

Research Article

# Continuous Sound Collection Using Smartphones and Machine Learning to Measure Cough

Lucia Kvapilova<sup>a</sup> Vladimir Boza<sup>a</sup> Peter Dubec<sup>a</sup> Martin Majernik<sup>a</sup>  
Jan Bogar<sup>a</sup> Jamileh Jamison<sup>a</sup> Jennifer C. Goldsack<sup>a</sup> Duncan J. Kimmel<sup>b</sup>  
Daniel R. Karlin<sup>a, c</sup>

<sup>a</sup>HealthMode Inc., San Francisco, CA, USA; <sup>b</sup>Albert Einstein College of Medicine, Bronx, NY, USA; <sup>c</sup>Tufts University School of Medicine, Boston, MA, USA

## Keywords

Algorithms · App-based digital biomarkers · Cough · Digital measurement · Machine learning · Respiratory

## Abstract

**Background:** Despite the efforts of research groups to develop and implement at least partial automation, cough counting remains impractical. Analysis of 24-h cough frequency is an established regulatory endpoint which, if addressed in an automated manner, has the potential to ease cough symptom evaluation over multiple 24-h periods in a patient-centric way, supporting the development of novel treatments for chronic cough, an unmet clinical need. **Objectives:** In light of recent technological advancements, we propose a system based on the use of smartphones for objective continuous sound collection, suitable for automated cough detection and analysis. Two capabilities were identified as necessary for naturalistic cough assessment: (1) recording sound in a continuous manner (sound collection), and (2) detection of coughs from the recorded sound (cough detection). **Methods:** This work did not involve any human subject testing or trials. For sound collection, we designed, built, and verified technical parameters of a smartphone application for sound collection. Our cough detection work describes the development of a mathematical model for sound analysis and cough identification. Performance of the model was compared to previously published results of commercially available solutions and to human raters. The compared solutions use the following methods to automatically or semi-automatically assess cough: 24-h sound recording with an ambulatory device with multiple microphones, automatic silence removal, and manual recording review for cough count. **Results:** Sound collection: the application demonstrated the

Daniel R. Karlin, MD, MA  
HealthMode Inc.  
548 Market Street 57448  
San Francisco, CA 94104-5401 (USA)  
E-Mail dan@healthmode.co

ability to continuously record sounds using the phone's internal microphone; the technical verification informed the configuration of the technical and user experience parameters. Cough detection: our cough recognition sensitivity to cough as determined by human listeners was 90 at 99.5% specificity preset and 75 at 99.9% specificity preset for a dataset created from publicly available data. **Conclusions:** Sound collection: the application reliably collects sound data and uploads them securely to a remote server for subsequent analysis; the developed sound data collection application is a critical first step toward future incorporation in clinical trials. Cough detection: initial experiments with cough detection techniques yielded encouraging results for application to patient-collected data from future studies.

© 2019 The Author(s)  
Published by S. Karger AG, Basel

## Introduction

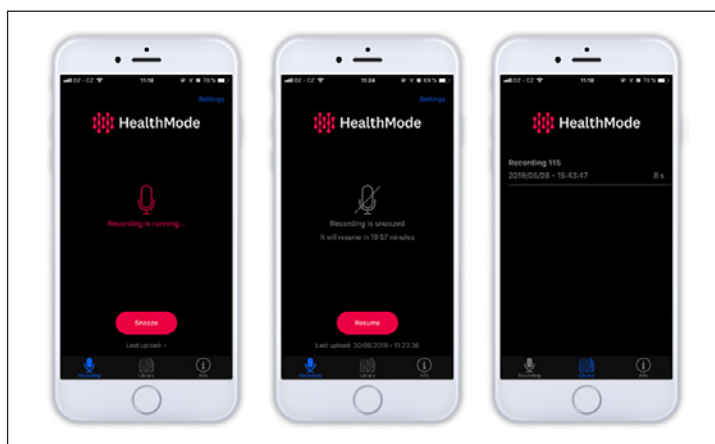
Cough is a common and meaningful symptom in many respiratory diseases, with both characteristic sounds and movements. Monitoring and measurement of cough in clinical trials and routine care typically relies upon patients self-reporting the frequency, severity, and quality of their own coughs [1]. However, individuals' self-perception is unavoidably influenced by perception bias (such as over- and under-perception of respiratory symptoms) [2]. This presents an opportunity for the development of objective cough monitoring systems. Research on cough monitoring to date indicates that the audio signal is the most meaningful data point [3]. As cough is most often paroxysmal and almost always episodic, continuous audio signal collection provides important information when compared to more episodic monitoring. For this reason, 24-h continuous cough count has become the exclusive endpoint for the approval of drugs that target cough [4]. However, the manual analysis of the coughs present in the sample is a task that can take nearly as long as the recording period itself.

