Validação de um chatbot como ferramenta para acompanhamento da dor crônica

Validation of a chatbot as a tool for monitoring of chronic pain

Validación de un chatbot como herramienta de seguimiento del dolor crónico

Vagner Zeizer Carvalho Paes[1], Thaís da Silva Veriato[1], Ana Carolina de Magalhães[1,4], Nathali Cordeiro Pinto[1,5], Marcelo Victor Pires de Sousa[1, 2, 3], Thereza Cury Fortunato[1]

1 Bright Photomedicine Ltda., Edifício CIETEC, Av. Prof. Lineu Prestes, 2242, São Paulo-SP, Brasil.
2 Departamento de Física, Universidade Federal do Ceará (UFC), Fortaleza-CE, Brasil.
3 D'Or Instituto de Pesquisa e Ensino (IDOR), Rio de Janeiro-RJ, Brasil.
4 Insper Instituto de Ensino e Pesquisa, São Paulo-SP, Brasil.
5 Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, Dr Enéas de Carvalho Aguiar, 44, São Paulo-SP, Brasil.
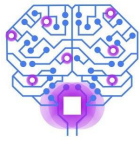
Autor correspondente: Vagner Zeizer Carvalho Paes.
E-mail: vagner.paes@brightmed.com.br

Resumo

Neste trabalho estudamos o uso de um chatbot como ferramenta de monitoramento da dor crônica. Foi realizada anamnese em 28 pacientes e, posteriormente, a resposta da intensidade da dor de cada paciente foi coletada pelo terapeuta e pelo chatbot. Foi obtida uma forte correlação de 0,94 entre a intensidade da dor coletada pelo terapeuta e o chatbot. Observamos que 50% das respostas ao chatbot foram registradas cerca de 30 minutos após o envio da mensagem. Pacientes na faixa etária de 30-60 anos responderam mais rápido do que outros. Com relação ao gênero, os pacientes do sexo masculino responderam mais rápido ao chatbot em média. Na medida em que as respostas foram registradas ao longo dos dias pelo chatbot, as respostas registradas diminuíram de forma aproximadamente linear. Validamos o chatbot como um monitoramento eficiente da intensidade da dor, sendo uma ferramenta de fácil interação e boa adesão por parte dos pacientes.

Descritores: Dor Crônica; Evolução Clínica; Sistemas Computacionais.
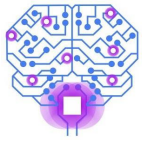
Abstract

In this work we studied the use of a chatbot as a monitoring tool for chronic pain. An anamnesis was performed on 28 patients, and afterward, the pain intensity response of each patient was collected by the therapist and the chatbot. It was obtained a strong correlation of 0.94 between the pain intensity collected by the therapist and the chatbot. We noted that 50% of the answers to the chatbot were recorded about 30 minutes after the message was sent. Patients in the age range of 30-60 years responded quicker than others. With respect to gender, male patients answered quicker the chatbot on average. As far as the answers were recorded over the days by the chatbot, the answers recorded diminished approximately linearly. We validated the chatbot as an efficient monitoring of the pain's intensity, being an easy interactive tool showing good adhesion by the patients.

Keywords: Chronic Pain; Clinical Evolution; Computer Systems.

Resumen

En este trabajo estudiamos el uso de un chatbot como herramienta de monitorización del dolor crónico. Se realizó una anamnesis a 28 pacientes, y posteriormente, el terapeuta y el chatbot recogieron la respuesta de intensidad del dolor de cada paciente. Se obtuvo una fuerte correlación de 0,94 entre la intensidad del dolor recogida por el terapeuta y el chatbot. Notamos que el 50% de las respuestas al chatbot se registraron unos 30 minutos después de que se envió el mensaje. Los pacientes en el rango de edad de 30 a 60 años respondieron más rápido que otros. Con respecto al género, los pacientes masculinos respondieron más rápido al chatbot en promedio. En cuanto a las respuestas registradas a lo largo de los días por el chatbot, las respuestas registradas disminuyeron aproximadamente de forma lineal. Validamos el chatbot como un eficiente monitor de la intensidad del dolor, siendo una herramienta interactiva fácil que muestra una buena adherencia por parte de los pacientes.

