Eur Respir J 2008; 31: 238–240
DOI: 10.1183/09031936.00163107

EDITORIAL

COPD dashboard: about official Documents, Authors, Science, Health status and BOARDs

H.J. Schünemann

Readers of this issue of the *European Respiratory Journal* (*ERJ*) should review the important document that it features, which is a result of a continuing and successful tradition of joint Task Forces of the European Respiratory Society (ERS) and the American Thoracic Society (ATS) [1]. The document will guide investigators and others making decisions about the use and application of outcome measures in order to reduce the burden of chronic obstructive pulmonary disease (COPD) in individual patients and groups of patients alike. The purpose of this editorial is to describe some general aspects related to document development for joint Task Forces and to comment on the content of the document by CAZZOLA *et al.* [1]. While the opinion, expertise and experiences described here have led me to apply for the position of documents editor for the ATS and perhaps influenced, at least in part, my selection for this position, they should not be equated with official ATS policy. This editorial reflects an individual's view on documents including those that guide clinical practice.

Given the standing of the ATS and the ERS in the field of respiratory, critical care and sleep medicine, the document featured in this issue of the *ERJ* is an approved statement of the Societies and, thus, becomes official and authoritative by default. As a result of the responsibility that the Societies therefore carry, but also just by nature of their stature, the principle aim of these official documents must be to provide evidence-based information to support evidence-based decision-making in healthcare. In the current climate of increased scrutiny of clinical practice guidelines and other official statements, what are the best strategies to ensure that documents are indeed based on the best available evidence?

Evidence can be of very low-to-high quality, which is determined by the risk of bias and the directness in relation to the question at hand. Whatever the underlying quality of evidence is, evaluation of evidence requires judgments and consensus by those evaluating it. The term ''best'' evidence refers to the (highest quality) evidence that is available to inform the answer to the question. This may be evidence from unsystematic observations, evidence from high-quality randomised trials or lie between these extremes. The key information that readers require is to lay out the basis for the answer transparently and adhere to predefined methods that minimise the risk of bias in both conduct and interpretation of science. Methods for evaluating information regarding the likelihood of bias and avoiding bias during document development are available and offer guidance for societies in the near future [2, 3]. In applying these methods, we expect a learning process to occur, as the science of document development and implementation will develop further just like any other scientific area in healthcare. Just as in evidence-based clinical practice, the more critical and challenging part is to ensure that the information about appropriate methods (those that are currently available and those that will develop in the future) is put into practice during the development of official documents. This process, which is one of systematic methods, support and supervision, is complicated but we have no choice if we want to maintain the high quality of official documents. Unique skills, currently not taught in research or clinical fellowships, which supplement the generated research evidence, are required to produce high-quality documents. In addition, those underwriting the content, such as professional societies, including the ERS and the ATS, must be armed against criticism that could arise as a consequence of publication of official documents. Indeed, the ATS and the ERS have increasingly recognised the need to do so. The vision of John Heffner, the immediate Past-President of the ATS, helped move the ATS to devote greater attention to this issue, in a similar manner to other professional organisations. The inauguration of the ATS Documents Development and Implementation Committee and the position of the ERS Guidelines Director are consequences of his vision. This vision includes respiratory societies having the responsibility of becoming better at working with methodologists whose careers focus on research methodology and document development. I would add to this vision that organised curricula providing the necessary skills of guideline and other document development are required. The necessary procedures that integrate these methodologists with the societies' mission, approved years ago by the societies, include peer review and editorial review before approval by bodies of the society. The Executive Committee of the ERS and the Board of Directors of the ATS are the determined final approval bodies of the Societies and must appropriately guard the editorial work of documents editors and alike in order to protect established procedures.

In the spirit of harmonising efforts and processes, as well as increasing efficiency, the ATS and the ERS have developed

STATEMENT OF INTEREST: A statement of interest for H.J. Schünemann can be found at www.erj.ersjournals.com/misc/statements.shtml

CORRESPONDENCE: H.J. Schünemann, Dept of Epidemiology, CLARITY Research Group, Italian ?National Cancer Institute Regina Elena, Via Elio Chianesi 53, 00144 Rome, Italy. Fax: 1 4433390565. E-mail: schuneh@mcmaster.ca

plans that will further facilitate the work of joint Task Forces. Joint Task Forces develop documents that range from clinical practice guidelines to policy statements. Authors of these documents will benefit from a process that assigns parallel tasks and responsibility to individuals and committees within the societies, and includes a single common review prior to approval by the societies bodies. This type of standardisation will, again, support the work of Task Forces. Task Forces are of utmost importance to the societies because they are built on the academic contributions of members of the societies who invest time and effort in the development of official documents. The societies are aware of the, sometimes immense, contributions required from its members during document development. Reputation and the other indirect and direct (academic) credit that Task Force members receive for authorship on documents are, therefore, important outcomes for Task Force members.

