change	3 ≤ ΔACT	*66 (17.8)*	**0.25**	0.021	0.037	0.464
(T3–T2)	0 < ΔACT< 3	*87 (23.5)*	**0.03**	0.776	−0.171	0.229
	ΔACT = 0	*54 (14.6)*	**Ref. cat.**	.	.	.
	−3 < ΔACT< 0	*93 (25.1)*	**−0.34**	0.0006	−0.532	− 0.145
	ΔACT≤−3	*71 (19.1)*	**−0.84**	< 0.0001	−1.047	−0.632
EQ-5D index change	3 ≤ ΔACT	*185 (49.9)*	**0.04**	0.089	−0.006	0.090
(T2–T1)	0 < ΔACT< 3	*66 (17.8)*	**0.01**	0.663	−0.039	0.061
	ΔACT = 0	*41 (11.1)*	**Ref. cat.**	.	.	.
	−3 < ΔACT< 0	*50 (13.5)*	**0.01**	0.637	−0.039	0.064
	ΔACT≤−3	*29 (7.8)*	**−0.02**	0.419	−0.084	0.035
EQ-5D index change	3 ≤ ΔACT	*66 (17.8)*	**0.02**	0.295	−0.022	0.071
(T3–T2)	0 < ΔACT< 3	*87 (23.5)*	**−0.01**	0.691	−0.052	0.034
	ΔACT = 0	*54 (14.6)*	**Ref. cat**	.	.	.
	−3 < ΔACT< 0	*93 (25.1)*	**−0.02**	0.398	− 0.060	0.024
	ΔACT≤−3	*71 (19.1)*	**−0.08**	0.0008	−0.123	−0.033
VAS change	3 ≤ ΔACT	*185 (49.9)*	**5.82**	0.024	0.782	10.850
(T2–T1)	0 < ΔACT< 3	*66 (17.8)*	**−1.62**	0.540	−6.809	3.572
	ΔACT = 0	*41 (11.1)*	**Ref. cat.**	**.**	**.**	**.**
	−3 < ΔACT< 0	*50 (13.5)*	**−5.29**	0.054	−10.658	0.084
	ΔACT≤−3	*29 (7.8)*	**−7.77**	0.014	−13.946	−1.592
VAS change	3 ≤ ΔACT	*66 (17.8)*	**5.96**	0.009	1.507	10.423
(T3–T2)	0 < ΔACT< 3	*87 (23.5)*	**2.26**	0.286	−1.900	6.424
	ΔACT = 0	*54 (14.6)*	**Ref. cat.**	.	.	.
	−3 < ΔACT< 0	*93 (25.1)*	**−0.69**	0.736	−4.726	3.340
	ΔACT≤−3	*71 (19.1)*	**−8.97**	< 0.0001	−13.298	−4.650

All results are adjusted for group (intervention vs. control), age, sex, smoking status, BMI, employed (yes/no), and HRQL at T0, T1, or T2 respectively. MID for ACT = 3. Abbreviations: *ACT* Asthma Control Test, *AQLQ* Asthma Quality of Life Questionnaire, *Ref. cat*. reference category, *VAS* Visual Analog Scale

Table 4 Responsiveness of the HRQL measures to continuous changes in ACT

		T1-T0			T2-T1			T3-T2		
		Estimate	**Stand. error**	***p*-value**	**Estimate**	**Stand. error**	***p*-value**	**Estimate**	**Stand. error**	***p*-value**
AQLQ	Q0.1	0.03	0.01	0.125	0.08	0.02	< 0.0001	0.13	0.01	< 0.0001
	Q0.5	0.05	0.01	< 0.0001	0.09	0.01	< 0.0001	0.10	0.01	< 0.0001
	Q0.9	0.05	0.02	0.003	0.09	0.02	< 0.0001	0.01	0.02	< 0.0001
EQ-5D index	Q0.1	0.003	0.004	0.414	0.009	0.002	< 0.0001	0.014	0.004	0.0002
	Q0.5	0.005	0.001	0.0005	0.005	0.001	0.0001	0.009	0.002	< 0.0001
	Q0.9	0.005	0.002	0.03	0.0014	0.001	0.171	0.004	0.001	0.008
VAS	Q0.1	1.12	0.40	0.005	1.65	0.31	< 0.0001	1.74	0.35	< 0.0001
	Q0.5	1.00	0.24	< 0.0001	1.1	0.23	< 0.0001	1.37	0.25	< 0.0001
	Q0.9	1.08	0.34	0.002	0.68	0.24	0.006	1.13	0.27	< 0.0001

