Chronic Obstructive Pulmonary Diseases:
Journal of the COPD Foundation

Review

Progress in Characterizing Patient-Centered Outcomes in COPD, 2004-2014

Paul Jones, PhD, FRCP[1]

Abbreviations: patient-reported outcomes, **PRO**; Global Initiative for chronic Obstructive Lung Disease, **GOLD;** logarithum of the odds, **log-odds**; Medical Research Council, **MRC**; Shortness of Breath with Daily Activities questionnaire, **SOBDA**; excacerbation of chronic pulmonary disease tool, **EXACT**; forced expiratory volume in one second, **FEV1** ; Chronic Respiratory Questionaire, **CRQ**; St. George's Respiratory Questionaire, **SGRQ**; COPD Biomarker Qualification Consortium, **CBQC**; COPD Assessment Test, **CAT**; Clinical COPD Questionaire, **CCQ**

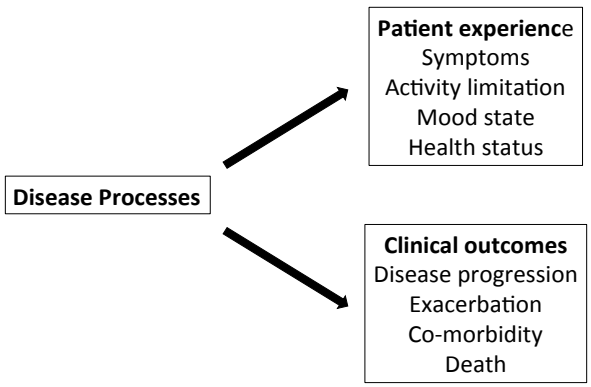
Funding Support: Not applicable
Date of Acceptance: February 28, 2014
Citation: Jones P. Progress in characterizing patient-centered outcomes in COPD, 2004-2014. *J COPD F.* 2014;1(1):17-22.doi:http://dx.doi.org/10.15326/jcopdf.1.1.2014.0121 .

Introduction

It is appropriate that the first edition of a new journal published by a patient advocacy organization should contain a review about measurements that put the patient at the center of the process. By 2004, the battle to give patient-reported outcomes (PRO's) their due place in the assessment of COPD patients for research studies and trials had been won. The strengths and limitations of forced expiratory volume in 1 second (FEV1) measurement had become recognized and even the terminology was being used. For example the Global Initiative for chronic Obstructive Lung Disease (GOLD) guidelines referred to *improving health status* (implying something measurable) as a treatment objective rather than *quality of life,* which is a much looser concept. In the last decade those gains have been consolidated and new developments have moved the research assessment and clinical practice much further. This review will outline some of highlights.

It is important to understand clearly the relationship between disease processes, disease experience as measured by PROs, and clinical outcomes.[1] This is illustrated in Figure 1. The effects of the underlying disease processes will affect both the symptomatic manifestation of the disease and the clinical outcomes, however, it does not mean that a treatment that improves symptoms will automatically improve clinical outcomes. It might, if the treatment reduces a disease process that is tightly linked to both, but on the other hand, it might not. For example, it is known that the Medical Research Council (MRC) dyspnoea score is a predictor of mortality, but that does not mean that reducing breathlessness will extend life.

Figure 1. Relationship Between Disease Processes, Clinical Outcomes and Patient Experience of COPD



1 Infection and Immunity Institute
St. George's University of London, UK

Address correspondence to:

Paul Jones, PhD, FRCP
Infection and Immunity Institute
St George's University of London
Cranmer Terrace
London SW170RE
UK
pjones@sgul.ac.uk

Keywords:

symptom measurement; patient reported outcome; questionnaires

Rasch Modelling

One of the major advances in the assessment of patient-centered outcomes in COPD during this 10-year period (2004-2014) was the introduction of item-response theory to questionnaire development, Rasch modelling being the technique most commonly used.[2,3] (Note: Simple descriptions of Rasch measurement can be found in some of the references cited in Table 1.)This approach is a major step forward compared to the previously used methods, which were comprised of a suite of techniques broadly grouped under the title of classical test theory. In principle, Rasch methodology permits the creation of an instrument with built-in invariant measurement properties, meaning that it works essentially like a ruler. This is important. Consider the situation of Patient A in Figure 2.

During the stable state he appears to have quite mild symptoms, but then has a period of worsening which is of defined magnitude. Patient B is much worse than Patient A when stable and she worsens by the same amount as Patient A. Using a questionnaire that was developed and tested for invariant measurement properties during its creation, some confidence can be placed on the conclusion that the size of change was the same in both patients. [Note: This does not mean that the consequence of the deterioration is the same for the two patients – that will be dependent on their initial state].

