The Leicester Cough Monitor: preliminary validation of an automated cough detection system in chronic cough

S.S. Birring*, T. Fleming*, S. Matos#, A.A. Raj¶, D.H. Evans# and I.D. Pavord¶

ABSTRACT: Chronic cough is a common condition that presents to both primary and secondary care. Assessment and management are hampered by the absence of well-validated outcome measures. The present study comprises the validation of the Leicester Cough Monitor (LCM), an automated sound-based ambulatory cough monitor.

Cough frequency was measured with the LCM and compared with coughs and other sounds counted manually over 2 h of a 6-h recording by two observers in nine patients with chronic cough in order to determine the sensitivity and specificity of the LCM. Automated cough frequency was also compared with manual counts from one observer in 15 patients with chronic cough and eight healthy subjects. All subjects underwent 6-h recordings. A subgroup consisting of six control and five patients with stable chronic cough underwent repeat automated measurements $\geqslant 3$ months apart. A further 50 patients with chronic cough underwent 24-h automated cough monitoring.

The LCM had a sensitivity and specificity of 91 and 99%, respectively, for detecting cough and a false-positive rate of 2.5 events·h^{-1}. Mean$\pm$SEM automated cough counts·patient·h^{-1} was 48$\pm$9 in patients with chronic cough and 2$\pm$1 in the control group (mean difference 46 counts·patient·h^{-1}; 95% confidence interval (CI) 20–71). The automated cough counts were repeatable (intra-subject SD 11.4 coughs·patient·h^{-1}; intra-class correlation coefficient 0.9). The cough frequency in patients undergoing 24-h automated monitoring was 19 coughs·patient·h^{-1}; daytime (08:00– 22:00 h) cough frequency was significantly greater than overnight cough frequency (25 versus 10 coughs·patient·h^{-1}; mean difference 15 coughs·patient·h^{-1}, 95% CI 8–22).

The Leicester Cough Monitor is a valid and reliable tool that can be used to assess 24-h cough frequency in patients with cough. It should be a useful tool to assess patients with cough in clinical trials and longitudinal studies.

KEYWORDS: Chronic cough, cough counts, cough frequency, cough monitor, Leicester Cough Monitor

Chronic cough is a common reason for referral to respiratory physicians. The assessment of patients with chronic cough is often based on the anatomical diagnostic protocol, which is a systematic evaluation based on the understanding that most cases are due to disease of the upper respiratory tract where cough receptors are most plentiful [1]. The most common conditions implicated in causing chronic cough in nonsmokers are asthma, gastro-oesophageal reflux and rhinitis, or a combination of these [2].

The identification of an important contribution by the aforementioned conditions is largely based on the evaluation of treatment trials [3]. However, there are few well-validated outcome measures to assess cough severity and treatment efficacy. Cough visual analogue scores, diary score cards, quality-of-life questionnaires, cough-reflex sensitivity measurement and cough monitors have been proposed as potential tools to assess cough severity [2]. The subjective nature of symptom scores and quality-of-life questionnaires [4, 5] and the poor specificity of cough-reflex sensitivity measurement [6] to identify patients with chronic cough have led to a renewed interest in the development of automated ambulatory cough monitors [7–15]. BIRRING et al. [7] have previously shown that there are marked differences in cough frequency between patients with chronic cough and healthy subjects and that these measurements are repeatable, and have suggested that cough frequency

measurement is potentially useful in the assessment of patients with chronic cough. Currently available cough monitors are limited by difficulty in achieving unrestricted ambulatory measurement in the patients' own environment, an inability to perform 24-h recording and a lack of automated cough detection systems. The aim of the present study was to develop an automated cough monitor (the Leicester Cough Monitor (LCM)) capable of recording cough sounds for 24 h. The present study shows the validation of the LCM and preliminary findings of 6- and 24-h recordings in patients with chronic cough.