Over the past decade, the biomedical research industry has seen development of automated cough recording and cough-counting technologies [5]. These include the Leicester Cough Monitor [6], Hull Automated Cough Counter [7], and VitaloJak [8, 9]. Common features of these cough monitors are: (1) the presence of one or more microphones, or a combination of microphones, and (2) recording capabilities delivered by a microphone, or a combination of various microphones. The microphones used for sound acquisition include those that are: internal to the recording device; externally mounted (often lapel-style); body-attached; or some combination of these.

In currently available cough recording and/or counting solutions, data acquired by the microphone are processed by algorithms that are trained to recognize and discard silence and, in the best case, to identify likely coughs [7]. While useful, these methods still require significant human input. Despite validation efforts by manufacturers, commercial solutions for automated cough monitoring remain dependent on manual counting, largely as a result of the constraints of their technological foundations.

The possibilities of real-time patient monitoring are expanding. This includes the detection of changes in an individual's health status outside of the clinic. The increasing ubiquity of smartphones is of particular interest, as they are natively equipped with the capability for precise sensing, on-board analysis, and connectivity. Sensors embedded in smartphones now offer researchers the hope for providing a more holistic and objective view of sickness and health. We propose a system based on the use of smartphones and their internal microphones for objective and continuous audio data collection suitable for automated cough analysis.

**Fig. 1.** HealthMode Cough application screens. The first screen shows the recording running in the application. On the second screen, the recording is snoozed by the user. The third screen shows the library of recorded samples. After the samples are uploaded, the library clears the entries. The Info button on each screen is a placeholder for an additional screen (not yet in use) containing study or other information for users in the future.