Descriptores: Dolor Crónico; Evolución Clínica; Sistemas de Computación.

Introduction

Pain is an unpleasant sensory symptom caused by several factors, and it is the main reason for seeking medical care and can be classified as acute or chronic pain (1). Acute pain is defined as a dynamic symptom that arises because of an inflammatory process after tissue injury, and usually this pain does not last more than three months. However, when this pain evolves into a condition of persistence which lasts more than three months, according to the International Classification of Diseases (11th edition), it is defined as a chronic pain associated with changes in the peripheral nervous system and in the central nervous system, which can cause disability, morbidity and generate costs in the hospital environment (2).

Related to disability, chronic pain is associated with the pathologies which can cause this condition, namely back, neck, and musculoskeletal pain. The center for Control and Prevention of Diseases in the United States of America reported a study of the variable prevalence from 11% to 40% of chronic pain (3). A review performed in the United Kingdom reported a prevalence rate of 43.5%, and when associated with moderate pain to extreme pain, it varied from 10.4% to 44.3% (4). In Brazil, Aguiar and colleagues performed a systematic review and found an average prevalence of 45.59%, predominantly in females (5).

Worldwide the prevalence of chronic pain is high and the study of gender, age, pain intensity, and therapeutic response in individuals affected by this condition is important developing new perspectives of treatments focusing on the care and management of pain. In this context, technological interventions have obtained visibility due to the use of the internet, mobile applications, or the fact that computers provide easy access, low cost, flexibility, and self-management in healthcare (6,7). Moreover, these interventions can provide for the patient a better management of his health, allowing the patient to become more proactive in the treatment when monitoring the evolution of his clinical condition (8).

Among the technological interventions, computer programs that interact with humans through chats (conversational agents) are appealing in the healthcare sector due to their acceptance and association with other hospital branches, such as the therapist (9). One of the conversational agents used in the chatbot is a system capable of interacting
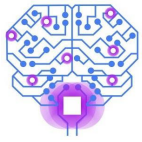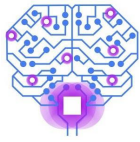
with the user by employing a chat connected to a message sender. The use of this technology in the healthcare sector for mapping the condition of the patient occurs to improve the health service (10). Hauser-Ulrich and colleagues developed the painSELFMANagement (SELMA) tool, a chatbot used for self-monitoring of chronic pain, which helped healthcare professionals to deliver psycho-therapeutic interventions in an empathetic way based on texts and media. The results related to the change in behavior and pain intensity were both positive. Moreover, the interaction between the SELMA's application and the patients has shown good adhesion, and usage by the patients (6). Goldenthal and colleagues also evaluated the usage of a chatbot by patients submitted to ureteroscopy due to symptoms and complications after this intervention. In this study, the chatbot has shown benefits related to the release of the concerns related to the symptoms (11).

Therefore, this work aimed to validate the use of chatbot as a tool for monitoring the evolution of the pain of the patients, correlate the answer obtained by the therapist and the ones obtained by the chatbot, analyze the delay to answer for each patient, the profile of adhesion concerning patient's gender, age, and skin phototype as well. So, by knowing the behavior of the patients' pain level and their clinical profiles, it is possible to improve the therapeutical potential according to clinical conditions of the patient. Therefore, providing a better management of the patient's pain along the treatment.