METHODS
Subjects
A total of 15 consecutive patients with an isolated chronic cough (>3 weeks duration) were recruited from a specialised cough clinic. The clinic receives referrals from primary and secondary care largely confined to a population of 970,000 within Leicestershire, UK. The causes of cough in patients with chronic cough were: cough variant asthma (n=4); gastro-oesophageal reflux (n=3); eosinophilic bronchitis (n=1); idiopathic (n=3); post-viral (n=1); bronchiectasis (n=1); chronic obstructive pulmonary disease (n=1); and chronic bronchitis (n=1). Nine out of these patients were randomly selected for the first stage of validation (cough variant asthma (n=4), eosinophilic bronchitis (n=1), idiopathic (n=2), post-viral (n=1), bronchiectasis (n=1)) and all patients were included in the second validation stage. Investigations were carried out according to a standardised algorithm [16]. The protocol for investigation and treatment and the criteria for accepting diagnosis were as previously described [16]. Eight controls were recruited from healthy volunteers responding to local advertising. Control subjects were asymptomatic, non-smokers and had normal spirometry and a concentration of methacholine required to cause a 20% decrease in forced expiratory volume in one second (FEV_1) of $>16 mg \cdot mL^{-1}$. No patients had received corticosteroids or other specific treatment for the condition causing cough for $\geqslant 6$ weeks prior to the study. A randomly selected subgroup of six control subjects and five patients with a stable chronic cough and stable treatment requirements (three with cough variant asthma, one with gastro-oesophageal reflux-associated cough and one with idiopathic chronic cough) participated in cough frequency repeatability studies 3–6 months after the first, at the same time of day in order to avoid possible bias from diurnal variations. A total of 50 further consecutive patients with chronic cough were recruited in order to evaluate 24-h recordings with the LCM (idiopathic cough (n=26), asthma (n=8), eosinophilic bronchitis (n=2), rhinitis (n=2), sarcoidosis (n=2), gastro-oesophageal reflux (n=3), bronchiectasis (n=2), chronic obstructive pulmonary disease (n=2), enlarged tonsils (n=2) and obstructive sleep apnoea (n=1)). All subjects gave full informed written consent to participate. The protocol for the present study was approved by the Leicestershire Research Ethics Committee.

Cough monitor
The LCM (fig. 1) is a digital ambulatory cough monitor that records sound continuously from a free-field microphone necklace (Sennheiser MKE 2-5; Sennheiser electronic GmbH & Co. KG, Woedemark, Germany) onto a digital sound recorder

(dimensions $26.7 \times 87 \times 32$ mm; iRiver iFP-799; iRiver Europe GmbH, Eschborn, Germany) at a sampling frequency of 16 kHz and with an encoding bit rate of 64 $kbit \cdot s^{-1}$. The cough monitor was attached at 09:00 h in all subjects and returned 6–24 h later. Subjects were told that the LCM was a new investigative tool being developed to assess the nature of the cough and were encouraged to resume their normal activity in their usual environment. When the recording was complete, data stored on the recorder was downloaded onto a computer, where it was analysed by an automated cough detection algorithm (the Leicester Cough Algorithm).

