Lena: a Voice-Based Conversational Agent for Remote Patient Monitoring in Chronic Obstructive Pulmonary Disease

David Cleres[a], Frank Rassouli[c], Martin Brutsche[c], Tobias Kowatsch[a,b] and Filipe Barata[a]

[a]*Center for Digital Health Interventions, Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland*
[b]*Center for Digital Health Interventions, Institute of Technology Management, University of St. Gallen, St. Gallen, Switzerland*
[c]*Lung Center, Cantonal Hospital St. Gallen, St. Gallen, Switzerland*

Abstract

Chronic obstructive pulmonary disease (COPD) is one of the leading causes of death worldwide. To manage the increasing number of COPD patients and reduce the social and economic burden of treatment, healthcare providers have sought to implement remote patient monitoring (RPM). Screen-based RPM applications, such as filling self-reports on the smartphone or computer, have been shown to increase the quality of life, reduce the frequency and severity of exacerbations, and increase physical activity in patients with COPD. These applications, however, are not without challenges for the elderly target population. They are often used on devices designed by and for a different age group, which makes filling out self-reports prone to error and induces fears of technology malfunctions. Voice-based conversational agents (VCAs) are available on more than 2.5 billion devices and are increasingly present in homes worldwide. Aside from their commercial success, VCAs are also credited with several functionalities, such as hands-free use, that make their adoption in healthcare attractive, especially for the elderly. In this work, we investigate the potential of VCAs for RPM of COPD. Specifically, we designed and evaluated *Lena*, a single-board computer-based VCA framed as a digital member of the medical team. Lena acts as RPM for the early prediction of COPD exacerbations by asking ten symptom-related questions to determine the patient's daily health status. This paper presents the patients' feedback after their interaction with Lena. Patients evaluated the acceptability of the system. Notably, all patients could imagine using the system once a day in the context of a larger study and wished to integrate Lena into their daily routine.

Keywords

Chronic obstructive pulmonary disease, Voice-based Conversational Agents, Remote Patient Monitoring, Single-Board Computer, Ubiquitous Computing

1. Introduction

COPD is the third leading cause of death worldwide [1]. It is responsible for 251 million reported deaths per year [2]. COPD is a chronic, progressive disease caused by airway inflammation due to smoking or long–term exposure to pollutants (e.g., dust, fumes, poor air quality) [2].

To cope with the increasing number of COPD patients and to reduce the pressure on health services, providers have sought to introduce RPM for COPD patients [3]. RPM is the automatic, continuous transmission and processing of physiological data, decision support, prediction of deterioration, and alerting.

COPD patients are usually treated as outpatients, except in cases of hospitalization due to an exacerbation [4]. COPD exacerbations occur on average one to four times per year [4] and represent the acute worsening of symptoms such as shortness of breath and cough [5]. Approximately 70 % of health care costs associated with COPD are due to emergencies and hospitalizations for treatment of exacerbations. RPM has the potential to reduce the frequency and severity of COPD exacerbations [6], thereby reducing healthcare costs [7]. While recent studies have shown potential benefits (e.g., increased quality of life [8], reduced frequency and severity of exacerbations [6], improved physical activity [9]) of RPM for patients with COPD, physiological parameters of COPD patients are not continuously monitored outside of hospitals, except for research purposes [10]. Moreover, symptoms are usually self–reported by the patients using pen and paper diaries [10]. Considering that COPD patients belong to the older part of the population, current screen–based applications (e.g., filling self–reports on smartphones or computers) do not come without challenges. Older adults often have low IT–literacy, fear of malfunction [11], and a lack of confidence in their ICT abilities [12]. In addition, the fact that most commercially available software is developed by and for a different age group limits the inclusion of more digital applications in the lives of older

---

Joint Proceedings of the ACM IUI 2021 Workshops, April 13–17, 2021, College Station, USA

✉ dcleres@ethz.ch (D. Cleres); frank.rassouli@kssg.ch (F. Rassouli); martin.brutsche@kssg.ch (M. Brutsche); tkowatsch@ethz.ch (T. Kowatsch); fbarata@ethz.ch (F. Barata)
0000-0002-7534-7960 (D. Cleres); 0000-0003-2426-3545 (F. Rassouli); 0000-0002-1612-3609 (M. Brutsche); 0000-0001-5939-4145 (T. Kowatsch); 0000-0002-3905-2380 (F. Barata)

CEUR Workshop Proceedings (CEUR-WS.org)

people, even though they recognize the benefits that come with increased ICT capabilities [13]. Further, there is evidence that older generations prefer voice–based to screen–based communication [14].

