Development and Evaluation of an Automated, Home-Based, Electronic Questionnaire for Detecting COPD Exacerbations

Francisco de B. Velazquez-Peña1, Daniel Sanchez-Morillo2*, Mario Crespo-Miguel3, Sonia Astorga-Moreno2, MJ Santi-Cano1, MA Fernandez-Granero2 and Antonio Leon-Jimenez4

1Department of Nursing and Physiotherapy, University of Cádiz, Cádiz, Spain
2Biomedical Engineering and Telemedicine Research Group, University of Cádiz, Cádiz, Spain
3Department of Philology, University of Cádiz, Cádiz, Spain
4Pulmonology and Allergy Unit, Puerta del Mar University Hospital, Cádiz, Spain

Submitted April 2015. Accepted for publication September 2015.

ABSTRACT
Collaboration between patients and their medical and technical experts enabled the development of an automated questionnaire for the early detection of COPD exacerbations (AQCE). The questionnaire consisted of fourteen questions and was implemented on a computer system for use by patients at home in an un-supervised environment. Psychometric evaluation was conducted after a 6-month field trial. Fifty-two patients were involved in the development of the questionnaire. Reproducibility was studied using 19 patients (ICC = 0.94). Sixteen out of the 19 subjects started the 6 month-field trial with the computer application. Cronbach’s alpha of 0.81 was achieved. In the concurrent validity analysis, a correlation of 0.80 (p = 0.002) with the CCQ was reported. The results suggest that AQCE is a valid and reliable questionnaire, showing that an automated home-based electronic questionnaire may enable early detection of exacerbations of COPD.

Keywords: COPD, diary, early detection, exacerbation, prediction, questionnaire, remote monitoring, symptom, telehealth, telemonitoring

1. INTRODUCTION
Chronic Obstructive Pulmonary Disease (COPD) affects a large number of people and is associated with significant morbidity, disability and mortality [1–3]. The natural history of the disease is characterized by periods of worsening symptoms, which vary in severity and frequency. Such periods, defined as exacerbations [2], lead to major health care costs and high morbidity [4], causing an adverse impact on quality of life related to health [5], a more rapid decline in lung function [6], increased use of health services [7], and decreased survival [2].

*Corresponding author: Daniel Sánchez Morillo, Escuela Superior de Ingeniería, Avda. de la Universidad de Cádiz, 10, Puerto Real, Cádiz, Spain. CP 11519. Phone: (34) 956-015709. E-mail: daniel.morillo@uca.es. Other authors: borja.vp@hotmail.es; mario.crespo@uca.es; sonia.asmo@gmail.com; mariajose.santi@uca.es; ma.fernandez@uca.es; anleji@hotmail.es.
Daily recording of symptoms have led to the observation that more than 50% of exacerbations are not reported to physicians (non-reported exacerbations) [8]. Reducing the frequency, severity, and duration of exacerbations is of great interest to patients and healthcare providers since early treatment improves exacerbation recovery and reduces the risk of hospitalization [9].

Currently, methods for detecting, assessing, and recording acute exacerbations of COPD (AECOPD) are being developed [9,10]. In this regards, two main criteria for defining AECOPD are being used. Event-based definition is defined as doctor visits to emergency rooms or unplanned visits to a physician for acute worsening of symptoms [8,11]. Events are therefore indirect measures of exacerbations, whose validity is not entirely clear, owing to international discrepancies in health care coverage and the fact that patients do not seek medical care for 50–70% of events of exacerbations [12]. The second definition is symptom-based and it evaluates AECOPD through daily symptom cards which the patient completes providing information about the progression of symptoms [10,13]. The cards are direct measures of the signs or symptoms of AECOPD. The latest approach allows for standardization of the data, reduction in recall bias, identifying the events, and determining the resolution based on a pre-defined algorithm score [10].

Most daily symptom cards and other quality of life questionnaires that have been used to detect exacerbations are paper-based questionnaires [8]. With improvements in technology, online automated card and questionnaire completion could help to detect AECOPD earlier. This work aims to develop and validate a home-based automated questionnaire for the early detection of COPD exacerbations (AQCE). This multi-disciplinary research challenge was addressed by building on and updating knowledge in specific domains through a cross-disciplinary collaboration. AQCE has been especially designed for domiciliary use on a portable device, allowing the standardized recording of symptoms and enabling the use of statistics and artificial intelligence techniques to produce early detection of AECOPD.