A general outline of the Leicester Cough Algorithm has been described previously [8, 17]. Briefly, the detection algorithm is based on Hidden Markov Models, a statistical method that can be used to characterise the spectral properties of a time-varying pattern. The cough detection algorithm was implemented based on the keyword-spotting approach, as defined in speech recognition, in which the objective is to detect the occurrence of a particular set of keywords in a sequence of continuous speech. Continuous ambulatory recordings in patients with chronic cough were used to train statistical models of the characteristics of cough sounds and of the audio background. During the detection process, the recorded audio signal was divided into contiguous 10-s segments to be analysed by the Hidden Markov Models-based algorithm. Each 10-s audio segment was recognised by the detection algorithm as a sequence of variable-length audio sections, each in turn classified either as background audio or as a possible cough sound, depending on its statistics [8]. A second algorithm phase then used brief operator input to facilitate the automated algorithm in order to eliminate sounds that
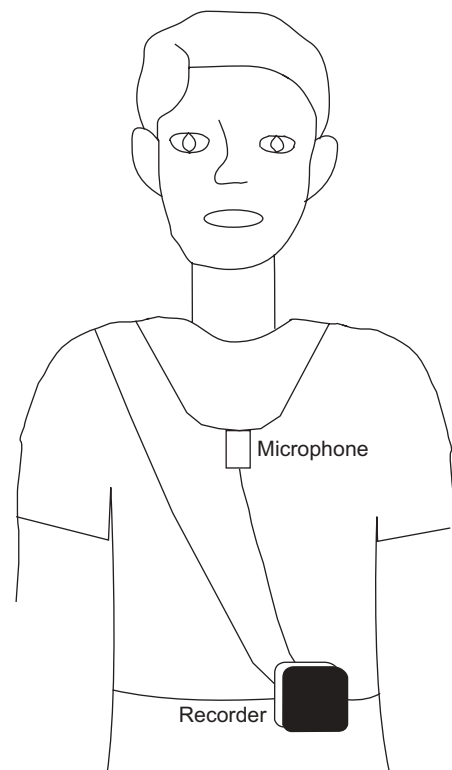

FIGURE 1. The Leicester Cough Monitor.

might have been wrongly classified as cough events in the first phase. For this, the operator is asked to classify, as cough or otherwise, a small fraction of the sounds detected in the first phase as possible cough sounds (the second phase takes 5 min to carry out for a 24-h recording and ~50 sounds are classified). The information is then used to create statistical models that are adapted to the characteristics of the cough sounds for that particular recording and the remaining sounds that were not shown to the operator are classified using these models.

Cough was defined as a characteristic explosive sound. The algorithm identifies coughs as single events whether they occur as isolated events or in a cluster (*i.e.* attempts were made to determine how many coughs occurred in paroxysms).

Validation
Stage 1

The first stage of validation compared automated cough counts against those identified by manual sound analysis of the first and fourth recorded hours of nine randomly selected patients with chronic cough. Manual analysis of sound recordings consisted of three blinded observer counts (observer one twice and observer two once) and cough or noncough sounds were positively identified when all three counts were in agreement based on sound and visual inspection of the acoustic trace. Each cough sound was identified separately, whether it occured singularly or in a cluster or epoch of coughs. Intra- and interobserver variability in cough counts was established from the two blinded analyses performed by observer one and by comparing the mean of observer one's counts for these periods with counts obtained by observer two. The recordings were analysed twice using the Leicester Cough Algorithm to assess intra-recording repeatability. In order to classify non-cough sounds and determine whether particular sounds were wrongly classified by the automated system, observer one listened to all 6 h of the nine patients' recordings and classified all recognisable sounds. The results of this analysis were compared with the automated classification.

Stage 2

The second stage of validation was extended to all recordings and compared automated cough counts against coughs identified manually by observer one, who analysed the entire 6-h recording. A further 50 patients with chronic cough underwent 24-h automated cough frequency measurement. Automated cough frequency was compared for repeatability studies.

Analysis

Subject characteristics were described using descriptive statistics and expressed as mean±SEM for parametric data and median for non-parametric data. Cough frequency was expressed as individual coughs·patient·h^{-1} for the duration of the recording. The validity of the LCM was presented as sensitivity, specificity and false-positive rate of the automated algorithm for detecting coughs as measured by observer manual analysis. Intra- and interobserver variability of manual cough counts and repeatability data was assessed as intra-class correlation coefficients and intra-subject SD.

RESULTS
The subject characteristics are shown in table 1.