From *The Voder*, the first attempt to electronically synthesize human speech by using a console with fifteen touch–sensitive keys and a pedal to select the appropriate bandpass filters to convert the hisses and tones into vowels, unveiled at the 1939 New York World's Fair, to today's *Alexa*, *Siri*, or *Google Assistant*, the way and frequency humans interact with VCAs have evolved dramatically. VCAs are now available on more than 2.5 billion devices such as smartphones and tablets, smart speakers, computers and have even been embedded in wearables or cars, thus nearly becoming ubiquitous in our daily lives [15]. One in four Americans owns a smart speaker, and in 2018 alone, their ownership increased by 40 percent [16]. Beyond their commercial success, VCAs are also credited with several functionalities that make their use in healthcare attractive. VCAs enable hands–free interaction, allow input from individuals with low literacy or with intellectual [17], motor and cognitive disabilities [18], or provide more natural support for routine health tasks when in–person healthcare is not possible [19]. Voice interaction also enables passive monitoring and analysis of audio samples for healthcare applications, such as Alzheimer's [20], depression [21], and schizophrenia [22]. In addition, recent work suggested the potential of speech (e.g., pause time, pause frequency, prosodic features, among others) as a marker of exacerbations in patients with COPD [23].

With this in mind, we argue that VCAs have the potential to enable RPM for COPD patients by facilitating the collection and sharing of health–related information with healthcare professionals, thereby improving quality of life, reducing exacerbations, and thus, reducing the costs in COPD care.

For these reasons, the authors developed and evaluated a single–board computer (SBC)–based VCA framed as *Lena*, a digital member of the medical team. This pilot study investigates the acceptability of a voice–based approach to alleviating patients' communication burden while filling questionnaires. More specifically, the VCA's ability to formulate its questions simply and understandably, as well as its ability to understand the patient's answer and respond accordingly, were assessed with four COPD patients.

2. Related Work

2.1. RPM for COPD

In 2015, the World Health Organization estimated that 3.17 million people died from COPD, accounting for five percent of all global deaths [2]. The disease's prevalence is on the rise due to higher smoking prevalence and the aging population [2]. In consequence, the burden on health care providers is expected to increase in the coming years. Evidence suggests that RPM for COPD patients can reduce hospitalizations [6] and costs associated with this disabling disease by at least 14 percent [7]. Moreover, McLean et al. showed in a Cochrane review that RPM can increase the quality of life of COPD patients and reduce the number of exacerbations [8]. There is also evidence in a review by Lundell et al. that RPM can improve physical activity levels in COPD patients [9]. Pedone and Lelli, in their review [24], reported a positive but nonsignificant effect of remote care on hospital admissions and emergency department visits. Similarly, McDowell et al. [25] showed that RPM with self–management improves the quality of life but does not significantly improve emergency care.

More recently, Rassouli et al. investigated the association between the COPD Assessment Test (CAT) and the risk of exacerbation of COPD [26]. In this study, patients completed an online questionnaire focused on detecting an acute exacerbation of COPD (daily) and the CAT (weekly). The authors found that by completing the questionnaires regularly, patients could assess their health status more accurately. Also, the evolution of the CAT could help assess the risk of future exacerbations.

2.2. VCAs for healthcare

Humans apply social rules in interactions with computers [27, 28]. VCAs enable this behavior by mimicking interpersonal conversation. Therefore, compared to text–based interaction, they are perceived as more socially present (i.e., the perception of interacting with an intelligent being) [29, 30] and even more believable than humans when it comes to information retrieval [31]. Similar to existing screen–based conversational agents [32, 33, 34], VCAs can form a working alliance [35] with patients, which has a positive impact on treatment outcomes [36, 37]. Recent works by Balasuriya et al. [38], Ireland et al. [39], and Kadariya et al. [40] also suggest that non–commercial dedicated VCAs can meet user expectations when supporting the prevention and management of chronic or mental illnesses. Furthermore, speech interaction also enables passive monitoring and analysis of voice samples for health applications, such as Alzheimer's [20], depression [21], schizophrenia [22], and COPD [23].

3. Methods

Lena's design and evaluation follow an iterative design science approach, in which single capabilities are con-

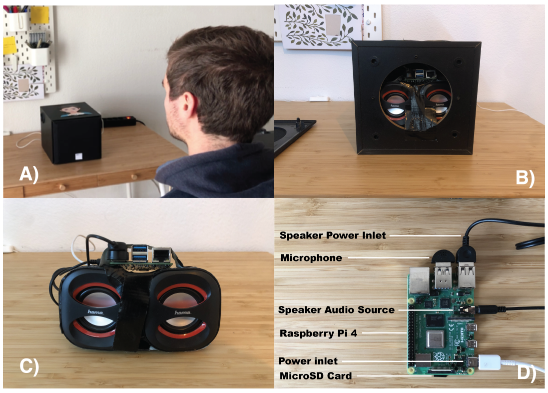

Figure 1: Detail view of Lena – A) Patient interacting with Lena. B) Front–view of Lena with open front. C) Front–view of Lena without its housing. D) Raspberry Pi 4 with all the connected components.

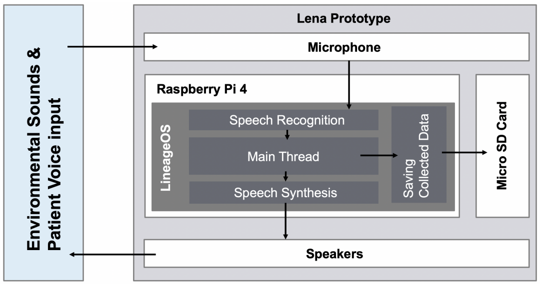

Figure 2: Illustration of Lena's embedded system. The sky blue box represents the external inputs coming from the patient. The white boxes show the hardware components. The light gray color highlights all the elements from which Lena is built. Grey represents the operating system. The features of the developed Android app appear in dark grey.

tinuously improved and eventually integrated into one system.