Having emphasised the importance of following due process, the document by CAZZOLA *et al.* [1] does what is of immense importance in medical science (and other sciences): (improving) standardisation. I recently overheard a great health services researcher, who dedicated the last decade of his career to the advancement of the UK National Health Service, tell a brief story, which I paraphrase here: ''Hospitals use different norm values for laboratory parameters such as sodium levels. How can one deal with this incomprehensible problem? The suggestion of bringing scientists together to discuss norm values would be a mistake. The scientists would not agree; the public would laugh and revolt. The better strategy to attain standardisation is to make an administrative, authoritative health system decision based on the best evidence.'' This statement describes the very goal of our documents. Standardisation will improve comparability and move science and healthcare forward more efficiently. Transparent judgments and implementation of what is considered best evidence are required and present the first step. The procedures require that individuals and organisations minimise the influence of personal and academic goals, and other personal interests, to serve the public best. The work of CAZZOLA *et al.* [1] provides long-needed progress, perhaps an inauguration in the field of COPD, in that regard. However, there are some points that are worth discussing and that relate to the use of best evidence in the document.

APPLYING BEST EVIDENCE IN OFFICIAL SOCIETY DOCUMENTS

An extremely accomplished group of authors compiled the cited evidence and prepared this statement. In all likelihood, the members of this Task Force will indeed know about all of the available evidence. Therefore, it is possible that none of the important studies remained undiscovered, despite the lack of more formal search methods and full descriptions of how the evidence was compiled. The authors also attempted to use predefined criteria to uniformly assess the outcomes. However, a direct link from what the authors reviewed to how they addressed these issues of validity is not as transparently described as readers may desire. Revisions of this document should therefore include systematic reviews of the outcome measures described here and a systematic assessment of their validity after defining clear criteria for inclusion and exclusion of the available evidence. The evaluation of surrogate outcomes using validated or common sense tools are steps that could also be considered [4].

THREE EXAMPLES WORTH MORE METHODOLOGICAL EFFORT

Investigators also need to tackle other areas. Much debate has focused on pulmonary function measures, *e.g.* forced expiratory volume in one second (FEV_1), as an outcome in clinical trials. One criterion required for their use is knowledge of what constitutes an important difference that leads decision makers to alter management in the absence of important harms from alternative management. This concept is known as ''interpretability'' and is often expressed by minimal important difference (MID), previously referred to as minimally clinical important difference (MCID) [5]. We dropped the ''C'' as it emphasised focus on the clinician when the real decision makers are patients or informed proxies [5]. Rarely, if ever, should decision makers base treatment decisions on FEV_1 or changes alone and, thus, it is doubtful that we can adopt the concept of the MID for pulmonary function measures. This concern becomes even clearer when we consider pulmonary function measures, such as diffusion capacity of the lung for carbon monoxide, which have no demonstrated property as surrogate for patient-important outcomes in clinical trials. Clarification and separation of these concepts is required. Interpretability for pulmonary function measures has different conceptual underpinnings compared with that for health-related quality of life (HRQoL) assessment, a field that fostered the development of the concept of the MID.

The use of combined end-points, such as the BODE (body mass index (BMI), degree of airway obstruction, dyspnoea, airway obstruction) index, also requires careful consideration before adopting it as an outcome measures in clinical trials. The reasons for caution concerning end-points have been described in work evaluating their validity and suggesting strategies as to dealing with their use [6, 7]. Only end-points that show similar effect sizes, are of similar importance to patients and occur with similar frequency should be combined [6, 7]. The constituents of the BODE index, while useful for prognostic purposes, have not demonstrated this property, yet. What are clinicians to tell their patients if their FEV_1 deteriorates but their BMI and dyspnoea improve, leaving no overall effect on the combined end-point? Communication of these treatment effects to patients would be extremely complicated: should we focus on the combined end-point and how would we express it? Or should we focus on the individual end-points?

CAZZOLA *et al.* [1] describe health status as a marker of HRQoL. Thus, it appears the true measure of importance would be HRQoL. One could ask why researchers are not focusing on instruments that have been described as measuring HRQoL, to omit the indirectness of a marker or surrogate. It is an impossible task to develop instruments that apply to all patients in the same way (for example, how would a question about ''walking a block'' apply to a wheelchair-bound COPD patient?), a requirement often put forward for health status instruments. Both instruments described as measuring health status and those developed to measure HRQoL perform well. While statistical methods can help us analyse complex health status and HRQoL data, the target is to use simple validated questionnaires that most patients can relate to. In the spirit of

standardisation, an important question is whether the separation of the concepts of health status and HRQoL is solid, necessary and justified in the context of COPD.

