Assessment of the Implementation of a Chatbot-Based Screening for Burnout and COVID-19 Symptoms Among Residents During the Pandemic

Bruno Nascimento Moreira, MD, MSc
Alexandre Sampaio Moura, MD, MPH, PhD
Aleida Nazareth Soares, PhD
Zilma Silveira Nogueira Reis, MD, PhD
Rosa Malena Delbone, MD, PhD

ABSTRACT

Background Early identification of COVID-19 symptoms and burnout among residents is essential for proper management. Digital assistants might help in the large-scale screening of residents.

Objective To assess the implementation of a chatbot for tele-screening emotional exhaustion and COVID-19 among residents at a hospital in Brazil.

Methods From August to October 2020, a chatbot sent participants' phones a daily question about COVID-19 symptoms and a weekly question about emotional exhaustion. After 8 weeks, the residents answered the Maslach Burnout Inventory-Human Services Survey (MBI-HSS). The primary outcome was the reliability of the chatbot in identifying suspect cases of COVID-19 and burnout.

Results Among the 489 eligible residents, 174 (35.6%) agreed to participate. The chatbot identified 61 positive responses for COVID-19 symptoms, and clinical suspicion was confirmed in 9 residents. User error in the first weeks was the leading cause (57.7%, 30 of 52) of nonconfirmed suspicion. The chatbot failed to identify 3 participants with COVID-19 due to nonresponse. Twelve of 118 (10.2%) participants who answered the MBI-HSS were characterized as having burnout by the MBI-HHS. Two of them were identified as at risk by the chatbot and 8 never answered the emotional exhaustion screening question. Conversely, among the 19 participants identified as at risk for emotional exhaustion by the chatbot, 2 (10.5%) were classified with burnout, and 5 (26.3%) as overextended based on MBI-HHS scores.

Conclusions The chatbot was able to identify residents suspected of having COVID-19 and those at risk for burnout. Nonresponse was the leading cause of failure in identifying those at risk.

Introduction

During the COVID-19 pandemic, frontline health care workers were at increased risk for COVID-19 and mental health problems such as anxiety, depression, and burnout.[1-4] Early detection of COVID-19 and burnout via a systematic approach would support clinician well-being and could positively impact health care.[5,6] Chatbots are virtual assistants that establish a dialogue using natural language and can help evolve screening processes.[7,8] Chatbots were helpful during the COVID-19 pandemic, particularly in telemonitoring patients, but their use to screen for burnout is lacking.[9-12]

We aimed to assess the reliability of a chatbot to identify suspect cases of COVID-19 and burnout among residents.

DOI: http://dx.doi.org/10.4300/JGME-D-22-00920.1

Editor's Note: The online version of this article contains the questionnaires used in the study and further data.

Methods

This study was conducted at Santa Casa Belo Horizonte Hospital, Brazil, between August and October 2020. All 489 medical and multiprofessional residents in 36 specialties were invited to participate.

The chatbot sent a daily WhatsApp message to participants' phones to screen for COVID-19 symptoms. Then, using an adaptation of the Brazilian Ministry of Health COVID-19 recommendations,[13] participants with a positive answer to fever or any combination of 2 other symptoms were referred to the occupational health service for clinical examination and COVID-19 testing.

A weekly multiple-choice question with validity evidence for detecting burnout was also sent by the chatbot (provided as online supplementary data).[14] Participants considered at risk for burnout were referred to psychological support. After an 8-week monitoring period, the chatbot sent a link to an online version of the Maslach Burnout Inventory-Human Services Survey (MBI-HSS).[15] The subscales were

TABLE 1
Distribution of Participants According to Latent Burnout Profile[16]

Profile	Burnout, n (%)	Overextended, n (%)	Ineffective, n (%)	Disengaged, n (%)	Engaged, n (%)
All participants who answered the MBI-HSS (n=118)	12 (10.2)	20 (17.0)	6 (5.1)	6 (5.1)	58 (49.2)
Participants who answered the MBI-HSS and were identified as at risk by the chatbot (n=19)	2 (10.5)	4 (21.1)	0 (0)	0 (0)	8 (42.1)

scored as recommended by the developers.[15] In addition, participants were categorized into 5 latent profiles ranging from "burnout" to "engagement."[16]

The attitudes toward the chatbot and privacy concerns were assessed in a convenience sample selected based on adherence to the screening process (adherents responded to >90% of the chatbot questions, and nonadherents did not answer any of the questions). The instrument used to assess attitudes was developed by B.N.M. and A.S.M. (online supplementary data).