3.1. Hardware

Lena consists of three hardware components: a Single–Board Computer (SBC), a USB microphone, and a pair of USB speakers. For aesthetics and protection, a wooded loudspeaker housing contains Lena's core components, as shown in Figure 1. More precisely, Lena's core consists of a Raspberry Pi 4 Model B (cf. Figure 1. D.; Specifications: 8 GB RAM and 1.5 GHz processor) and a SanDisk Ultra microSDHC of 16 GB. The USB microphone has a sensitivity of −67 dBV/pBar, −47 dBV/Pascal ±4 dB, the frequency ranges from 100 – 16kHz, and on–device noise–canceling filters out the background noise. Finally, the USB–powered speakers are connected to the 3.5 mm jack port of the SBC. Their frequency ranges from 50 to 20′000 Hz.

3.2. Software

3.2.1. Operating system

The SBC used to power Lena runs LineageOS, a free and open–source operating system based on the Android mobile platform. This has the advantage that all apps already available in the Google Play Store can also be used on the Raspberry Pi after installing Open GApps. The Open GApps Project is an open–source effort that provides pre–built Google Apps packages. By installing Open GApps, the Google Play Store became available, as well as Google's speech functionalities and APIs, which will be discussed in more detail below.

3.2.2. How Lena communicates

To engage with a patient, Lena must be able to (i) recognize when the patient talks with it, (ii) understand the patient's speech, and (iii) to respond in an intelligible way. The next three paragraphs describe the implementation of these points.

Key phrase recognition To activate Lena, a patient must utter the key phrase "Hello, Lena". The developed app recognizes this key phrase in real–time. After recognizing the key phrase, Lena starts speaking, explains the conversation's goal to the patient, and asks the first question.
The key phrase recognition uses Vosk, an offline open–source speech recognition toolkit that understands 17 languages and dialects.

Speech recognition To reliably understand the patient's responses, Lena uses Android's API. More specifically, the *SpeechRecognizer*, *RecognitionService*, and *RecognizerIntent* Android speech classes handle the speech–to–text transformation of the patient's speech. We implemented speech recognition to operate purely offline. In this way, the patient's personal and sensitive data remains on the device and does not need to be anonymized by another algorithm, which could affect the quality of the recording. To enable this functionality, the authors had to specify the flag *EXTRA_PREFER_OFFLINE* in the Android *RecognizerIntent* class. Furthermore, the offline speech recognition package must be downloaded on the device from the *Language and Input* section of the LineageOS system settings. Finally, the Android speech recognition framework must be set to the one provided by Google.

Speech synthesis Lena's offline text–to–speech (TTS) capabilities are also based on the Android API. The Android class *TextToSpeech* converts Lena's predefined conversational turns from text to speech. Other TTS APIs (e.g., Amazon Polly, Google Cloud TTS, IBM Watson TTS) could have been used to make Lena's speech more melodic and less monotone. However, the authors decided to use the Android API, which allows speech to be synthesized offline without installing additional software packages while delivering an intelligible speech.

3.3. Experimental set–up

3.3.1. Patient recruitment & interaction with the VCA

Four COPD patients (three male, one female) of age 69 ± 5 interacted with *Lena*. In practice, Lena was framed as a digital member of the medical team.

Three different hospital rooms served to conduct the interviews (a furnished hospital bedroom, a doctor's office, and an examination room) to evaluate Lena under various circumstances. Also, one patient received oxygen therapy while interacting with Lena. Before the interaction with Lena began, the authors instructed the patients on the nature of the questions that would be asked. The patients were randomly recruited based on their previous or current participation in a COPD–RPM study [41, 26]. In this study, patients completed a daily questionnaire via their personal computer or smartphone. Rassouli et al. were able to detect 60 out of 63 acute exacerbations of COPD with a sensitivity, specificity, positive and negative predictive value of 95, 98, 26, and 99.9 %, respectively [41]. During the interaction, Lena asked and the patients answered the following questions in German: (i) *Do you have more dyspnoea today, exceeding your usual variation?*, (ii) *Do you have more sputum today, exceeding your usual variation?*, (iii) *Is your sputum today more yellow or green, exceeding your usual variation?*, (iv) *Do you have more cough today, exceeding your usual variation?*, (v) *Do you feel febrile today?*, (vi) *Do you feel like having a common cold today?*, (vii) *Do you feel unwell today, exceeding your usual variation?* and (viii) *Did you start to take your emergency medication within the last 24 h?*. This time, however, the questions were asked and answered verbally. By selecting these patients, it was possible to objectively compare the speech–based solution with the previously used screen–based solution.

Concretely, the patients sat in front of Lena, as shown in Figure 1. After being triggered with the key phrase *"Hello Lena"*, Lena began asking questions to capture a questionnaire of the patient's perceptions of current COPD severity. The questionnaire consisted of nine binary questions and one open–ended question. The latter referred to the patient's mood. Depending on the patient's mood, Lena would react differently, e.g., by telling a joke if the patient was not feeling well. At the beginning of the interview, Lena would ask the patient whether to use formal or casual language throughout the interview. Finally, the eight other binary questions were aimed at evaluating the patient's health status.