Validation stage 1 (First and fourth recorded hour)
Mean cough counts were 39 coughs·patient^{-1}·h^{-1} by automated analysis compared with 43 coughs·patient^{-1}·h^{-1} by manual analysis (mean difference -4 coughs·patient^{-1}·h^{-1}, 95% confidence interval (CI) -6–13; p=0.4). The intra- and interobserver intra-class correlation coefficients for manual analysis of sound recordings (between observers one and two) were 0.99 and 0.98, respectively (both p<0.001). The intra- and interobserver (*i.e.* between mean of observer one and observer two) intra-class correlation coefficients for manual analysis of sound recordings were 0.99 and 0.98, respectively (both p<0.001). The intra-class correlation coefficient between automated and manual observer counts was 0.9 (p<0.001; fig. 2a). The accuracy of manual and automated cough counts appeared similar in recordings containing paroxysms and those with isolated coughs. The Leicester Cough Algorithm had a sensitivity and a specificity of 91% and 99%, respectively, for detecting cough sounds and a median false positive rate of 2.5 events·patient^{-1}·h^{-1} against the gold standard of coughs detected manually by observer one twice and observer two once. There was no evidence that any particular sound was more likely to be classified as a false positive (fig. 3).

Validation stage 2 (6-h recordings)
Mean±SEM automated cough counts were 48±9 coughs·patient^{-1}·h^{-1} in patients with chronic cough and 2±1 coughs·patient^{-1}·h^{-1} in control subjects (mean difference 46 coughs·patient^{-1}·h^{-1}, 95% CI 20–71 coughs·patient^{-1}·h^{-1}; p<0.001; fig. 4). There were no significant differences in cough frequency between diagnostic groups. The cough analysis took 2 h to complete, comprising 5 min for data download, 105 min for computer automated analysis (an operator was not required to be present during phase 1) and 10 min for operator input (phase 2) and printing results.

The intra-class correlation coefficient between automated and observer counts was 0.93 (p<0.001; fig. 2b). The LCM had sensitivity and specificity of 86 and 99%, respectively, for detecting cough sounds and a median false-positive rate of 1.0 events·patient^{-1}·h^{-1}.

The automated cough counts were repeatable in the 11 subjects who underwent repeatability testing (intra-subject SD 11.4 coughs·patient^{-1}·h^{-1}, intra-class correlation coefficient 0.9; fig. 5).

TABLE 1	Subject characteristics	
	Control	**Chronic cough**
Subjects (male) n	8 (0)	15 (5)
Age yrs	48±3	55±4
Cough duration yrs		5±2
FEV1 % pred	91±5	89±7
FEV1/FVC %	79±1	75±4

Data are expressed as mean±SEM, unless otherwise stated; FEV1: forced expiratory volume in one second; pred: predicted; FVC: forced vital capacity.
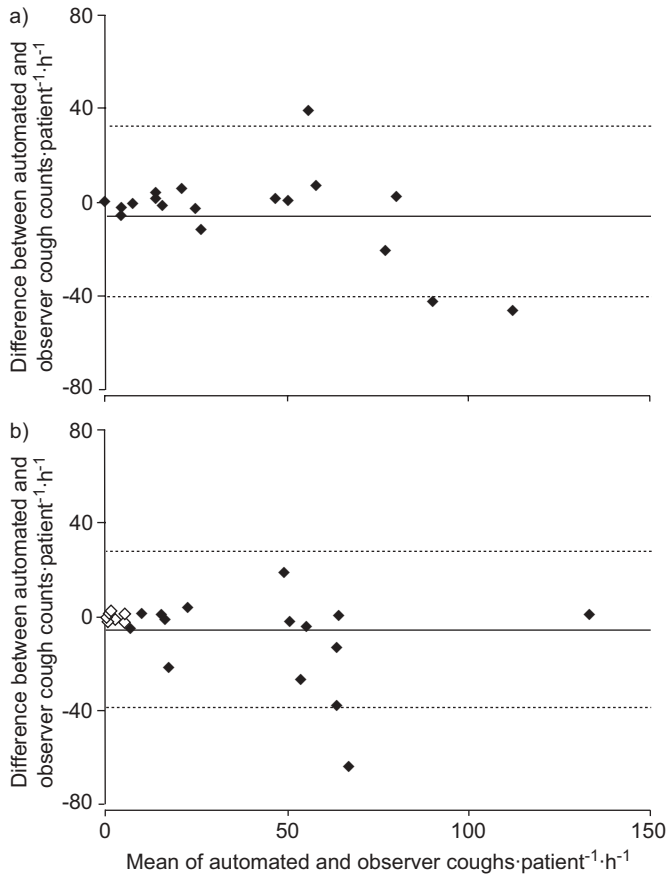
FIGURE 2. Bland-Altman plot of automated versus manual observer cough counts·patient⁻¹·h⁻¹. a) Validation stage 1 at which first and fourth recorded hour (n=9) were analysed. Each hour is depicted individually. b) Validation stage 2 at which 6-h recordings (n=23) were analysed in their entirety. The complete cough detection algorithm (phases 1 and 2) was tested in each validation stage. –––: mean difference; ······: 95% limits of agreement (2× intra-subject SD); ◇: control subjects; ◆: chronic cough patients.