2. METHODS
2.1. Questionnaire Development
Patients with a diagnosis of COPD and a history of at least one hospitalization in the previous year were selected to participate in the development of an AQCE. Inclusion criteria were the capacity to interpret a Likert scale. There were no age restrictions to participating in the study. All patients provided informed consent and agreed to participate voluntarily. The study was approved by the Research Ethics Committee of Puerta del Mar University Hospital, Cádiz, Spain.

The process for the questionnaire development was twofold. In the first iterative phase, a literature review on AECOPD was carried out, and prodromal symptoms were identified and characterized. Physicians then developed the items and a computational linguist performed the review of all aspects of language structure and appropriate length of the items, resulting in a questionnaire of 15 questions. Eight out the 15 questions were dichotomous, 6 were Likert scale type, and the last one involved a
coordination test. This coordination test consisted of a bull’s eye that the patient had to hit using his forefinger. After a first draft, the terms and interpretability by patients were checked through semi-structured interviews with 52 patients diagnosed with COPD. Keywords and expressions used by patients were considered to improve both the way the questions were presented and the possible answers. In the final draft, linguistic and medical structures were checked for consistency. The questionnaire included questions related to general health, cough, sputum, dyspnea, sleep conditions, cold symptoms, and lung sounds.

In the second phase, the questionnaire was installed on a laptop equipped with an 8.9-inch touch screen and voice recognition system, which was previously evaluated for effectiveness in collecting verbal commands. The font size, sound, and color scheme were selected according to basic usability criteria. The questionnaire was to be completed by pressing an answer on the touch screen or by voice commands.

To assess the degree of comprehension (interpretability) of the questionnaire installed on the computer, 35 out of the initial 52 participants indicated, using a Likert scale, their level of understanding of each item (1 for a low level of understanding and 5 for a very good level of understanding). After that, they were asked to explain the perceived meaning of each item. Two independent observers rated the patients’ responses, assessing the degree of comprehension of each item and detailing if the explanations provided by the patients matched the true meaning of the question. Finally, 3 ratings quantifying interpretability were obtained (one from the patient and two from the physicians).

2.2. Electronic Implementation with Multimodal Interface

The developed questionnaire was implemented on a portable computer for home use. The device was equipped with a multimodal interface, HSDPA (high-speed down-link packet access) connection, touch screen, speaker, and microphone.

Requirements of potential users (25 users from elderly homes) were considered during design and development stages, in an iterative process performed using a user-centered design methodology and aimed to optimize usefulness and usability of the solution. Pulmonologists, a computational linguist, and usability and software engineers planned the whole development process and gathered users’ requirements. According to these requirements, software engineers developed prototypes iteratively using feedback on performance and usability. Content layout, visual aspects, and interface control evolved during the iterations. After three incremental versions, a final prototype that maximized user satisfaction effectiveness and efficiency was released. In this final prototype, two computer-patient interaction modes were implemented: a visual-touch mode, allowing the patient to answer the questions touching the screen, and automatic speech recognition mode, which empowered the user to give commands or responses verbally. During the operation of the device, different dialogue-based screens were displayed to users. Each screen was designed to represent a simple conversation with a threefold structure, namely: a) presentation of the information (i.e. question) by the system b) options selectable to be answered and c) confirmation of the
selected option. This conversational structure was presented to users either graphically and/or orally. After completing the test, responses were sent to the remote central server using a XML file. For further details on the design and implementation procedures, refer to reference [14].

2.3. Questionnaire Validation
To validate the questionnaire, a prospective follow-up study using a group of adult patients diagnosed with COPD was carried out. The coordination test was excluded from the validation analysis and psychometric analysis was performed using the resulting 14-item questionnaire. Two clustering alternatives arose:

- Grouping of symptoms according to the criteria of Anthonisen et al. [15], who defined two groups of symptoms: major (dyspnea, sputum volume, sputum purulence) and minor (cough, fever, lung sounds, and cold symptoms). A third group was added. We named it “complementary symptoms” and included subjective health, sleep problems, and the use of rescue inhalers.
- Grouping of domains in 3 clusters, namely: a) dyspnea, which included questions related to shortness of breath on exertion, lung sounds, and breathing problems during sleep; b) cough and sputum, with questions about the frequency of cough and sputum volume and purulence, and c) general symptoms with questions about subjective health status and flu-like symptoms.