At the end of the dialog, Lena would thank the patient for sharing today's symptoms with the study team and switch back into an idle mode, waiting for the patient to say the key phrase.

3.3.2. Evaluation and interview details

After the interaction with Lena, the patients were interviewed by the study team for 20 minutes.

The first part consisted of filling out a pen and paper survey with six questions evaluated on a 7–point Likert scale (see Table 1). The questions aimed to evaluate Lena and were based on the technology acceptance model [42]. Perceived enjoyment was defined as the degree to which the activity of using technology is perceived to be enjoyable [43]. Finally, relative advantage represented the degree to which a novel application is perceived as being better than the state of the art [44]. The second part of the interview consisted of a face–to–face interview between the patient and a member of the study team (see Table 2).

4. Results

4.1. Acceptability

The patients understood Lena's questions accurately (EOU1) and felt that Lena understood their responses (EOU2). They enjoyed conversing with Lena (ENJ1) and could imagine using Lena at least once a day to complete the questionnaire (ITU1), and found Lena useful (USE1). Last but not least, patients would prefer to use Lena than the existing screen–based solution (RA1). The entire conversation lasted an average of two minutes.

After the conversation with Lena, the authors conducted face–to–face interviews with patients to understand the current system's strengths and weaknesses. Table 2 shows a condensed summary of these interviews. All patients owned a smartphone (TECH1). Half of the patients were already intentionally using a VCA in the context of driving or reading a text aloud (TECH2). All patients could imagine using Lena instead of the screen–based application for a one–year study (ITU2). When asked about the advantages and disadvantages of completing the questionnaire orally, patients emphasized the ease of use of Lena and that no login is required, which is the case with the screen–based study (ENJ3). Another patient responded that interacting verbally with Lena would increase compliance as it felt more engaging (ENJ3). Two

Table 1
Pen and paper survey questions and aggregated patient responses. SD: Strongly Disagree, SA: Strongly Agree

Construct	Code	Question–Item	Answer Options	Results (M ± SD)
Perceived ease of use	EOU1	I understand Lena's questions	SD (−3) – SA (3)	2.75 ± 0.5
Perceived ease of use	EOU2	Lena understands my answers	SD (−3) – SA (3)	2.25 ± 0.96
Perceived enjoyment	ENJ1	I was happy to respond to Lena's questions	SD(−3) – SA (3)	2.5 ± 0.58
Intention to use	ITU1	I can imagine Lena to answer a few questions daily	SD(−3) – SA (3)	2.00 ± 1.41
Perceived usefulness	USE1	I find Lena useful to give my daily assessment on COPD control study	SD (−3) – SA (3)	2.25 ± 0.96
Relative Advantage	RA1	In the context of the planned study, I would rather answer Lena once a day than fill a questionnaire on a smartphone for this purpose	SD (−3) – SA (3)	1.50 ± 1.73

Table 2
Face–to–face interview questions and patient answers. P: Patient, NI: No Idea

Construct	Code	Question–Item	Answers (P1, P2, P3, P4)
Tech. Affinity	TECH1	Do you possess a smartphone?	Yes, Yes, Yes, Yes
Tech. Affinity	TECH2	Do you already use a voice assistant (Alexa, Cortana, Siri)?	No, In car, No, Read aloud
Intention to use	ITU2	Would you be willing, as part of a study on self–management of COPD, to answer Lena's questions by voice before going to bed or when getting up for a period of one year?	Yes, Yes, Yes, Yes
Perceived enjoyment	ENJ3	What did you like and dislike about Lena, especially in comparison to the questionnaire on the smartphone?	Engagement, Ease–of–use, NI, No login process
Perceived enjoyment	ENJ4	Would you prefer to interact in Swiss German dialect instead of German?	Yes, No, No, Yes
Improvements	IMP1	What would definitely need to be improved about Lena?	Intonation, NI, NI, NI
Improvements	IMP2	What could possibly go wrong with Lena's questions or with your answers? What should we consider in the development of Lena?	Power outage, NI, NI, Holidays

shortcomings were that Lena sometimes took too long to interpret her utterances, and one of the two testers highlighted Lena's intonation as a disadvantage (IMP1). In their opinion, these two factors made the dialogue sluggish at some moments. One patient was concerned that Lena would not fit in his luggage when traveling on vacation, and that in this case, completing the questionnaire would require a different solution (IMP2). P4 also pointed out possible risks caused by power outages since the system needs to be plugged in (IMP2). Finally, whether Lena should rather address the patients in Swiss German dialect instead of German was answered positively by two patients (ENJ4).

4.2. Speech recognition

To evaluate Lena's speech recognition capabilities, the authors recorded and transcribed the interviews using a smartphone phone. Lena saved the recognized speech of each interview in JSON format. The patients could freely answer Lena's questions. However, to understand the positive or negative connotation of a patient's answer, the algorithm searched for a keyword such as "yes" or "no".