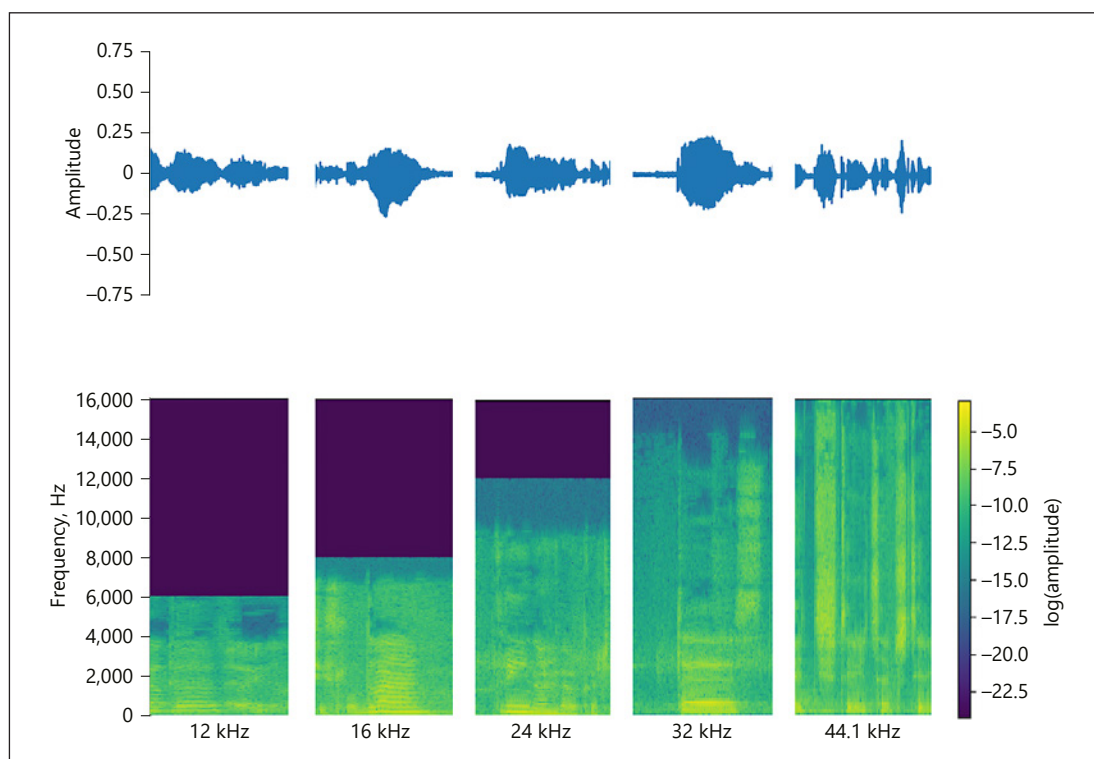
## Materials and Methods

### *Sound Collection: Software Development*

We designed and developed the audio collection application “HealthMode Cough” to collect audio data in a continuous manner via smartphone. The application was programmed using XCode developer tools (XCode version 10.1 for macOS 10.13.6+) and runs on an iOS platform (iOS 11+). The essential feature of this application, and the one for which it is optimized, is continuous sound recording. The recordings are captured locally and subsequently sent to the secure cloud server. The recording and uploading schedule is set to 5-min-long recordings and 30-min-long upload intervals. All data are encrypted at rest and during transfer, and only authorized research personnel can access the recordings.

Other features of the application include a screen with instructions for the participant, information about the system, details of data protection, a snooze button allowing the participant to pause the recording for a specified interval, and notifications when low battery is detected that encourages the user to recharge the phone battery. Figure 1 presents the main HealthMode Cough application screens in more detail.

Next, we tested parameters of the recording to optimize for captured sound quality, data yield, and battery consumption. For optimization of the recording quality, we created 5-s recordings with 5 sampling frequency presets: 12, 16, 24, 32, and 44.1 kHz. We evaluated the recordings based on two criteria – recording file size to inform the amount of data collected and transferred, and sound quality necessary to detect and classify cough sounds evaluated via literature review of preceding cough sound classification experiments. Based on the respective file size of the recordings, we calculated the final amount of data that would be collected and transmitted over a 2-week monitoring period. The sound quality was evaluated based on previous experiments with cough sounds classification with the use of spectrograms which adopted sampling frequencies that produced sounds within the human hearing range of 20 Hz to 20 kHz. These experiments showed results in diagnosing the presence of excess mucus in recordings resampled to 8 kHz, with the information-containing frequency being under 4 kHz [10], in recordings resampled to 20 kHz with vocal sound observed between 4 and 10 kHz [11], or classifying respiratory diseases at frequencies under 1.7 kHz [12]. It has been established that the frequencies of cough are widely spread up to 20 kHz, but the information-containing signal appears in lower frequencies [11]. Figure 2 shows spectrograms of our five test recordings. For the sampling frequency presets 12, 16, 24, 32, and 44.1 kHz, we calculated the amount of data that a 2-week uninterrupted recording would yield.



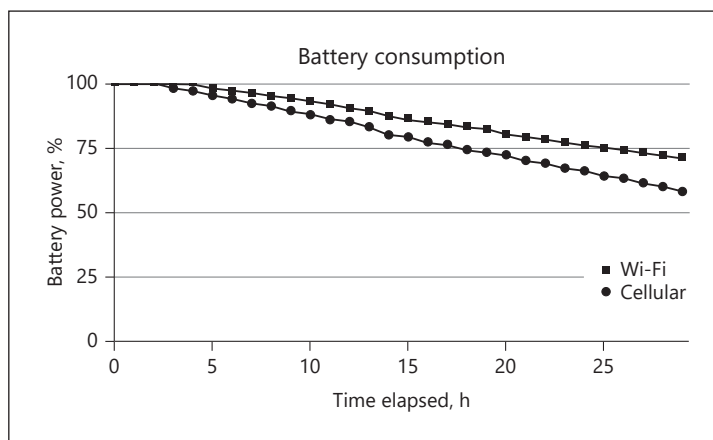
**Fig. 2.** Waveforms and their respective spectrograms of five test recordings at 12, 16, 24, 32, and 44.1 kHz.