2.3.1. Subjects
For the validation study, patients were recruited from the Pulmonology, Allergy, and Thoracic Surgery Unit in Puerta del Mar University Hospital. Inclusion criteria were former smoker patients who had attempted smoking cessation at least 6 months prior to the study. Participants were in a stable phase of the disease, aged over 60, with FEV1/FVC < 0.7, and good cognitive and motor skills. Subjects had a history of prolonged exposure to smoking and cumulative cigarette consumption greater than 20 packs-years. Some of these patients participated in the development phase. Patients with significant associated respiratory disease and/or significant cardiovascular comorbidity were excluded. All patients provided their informed consent form and agreed to participate voluntarily. The study was approved by the Research Ethics Committee of Puerta del Mar University Hospital.

2.3.2. Procedure
Figure 1 details the process and factors involved in the development and evaluation of the questionnaire. Nineteen patients completed two personal interviews carried out in a 15- to 20-day interval during the test-retest reliability study for the reproducibility analysis. Clinical stability was assessed by spirometry, medical history, and two surveys of quality of life: the Clinical COPD Questionnaire (CCQ) [16,17] and COPD and Asthma Sleep Impact Scale (CASIS) [18,19].

Sixteen out of 19 patients were provided with a portable computer device to complete the validation analysis. The study period with the computer application was 6 months. During this time, patients continued with treatment and clinical follow-up as usual. One patient dropped out halfway through the study, but was considered for the validation analysis.
Figure 1. Stages of development and validation for the automated questionnaire for detecting COPD exacerbations (AQCE).
2.3.3. Statistical Analysis
The Statistical Package SPSS v19 was used. Mean and standard deviation for quantitative variables and absolute and relative frequencies for each category’s qualitative variable were used to describe the sample. For the questionnaire development, the agreement between the interpretability evaluation scores obtained from patients and observers was estimated by the proportion of overall agreement. For the questionnaire validation, validity and reliability were analyzed. Both total score and score for the two cluster systems (domains and symptoms) were used.

Reliability analysis was performed by a study of reproducibility via intra-class correlation coefficient (ICC) with the data obtained for the test-retest in the initial group of 19 patients.

The group of 16 patients who had been provided with the device was considered for the rest of the validation study. The internal consistency test was performed on these patients applying Cronbach’s alpha using data from the first day. Concurrent validity was assessed by comparing the results obtained from the AQCE, CCQ, and CASIS questionnaires, which were all completed on the same day. Pearson or Spearman correlation was used depending on the nature of the data. For content validity, floor and ceiling effects were tested [20].

3. RESULTS
3.1. Questionnaire Development and Implementation
Thirty-five patients completed the interpretability study. Demographic and clinical characteristics of participants are shown in Table 1(a). Interpretability degree of eleven of the items was acceptable (values 4 to 5) in 100% of patients. In the remaining three items, one for cold symptoms (item 9), and two for problems or symptoms during sleep (items 12 and 13), interpretability degree was acceptable to 97.1% of patients. The average percentage of agreement found in the interpretability study between patients and observer 1 was 0.95, 0.97 between patients and observer 2, and 0.96 between the two observers.

As a consequence, the 15 items of the initial draft were retained for the subsequent validation study. A structured questionnaire that included these 15 items was installed on the portable device. The iterative design process, which started with an initial conceptual design planned by the design and usability experts from the technical and clinical requirements [14], led to a final system which was satisfactory from the user’s point of view.

3.2. Questionnaire Validation
The characteristics of patients who used the computer device are described in Table 1(b). Internal consistency of AQCE is shown in Table 2, where total and cumulative scores for the two grouping criteria previously described are detailed. Cronbach’s alpha ranged from 0.60 in the minor symptoms domain, to 0.85 in the cough and sputum domain. For the total score, Cronbach’s alpha was 0.81.

Intraclass correlation coefficients (ICC) are shown in Table 2. The floor (lower score) and ceiling (higher score) effects for the AQCE questionnaire that were studied
in the validation sample are also shown in Table 2. The dyspnea domain had the greatest floor effect (20.1%) while floor was lower than 7% in the rest of cases.