All the interviews could be initiated with the key phrase and completed by the patients. It took, on average, 2±1.41 attempts to trigger the interaction. Finally, the patient's speech transcribed by Lena perfectly matched the authors' transcription; all responses were correctly identified and transcribed.

5. Discussion

All COPD patients succeeded in triggering, understanding, and interacting with Lena to complete the questionnaire. The speech recognition results indicate that the prototype can understand the patients' responses with perfect accuracy and lead the discussion accordingly.

Also, all patients declared their willingness to use Lena daily over a longer period (e.g., 12 months) (ITU1, ITU2) and favored this voice–based solution over the original screen–based application [41] (RA1). This suggests that Lena qualifies for validation in a longitudinal study with COPD patients.

Lena's voice assistant capabilities rely on Android's speech recognition and speech synthesis and Vosk's speech recognition APIs. All three can be used offline without transferring patients' recorded speech samples to an external server. With this approach, we recognize the sensitive nature of medical and speech data and ensure privacy and security.

Although the healthcare sector already uses VCAs [45, 46], this study provides the first insights regarding the feasibility, relevance, and acceptability for such an application for COPD patients. In contrast to the proposed pen and paper or screen–based applications to collect self–reports' information, Lena provides a tailored application for the elderly target population by verbally capturing patient information without requiring interaction with another person via a phone call or in–person visit [41, 24, 9]. Moreover, recent work has suggested the potential of speech (e.g., pause time, pause frequency, prosodic features, a.s.o.) and cough [47] as a marker of COPD exacerbations [23]. Considering that recent research has also shown the feasibility to detect cough events with high accuracy on devices with limited computing power (e.g., smartphones [48]) and the ability to continuously listen in the background, detect and count coughs comparable to human annotators [49], we argue that such a smart speaker system could not only collect subjective information but also provide passively and objectively measured parameters related to the health status of patients with COPD.

The study population received Lena well in terms of usefulness, ease and intention of use, and enjoyment. Nevertheless, there is a need for a systematic and comparative evaluation of the system (e.g., test the distance and direction in which the microphone was effective, test on older adults with an absence of basic digital literacy skills) to understand which aspects of Lena are best for patient acceptance and better outcomes than standard in–person care.

Patients welcomed completing the questionnaire without having to sign in. However, for Lena to meet the same data security standards as the state of the art screen–based system, voice–based authentication could be used, as it prevents unauthorized use and protects patient privacy [38].

Finally, Lena's current version did not allow patients to freely express their symptoms since the questionnaire consisted only of binary questions. Nonetheless, using a voice–based system allows patients to communicate more naturally to provide more information about their symptoms and make information exchange seamless with the medical team. The combination of the voice–based questionnaires and the passively recorded data should not only open a wide range of new research directions but, more importantly, provide better support for COPD patients.

6. Conclusion

This pilot study proposes Lena, a state of the art VCA for COPD RPM. Lena interacts with the patient in a spoken natural language to collect daily self–reports. This first evaluation yielded promising acceptance results of a VCA–based RPM application for COPD. All patients also showed a willingness to integrate Lena into their daily routine and saw its potential to improve future RPM solutions. We plan to integrate and evaluate Lena in a longitudinal observational study with COPD patients.

This research is a first step towards enabling scalable and natural–language–delivered RPM to facilitate access to health–related self–management services. It may also help overcome limitations of text–based systems, such as the lack of literacy of users in countries with low education index, accessibility for the elderly population, or even empowerment of patients with mental, motor, or cognitive disabilities.

References

[1] T. G. I. for Chronic Obstructive Lung Disease, 2020 GOLD REPORTS, 2020. URL: https://goldcopd.org/gold-reports/.

[2] WHO, Chronic respiratory diseases, 2017. URL: https://bit.ly/39y6qQN.

[3] C. E. Bolton, C. S. Waters, S. Peirce, G. Elwyn, Insufficient evidence of benefit: a systematic review of home telemonitoring for copd, J Eval Clin Pract 17 (2011) 1216–22. doi:`10.1111/j.1365-2753.2010.01536.x`.

[4] I. Tomasic, N. Tomasic, R. Trobec, M. Krpan, T. Kelava, Continuous remote monitoring of copd patients—justification and explanation of the requirements and a survey of the available technologies, Medical & Biological Engineering & Computing 56 (2018) 547–569. URL: https://dx.doi.org/10.1007/s11517-018-1798-z. doi:`10.1007/s11517-018-1798-z`.

[5] S. D. Sullivan, S. D. Ramsey, T. A. Lee, The economic burden of copd, Chest 117 (2000) 5S – 9S. URL: http://www.sciencedirect.com/science/article/pii/S0012369215527487. doi:`https://doi.org/10.1378/chest.117.2_suppl.5S`.

[6] F. C. Cordova, D. Ciccolella, C. Grabianowski, J. Gaughan, K. Brennan, F. Goldstein, M. R. Jacobs,

G. J. Criner, A telemedicine-based intervention reduces the frequency and severity of copd exacerbation symptoms: A randomized, controlled trial, Telemed J E Health 22 (2016) 114–122. doi:10.1089/tmj.2015.0035.