All results are adjusted for group (intervention vs. control), age, sex, smoking status, BMI, employed (yes/no), and HRQL at T0, T1, or T2 respectively. Abbreviations: Stand. error: standard error, ACT: Asthma Control Test, AQLQ: Asthma Quality of Life Questionnaire, VAS: Visual Analog Scale

Table 5 Mean change in HRQL scores stratified by GROC

T1–T2 *N* = 371	Distinction between deterioration and improvement							No distinction			
N = 371	Large neg *n* = 6	Mod. neg *n* = 13	Small neg *n* = 43	No change *n* = 107	Small pos *n* = 52	Mod. pos *n* = 96	Large pos *n* = 54	No change *n* = 107	Small change *n* = 95	Mod. change *n* = 109	Large change *n* = 60
AQLQ change *Mean [CI]*	0.02 [− 0.25; 0.29]	−0.35 [− 0.69; − 0.01]	−0.04 [− 0.2; 0.12]	0.14 [0.05; 0.24]	0.62 [0.40; 0.83]	1.13 [0.99; 1.27]	1.78 [1.58; 1.97]	0.38 [0.32; 0.44]	0.60 [0.49; 0.71]	1.05 [0.92; 1.19]	1.62 [1.4; 1.83]
EQ-5D index change *Mean [CI]*	0.04 [−0.35; 0.44]	−0.07 [− 0.17; 0.03]	−0.02 [− 0.06; 0.02]	0.00 [− 0.02; 0.03]	0.04 [0.00; 0.08]	0.08 [0.06; 0.11]	0.08 [0.05; 0.11]	0.08 [0.06; 0.1]	0.08 [0.06;0.11]	0.11 [0.09;0.13]	0.11 [0.08; 0.14]
VAS change *Mean [CI]*	−6.67 [−16.98; 3.65]	−9.62 [−22.24; 3.01]	−7.14 [− 11.72; − 2.56]	− 0.45 [−3.22; 2.32]	9.54 [5.96; 13.12]	17.52 [14.37; 20.67]	21.98 [18.10; 25.87]	9.42 [7.32; 11.51]	12.2 [10.07; 14.37]	18.4 [15.7; 21.12]	20.6 [16.92; 24.31]
T1–T3 *N* = 371	Distinction between deterioration and improvement							No distinction			
	Large neg *n* = 3	Mod. neg *n* = 18	Small neg *n* = 40	No change *n* = 119	Small pos *n* = 63	Mod. pos *n* = 85	Large pos *n* = 43	No change *n* = 119	Small change *n* = 103	Mod. change *n* = 103	Large change *n* = 46
AQLQ change *Mean [CI]*	−0.10 [−1.0; 0.77]	−0.03 [− 0.35; 0.29]	0.01 [− 0.17; 0.19]	0.34 [0.23; 0.45]	0.73 [0.57; 0.90]	1.17 [1.0; 1.35]	1.62 [1.34; 1.91]	0.53 [0.45; 0.61]	0.69 [0.6; 0.79]	1.14 [1.02; 1.27]	1.57 [1.3; 1.84]
EQ-5D index change *Mean [CI]*	−0.18 [−1.11; 0.76]	− 0.09 [− 0.22; 0.03]	0.00 [− 0.04; 0.04]	0.02 [− 0.01; 0.04]	0.04 [0.02; 0.07]	0.06 [0.03; 0.09]	0.10 [0.03; 0.16]	0.09 [0.07; 0.11]	0.08 [0.07; 0.1]	0.12 [0.09; 0.14]	0.15 [0.09; 0.2]
VAS change *Mean [CI]*	−23.33 [−37.68; −8.99]	−7.56 [−16.59; 1.48]	−4.88 [− 10.85; 1.10]	0.61 [− 2.16; 3.39]	7.40 [3.95; 10.85]	17.07 [13.63; 20.51]	21.60 [17.12; 26.1]	9.6 [7.49; 11.79]	12.3 [10.0; 14.6]	18.1 [15.47; 20.79]	22.4 [18.49; 26.25]

Abbreviations: *HRQL* health-related quality of life, *GROC* Global Rating of Change, *ACT* Asthma Control Test, *AQLQ* Asthma Quality of Life Questionnaire, *VAS* Visual Analog Scale, *CI* 95% confidence interval, *mod* moderate, *neg* negative, *pos* positive

which increases the volatility of the measurements. Nevertheless, all instruments are suitable for repeated measurements.