Methods

In this study, 28 patients diagnosed with chronic back pain were recruited at NotreDame Intermédica (SP, Brazil) to perform customized photobiomodulation therapy sessions using the Light-Aid device (Bright Photomedicine, SP, Brazil). This equipment is composed of LEDs with a wavelength corresponding to 850±15 nm. Invitare Pesquisa Clinica Auditoria e Consultoria LTDA Ethics Committee approved the study under the number CAAE 51366221.0.0000.8098. Patients agreed to participate in the study and signed a Free and informed consent form. First, the patients underwent anamnesis and, later, to monitor the evolution of the patient's pain, the therapist collected the answers regarding the patients' pain intensity through a form and by the chatbot service. Daily, the

chatbot sent a message in WhatsApp Messenger at 9 AM to the patients asking for their pain level at that moment, on a scale level ranging from 0 to 10. If the patient did not provide a valid answer, i.e., an integer number between 0 and 10, the chatbot would send again a message asking for their pain level. Figure 1 shows an example dialog of the chatbot with a patient asking for her pain level at the moment the message was sent. Then, the data were processed and analyzed using the Python 3 programming language and its libraries.
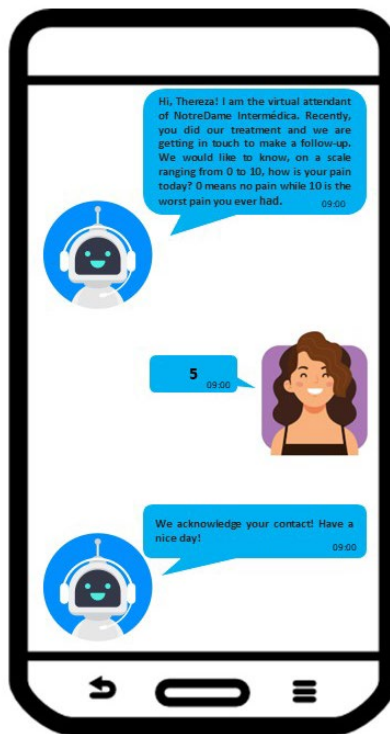

Figure 1 – Illustration of a dialog of the chatbot with a patient asking for the patient's pain level at the moment the message was sent. The conversation presented here was translated from Brazilian Portuguese into English.

- **Patient's Profiles**

Figure 2 shows the patient's profile, acquired during the anamnesis of these patients. Graphs show distributions concerning age, skin phototype (in the Fitzpatrick scale (12)), gender, and BMI. With respect to the patient's age, the overall average was
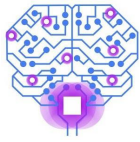
48.8 ± 11.0 years, as shown in Figure 2A. The most predominant skin phototypes were 3, 4, and 5, ranked according to the Fitzpatrick scale, as shown in Figure 2B. The study consisted of 15 male and 13 female patients (Figure 2C). Using the World Health Organization criteria to categorize BMI (13), the current study's patients were classified as shown in Figure 2D.
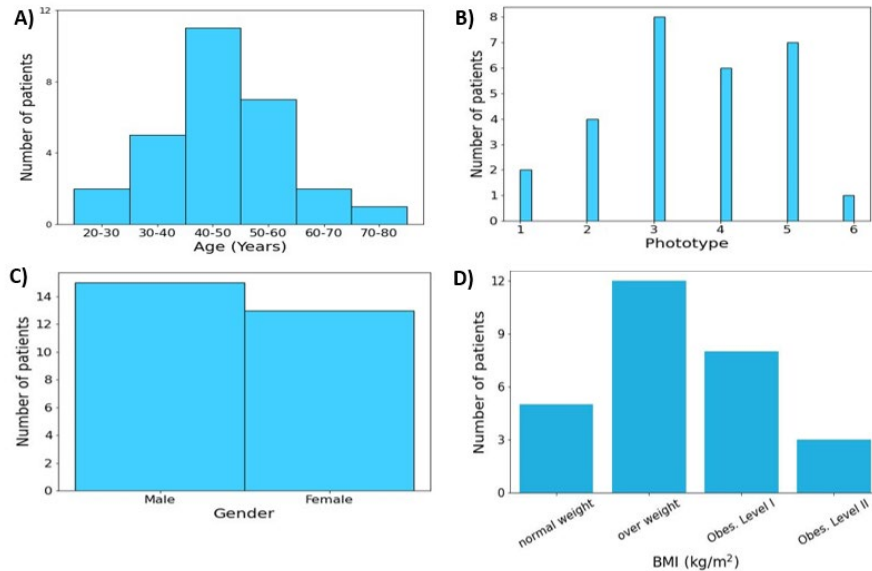

Figure 2 – Anamnesis of the patients of the current study. A) Age. B) Phototype. C) Gender. D) BMI.