QUALITY OF EVIDENCE FROM RANDOMISED CONTROLLED TRIALS DOES NOT APPLY TO PROGNOSTIC OR RISK FACTORS

Confusion exists among investigators concerning whether prognostic markers can be evaluated using common evidence grading systems. Quality of evidence from randomised controlled trials does not apply to prognostic markers. Therefore, for evaluation of prognostic markers, commonly used evidence-grading systems are not applicable. Consider the use of FEV1. The study design that would be required to reflect what has been described as ''evidence level A'' (from randomised trials) would be a randomisation of patients to different levels of FEV1, then a comparison of the outcomes of groups who are randomised to different levels of FEV1. To assign a label of evidence from randomised trials, one would then need to show that these groups have different outcomes. Such studies do not exist because in biomedical research, risk factors or prognostic factors (that are caused by picking up or quitting smoking to alter FEV1, altering BMI or other prognostic factors) are not randomised. All that is available is that these indices were used in randomised trials as outcomes, but this is a fundamentally different concept. The key message is that randomisation does not directly apply to predictors, prognostic factors or risk factors. What can be used is evidence grading for modification of these risk or prognostic factors (*i.e.* management or interventions that lead to change in the predictors, such as interventions to modify FEV1 or BMI). However, the evidence grading will relate to the intervention and not to the marker. Thus, the document by CAZZOLA *et al.* [1] appropriately avoids such classification for prognostic markers, and other documents, having wrongly applied this concept, should follow.

FOCUS ON PATIENT-IMPORTANT OUTCOMES

The document describes a large array of outcomes and appropriately acknowledges that there is a lack of data suggesting validity (*i.e.* that they measure what they are intended to measure) for many of the markers as an outcome in clinical trials. Validation will include a proven track record of the surrogate being important to patients or showing an indisputable relationship of the surrogate to a patient-important outcome. Nevertheless, surrogates will remain what they intend to be: ''surrogates'' for what matters to patients. Thus, the primary focus should remain the outcomes that matter to patients: mortality, morbidity and HRQoL. Investigators should accept this and will have to design and power trials appropriately.

SUMMARY

The document by CAZZOLA *et al.* [1] provides excellent and long needed progress for guiding and standardising use of appropriate outcome measures in clinical trials of patients with chronic obstructive pulmonary disease. As the science evolves in the area of chronic obstructive pulmonary disease outcomes, users should look out for important developments in the field, including revisions of the document. The input from scientists who focus on evidence appraisal and knowledge translation, and thorough review by societies in collaboration with Task Forces aims to further improve documents. Thus, we need to acknowledge and recognise that we are trying to ''hit a moving target'' where the ''target'' is the best evidence and ''hitting'' refers to the decision by societies whether the document is appropriate for official adoption. The primary outcomes for clinical trials are those that are patient-important: mortality, morbidity and quality of life. Trials must focus on their use. Surrogates can help to draw conclusions about these patient-important outcomes but do require careful validation before they are used in clinical trials. Finally, we must realise that there are many and different skills required for sitting at the chronic obstructive pulmonary disease and other dashboards. Skill training in the identification and assessment of evidence as well as document development are urgently needed to compliment the expertise of others. As always, open collaboration focusing on the best evidence will move us in the right direction.

REFERENCES

1 Cazzola M, MacNee W, Martinez FJ, *et al.* Outcomes for COPD pharmacological trials: from lung function to biomarkers. *Eur Respir J* 2008; 31: 416–468.
2 Schunemann HJ, Jaeschke R, Cook DJ, *et al.* An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am J Respir Crit Care Med* 2006; 174: 605–614.
3 Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med* 1997; 126: 376–380.
4 Lassere MN, Johnson KR, Boers M, *et al.* Definitions and validation criteria for biomarkers and surrogate endpoints: development and testing of a quantitative hierarchical levels of evidence schema. *J Rheumatol* 2007; 34: 607–615.
5 Schunemann HJ, Guyatt GH. Commentary – goodbye M(C)ID! Hello MID, where do you come from? *Health Serv Res* 2005; 40: 593–597.
6 Ferreira-Gonzalez I, Busse JW, Heels-Ansdell D, *et al.* Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. *BMJ* 2007; 334: 786.
7 Montori VM, Permanyer-Miralda G, Ferreira-Gonzalez I, *et al.* Validity of composite end points in clinical trials. *BMJ* 2005; 330: 594–596.