The cough frequency in patients undergoing 24-h monitoring was 19 coughs·patient⁻¹·h⁻¹; daytime (08:00–22:00 h) cough frequency was significantly greater than overnight cough frequency (25 versus 10 coughs·patient⁻¹·h⁻¹; mean difference 15 coughs·patient⁻¹·h⁻¹, 95% CI 8–22 coughs·patient⁻¹·h⁻¹; p<0.001; fig. 6).

DISCUSSION

The LCM is a lightweight 24-h automated ambulatory cough monitor that is easy to use and measures cough in the subjects' own environment. The present study has shown that it is a valid and reliable tool for objectively measuring cough frequency. The high sensitivity and specificity for the detection of cough sounds is comparable to other routine diagnostic clinical tools and superior to that reported for other more cumbersome cough detection systems. Preliminary data is presented in the current study showing that the cough frequency measured with the LCM is repeatable over ⩾3 months, a period relevant to the duration of treatment trials that form an important part of the assessment of patients with chronic cough. Repeatability was marginally better than that of recordings analysed manually [7]. The present data
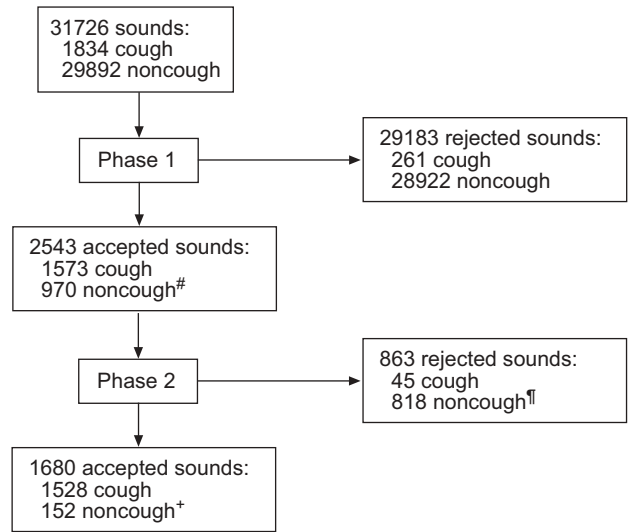

FIGURE 3. False positives characterised manually that were detected by the automated Leicester Cough Monitor in nine 6-h recordings of patients with chronic cough. Automated analysis of sound recording is performed during phase 1 and automated analysis following operator input is performed during phase 2. The sensitivity for cough detection is slightly lower than that from the gold standard validation stage 1, since the comparator was a 6-h manual counting by observer one only. #: noncough sounds were: speech (n=351), impulsive noise (n=222) throat clearing (n=119), environmental noise (n=111), laugh (n=80), other person coughing (n=31), incomplete coughs (n=22), child talking/shouting (n=21), sneeze (n=6), telephone ringing (n=3), burp (n=2) and dog barking (n=2). ¶: noncough sounds were: speech (n=293), impulsive noise (n=214), throat clearing (n=96), environmental noise (n=87), laugh (n=67), other person coughing (n=15), incomplete coughs (n=18), child talking/shouting (n=17), sneeze (n=5), telephone ringing (n=3), burp (n=2) and dog barking (n=1). +: noncough sounds were: speech (n=58), impulsive noise (n=8), throat clearing (n=23), environmental noise (n=24), laugh (n=13), other person coughing (n=16), incomplete coughs (n=4), child talking/shouting (n=4), sneeze (n=1) and dog barking (n=1).