In a recent study, David and colleagues showed that the prevalence of lumbar pain in Brazil concerning gender seems to be more common in women than in men, however, there was not a significant statistical difference in the population studied (14). In the current work, our sample is composed of 15 men and 13 women, and the same behavior was observed. The higher prevalence of this pathology was also associated with aging and weight, which corroborates the results that we found that most of the patients were in the range of 40-60 years, and about BMI, the incidence observed was between overweight and obesity level I. Bento and colleagues analyzed 600 individuals on the prevalence of lumbar pain, and regarding the skin color of the patients studied, most of the patients had white skin, what was not observed in this study, in which most of the patients were not white (phototype 3, 4 and 5) according to the Fitzpatrick scale (12,15).
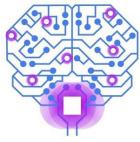
Results and Discussion

After anamnesis was performed in all patients, they were submitted to the evaluation of pain intensity by using Verbal Numerical Rating Scale (vNRS) during the sessions, as well as in a daily process utilizing a chatbot service, to evaluate the pain intensity during the treatment and to verify the accuracy of the patient's response to the therapist during the sessions. Figure 3 shows the correlation between the responses collected by the therapist and by the chatbot with respect to the patient's vNRS. The Pearson's r correlation coefficient was 0.94, indicating a strong correlation between the response obtained by the therapist and the chatbot service.
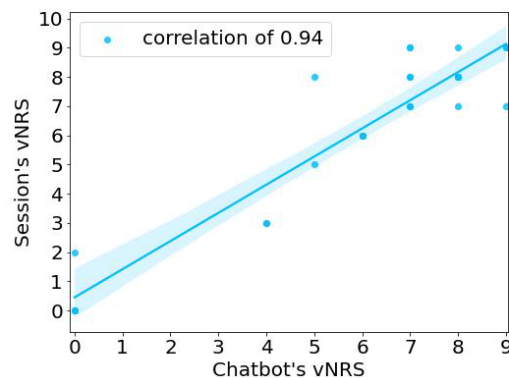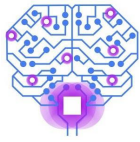

Figure 3 – Correlation between the vNRS responses assessed by the therapist and chatbot.

Cateb and colleagues reported a pilot study of the validation of a chatbot in teleassistance of COVID-19, in which the symptoms of the patients were evaluated before the teleconsulting and compared to the respective medical report. Dyspnea was the most frequent symptom and the reliability of the response for the chatbot as well as for the medic was significant and moderate, with a Kappa equal to 0.605 (16). This result corroborates the high linear correlation (0.94) found here between the patient's vNRS collected by the therapist and the chatbot's service, indicating that the chatbot can be an adequate tool to monitor the pain of the patients during the treatment of low back pain.

Figure 4A shows a normalized histogram (bin equal to ten minutes) of patients' delay to respond up to 12 hours after sending the message via the chatbot, which corresponds to 63% of the responses obtained, and only 0.65% of the responses were collected between 12 and 24 hours after the chatbot sends the message. The remaining