Using IBM SPSS Statistics 27, we summarized descriptive data using frequencies for categorical variables and measures of central tendency for continuous variables. Categorical variables were compared using the chi-square test. Statistical significance was defined by $P<.05$. Ethical approval was obtained from the Institutional Review Board of Santa Casa Hospital, and all participants signed informed consent.

Results

Among 489 residents, 174 (35.6%) agreed to participate. The mean age of participants was 29 (±3.31) years, and 70.1% (122 of 174) were female. The mean number of weekly responses to the chatbot was 340 (±204.7), with the larger number of responses occurring in the first week (n=825).

The chatbot identified 107 (3.3%) positive responses for COVID-19 symptoms of 3274 total responses, corresponding to 61 unique responders. Thirty users reported selecting the wrong option in the chatbot. Excluding the wrong answers, the specificity of the chatbot relative to the occupational health assessment was 29.0% (9 of 31).

Occupational health services quarantined 3 participants with COVID-19 symptoms not identified by the chatbot due to nonadherence. Although the chatbot was able to identify all suspect cases in those who were actively interacting with the tool, its overall sensitivity (including nonresponder participants) was 75% (9 of 12).

Regarding burnout screening, the chatbot registered 301 responses throughout the monitoring period (mean weekly response 37.1 [±51.9]), with a predominance (41.9%, 126 of 301) of participants reporting "some stress, but not feeling burned out" (online supplementary data). Nineteen of the 31 participants identified by the chatbot as at risk for burnout answered the MBI-HSS; 2 participants were classified as having burnout and 4 as overextended, with overall specificity 31.6% (6 of 19; TABLE 1).

The MBI-HSS classified as having burnout 10 additional participants who were not identified as at risk by the chatbot. Among those, 8 never answered the weekly burnout question, and 2 answered the chatbot only in the first 2 weeks of the study. As a result, the overall sensitivity of the chatbot relative to the MBI-HSS was 16.7% (2 of 12). However, among participants who answered at least once to the burnout question of the chatbot, the sensitivity increased to 50% (2 of 4).

A larger proportion of participants identified as at risk for burnout by the chatbot had high levels of emotional exhaustion compared to those not identified at risk (57.9% [11 of 19] vs 34.0% [33 of 97], respectively; $P=.034$; TABLE 2). However, for the other 2 domains there was no statistical difference between groups.

The assessment of attitudes toward the chatbot and privacy concerns was stratified according to participants' adherence to the chatbot (online supplementary data). Both groups understood the objective of the chatbot and recognized the importance of telemonitoring but felt that data was unsafe in a chatbot. On the other hand, participants adhering to the chatbot rated more favorably interacting with a chatbot and using a machine interface for telehealth monitoring. Another difference was the nonadherents' opinions about using WhatsApp strictly for personal issues.

Discussion

We evaluated the implementation of a chatbot to screen for COVID-19 and burnout among medical and multiprofessional residents during the COVID-19 pandemic. The chatbot correctly identified 9 COVID-19 suspect cases, and the 3 missed cases were among nonresponders. However, the specificity of the screening criteria was low, and clinical symptom-based scoring systems might help improve it in the future.

Using a single-item question to screen for burnout might oversimplify the complex relationship between

TABLE 2
Distribution of Participant Score Levels in the Different Burnout Domains

Responders to MBI-HSS	Level	Emotional Exhaustion, n (%)	Depersonalization, n (%)	Personal Accomplishment, n (%)
All participants (n=116)	High	44 (37.9)	18 (15.5)	32 (27.6)
	Moderate	23 (19.8)	19 (16.4)	33 (28.4)
	Low	49 (42.2)	79 (68.1)	51 (44.0)
	Mean score (σ)	21.9 (13.4)	3.8 (5.7)	37.4 (7.6)
Participants identified as at risk by the chatbot (n=19)	High	11 (57.9)[a]	5 (26.3)	4 (21.1)
	Moderate	5 (26.3)	3 (15.8)	9 (47.4)
	Low	3 (15.8)	11 (57.9)	6 (31.6)
	Mean score (σ)	30.8 (12.4)	4.7 (6.3)	37.1 (6.4)