Rasch methodology also allows direct tests of whether different groups of patients answer the constituent items differently. An example is given in Figure 3 of the Rasch Item Characteristic Curve for an item concerning sputum colour.

Figure 2. Hypothetical Symptom Scale Illustrating Two Patients

Hypothetical symptom scale illustrating two patients, one more severe when stable than the other, but both experiencing an acute deterioration of the same magnitude. Only a scale with invariant measurement properties allows the assumption that the deterioration in Patient A is truly the same as in Patient B.

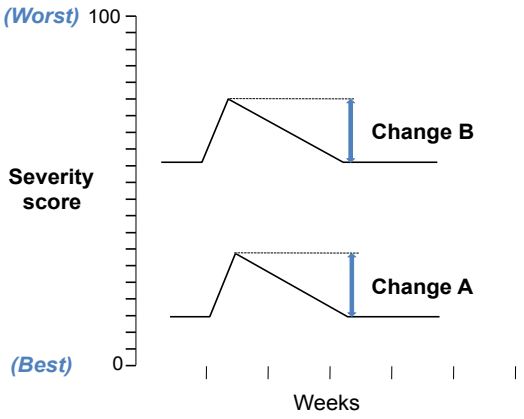
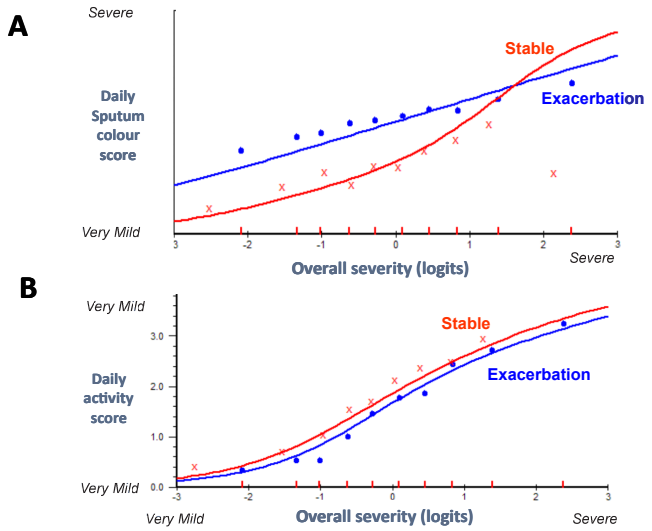


Figure 3. Item Characteristic Curves From a Rasch Analysis of Items Considered for Inclusion in an Exacerbation Diary Card

A) An item concerned with sputum colour and B) An item about physical activity. Red is obtained in the stable state, blue is during exacerbation. It can be seen that the relationship between the item score (y axis) and the overall severity score (x axis) differs between stable and exacerbation in the case of the sputum colour item, but not in the case of physical activity.



Upon viewing the relationship between the response to the item (on the y axis) and the patients' responses to all the other items in the item set on the x axis (measured in the core units of Rasch analysis logits – the logarithm of the odds or *log-odds* of responding), it can be concluded that the relationship of sputum colour to overall severity is different between the stable and exacerbation state. By contrast, another item (daily physical activity) is very reliable as a measure of severity between states. This methodology also allows a direct test for the potential for bias in responses between males and females and between languages. Finally it allows the creation of small and efficient questionnaires that are much quicker and easier to use than those created using older methods. Brief details of some of these new instruments are contained in Table 1.

Dyspnea

Measurement of dyspnea has been available for over half a century in the form of the MRC scale. This is an instrument created by physicians that has stood the test of time. Building on work from Japan that showed that the MRC grade could predict mortality,[4] the modified version (mMRC) was built into a multi-dimensional prognostic scale – the BODE.[5] However, the MRC scale

Table 1. Questionnaires Developed Using Rasch Methodology

Instrument	What's Measured	No. of Items
COPD Assessment Test (CAT)	Overall health status	8
Dyspnoea-12 Questionnaire	Breathlessness	12
EXACT Diary Card	Exacerbation onset, magnitude and recovery	14
Shortness of Breath with Daily Activities (SOBDA) Questionnaire	Effect of breathlessness on daily activiy	13

illustrates one of the problems in assessing breathlessness – reliable measurement of symptoms requires a standardized stimulus, but the level of breathlessness will depend on the level of activity being undertaken. As a result, many dyspnea instruments (with the exception of Visual Analogue and Borg Scales) relate breathlessness to activity. In fact, they are measures of dyspnea-related disability.