For testing of the battery consumption, we ran the application uninterrupted on two iPhone 8 devices with 256 GB of onboard storage. The devices remained connected either exclusively to Wi-Fi or exclusively to cellular networks for the duration of the experiment. We recorded charge levels at 30-min intervals. Both phones started recording at the same time with the battery fully charged, continuing to record and upload the sound data for 29 h.

#### *Cough Detection: Dataset Creation*

To train the cough recognition model, we used publicly available data from internet sources: we downloaded 41 YouTube videos and 5 cough examples from the SoundSnap website (links to these datasets are available in online suppl. material 1; for all online suppl. material, see [www.karger.com/doi/10.1159/000504666](http://www.karger.com/doi/10.1159/000504666)). Combined, these tracks contained cough sounds from 20 different people; 7 male and 13 female (the gender was identified by 2 independent data annotators). These sample videos and their respective audio tracks contained only cough sounds without any additional background noises. As such, annotators were able to separate the cough sounds by simply looking for a loud noise after a period of silence to yield a dataset of approximately 1,500 coughs.

To train the recognition model, we collected background noises, including recordings from loud streets, open offices, a train station, a crowded market, and a bar, which contained various noises as well as human speech. These sounds were collected from publicly available videos on YouTube (online suppl. material 2). We used sound mixing techniques to prepare a large dataset of 1-s samples with coughs in various environments by recombining samples of cough and background noises in various permutations, enabling us to train our models to recognize cough even in the presence of background noise.



**Fig. 3.** Comparison of battery depletion using a Wi-Fi connection and cellular connection over the course of 29 h.

Both cough sounds and background noises were split into non-overlapping training and test sample datasets (each cough-originator from the recordings was assigned to either the training or test set) and mixing was performed in each set separately. A small portion of the training set was used as a validation set for hyperparameter tuning.

#### *Cough Detection: Recognition Models*

The mathematical modeling experiments were designed for automatic detection of cough sounds from the created dataset. Before processing, we resampled all audio to 16 kHz mono. To create our model, we split the sound recordings into 1-s samples, which we preprocessed using the Fourier transform with a window size of 25 ms, a window hop of 10 ms, and a periodic Hann window to generate spectrograms of the sound. Then we used convolutional neural networks to classify the spectrograms for model training. We followed the approach of Hershey et al. [13], using convolutional neural network architectures designed for large-scale audio classification, which concludes that image classification analogs of the convolutional neural networks outperform raw features' classifiers on audio classification tasks.

We measured the performance of the cough detector using audio datasets created from online sources. We tested two sensitivity presets: 98 and 99%. We experimented with various techniques for creating the recognition models, but ultimately found success with a simple deep convolutional neural network that classifies relatively small sound samples individually. The final output estimates in each small sample whether a cough sound is present or not. We calculated the sensitivity and specificity of our models and compared our results with commercially available cough monitors, as well as interrater agreement with and between human raters.

To ensure protection of privacy of any voice content captured incidentally, all recordings were clipped to 1-s final sample files. This minimizes the possibility of extracting meaningful information from spoken words, while maintaining the ability to listen to recorded cough sounds to review model performance.

## Results

#### *Sound Collection: Software Development*

Based on the experiment, 2 weeks of continuous patient recording with selected frequency presets would yield 3.04 GB (at 12 kHz), 3.61 GB (at 16 kHz), 4.68 GB (at 24 kHz), 6.98 GB (at 32 kHz), and 9.27 GB (at 44.1 kHz) of data. From the literature review and our testing, we

**Table 1.** Comparison of the performance of HealthMode’s cough recognition models, human inter-annotator agreement, and commercially available cough-counting solutions