The AQCE scores showed a normal distribution. Agreement between the results obtained by the AQCE and the results obtained for the same population by CCQ and CASIS was measured. Pearson correlation coefficients, along with p values parenthesized, are shown in Table 3. Significant correlations between the CCQ questionnaire and its domains with the AQCE total score and the major symptoms domain were obtained. For the CASIS questionnaire, a good correlation was estimated with the dyspnea, general symptoms, and complementary symptoms domains.
The AQCE questionnaire was proposed as an initiative to address the need to develop and validate an instrument for early detection of AECOPD. The process of development and validation was performed entirely from a computer with multimodal interaction for

### Table 2. Reliability results for the AQCE questionnaire (a). Floor and ceiling effects for validation study of the AQCE questionnaire (n = 2194) (b)

<table>
<thead>
<tr>
<th>Domain</th>
<th>Cronbach’s alpha</th>
<th>ICC (n − 19) CI 95%</th>
<th>p value</th>
<th>Ceiling %</th>
<th>Floor %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>0.81 0.94 0.85-0.97 &lt;0.001</td>
<td>0 0 7 0.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyspnea</td>
<td>0.75 0.91 0.76-0.96 &lt;0.001</td>
<td>28 1.3 440 20.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cough and Sputum</td>
<td>0.85 0.87 0.67-0.95 &lt;0.001</td>
<td>0 0 142 6.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Symptoms</td>
<td>0.69 0.36 -0.63-0.75 0.173</td>
<td>0 0 15 0.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major Symptoms</td>
<td>0.63 0.91 0.77-0.96 &lt;0.001</td>
<td>0 0 45 2.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minor Symptoms</td>
<td>0.60 0.67 0.17-0.87 0.01</td>
<td>1 0.1 121 5.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complementary Symptoms</td>
<td>0.80 0.77 -0.42-0.91 0.001</td>
<td>7 0.3 16 0.7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ICC: Intraclass correlation coefficient. CI: confidence interval.

### Table 3. Concurrent validity of the AQCE questionnaire (n = 16)

<table>
<thead>
<tr>
<th>Domain</th>
<th>CCQ Symptoms</th>
<th>CCQ Functional State</th>
<th>CCQ Mental State</th>
<th>CASIS</th>
<th>FEV₁</th>
<th>FVC</th>
<th>FVC/FEV₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>0.8 (-0.002)</td>
<td>0.72 (-0.009)</td>
<td>0.7 (-0.012)</td>
<td>0.7 (-0.012)</td>
<td>0.4 (0.193)</td>
<td>-0.63 (0.028)</td>
<td>-0.51 (0.092)</td>
</tr>
<tr>
<td>Dyspnea</td>
<td>0.52 (0.052)</td>
<td>0.32 (-0.307)</td>
<td>0.56 (-0.307)</td>
<td>0.45 (0.142)</td>
<td>0.61 (0.034)</td>
<td>-0.39 (0.214)</td>
<td>-0.24 (0.461)</td>
</tr>
<tr>
<td>Cough and Sputum</td>
<td>0.58 (-0.086)</td>
<td>0.67 (-0.307)</td>
<td>0.45 (-0.060)</td>
<td>0.4 (-0.142)</td>
<td>0.56 (-0.034)</td>
<td>-0.2 (-0.214)</td>
<td>-0.12 (-0.461)</td>
</tr>
<tr>
<td>Sputum</td>
<td>0.295 (-0.048)</td>
<td>0.24 (-0.016)</td>
<td>0.19 (-0.144)</td>
<td>0.38 (-0.195)</td>
<td>0.79 (-0.061)</td>
<td>-0.58 (-0.540)</td>
<td>-0.61 (-0.715)</td>
</tr>
<tr>
<td>General</td>
<td>0.351 (-0.351)</td>
<td>0.460 (-0.506)</td>
<td>0.553 (-0.553)</td>
<td>0.223 (-0.002)</td>
<td>0.04 (-0.048)</td>
<td>-0.036 (-0.048)</td>
<td>-0.278 (-0.036)</td>
</tr>
</tbody>
</table>

Pearson’s coefficients (p-values in parentheses) are shown. CCQ: Clinical COPD Questionnaire; CASIS: COPD and Asthma Sleep Impact Scale. FVC: Forced vital capacity. FEV₁: Forced expiratory volume in one second.

### 4. DISCUSSION

The AQCE questionnaire was proposed as an initiative to address the need to develop and validate an instrument for early detection of AECOPD. The process of development and validation was performed entirely from a computer with multimodal interaction for
This research was conducted by a cross-disciplinary team that included researchers from the following disciplines: medical, biomedical engineering, information and communication technologies, and social science.

The methodology followed for the development of AQCE was similar to the one used in other COPD questionnaires [16,18]. However, the difference in this study was the inclusion of an evaluation of both patient and physician interpretability of the items used in the questionnaire.

After an interpretability study, the questionnaire was installed on a portable computer and evaluated for 6 months in a study on patients who had a history of AECOPD in the previous year. AQCE showed a high internal consistency for the total score and the domains of cough and sputum and complementary symptoms. The results confirm that the AQCE questionnaire is capable of assessing the different aspects of the disease without overlap.