suggest that the LCM might be a particularly useful outcome measure in assessing patients with cough and measuring the response to therapy in the clinic and in clinical trials.

A limitation of the present study is that evaluation of cough frequency was based on 6-h daytime cough recordings owing
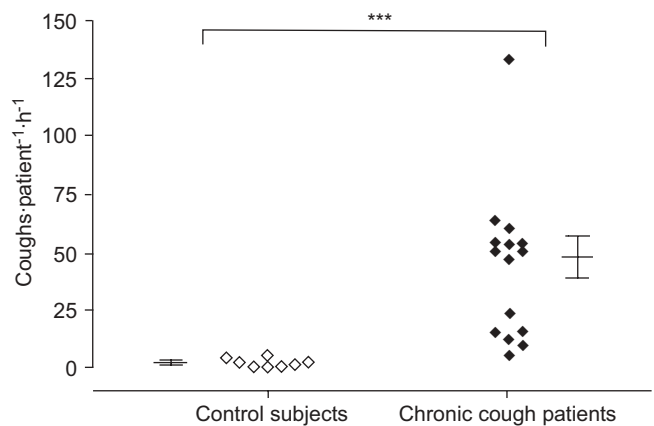

FIGURE 4. Mean±SEM automated cough counts·patient⁻¹·h⁻¹ in control subjects and chronic cough patients (6-h recordings). ***: p<0.001.
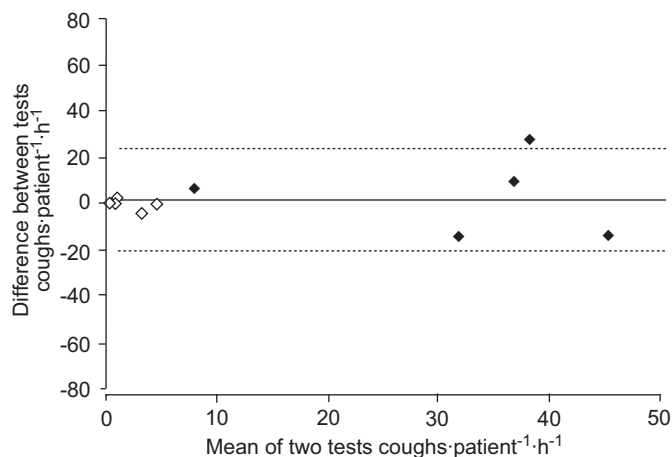
FIGURE 5. Bland-Altman plot of automated cough counts·patient^{-1}·h^{-1} repeated over 3–6 months in chronic cough patients and control subjects. ——: mean difference; ·······: 95% limits of agreement (2 × intra-subject SD). ◇: control subjects; ♦: chronic cough patients.

FIGURE 6. 24-h ambulatory, automated cough frequency recordings in 50 chronic cough patients. Data are presented as mean+SEM.

to limited battery life (6–8 h) at the inception of the study. Advances in battery technology, since then, have allowed the extension of recordings to ⩾24 h. The automated system has allowed these recordings to be analysed relatively quickly and accurately and should facilitate the investigation of potential diurnal variations in cough frequency and the effects of aggravators of potential cough, such as environmental pollution, cigarette smoking and gastro-oesophageal reflux. In the present study, a range of sounds including speech, throat clearing and environmental noises caused false-positive detected coughs. There was little evidence that any of these sounds caused particular difficulties with detection, nor did cough paroxysms appear to present problems for accurate manual and automated cough counts. However, greater experience with the monitor may identify sounds or cough paroxysms that present particular problems for the algorithm to classify and allow further refinement of the algorithm. The present study involved small numbers of subjects and it will be important to study a larger population of control subjects and patients with well-defined respiratory disease, before and after treatment, and in different environments, to fully validate the cough monitor. The present preliminary work suggests that such a study will be feasible.