Solution	Specificity, %	Sensitivity, % of detected coughs	Subjects, <i>n</i>	Solution details	Ref.
HealthMode Cough (public dataset)	99.90	75	22	Mobile phone microphone	–
HealthMode Cough (public dataset)	99.50	90	22	Mobile phone microphone	–
Inter-annotator agreement (recorded child cough)	na	75–80	78	Manual counting, bedside non-contact microphone	13
Lifeshirt™	99.60	78	8	A vest detecting chest wall movements	10
VitaloJak	98	99.7	10	Contact microphone	12
Leicester Cough Monitor	99	91	9	Microphone at chest	6
The CoughCOUNT™	94	96	12	Monitors chest wall motion	11
Hull Automated Cough Counter	96	80	33	Lapel microphone	7
LR100 Cough Monitor	na	81	30	Manual counting	15

adopted the sound sampling frequency of 16 kHz. This setting ensures reliable cough modeling, while limiting the total amount of collected data and decreasing risks connected with large file sizes during storage and upload, such as insufficient storage space and long transmission/upload times. This low frequency audio recording approach results in smaller file sizes, thus decreases the burden of the data transfer to remote servers.

In the battery consumption testing experiment, after 29 h of sound collection and upload via Wi-Fi or cellular network, the difference between battery power level was 13% – with 42% of battery power remaining in the phone using the cellular network (58% used) and 29% battery power remaining in the phone using the Wi-Fi network (71% used) for data upload, as shown in Figure 3.

We measured the reliability of the application by leaving it running over a long period of time. At the time of writing, the application has been collecting data reliably with no unexpected shutdowns (crashes) through a 24-week period. The system is continuously recording and uploading data as well as logging predefined events, such as hourly recording coverage, hourly battery level, snooze events, and errors. These events are displayed on a web-application dashboard visible to the research team, enabling the research personnel to monitor these variables in real-time.

#### *Cough Detection*

The final results obtained for the cough recognition models were 90 at 99.5% specificity preset (Cohen’s kappa 0.5) and 75 at 99.9% specificity preset (Cohen’s kappa 0.72). The standard to compute sensitivity and specificity was manual counting from the recordings.

We compared our results of performance characteristics with published results of the specific solutions in Table 1. The data on performance characteristics of specific solutions were obtained from previously published independent study findings, which are referenced within the table. In this comparison, our initial cough recognition models trained on publicly available datasets reach comparable levels of specificity and sensitivity. The results show that our initial solution surpasses the referenced inter-annotator agreement for manual counting as well as several other described commercial solutions, for example the Leicester Cough Monitor [6]. However, VitaloJak [8, 9] maintains superior performance at comparable specificity presets.



## Discussion

### *Sound Collection: Software Development and User Testing*

At the time of writing, we have collected over 5,000 h of continuous data without unexpected shutdowns, indicating the suitability of this technology for a long-term passive recording task. Based on minimum and maximum battery consumption, the battery life of the smartphone will be sufficient for 24 h per day of continuous recording and data transfer, the minimum desired data collection period required for use in clinical research. The uniformity of this performance can be secured using phones of the same brand and model, all provisioned to restrict the use of applications other than the cough-recording application. Over the recording period, we were able to monitor in a real-time manner hourly recording coverage, hourly battery level, snooze events, and errors via a web-application dashboard. This is a useful part of the system, enabling the future research team to identify technical issues that may occur, as well as potential sources of non-compliance to the research instructions for the patient.

The HealthMode Cough application verification and validation test results informed decisions about the system's setup and optimization for the frequency of sound recording, anticipated amount of data to be transferred, and device battery consumption. Optimal parameters for the sampling frequency were a trade-off between the size of the output files versus sufficient frequency of recording to assess coughs. The file size information informed the selection of the most suitable monthly data plan for the future clinical study, whereas the literature review informed about the adequacy of the selected recording frequency for future cough classification modelling experiments.

A useful point for discussion is anticipated data quality collected over multiple days of the application use. We expect that the data quality might vary based on where on the patient's body the phone is worn, such as having the phone in a bag, pocket, or further distance from the participant.

Another challenge presented by the use of mobile application rather than ambulatory devices with microphones mounted on the patient's body is possible non-adherence caused by the participant not carrying the phone in close proximity throughout the study period. For example, the phone may be left lying on a table while the patient moves away from it. We intend to implement a detection system for this in future versions of the application. In the near term, for ongoing research with this application, we plan to mitigate this risk of data loss or diminished quality by providing a body-worn case for the smartphone – belt clips, running belts, or armbands that would hold the phone on or near the subject's body during the study.