Abbreviation: MBI-HSS, Maslach Burnout Inventory-Human Services Survey.
[a] Thirty-four percent (33 of 97) of those not identified as at risk by the chatbot had a high level of emotional exhaustion.

the individual and their workplace.[17] However, Rohland et al state that screening with a single-item question simplifies the process, making it quicker for participants and increasing response rates.[14] Indeed, the participants who were flagged as at risk by the chatbot scored higher for emotional exhaustion in line with previous results, showing that this single-item question correctly identified emotional exhaustion among physicians.[14]

After reviewing the 10 participants missed by the chatbot in the burnout screening process, 8 were deemed as nonadherent to the chatbot. The other 2 could be explained by the fact that the MBI-HSS was filled out approximately 6 weeks after the initial chatbot questionnaire, meaning the participant may have had changes in their mental health over that period.

We observed higher adherence rates to the chatbot at the beginning of the monitoring process that declined throughout the 8 weeks, which may have been secondary to response fatigue. In addition, our screening process was not mandatory, which could result in lower adherence rates.[18,19] However, we must consider that forcing participants to interact with the chatbot could add to burnout.

Finally, both adherent and nonadherent participants understood the goal and importance of telemonitoring. Still, a more significant proportion of nonadherents felt that WhatsApp should be restricted to personal use, leading us to wonder if those participants were apprehensive about using telehealth tools.

This study does have limitations. First, it was conducted in a single hospital during a pandemic when health care professionals were overwhelmed.[20] In such a context, adding the extra task of interacting with a chatbot might be undesired, particularly when there was no clear individual benefit. Another limitation is that we cannot rule out survival bias. Indeed, as the chatbot represented extra work, this could have resulted in a dropout of participants who were burned out.

Conclusions

The chatbot could identify and automatically refer users identified as having either COVID-19 symptoms or those at increased risk for burnout. Nonadherence was the leading cause of the failure of the chatbot to identify some of those at risk.

References

1. Koh WC, Naing L, Chaw L, et al. What do we know about SARS-CoV-2 transmission? A systematic review and meta-analysis of the secondary attack rate and associated risk factors. *PLoS One*. 2020;15(10):e0240205. doi:10.1371/journal.pone.0240205
2. Marvaldi M, Mallet J, Dubertret C, Moro MR, Guessoum SB. Anxiety, depression, trauma-related, and sleep disorders among healthcare workers during the COVID-19 pandemic: a systematic review and meta-analysis. *Neurosci Biobehav Rev*. 2021;126:252-264. doi:10.1016/j.neubiorev.2021.03.024
3. Ghahramani S, Lankarani KB, Yousefi M, Heydari K, Shahabi S, Azmand S. A systematic review and meta-analysis of burnout among healthcare workers during COVID-19. *Front Psychiatry*. 2021;12:758849. doi:10.3389/fpsyt.2021.758849
4. Fernandez R, Sikhosana N, Green H, et al. Anxiety and depression among healthcare workers during the COVID-19 pandemic: a systematic umbrella review of the global evidence. *BMJ Open*. 2021;11(9):e054528. doi:10.1136/bmjopen-2021-054528
5. National Academies of Sciences, Engineering, and Medicine; National Academy of Medicine; Committee on Systems Approaches to Improve Patient Care by Supporting Clinician Well-Being. *Taking Action Against*

Clinician Burnout: A Systems Approach to Professional Well-Being. National Academies Press; 2019.