We assume that PR improves asthma control and clinical parameters and thus positively affects (at least disease-specific) HRQL. Therefore, in our pooled analysis, we had subgroups experiencing improvement (mostly in the IG) and patient groups staying relatively stable (mostly in the CG). This allowed us to examine HRQL changes in a heterogeneous study population. AQLQ was sensitive to big positive and negative changes (changes ≥|MID|). VAS was also able to differentiate between patients with deteriorating or improving HRQL by more than the MID-ACT, but not between small negative or positive changes. Given that the reference group for all HRQL tools is "no change", a detection of changes below MID is very challenging because of the slight differences from the reference level. The EQ-5D index in our sample could not differentiate significantly between patients reaching a clinically relevant change on ACT (MID) or not, except for one case. This might be an issue regarding cost–utility studies using QALYs as the primary outcome, as suggested by the National Institute for Health and Care Excellence guidelines because, even if patients reach a clinically relevant increase in ACT (MID) through an intervention, it might be overlooked by the EQ-5D index. Thus, the intervention would not be considered cost effective. Looking at the quantile regression approach, a slightly different pattern emerged, where the EQ-5D index detects changes. However, we believe that the magnitude of the change on the EQ-5D index does not match the change on the ACT (e.g. at quantile 0.5 a MID change on ACT only changes the EQ-5D index by approximately 20% of its estimated MID), and leaves a significant improvement on the ACT undetected. Cost–utility studies should thus consider other secondary outcomes, which can potentially evaluate these changes. Similar results were reported from Sullivan et al. [14]; however, the comparison is hindered to some extent, as Sullivan et al. used the previous 3L version of EQ-5D. Therefore, a direct comparison is difficult. VAS and the AQLQ could be used to complement the EQ-5D index, as they showed better (although not perfect) responsiveness to changes. However, AQLQ and VAS are not appropriate measures for cost–utility analysis, but for cost-effectiveness analyses only. In our sensitivity analysis, we confirmed that all measurements react positively to an improvement in ACT. Nevertheless, we think that regarding the magnitude of change, teh EQ-5D index does not react sufficiently sensitive to detect important changes in asthma control. Indeed observed changes in EQ. 5D are rather small and might hence mask the parallel substantial changes in ACT.

Using the GROC to identify the MID for the AQLQ resulted in a slightly higher MID than previous literature would suggest (0.65 vs. 0.5) [24]. However, MID calculations usually differ depending on the study population and the calculation method used. As expected, in the case of deterioration, a smaller change is considered clinically relevant than in the case of improvement. This suggests the existence of different MIDs depending on the direction of change. However, the consideration of different MIDs might not be manageable in a clinical setting. Thus, for most indications, a single MID is used. In the combined analysis, the EQ-5D index characterized no change and minimal change with similar values. Consequently, we can assume that the EQ-5D-5L is less suitable to detect changes in the HRQL of patients, as the previous calculations show. Probably, the dimensions are covering life aspects broadly, but they might miss other important aspects related to asthma. To overcome this issue, Whalley et al. suggest, for example, the addition of a respiratory domain to the EQ-5D [15]. Nevertheless, the calculated value (0.08) was close to the simulation-based values from McClure et al. (0.07) [32]. This suggests the validity of our results; however, the low responsiveness to changes in the utilities should be kept in mind. Furthermore, there is an ongoing debate about the use of MID in economic evaluations, because of its narrow definition [33]. Additionally, there are also concerns about the methodological challenges to incorporate HRQL into RCTs (e.g., HRQL tools being preference based), which also have to be kept in mind during interpretation [34]. These results contribute to the controversy described in the introduction about the use of the EQ-5D in asthma patients. Our study cannot comment on the content validity of the EQ-5D, but we can agree that there might be a need to reconsider the five dimensions in this setting, although further research is necessary on this topic. Another possible solution might be the use of a bolt-on method, which amends the EQ-5D with information on the initially missing dimension [35]. However, there is no scientific consensus about the most suitable bolt-on method yet.

Szende et al. [36] used the previous 3L version and showed evidence of ceiling effects [36]. This implies that the discriminative properties of the EQ-5D in patients experiencing good health may not be sufficient. McTaggart-Cowan et al. are addressing similar aspects, questioning the ability of the EQ-5D to discriminate across different disease severity [13]. Although we experienced similar issues, the use of the 5 L version seemed to lower the magnitude of these.