Further use of the solution will include a provisioned second smartphone, rather than the participants primary phone, and study staff reminders to carry the phone at all times during the study period. We believe that these occasional drops in recording quality will be balanced by the prolonged monitoring time over multiple days. Future studies will examine this hypothesis.

Another approach to reduce the burden of carrying an additional phone would be to adopt the BYOD (bring your own device) approach and having the application installed on the participant's primary phone. Although easing the user experience burden, this design would require extensive verification and testing of the system on various mobile devices that differ in manufacturers, operating systems, and sensors, specifically microphones. To ensure unified sample quality in our initial studies, we lean towards provisioning a single type of device for any research that would utilize this application.

### *Cough Detection*

The experiments performed with publicly available sound data yielded results that suggest our approach is on the right course towards fully automated, multi-day, 24-h cough counts measured in real-time with performance comparable to existing, accepted solutions.

Taking a closer look at confounding factors and types of sounds that produced some of the false positives in the assessment, we found that the closest sounds to coughs were various sounds of throat clearing. This is an interesting finding, as such sounds may also be the source of inter-rater disagreement. Some other confounding sounds that were often classified as cough by the model were various door slams, object thuds or falls, sneezes, parts of speech, or parts of distorted voice in the background.

Although this experiment is only a first step, the future of automated cough detection offers exciting possibilities. The collected sound samples contain timestamp metadata, so it may be possible to construct a daily cough map to observe the temporal distribution and duration of one's coughs. From the spectrograms it may be possible to assess the intensity or type of cough in the future. With additional data about the application user, it may be possible to extend the models to classify productive/non-productive cough [11], or cough specific for various respiratory diseases [12]. The clinical applications lie in various areas of clinical research and practice. It also has the potential to be used as an efficacy (advantages of continuous vs. snapshot monitoring) or safety (early symptoms detection) assessment tool in clinical research.

#### *Privacy and Continuous Sound Recording*

Developments in artificial intelligence promise new opportunities for data analysis, but they present their own ethical concerns. We must consider the best practices for collecting and using health data amidst the complexities that arise in an age of generally reduced privacy. Current and commonly used methods of cough detection require extensive human review of recordings. While continuing to consider and improve upon privacy protections is our goal, our system provides a greater level of privacy by ensuring that, in production use, no human must routinely listen to the recordings. We are working toward a system that minimizes human interaction with recorded audio data, and we record and store data to limit risk of exposure; each individual data packet is as short as possible, decontextualized, and ambiguous on its own. We believe that maximizing privacy is an ethical foundation of technological development in clinical trials, and it is always a factor guiding our design.

In the past, device monitoring in clinical trials has led to the loss of participants over privacy concerns [14]. Our efforts to mitigate these concerns will improve the user experience, yet we anticipate that there may be limits to the sense of security we instill. Discomfort may be inherent in continuous recording, and no matter how de-identified and safe the data are, the understandable apprehension around the presence of a continuous audio recording device is likely to be a lingering privacy concern among participants.

#### **Conclusion**

We designed and developed a smartphone application for continuous audio data collection. The system's performance and parameters were optimized for ease of application in clinical research. We have taken extensive steps to maximize privacy and safety of the solution where possible.

Initial experiments with cough detection techniques from various audio samples yielded encouraging results for further application in patient-collected data from an upcoming naturalistic clinical study. Our future development goals are to collect large amounts of high-quality audio data from patients with chronic cough, accompanied by ePRO and clinRO data. This database will provide high-quality input for cough recognition modelling for our production research cough frequency measurement solution. Our goal performance characteristics include the ability to detect coughs continuously with a 92% sensitivity at 99% specificity in audio data from study participants in the real world.



Subsequent to the research described here, we are preparing a naturalistic clinical study of patients with refractory chronic cough to determine the reliability of smartphone use for continuous ongoing audio data collection in a real-world research setting. The results of this study will be reported in subsequent papers.