[7] G. Paré, P. Poba-Nzaou, C. Sicotte, A. Beaupré, E. Lefrançois, D. Nault, D. Saint-Jules, Comparing the costs of home telemonitoring and usual care of chronic obstructive pulmonary disease patients: A randomized controlled trial, European Research in Telemedicine / La Recherche Européenne en Télémédecine 2 (2013) 35–47. URL: http://www.sciencedirect.com/science/article/pii/S2212764X13000241. doi:https://doi.org/10.1016/j.eurtel.2013.05.001.

[8] S. McLean, U. Nurmatov, J. L. Liu, C. Pagliari, J. Car, A. Sheikh, Telehealthcare for chronic obstructive pulmonary disease, Cochrane Database Syst Rev (2011) Cd007718. doi:10.1002/14651858.CD007718.pub2.

[9] S. Lundell, A. Holmner, B. Rehn, A. Nyberg, K. Wadell, Telehealthcare in copd: a systematic review and meta-analysis on physical outcomes and dyspnea, Respir Med 109 (2015) 11–26. doi:10.1016/j.rmed.2014.10.008.

[10] F. Pitta, T. Troosters, V. S. Probst, M. A. Spruit, M. Decramer, R. Gosselink, Quantifying physical activity in daily life with questionnaires and motion sensors in copd, Eur Respir J 27 (2006) 1040–55. doi:10.1183/09031936.06.00064105.

[11] J. Kowalski, A. Jaskulska, K. Skorupska, K. Abramczuk, C. Biele, W. Kopeć, K. Marasek, Older adults and voice interaction: A pilot study with google home, in: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, CHI EA '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 1–6. URL: https://doi.org/10.1145/3290607.3312973. doi:10.1145/3290607.3312973.

[12] W. Kopeć, R. Nielek, A. Wierzbicki, Guidelines towards better participation of older adults in software development processes using a new spiral method and participatory approach, in: Proceedings of the 11th International Workshop on Cooperative and Human Aspects of Software Engineering, CHASE '18, Association for Computing Machinery, New York, NY, USA, 2018, p. 49–56. URL: https://doi.org/10.1145/3195836.3195840. doi:10.1145/3195836.3195840.

[13] A. Aula, Learning to use Computers at a Later Age, in: HCI and the Older Population, Univ. of Glasgow, Leeds, UK, 2004.

[14] D. T. Duncan, W. C. Goedel, J. H. Williams, B. Elbel, Acceptability of smartphone text- and voice-based ecological momentary assessment (ema) methods among low income housing residents in new york city, BMC Research Notes 10 (2017) 517. URL: https://doi.org/10.1186/s13104-017-2850-z. doi:10.1186/s13104-017-2850-z.

[15] B. Kinsella, A. Mutchler, Voice assistant consumer adoption in healthcare, Technical Report, 2019. URL: https://voicebot.ai/voice-assistant-consumer-adoption-report-for-healthcare-2019.

[16] B. Kinsella, A. Mutchler, Smart Speaker Consumer Adoption Report, Technical Report March, 2019. URL: https://voicebot.ai/wp-content/uploads/2019/03/smart_speaker_consumer_adoption_report_2019.pdf.

[17] S. S. Balasuriya, L. Sitbon, A. A. Bayor, M. Hoogstrate, M. Brereton, Use of voice activated interfaces by people with intellectual disability, in: Proceedings of the 30th Australian Conference on Computer-Human Interaction, OzCHI '18, Association for Computing Machinery, New York, NY, USA, 2018, p. 102–112. URL: https://doi.org/10.1145/3292147.3292161. doi:10.1145/3292147.3292161.

[18] F. Masina, V. Orso, P. Pluchino, G. Dainese, S. Volpato, C. Nelini, D. Mapelli, A. Spagnolli, L. Gamberini, Investigating the accessibility of voice assistants with impaired users: Mixed methods study, Journal of Medical Internet Research 22 (2020) e18431. URL: https://dx.doi.org/10.2196/18431. doi:10.2196/18431.

[19] E. Sezgin, L. K. Militello, Y. Huang, S. Lin, A scoping review of patient-facing, behavioral health interventions with voice assistant technology targeting self-management and healthy lifestyle behaviors, Translational Behavioral Medicine 10 (2020) 606–628. URL: https://dx.doi.org/10.1093/tbm/ibz141. doi:10.1093/tbm/ibz141.

[20] M. L. B. Pulido, J. B. A. Hernández, M. A. F. Ballester, C. M. T. González, J. Mekyska, Z. Smékal, Alzheimer's disease and automatic speech analysis: A review, Expert Systems with Applications 150 (2020) 113213. URL: https://dx.doi.org/10.1016/j.eswa.2020.113213. doi:10.1016/j.eswa.2020.113213.

[21] J. Wang, L. Zhang, T. Liu, W. Pan, B. Hu, T. Zhu, Acoustic differences between healthy and depressed people: a cross-situation study, BMC Psychiatry 19 (2019). URL: https://dx.doi.org/10.1186/s12888-019-2300-7. doi:10.1186/s12888-019-2300-7.

[22] N. B. Mota, M. Copelli, S. Ribeiro, Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance, npj Schizophrenia 3 (2017). URL: https://dx.doi.org/10.1038/s41537-017-0019-3. doi:10.1038/s41537-017-0019-3.