Although the EQ-5D index showed slightly worse properties than the AQLQ, we should be aware of the different approaches behind the questionnaires. Generic questionnaires cover broad life aspects and facilitate

comparisons among different disease groups, whereas disease-specific measures are for within-group comparisons. Furthermore, regarding the responsiveness of the tools to an ACT change is easier for the AQLQ, as it measures similar aspects and thus has overlapping content, whereas the EQ-5D index lacks asthma-specific content and can only indirectly measure such a construct [37, 38].

There are some limitations to this study. As the EQ-5D assesses current health, whereas the AQLQ has a timeframe of 2 weeks and the ACT of 4 weeks, there is a potential bias while comparing these measures directly. Because asthma has a varying intensity, depending on the asthma attacks, valuing health on a single day may lead to distorted results.

Additionally, there is a chance that HRQL tools behave differently in the control vs intervention CG vs. the IG, and a stratified analysis would be recommended. To achieve a sufficiently high n, we conducted a pooled analysis, but we think that our adjustment for the group variable best possibly accounted for this issue.

The generalizability of the results is not necessarily given for patients outside Germany. Furthermore, patients with initially controlled asthma were not included in this analysis; therefore, we might miss important aspects about mild asthma cases. Nevertheless, the number of patients in this randomized controlled setting was high, and we believe our results are still valuable for the examined disease group.

Conclusion

In conclusion, all presented HRQL tools had good discriminatory power and good reliability. However, EQ-5D-5L had difficulties in detecting (particularly small) changes in disease control. Nevertheless, EQ-5D is still an important tool to compare HRQL across disease areas and to facilitate health economic evaluations, also in the field of asthma. Therefore to draw a more comprehensive picture, we would suggest using supplementary measures (e.g., AQLQ) to EQ-5D-5L to evaluate asthma-specific interventions. Nevertheless, it is still an important tool to compare HRQL across disease areas and to facilitate health economic evaluations.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12890-020-01205-8.

Additional file 1.

Abbreviations

ACT: Asthma Control Test; AQLQ: Asthma Quality of Life Questionnaire; CG: Control group; GROC: Global Ratings of Change; HRQL: Health-related quality of life; ICC: Intra-class correlation; IG: Intervention group; MID: Minimal important difference; PR: Pulmonary rehabilitation; QALY: Quality adjusted life years; VAS: Visual Analog Scale

Acknowledgments

Not applicable.

Authors' contributions

KS initiated the project. BS analyzed the data and drafted the manuscript. LS and MS consulted on statistical methods, and DN and KS on medical questions and implications. All authors proofread the manuscript and approved the final version.

Funding

The study was funded by the Deutsche Rentenversicherung Bayern Süd, No. DRKS00007740.

Availability of data and materials

The datasets generated and analyzed during the current study are not publicly available because they containing information that could compromise research participant privacy, but are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

The study was approved by the ethics committee of Bayerischen Landesärztekammer (No. 15017). All participants provided written informed consent.

Consent for publication

Not applicable.

Competing interests

All authors declare no competing interests.

Author details

[1]Institute of Health Economics and Health Care Management, Helmholtz Zentrum München (GmbH) – German Research Center for Environmental Health, Comprehensive Pneumology Center Munich (CPC-M), Member of the German Center for Lung Research (DZL), Ingolstädter Landstraße 1, 85764 Neuherberg, Germany. [2]Klinik Bad Reichenhall, Center for Rehabilitation, Pulmonology and Orthopedics, Salzburger Str. 8, 83435 Bad Reichenhall, Germany. [3]LMU University of München, Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Ziemssenstraße 1, 80336 Munich, Germany. [4]University of Würzburg, Institute of Clinical Epidemiology and Biometry, Sanderring 2, 97070 Würzburg, Germany. [5]IFT – Institut für Therapieforschung, Leopoldstrasse 175, 80804 Munich, Germany.