6. Shanafelt TD. Physician well-being 2.0: where are we and where are we going? *Mayo Clin Proc*. 2021;96(10):2682-2693. doi:10.1016/j.mayocp.2021.06.005
7. Dale R. The return of the chatbots. *Nat Lang Eng*. 2016;22(5):811-817. doi:10.1017/S1351324916000243
8. Asch DA, Nicholson S, Berger ML. Toward facilitated self-service in health care. *N Engl J Med*. 2019;380(20):1891-1893. doi:10.1056/NEJMp1817104
9. Ravaud P, Ouay F le, Depaulis E, Huckert A, Vegreville B, Tran VT. Reconfiguring health services to reduce the workload of caregivers during the COVID-19 outbreak using an open-source scalable platform for remote digital monitoring and coordination of care in hospital Command Centres. Published March 12, 2020. Accessed April 20, 2023. https://arxiv.org/abs/2003.05873
10. Herriman M, Meer E, Rosin R, Lee V, Washington V, Volpp K. Asked and answered: building a chatbot to address Covid-19-related concerns. N*EJM Catalyst*. Published June 18, 2020. Accessed April 20, 2023. https://catalyst.nejm.org/doi/full/10.1056/CAT.20.0230
11. Almalki M, Azeez F. Health chatbots for fighting COVID-19: a scoping review. *Acta Inform Med*. 2020;28(4):241-247. doi:10.5455/aim.2020.28.241-247
12. Amiri P, Karahanna E. Chatbot use cases in the COVID-19 public health response. *J Am Med Inform Assoc*. 2022;29(5):1000-1010. doi:10.1093/jamia/ocac014
13. Ministério da Saúde, Brasil. Orientações Para Manejo De Pacientes Com COVID-19. Published April 22, 2021. Accessed April 20, 2023. https://www.gov.br/saude/pt-br/coronavirus/publicacoes-tecnicas/recomendacoes/orientacoes-para-manejo-de-pacientes-com-covid-19/@@download/file
14. Rohland BM, Kruse GR, Rohrer JE. Validation of a single-item measure of burnout against the Maslach Burnout Inventory among physicians. *Stress Health*. 2004;20(2):75-79. doi:10.1002/smi.1002
15. Maslach C, Jackson S, Leiter M. *Maslach Burnout Inventory*. 4th ed. Mind Garden; 2017.
16. Leiter MP, Maslach C. Latent burnout profiles: a new approach to understanding the burnout experience. *Burn Res*. 2016;3(4):89-100. doi:10.1016/j.burn.2016.09.001
17. Maslach C, Schaufeli WB, Leiter MP. Job burnout. *Annu Rev Psychol*. 2001;52(1):397-422. doi:10.1146/annurev.psych.52.1.397
18. Zhang H, Dimitrov D, Simpson L, et al. A web-based, mobile-responsive application to screen health care workers for COVID-19 symptoms: rapid design, deployment, and usage. *JMIR Form Res*. 2020;4(10):e19533. doi:10.2196/19533
19. Judson TJ, Odisho AY, Young JJ, et al. Implementation of a digital chatbot to screen health system employees during the COVID-19 pandemic. *J Am Med Inform Assoc*. 2020;27(9):1450-1455. doi:10.1093/jamia/ocaa130
20. Shoja E, Aghamohammadi V, Bazyar H, et al. COVID-19 effects on the workload of Iranian healthcare workers. *BMC Public Health*. 2020;20(1):1636. doi:10.1186/s12889-020-09743-w

Bruno Nascimento Moreira, MD, MSc, is the Head of Oryx Lab Innovation Center, Faculdade Santa Casa BH, Belo Horizonte, MG, Brazil; **Alexandre Sampaio Moura, MD, MPH, PhD,** is a Professor, Faculdade Santa Casa BH, Belo Horizonte, MG, Brazil; **Aleida Nazareth Soares, PhD,** is a Professor, Faculdade Santa Casa BH, Belo Horizonte, MG, Brazil; **Zilma Silveira Nogueira Reis, MD, PhD,** is Faculty of Medicine, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil; and **Rosa Malena Delbone, MD, PhD,** is Academic Director, Faculdade Santa Casa BH, Belo Horizonte, MG, Brazil.

Funding: This work was supported by the Foundation for Research Support of the State of Minas Gerais, Minas Gerais, Brazil (grant #APQ-00218-20).

Conflict of interest: The authors declare they have no competing interests.

Corresponding author: Alexandre Sampaio Moura, MD, MPH, PhD, Faculdade Santa Casa BH, Belo Horizonte, MG, Brazil, alexandremoura@faculdadesantacasabh.edu.br, Twitter @alexandresmoura

Received November 30, 2022; revisions received February 22, 2023, and April 14, 2023; accepted April 17, 2023.