[23] V. Nathan, K. Vatanparvar, M. M. Rahman, E. Ne-

mati, J. Kuang, Assessment of chronic pulmonary disease patients using biomarkers from natural speech recorded by mobile devices, in: 2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN), 2019, pp. 1–4. doi:10.1109/BSN.2019.8771043.

[24] C. Pedone, D. Lelli, Systematic review of telemonitoring in copd: an update, Pneumonol Alergol Pol 83 (2015) 476–84. doi:10.5603/PiAP.2015.0077.

[25] J. E. McDowell, S. McClean, F. FitzGibbon, S. Tate, A randomised clinical trial of the effectiveness of home-based health care with telemonitoring in patients with copd, J Telemed Telecare 21 (2015) 80–7. doi:10.1177/1357633x14566575.

[26] F. Rassouli, F. Baty, D. Stolz, W. Albrich, M. Tamm, S. Widmer, M. Brutsche, Longitudinal change of copd assessment test (cat) in a telehealthcare cohort is associated with exacerbation risk, International Journal of Chronic Obstructive Pulmonary Disease Volume 12 (2017) 3103–3109. URL: https://dx.doi.org/10.2147/copd.s141646. doi:10.2147/copd.s141646.

[27] C. Nass, J. Steuer, E. R. Tauber, Computers are social actors, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '94, Association for Computing Machinery, New York, NY, USA, 1994, p. 72–78. URL: https://doi.org/10.1145/191666.191703. doi:10.1145/191666.191703.

[28] B. Fogg, Persuasive computers: Perspectives and research directions, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '98, ACM Press/Addison-Wesley Publishing Co., USA, 1998, p. 225–232. URL: https://doi.org/10.1145/274644.274677. doi:10.1145/274644.274677.

[29] E. Cho, Hey google, can i ask you something in private?, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 1–9. URL: https://doi.org/10.1145/3290605.3300488. doi:10.1145/3290605.3300488.

[30] K. Kim, N. Norouzi, T. Losekamp, G. Bruder, M. Anderson, G. Welch, Effects of patient care assistant embodiment and computer mediation on user experience, IEEE, ???? URL: https://dx.doi.org/10.1109/aivr46125.2019.00013. doi:10.1109/aivr46125.2019.00013.

[31] A. Pradhan, A. Lazar, L. Findlater, Use of intelligent voice assistants by older adults with low technology use, ACM Transactions on Computer-Human Interaction 27 (2020) 1–27. URL: https://dx.doi.org/10.1145/3373759. doi:10.1145/3373759.

[32] B. Liu, S. S. Sundar, Should machines express sympathy and empathy? experiments with a health advice chatbot, Cyberpsychology, Behavior, and Social Networking 21 (2018) 625–636. URL: https://dx.doi.org/10.1089/cyber.2018.0110. doi:10.1089/cyber.2018.0110.

[33] C. M. Schwartzman, J. F. Boswell, A narrative review of alliance formation and outcome in text-based telepsychotherapy, Practice Innovations 5 (2020) 128–142. doi:10.1037/pri0000120.

[34] T. Bickmore, A. Gruber, R. Picard, Establishing the computer–patient working alliance in automated health behavior change interventions, Patient Education and Counseling 59 (2005) 21–30. URL: https://dx.doi.org/10.1016/j.pec.2004.09.008. doi:10.1016/j.pec.2004.09.008.

[35] A. O. Horvath, L. Luborsky, The role of the therapeutic alliance in psychotherapy, J Consult Clin Psychol 61 (1993) 561–73. doi:10.1037//0022-006x.61.4.561.

[36] D. J. Martin, J. P. Garske, M. Katherine Davis, Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review, Journal of Consulting and Clinical Psychology 68 (2000) 438–450. URL: /record/2000-15699-009. doi:10.1037/0022-006X.68.3.438.

[37] N. Mead, P. Bower, Patient-centred consultations and outcomes in primary care: a review of the literature, Patient Educ Couns 48 (2002) 51–61. doi:10.1016/s0738-3991(02)00099-x.

[38] U. U. Rehman, D. J. Chang, Y. Jung, U. Akhtar, M. A. Razzaq, S. Lee, Medical instructed real-time assistant for patient with glaucoma and diabetic conditions, Applied Sciences 10 (2020) 2216. URL: https://dx.doi.org/10.3390/app10072216. doi:10.3390/app10072216.

[39] D. Ireland, C. Atay, J. Liddle, D. Bradford, H. Lee, O. Rushin, T. Mullins, D. Angus, J. Wiles, S. McBride, A. Vogel, Hello harlie: Enabling speech monitoring through chat-bot conversations, in: Studies in Health Technology and Informatics, volume 227, IOS Press, 2016, pp. 55–60. doi:10.3233/978-1-61499-666-8-55.

[40] D. Kadariya, R. Venkataramanan, H. Y. Yip, M. Kalra, K. Thirunarayanan, A. Sheth, kbot: Knowledge-enabled personalized chatbot for asthma self-management, IEEE, ???? URL: https://dx.doi.org/10.1109/smartcomp.2019.00043. doi:10.1109/smartcomp.2019.00043.