Received: 12 December 2019 Accepted: 2 June 2020
Published online: 13 June 2020

References

1. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention 2018 [Available from: www.ginasthma.org.
2. Williams SA, Wagner S, Kannan H, Bolge SC. The association between asthma control and health care utilization, work productivity loss and health-related quality of life. J Occup Environ Med. 2009;51(7):780–5.
3. Hernandez G, Dima AL, Pont A, Garin O, Marti-Pastor M, Alonso J, et al. Impact of asthma on women and men: comparison with the general population using the EQ-5D-5L questionnaire. PLoS One. 2018;13(8):e0202624.
4. Skolnik NS, Carnahan SP. Primary care of asthma: new options for severe eosinophilic asthma. Curr Med Res Opin. 2019;35(7):1309–18.
5. Gibson PG, Powell H, Coughlan J, Wilson AJ, Abramson M, Haywood P, et al. Self-management education and regular practitioner review for adults with asthma. Cochrane Database Syst Rev. 2003;1:CD001117.
6. Bruton A, Lee A, Yardley L, Raftery J, Arden-Close E, Kirby S, et al. Physiotherapy breathing retraining for asthma: a randomised controlled trial. Lancet Respir Med. 2018;6(1):19–28.
7. Turner S, Eastwood P, Cook A, Jenkins S. Improvements in symptoms and quality of life following exercise training in older adults with moderate/severe persistent asthma. Respiration. 2011;81(4):302–10.
8. Mendes FA, Goncalves RC, Nunes MP, Saraiva-Romanholo BM, Cukier A, Stelmach R, et al. Effects of aerobic training on psychosocial morbidity and symptoms in patients with asthma: a randomized clinical trial. Chest. 2010;138(2):331–7.

9. Wacker ME, Jorres RA, Karch A, Wilke S, Heinrich J, Karrasch S, et al. Assessing health-related quality of life in COPD: comparing generic and disease-specific instruments with focus on comorbidities. BMC Pulmonary Med. 2016;16(1):70.
10. Rabin R, Gudex C, Selai C, Herdman M. From translation to version management: a history and review of methods for the cultural adaptation of the EuroQol five-dimensional questionnaire. Value Health. 2014;17(1):70–6.
11. Weinstein MC, Torrance G, McGuire A. QALYs: the basics. Value Health. 2009; 12(Suppl 1):S5–9.
12. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10(4): 407–15.
13. McTaggart-Cowan HM, Marra CA, Yang Y, Brazier JE, Kopec JA, FitzGerald JM, et al. The validity of generic and condition-specific preference-based instruments: the ability to discriminate asthma control status. Qual Life Res. 2008;17(3):453–62.
14. Sullivan PW, Ghushchyan VH, Campbell JD, Globe G, Bender B, Magid DJ. Measurement of utility in asthma: evidence indicating that generic instruments may miss clinically important changes. Qual Life Res. 2016; 25(12):3017–26.
15. Whalley D, Globe G, Crawford R, Doward L, Tafesse E, Brazier J, et al. Is the EQ-5D fit for purpose in asthma? Acceptability and content validity from the patient perspective. Health Qual Life Outcomes. 2018;16(1):160.
16. Hyland ME, Jones RC, Lanario JW, Masoli M. The construction and validation of the Severe Asthma Questionnaire. Eur Resp J. 2018;52:1.
17. Hernandez G, Garin O, Dima AL, Pont A, Marti Pastor M, Alonso J, et al. EuroQol (EQ-5D-5L) validity in assessing the quality of life in adults with asthma: cross-sectional study. J Med Internet Res. 2019;21(1):e10178.
18. Schultz K, Seidl H, Jelusic D, Wagner R, Wittmann M, Faller H, et al. Effectiveness of pulmonary rehabilitation for patients with asthma: study protocol of a randomized controlled trial (EPRA). BMC Pulmonary Med. 2017;17(1):49.
19. Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, et al. Development of the asthma control test: a survey for assessing asthma control. J Allergy Clin Immunol. 2004;113(1):59–65.
20. Thomas M, Kay S, Pike J, Williams A, Rosenzweig JR, Hillyer EV, et al. The asthma control test (ACT) as a predictor of GINA guideline-defined asthma control: analysis of a multinational cross-sectional survey. Prim Care Resp J. 2009;18(1):41–9.
21. Schatz M, Kosinski M, Yarlas AS, Hanlon J, Watson ME, Jhingran P. The minimally important difference of the Asthma Control Test. J Allergy Clin Immunol. 2009;124(4):719–23 e1.
22. Juniper EF, Guyatt GH, Epstein RS, Ferrie PJ, Jaeschke R, Hiller TK. Evaluation of impairment of health related quality of life in asthma: development of a questionnaire for use in clinical trials. Thorax. 1992;47(2):76–83.
23. Juniper EF, Buist AS, Cox FM, Ferrie PJ, King DR. Validation of a standardized version of the asthma quality of life questionnaire. Chest. 1999;115(5):1265–70.
24. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. J Clin Epidemiol. 1994;47(1):81–7.
25. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res. 2013;22(7):1717–27.
26. Ludwig K, von der Schulenburg JM G, Greiner W. German Value Set for the EQ-5D-5L. Pharmacoeconomics. 2018;36(6):663–74.
27. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther. 2009;17(3):163–70.
28. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res. 1995;4(4):293–307.
29. Grissom RJKJ. Effect Sizes for Research: Univariate and Multivariate Applications; 2012.
30. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86(2):420–8.
31. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.
32. McClure NS, Sayah FA, Xie F, Luo N, Johnson JA. Instrument-defined estimates of the minimally important difference for EQ-5D-5L index scores. Value Health. 2017;20(4):644–50.
33. Whitehurst DGT, Bryan S. Trial-based clinical and economic analyses: the unhelpful quest for conformity. Trials. 2013;14(1):421.
34. Drummond M. Introducing economic and quality of life measurements into clinical studies. Ann Med. 2001;33:344–9.
35. Hoogendoorn M, Oppe M, Boland MRS, Goossens LMA, Stolk EA, Rutten-van MM. Exploring the impact of adding a respiratory dimension to the EQ-5D-5L. Med Decision Making. 2019;39(4):393–404.
36. Szende A, Svensson K, Stahl E, Meszaros A, Berta GY. Psychometric and utility-based measures of health status of asthmatic patients with different disease control level. Pharmacoeconomics. 2004;22(8):537–47.
37. Schatz M, Mosen DM, Kosinski M, Vollmer WM, Magid DJ, O'Connor E, et al. The relationship between asthma-specific quality of life and asthma control. J Asthma. 2007;44(5):391–5.
38. Stucky BD, Sherbourne CD, Edelen MO, Eberhart NK. Understanding asthma-specific quality of life: moving beyond asthma symptoms and severity. Eur Respir J. 2015;46(3):680–7.