This pattern of inactivity dependence in dyspnea measurement was broken when a new instrument called the Dyspnoea-12 was created.[6] It produced reliable measurements of breathlessness based purely on the adjectives that patients used to describe their symptom. The methodology used was grounded firmly in patients' reports of their experience of breathlessness – unlike the MRC, which was based on the views of experienced physicians. It also used Rasch methodology, which enabled the developers to identify items that behaved the same way in patients with COPD, interstitial lung disease and heart failure.

Daily Activity

A recent development has been to further seperate activity from breathlessness. The challenge here is to develop a short instrument that captures, in an unbiased manner, all the essential activities that COPD patients complete. Here the distinction between *health status* and *quality of life* becomes important. It is perhaps best to consider that quality of life is a construct that applies to individuals: we are all unique, and so disease will affect us in ways that may be unique to each person. In contrast, health status is a constrained measurement that treats every unique individual as if he or she were a typical or standard patient (who does not exist, of course). The list of every conceivable activity that COPD patients may undertake for work, socially or for leisure is of course unlimited, so a very standardized narrow approach is required. The Shortness of Breath with Daily Activities Questionnaire, (SOBDA) a unidimensional instrument designed for electronic data capture, has been developed to meet this need.[7] Its developers used Rasch modelling in its development.

Exacerbations

COPD exacerbations are arguably the single most important aspect of COPD, yet despite that there had been little scientific study of the symptomatic characteristics of a COPD exacerbation until the creation of the exacerbation of chronic pulmonary disease tool (EXACT) diary.[8,9] In itself that creation process was a notable development, because the creation process was a partnership between industry, a regulatory authority (the Food and Drug Administration , FDA) and academia. Like the MRC, the criteria for determining the presence of an exacerbation (*the Anthonisen criteria*) were established a long time ago by an experienced physician [10] and later formed the basis of a very successful COPD diary.[11] However, these criteria were not based on the patient's experience of the symptoms of COPD. This was rectified by the EXACT's development, which not only identifies the traditional symptoms of cough, sputum production, chest discomfort and breathlessness, but also shows that items concerning tiredness, sleep disturbance and being scared have very good measurement properties.[8] The EXACT was developed using Rasch analysis.

Data using the EXACT in clinical trials are now emerging and potentially the most significant finding may be its ability to detect and measure the severity of unreported exacerbations. [12] This diary was designed to have very reliable measurement characteristics, but other exacerbation diaries have also provided important insights. For example, the London Chest diary card first demonstrated the existence of unreported exacerbations [13] and drew attention to the possibility of different time courses of exacerbation onset [14] and that aspects of the severity of exacerbations may show seasonal variation.[15]

Use of diary cards to detect unreported exacerbations may prove to have a profound influence on COPD management. Unreported exacerbations significantly outnumber reported ones and are known to have a similar medium-term effect on health status as those that are reported by patients to their clinicians, and therefore are

treated.[16] It is already known that reported exacerbations have a significant effect on rate decline in FEV1,[17] and it remains to be seen whether unreported ones have the same effect. If that is found to be the case, it may alter the entire COPD management paradigm, focussing attention on the detection and prevention of *chest infections* that the patients don't report.

The impact of reliable exacerbation measurement may be different from the other areas reviewed here. Other new PROs provide simple and valid measurements of constructs that were already known but not always easy to measure. Exacerbation diary cards provide new insights into the nature of COPD.

Respiratory Symptoms

A spin off from the development of the EXACT was the EXACT-RS, a diary card designed to provide a measure of respiratory symptoms for use as an outcome measure. A validation study has just been reported.[18]

Health Status Measurement

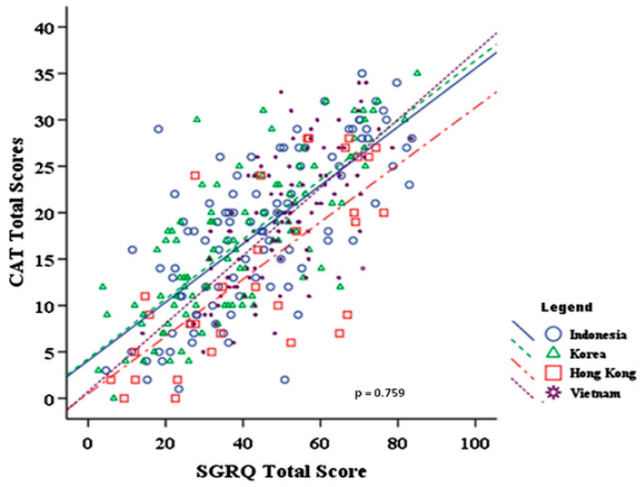
In 2004, the Chronic Respiratory Questionnaire (CRQ) and St. George's Respiratory Questionnaire (SGRQ) were well-established instruments and formed a standard part of clinical trials. A PubMed search up to 2004 using the text word *SGRQ* identified 23 hits; in 2014 there are 647 hits for the same term. By 2007, plots of SGRQ scores were appearing in *New England Journal of Medicine* papers describing major clinical trials.[19,20] In 2013, the FDA indicated its acceptance of the SGRQ, a process that is now being formalized through the COPD Biomarker Qualification Consortium (CBQC).[21]