[41] F. Rassouli, M. Pfister, S. Widmer, F. Baty, B. Burger, M. H. Brutsche, Telehealthcare for chronic obstructive pulmonary disease in switzerland is feasible and appreciated by patients, Respiration 92 (2016) 107–13. doi:10.1159/000448377.

[42] F. D. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology, MIS Quarterly 13 (1989) 319. URL: https://dx.doi.org/10.2307/249008. doi:10.2307/249008.

[43] F. D. Davis, R. P. Bagozzi, P. R. Warshaw, Extrinsic

and intrinsic motivation to use computers in the workplace 1, Journal of applied social psychology 22 (1992) 1111–1132.

[44] E. M. Rogers, Diffusion of Innovations, 5th ed., Free Press, 2003.

[45] Aiva Health, Virtual health assistant | aiva health, 2020. URL: https://www.aivahealth.com/.

[46] Mayo Clinic, Skills from mayo clinic, 2020. URL: https://www.mayoclinic.org/voice/apps.

[47] M. G. Crooks, A. den Brinker, Y. Hayman, J. D. Williamson, A. Innes, C. E. Wright, P. Hill, A. H. Morice, Continuous cough monitoring using ambient sound recording during convalescence from a copd exacerbation, Lung 195 (2017) 289–294. URL: https://doi.org/10.1007/s00408-017-9996-2. doi:`10.1007/s00408-017-9996-2`.

[48] F. Barata, K. Kipfer, M. Weber, P. Tinschert, E. Fleisch, T. Kowatsch, Towards device-agnostic mobile cough detection with convolutional neural networks, in: 2019 IEEE International Conference on Healthcare Informatics (ICHI), 2019, pp. 1–11. doi:`10.1109/ICHI.2019.8904554`.

[49] F. Barata, P. Tinschert, F. Rassouli, C. Steurer-Stey, E. Fleisch, M. A. Puhan, M. Brutsche, D. Kotz, T. Kowatsch, Automatic recognition, segmentation, and sex assignment of nocturnal asthmatic coughs and cough epochs in smartphone audio recordings: Observational field study, Journal of Medical Internet Research 22 (2020) e18082. URL: https://dx.doi.org/10.2196/18082. doi:`10.2196/18082`.

CEUR Workshop Proceedings http://ceur-ws.org ISSN 1613-0073

Version 1.04-ntp
2021-02-12

CEUR Workshop Proceedings
http://CEUR-WS.org
RWTH Aachen, Informatik 5, Aachen, Germany

Author Agreement to Publish a Contribution as Open-Access on CEUR-WS.org

Herewith I/we (the author(s) resp. the copyright holders) **agree** that my/our contribution:

Lena: a Voice-Based Conversational Agent for Remote Patient Monitoring in Chronic Obstructive Pulmonary Disease

authored by:

David Cleres , Frank Rassouli , Martin Brutsche , Tobias Kowatsch, and Filipe Barata

with corresponding author

Name:

Affiliation:

Address:

Email:

shall be made available as an open-access publication under the **Creative Commons License Attribution 4.0 International (CC BY 4.0)**, available at https://creativecommons.org/licenses/by/4.0/legalcode, and be published as part of the proceedings volume of the event

Name and year of the event:

Editors of the proceedings (editors):

I/we agree that my/our contribution is made available publicly under the aforementioned license on the servers of CEUR Workshop Proceedings (CEUR-WS). I/we grant the editors, RWTH Aachen, CEUR-WS, and its archiving partners the non-exclusive and irrevocable **right to archive** my/our contribution and **to make it accessible** (online and free of charge) for **public distribution**. This granted right extends to any associated metadata of my/our contribution. Specifically, I/we license the associated metadata under a Creative Commons CC0 1.0 Universal license (public domain). I/we agree that our author names and affiliations is part of the associated metadata and may be stored on the servers of CEUR-WS and made available under the CC0 license. I/we acknowledge that the editors hold the copyright for the proceedings volume of the aforementioned event as the official collection of contributions to the event.

I/we have not included any copyrighted third-party material such as figures, code, data sets and others in the contribution to be published.

I/we warrant that my/our contribution (including any accompanying material such as data sets) does not infringe any rights of third parties, for example trademark rights, privacy rights, and intellectual property rights. I/ we understand that I/we retain the copyright to my/our contribution. I/we understand that the dedication of my/our contribution under the CC BY 4.0 license is irrevocable.
I/we understand and agree that the **full responsibility/liability** for the content of the contribution rests upon me/us as the authors of the contribution. I/we **release and relieve** the aforementioned editors, RWTH Aachen, individuals providing the CEUR-WS service, and the archiving partners of CEUR-WS **of liability caused by ordinary/simple negligence** in the publication or archiving of my/our aforementioned contribution via the servers used by CEUR-WS.
I/we have read the conditions of the Creative Commons License Attribution 4.0 International (CC BY 4.0), and agree to apply this license to my/our contribution.

Zürich, 02.03.2021,

Location, Date, Signature of the corresponding author representing all authors
(Signature must be handwritten with a pen on paper)