Publisher's Note

Chapter 5 – Article 4

Cost-Effectiveness of Pulmonary Rehabilitation in Patients With Bronchial Asthma: An Analysis of the Effectiveness of Pulmonary Rehabilitation in Patients With Asthma Randomized Controlled Trial

Published in:

Böckmann D, Szentes BL, Schultz K, Nowak D, Schuler M, Schwarzkopf L. Cost-Effectiveness of Pulmonary Rehabilitation in Patients With Bronchial Asthma: An Analysis of the EPRA Randomized Controlled Trial. Value Health. 2021 Sep;24(9):1254-1262.
doi: 10.1016/j.jval.2021.01.017. Epub 2021 Apr 30. PMID: 34452704.

Acknowledgements

Ich möchte mich mit diesen Zeilen bei all den Personen bedanken, die mich bei der Ausarbeitung meiner Dissertation unterstützt haben.

Ein besonderer Dank gilt meiner Doktormutter PD Dr. Larissa Schwarzkopf und meinem Betreuer Prof. Dr. Rolf Holle für die hervorragende Betreuung und fachliche Unterstützung während der Erstellung der Dissertation.
Ich möchte mich auch bei meinen Koautorinnen und Koautoren ganz herzlich für die stetige Unterstützung und für die gute Mitarbeit bedanken.

Während der Fertigstellung meiner Dissertation konnte ich mich auch immer auf meine ehemalige Kolleginnen und Kollegen verlassen. Für diese Unterstützung will ich mich bei Julia Walter, Johanna Lutter, Phillen Maqhuzu, Manuel Huber, Florian Karl und Christoph Kurz bedanken. Besonders hervorheben möchte ich Julia Walter, die sich immer Zeit für alle meine Fragen genommen hat, egal ob diese fachliche, sprachliche, oder persönliche Natur waren.

Ich möchte mich bei meiner guten Freundin Edda bedanken, die mir immer motivierend zur Seite stand. Ich bedanke mich bei meinem Lebensgefährten Paul, der immer ein offenes Ohr für mich hat und für die Aufmunterungen in schwierigen Phasen meiner Arbeit.
Last but not least danke ich meiner Familie die ohne ihr Wissen mir viel Kraft und Antrieb gegegeben hat.

Mai 2021