The large increase in health status measurement over the last decade largely took place in the context of clinical trials and scientific studies, but health status measurement could not find a place in routine management because of the size and complexity of the CRQ and SGRQ. This was corrected by the development of the CCQ, now known as the Clinical COPD Questionnaire, although originally called the COPD Control Questionnaire.[22] This is a short 12-item instrument developed using classical test theory methodology. Perhaps of greater significance was the development of the COPD Assessment Test (CAT), since this 8-item instrument was developed using the modern rigour of Rasch analysis.[23] This short and simple instrument correlates well with the much more complex SGRQ.[24]

Since its first publication in 2009, this writer knows of 24 subsequent research papers concerning the CAT. Although developed specifically for routine practice, the CAT has strong measurement properties, making it possible to acquire reliable health status measurements very simply and cheaply in many languages and in many settings. It appears to behave consistently, in terms of its relationship to the SGRQ, in different countries.[25] The example given in Figure 4 is particularly notable because it was used in 4 languages that were not included in the CAT's development and certainly not the SGRQ's. This figure also has another significance: it illustrates that health status measurement, although originated in English speaking countries, has now become fully internationalized. Scores obtained in one language/culture should have the same meaning as scores obtained in another.

Figure 4. Relationship Between SGRQ and CAT Score in Four Asian Countries.

The p value of 0.759 applies to a test for a difference in slope between countries. Reproduced with permission from Chest.[25]



The other major development in health status measurement came about in 2011 when GOLD recommended use of the CAT in its new comprehensive assessment scheme. In 2014 it went further and recommended it as the assessment of choice, because it provides a comprehensive measure of the overall impact of the disease and better reflects the complexity of COPD than more narrow measures such as the MRC. Health status measurement is on the verge of becoming standard practice in routine clinical assessment of COPD patients.[26] It is no longer just a research tool.

Looking Forward

Developments over the last decade have brought reliable symptomatic measurement into routine clinical and research use, and it is worth noting that this has been done through the creation of simple instruments (8 items in the CAT vs 52 in the SGRQ). By necessity this forces all patients onto the same metric thereby squeezing out individuality. To capture the wide variety of impacts of COPD, the questionnaires must have many more items. Using traditional methods this would be unwieldy, but it can be achieved through a process known as *item banking* coupled with the sophisticated algorithms used in computer adaptive testing.[27] Potentially, this may allow precise estimates of symptom and health status scores within a very small number of keystrokes even less than those needed to complete the 8 items in the CAT. Progress in this very complex field has been slower than hoped but instruments are being developed for COPD – for example the 100-item Dyspnea Management Questionnaire (DMT-CAT).[28] An instrument for assessing asthma has also been published recently.[29]

Conclusions

The developments in patient-centered outcome measurement described here reflect broader shifts in the understanding of and approaches to COPD. No longer is it seen simply as a disease of airflow limitation assessed by the FEV1. Over the next decade, treatment for COPD will be individually *personalized* to each patient's needs. The outcome of developments in PRO measurement over the last decade will play a central role in that process.

References:

1. Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modification PRO Task Force Report. *Value Health*. 2009;12(8):1075-1083.
2. Andrich D. Rasch Models for Measurement. London, England: Sage Publications; 1988.
3. Andrich D. Controversy and the Rasch Model: a characteristic of incompatible paradigms? *Med Care*. 2004;42(1):1-7.
4. Nishimura K, Izumi T, Tsukino M, Oga T. Dyspnea is a better predictor of 5-year survival than airway obstruction in patients with COPD. *Chest*. 2002;121(5):1434-1440.
5. Celli BR, Cote CG, Marin JM, et al. The body-mass index, airflow obstruction, dyspnea and exercise capacity index in chronic obstructive pulmonary disease. *New Engl J Med*. 2004;350:1005-1012.
6. Yorke J, Moosavi SH, Shuldham C, Jones PW. Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. *Thorax*. 2010;65(1):21-26.
7. Howard K, Berry P, Petrillo J, et al. Development of the Shortness of Breath with Daily Activities questionnaire (SOBDA). *Value Health*. 2012;15(8):1042-1050.
8. Jones PW, Chen W-H, Wilcox TK, Sethi S, Leidy NK. Characterizing and quantifying the symptomatic features of COPD exacerbations. *Chest*. 2011;139(6):1388-1394.
9. Leidy NK, Wilcox TK, Jones PW, et al. Standardizing measurement of chronic obstructive pulmonary disease exacerbations. *Am J Respir Crit Care Med*. 2011;183(3):323-329.
10. Anthonisen NR, Manfreda J, Warren CPW, Hershfield ES, Harding GKM, Nelson NA. Antiobiotic therapy in exacerbations of chronic obstructive pulmonary disease. *Ann Intern Med*. 1987;106(2):196-204.
11. Seemungal TA, Donaldson GC, Bhowmik A, Jeffries DJ, Wedzicha JA. Time course and recovery of exacerbations in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2000;161(5):1608-1613.
12. Leidy MK, Murray LT, Jones PW, Sethi S. Performance of the EXACT Patient-Reported Outcome (PRO) measure in three clinical trials in COPD. *Annals ATS*. 2014; In Press.
13. Seemungal TAR, Donaldson GC, Paul EA, Bestall JC, Jefferies DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 1998;157:1418-1422.
14. Aaron SD, Donaldson GC, Whitmore GA, Hurst JR, Ramsay T, Wedzicha JA. Time course and pattern of COPD exacerbation onset. *Thorax*. 2012;67(3):238-243.
15. Donaldson GC, Goldring JJ, Wedzicha JA. Influence of season on exacerbation characteristics in patients with COPD. *Chest*. 2012;141(1):94-100.
16. Langsetmo L, Platt RW, Ernst P, Bourbeau J. Underreporting exacerbation of chronic obstructive pulmonary disease in a longitudinal cohort. *Am J Respir Crit Care Med*. 2008;177(4):396-401.
17. Celli BR, Thomas NE, Anderson JA, et al. Effect of pharmacotherapy on rate of decline of lung function in chronic obstructive pulmonary disease: results from the TORCH study. *Am J Respir Crit Care Med*. 2008;178(4):332-338.
18. Leidy MK, Sexton CC, Jones PW, et al. Measuring respiratory symptoms in clinical trials of COPD: reliability and validity of a daily diary. *Thorax*. 2014; doi:10.1136/thoraxjnl-2013-204428 .
19. Calverley PM, Anderson JA, Celli BR, et al. Salmeterol and fluticasone proprionate and survival in chronic obstructive pulmonary disease. *New Engl J Med*. 2007;356:775-789.
20. Tashkin DP, Celli B, Senn S, et al. A 4-year trial of tiotropium in chronic obstructive pulmonary disease. *N Engl J Med*. 2008;359(15):1543-1554.
21. Casaburi R, Celli B, Crapo, et al. The COPD Biomarker Qualification Consortium (CBQC). *COPD*. 2013;10(3):367-377.
22. van der Molen T, Willemse BWM, Schokker S, ten Hacken NHT, Postma DS, Juniper EF. Development, validity and responsiveness of the Clinical COPD Questionnaire. Health Qual Life Outcomes. 2003;1:13. doi:10.1186/1477-7525-1-13.
23. Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N. Development and first validation of the COPD Assessment Test. *Eur Respir J*. 2009;34(3):648-654.
24. Jones PW, Brusselle G, Dal Negro RW, et al. Properties of the COPD assessment test in a cross-sectional European study. *Eur Respit J*. 2011;38(1):29-35.
25. Kwon N, Amin M, Hui DS, et al. Validity of the Chronic Obstructive Pulmonary Disease Assessment Test (CAT) translated into local languages for Asian patients. *Chest*. 2013;143:703-710.
26. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease. GOLD Web site. http://www.goldcopd.org/guidelines-global-strategy-for-diagnosis-management.html Published January 2014. Accessed March 3, 2014.
27. Bjorner JB, Chang C-H, Thissen D, Reeve BB. Developing tailored instruments: item banking and computerized adaptive assessment. *Qual Life Res*. 2007;16 (suppl 1):95-108.
28. Norweg A, Ni P, Garshick E, O'Connor G, Wilke K, Jette AM. A multidimensional computer adaptive approach to dyspnea assessment. *Arch Phys Med Rehabil*. 2011;92(10):1561-1569. doi: 10.1016/j.apmr.2011.05.020.
29. Stucky BD, Edelen MO, Sherbourne CD, Eberhart NK, Lara M. Developing an item bank and short forms that assess the impact of asthma on quality of life. *Respir Med*. 2014;108(2):252-263.