### Statement of Ethics

This article does not involve work with any human subjects. We will submit a consolidated view of all work with human subjects with the appropriate IRB statement in future paper submissions. All sound data for cough recognition modeling purposes were obtained from publicly available data, which are listed in the online supplementary material. No human subject data were used for this research.

### Disclosure Statement

L.K., V.B., P.D., M.M., J.G., J.J., and D.R.K. are employees of and shareholders in HealthMode Inc.

### Funding Sources

The research and development described in this paper was funded by HealthMode Inc.

### Author Contributions

Lucia Kvapilova and Daniel R. Karlin contributed to the overview, research design, methods development, application testing and wrote the manuscript. Peter Dubec contributed to the smartphone application development and testing. Vladimir Boza and Jan Bogar contributed to the systems development, data analysis, and cough recognition models development. Martin Majernik contributed to the research design and methods development. Duncan J. Kimmel, Jennifer Goldsack, and Jamileh Jamison contributed to the writing the manuscript. All authors reviewed the manuscript and approved the final revision.

### References

- 1 Dicipingaitis PV. Cough: an unmet clinical need. *Br J Pharmacol*. 2011 May;163(1):116–24.
- 2 Steele AM, Meuret AE, Millard MW, Ritz T. Discrepancies between lung function and asthma control: asthma perception and association with demographics and anxiety. *Allergy Asthma Proc*. 2012 Nov-Dec;33(6):500–7.
- 3 Spinou A, Biring SS. An update on measurement and monitoring of cough: what are the important study endpoints? *J Thorac Dis*. 2014 Oct;6(7 Suppl 7):S728–34.
- 4 Bolser DC. Pharmacologic management of cough. *Otolaryngol Clin North Am*. 2010 Feb;43(1):147–55.
- 5 Smith J, Woodcock A. Cough recording technology. *Curr Respir Med Rev*. 2011;7(1):34–9.
- 6 Biring SS, Fleming T, Matos S, Raj AA, Evans DH, Pavord ID. The Leicester Cough Monitor: preliminary validation of an automated cough detection system in chronic cough. *Eur Respir J*. 2008 May;31(5):1013–8.
- 7 Barry SJ, Dane AD, Morice AH, Walmsley AD. The automatic recognition and counting of cough. *Cough*. 2006 Sep;2(1):8.
- 8 McGuinness K, Kelsall A, Lowe J, Woodcock A, Smith JA. Automated cough detection: a novel approach. *Am J Respir Crit Care Med*. 2007;175:A381.
- 9 McGuinness K, Holt K, Dockry R, Smith J. Validation of the VitaloJAK 24 Hour Ambulatory Cough Monitor. *Thorax*. 2012;67(Suppl 2):A131.

- 10 Chatzarrin H, Arcelus A, Goubran R, Knoefel F. Feature extraction for the differentiation of dry and wet cough sounds. *IEEE Int Symp*. 2011. DOI: 10.1109/MeMeA.2011.5966670.
- 11 Akira Murata, Yasuyuki Taniguchi, Yasushi Hashimoto, Yasuyuki Kaneko, Yuji Takasaki, and Shoji Kudoh. Discrimination of productive and non-productive cough by sound analysis. *Int Med*. 1998;37:732–5.
- 12 Knocikova J, Korpas J, Vrabec M, Javorka M. Wavelet analysis of voluntary cough sound in patients with respiratory diseases. *J Physiol Pharmacol*. 2008 Dec;59(Suppl 6):331–40.
- 13 Hershey S, Chaudhuri S, Ellis DP, Gemmeke J, Jansen A, Moore RC, et al. CNN architectures for large-scale audio classification. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2017; New Orleans*. <https://doi.org/10.1109/ICASSP.2017.7952132>.
- 14 Wu R, Liaqat D, De Lara E, Son T, Rudzicz F, Alshaer H, et al. Feasibility of using Android smartwatches for nearly continuous monitoring of patients with COPD. *Am J Respir Crit Care Med*. 2018;197:A4983.
- 15 Swarnkar V, Abeyratne UR, Chang AB, Amrulloh YA, Setyati A, Triasih R. Automatic identification of wet and dry cough in pediatric patients with respiratory diseases. *Ann Biomed Eng*. 2013 May;41(5):1016–28.