Cough frequency was stable through the day and was significantly reduced overnight compared with daytimes in accord with previous data, suggesting a diurnal variation in cough frequency [5, 9, 13]. Further work is required to determine the validity and the short- and longer-term repeatability of 24-h cough recordings.

A limitation of the present study is that only 2 h per recording were used to compare automated cough counts with those obtained from manual counting for the validation study. Manual cough counting is very time consuming and laborious so only the first and fourth hours of each recording were manually analysed for consistency. Each recording was manually counted three times in order to obtain a more robust measure of the true cough frequency. The LCM had a high sensitivi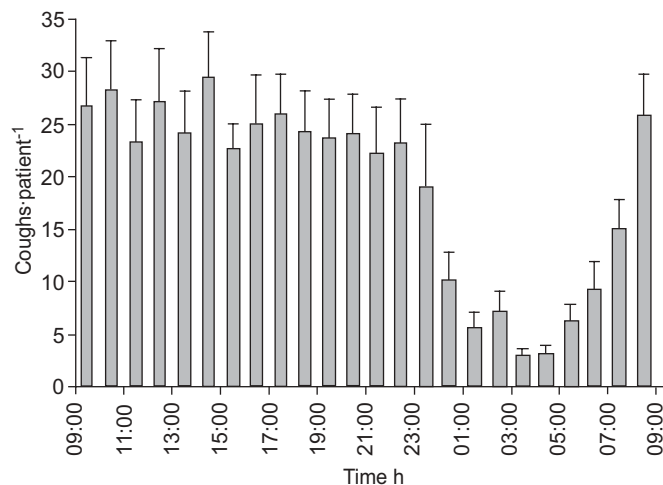ty and specificity for detecting cough against this gold standard. This was confirmed in the second part of the validation study where cough counts derived from automated analysis of 6-h recordings were compared with cough counts derived from a single manual observer. The sensitivity of the cough algorithm was slightly lower with 6-h recordings compared with 2-h recordings. This is most probably because of the more robust measure of true cough frequency used for the 2-h recordings, compared with single observer manual counts used for the 6-h recordings.

A potential criticism of cough counts derived from audio recordings is that they might not accurately reflect the true cough rate since it is not possible to visualise the act of coughing. However, a recent study [18] compared manual cough counts from audio with video recordings and found them to be very similar. That study concluded that manual cough counts from audio recordings should be regarded as the gold standard to validate cough monitors since audio recordings had superior sound quality to that from video recordings.

The LCM quantifies cough frequency as single episodes of cough rather than epochs or clusters of coughs and cough seconds (seconds containing cough), as used by others [12]. The present authors believe that single cough episodes are a more meaningful measure and are easier to interpret by physicians and patients. Cough frequency, rather than intensity, was measured since cough events are less influenced by microphone position and muffling of sounds by covering the mouth during the act of coughing. Furthermore, cough intensity determined with sound analysis lacks responsiveness compared with cough frequency in clinical trials of antitussive drugs [10]. Cough intensity determined by other parameters such as airflow or chest wall movement is less practical for routine clinical measurement. The LCM was validated in subjects with chronic cough due to a wide range of conditions so it reliably detects coughs with differing characteristics.

One of the challenges of developing cough monitors in the past has been differentiating cough sounds from throat clearing, sneezing, speech and other cough-like sounds. The Leicester Cough Monitor differentiates cough from other sounds reliably

as indicated by the high sensitivity and particularly high specificity for detecting cough. The Leicester Cough Monitor represents a potential advance over existing cough monitors in that it is portable outside a controlled environment, it does not require measurement of abdominal electromyography and it can be set to record for 24 h. The 24-h sound recording is discarded after the automated analysis process, thereby ensuring patient privacy. The Leicester Cough Monitor could be used to validate the presence of excessive cough, to assess cough severity and to objectively evaluate response to therapy. Further studies are required in order to assess the use of the Leicester Cough Monitor in clinical practice.