The total score and the scores of domains were repeatable over time, with an ICC above 0.90 in most cases except for the general symptoms domain (0.36) and the Anthonisen minor symptoms cluster (0.67) [21]. The former result can be attributed to the presence of comorbidities in a patient who presented difficulties in answering the question on night sleep accurately.

Higher correlations between AQCE and quality of life questionnaires were obtained for the total score and the Anthonisen major symptoms domain and the CCQ different domains. The CASIS questionnaire was better correlated with the general symptoms and complementary symptoms domains in AQCE. Regarding the correlation with spirometry data, no significant correlation was found. This was to be expected since spirometry is a weak health indicator [22–24]. Excluding dyspnea, the floor and ceiling effects indicated good validity.

While the evidence to support the effectiveness of integrating tele-monitoring into existing clinical services for patients with COPD is limited [25], there is a need for developing accurate algorithms for detecting AECOPD based on symptom scores. These algorithms rely on a questionnaire’s ability to measure what is expected. In this regards, the lack of a specific validated questionnaire for early detection of an AECOPD may explain previous poor performances on telehealth studies.

Explorative studies to evaluate the discriminative properties of questionnaires in detecting exacerbations are currently underway. CCQ, with significant evaluative properties in measuring health status, is being used as a tool to detect AECOPD in some of these studies. However, its reliability is still not clear [26]. Despite the EXACT study reporting an effective method of evaluating exacerbation severity, its ability for accurately detecting an AECOPD remains unclear [27]. A possible cause may be that the EXACT study was designed for evaluating the effect of treatment on COPD exacerbations, not for early detection of exacerbations.

A number of differences exist between our study and the EXACT study. First, AQCE was developed for use on a computer. To our knowledge, this is the first questionnaire developed and validated for this platform. This approach enables data mining and machine learning methods as well as the opportunity for easier translation in to clinical practice. Secondly, AQCE is designed for early detection of an AECOPD. Recently, AQCE has been used together with machine-learning techniques to predict both
symptom and event-based exacerbations of COPD. Symptom-based episodes were predicted with an average margin of 4.8 days prior to onset and a detection accuracy of 80.5%, with a low rate of false positives [28]. Event-based exacerbations were detected an average of 4.5 days prior to the event with an accuracy over 93% [29]. Finally, AQCE was designed for older patients. Low compliance with daily symptom reporting has been identified as the main reason for failure in previous tele-monitoring studies [30–31]. To improve compliance and accessibility, older patients were involved in the early design and evaluation stages, including the decision to use a multimodal computer interface [14]. While the EXACT and AQCE have shown good results for evaluating severity and detecting exacerbations, respectively, both tools could complement each other.

There were a number of limitations associated with this study. The study involved a relatively small and homogeneous cohort. The study was restricted to an analysis of the usefulness of AQCE for patients at the more severe end of the COPD spectrum. A larger sample size would ensure a better generalizability of the results. Another limitation was that the reliability and validation findings were based on data solely from Spain. Use of standardized techniques would enable cultural and linguistic validity across languages.

Further testing is envisaged before ACQE can be accepted as a fully validated tool for use in clinical practice. The next step should include testing the tool in a real environment setting. A composite measure that combines physiological measurements with symptoms to predict deterioration would address this aspect [25]. Measurements that could be included in future tests are respiratory rate, respiratory sounds, and general activity levels. These parameters are known to change with COPD exacerbations. These tests should continue to assess the tool as a tele-health intervention with objective physiological measurements included to improve early detection and monitoring recovery from exacerbations.

5. CONCLUSIONS
This study developed and validated an AQCE by taking a multidisciplinary approach. The results of this study show that AQCE is easily understood by patients and displays a good item interpretation agreement between patients and doctors.

The questionnaire appears to be a valid and reliable method for early detection of exacerbations in patients with COPD. Further testing is required before the questionnaire can be used in clinical practice.

ACKNOWLEDGEMENTS
This work was supported in part by the Ambient Assisted Living (AAL) E.U. Joint Programme, by grants from Ministerio de Educación y Ciencia of Spain and Instituto de Salud Carlos III under Projects PI08/90946 and PI08/90947.

CONFLICT OF INTEREST
The authors indicated no potential conflicts of interest.
REFERENCES


Development and Evaluation of an Automated, Home-Based, Electronic Questionnaire for Detecting COPD Exacerbations