REFERENCES
1 Irwin RS, Madison JM. The diagnosis and treatment of cough. *N Engl J Med* 2000; 343: 1715–1721.
2 Irwin RS. Assessing cough severity and efficacy of therapy in clinical research: ACCP evidence-based clinical practice guidelines. *Chest* 2006; 129: Suppl. 1, 232S–237S.
3 Morice AH, Fontana GA, Sovijarvi AR, *et al.* The diagnosis and management of chronic cough. *Eur Respir J* 2004; 24: 481–492.
4 Birring SS, Prudon B, Carr AJ, Singh SJ, Morgan MD, Pavord ID. Development of a symptom specific health status measure for patients with chronic cough: Leicester Cough Questionnaire (LCQ). *Thorax* 2003; 58: 339–343.
5 Raj AA, Birring SS. Clinical assessment of chronic cough severity. *Pulm Pharmacol Ther* 2007; 20: 334–337.
6 Prudon B, Birring SS, Vara DD, Hall AP, Thompson JP, Pavord ID. Cough and glottic-stop reflex sensitivity in health and disease. *Chest* 2005; 127: 550–557.
7 Birring SS, Matos S, Patel RB, Prudon B, Evans DH, Pavord ID. Cough frequency, cough sensitivity and health status in patients with chronic cough. *Respir Med* 2006; 100: 1105–1109.
8 Matos S, Birring SS, Pavord ID, Evans DH. Detection of cough signals in continuous audio recordings using hidden Markov models. *IEEE Trans Biomed Eng* 2006; 53: 1078–1083.
9 Coyle MA, Keenan DB, Henderson LS, *et al.* Evaluation of an ambulatory system for the quantification of cough frequency in patients with chronic obstructive pulmonary disease. *Cough* 2005; 1: 3.
10 Pavesi L, Subburaj S, Porter-Shaw K. Application and validation of a computerized cough acquisition system for objective monitoring of acute cough: a meta-analysis. *Chest* 2001; 120: 1121–1128.
11 Chang AB, Newman RG, Phelan PD, Robertson CF. A new use for an old Holter monitor: an ambulatory cough meter. *Eur Respir J* 1997; 10: 1637–1639.
12 Smith J, Owen E, Earis J, Woodcock A. Effect of codeine on objective measurement of cough in chronic obstructive pulmonary disease. *J Allergy Clin Immunol* 2006; 117: 831–835.
13 Hsu JY, Stone RA, Logan-Sinclair RB, Worsdell M, Busst CM, Chung KF. Coughing frequency in patients with persistent cough: assessment using a 24 hour ambulatory recorder. *Eur Respir J* 1994; 7: 1246–1253.
14 Barry SJ, Dane AD, Morice AH, Walmsley AD. The automatic recognition and counting of cough. *Cough* 2006; 2: 8.
15 Paul IM, Wai K, Jewell SJ, Shaffer ML, Varadan VV. Evaluation of a new self-contained, ambulatory, objective cough monitor. *Cough* 2006; 2: 7.
16 Brightling CE, Ward R, Goh KL, Wardlaw AJ, Pavord ID. Eosinophilic bronchitis is an important cause of chronic cough. *Am J Respir Crit Care Med* 1999; 160: 406–410.
17 Matos S, Birring SS, Pavord ID, Evans DH. An automated system for 24-hour monitoring of cough frequency: the Leicester cough monitor. *IEEE Trans Biomed Eng* 2007; 54: 1472–1479.
18 Smith JA, Earis JE, Woodcock AA. Establishing a gold standard for manual cough counting: video *versus* digital audio recordings. *Cough* 2006; 2: 6.