36% corresponds to patients who did not answer the chatbot during the study period. It is important to note from Figure 4A that approximately 17.5% of the answers were recorded in the first 10 minutes and 50% of the answers were recorded until 34 minutes after the chatbot's message was sent to the patients. Figure 4B shows the violin plot of the delay to answer of the patients, showing that the answers are mostly recorded around 30 minutes, and answers after approximately 4 hours are quite rare.
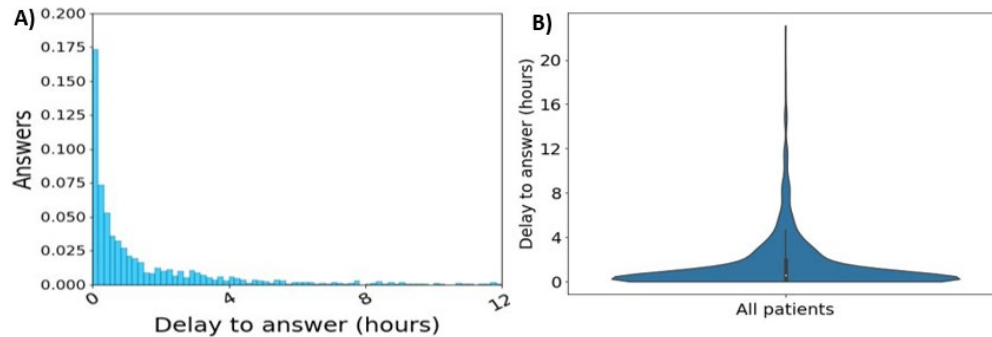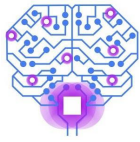

Figure 4 – A) Normalized histogram of the delay to answer in hours of all the patients; B) Violin plot of the delay to answer of all patients.

In this study, the interaction between the patients and the chatbot was in most times performed in a fast way, suggesting that the interest in the patient's pain can demonstrate some special care for his condition, making him more comfortable and, consequently, this augments the interest of the patient to use this tool.

Figure 5A shows a histogram of the answers given by the age range, as well as the number of days in which no answer was recorded for each patient. From this figure, it can be seen that younger patients tend to give more responses to the chatbot's system, and older fewer responses, especially patients in the 70-80 years range. Figure 5B shows a violin plot of the delay to answer in hours of the patients to answer the chatbot's service, showing that patients with ages in the range of 30-60 tend to respond quicker, but patients with ages in the range of 20-30 and 70-80 years take longer to answer. It is noteworthy that the results presented here for the age ranges of 20-30 and 70-80 might not represent the true behavior of these patients, because of the very low number of patients within these age ranges. In Figure 5B, when comparing males and females in the same age range, it can be seen that female patients seem to be responding in average later than

male patients. This comparison is not possible in the range of 20-30 years as this group has only males. The same happens to the 70-80 years group, which has only females.
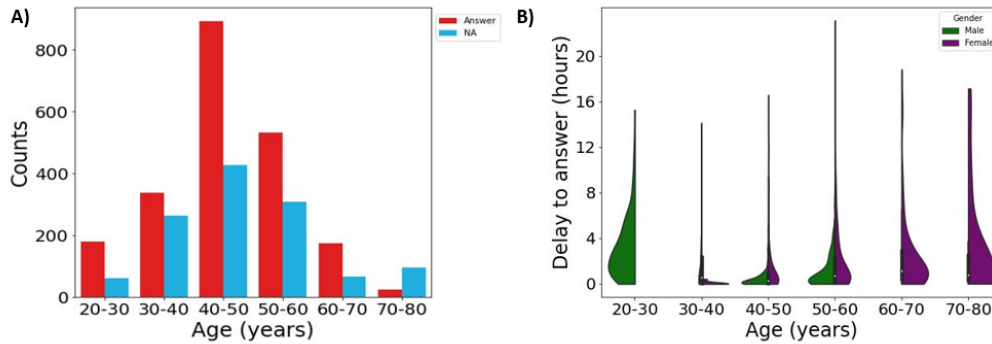

Figure 5 – A) Histogram of the answer per age range, as well as number of days in which there was no response recorded for each patient. **B)** Violin plot of the delay to answer to the chatbot's message per age range and separated by patient's gender.

Figure 6A shows a histogram of the answers received by the chatbot's service and the absence of answers recorded for each patient as a function of gender. It can be seen that male patients are more prone to respond to the chatbot's service than female patients. Figure 6B shows a violin plot of the delay in answering the chatbot's service message, from which it can be seen that male patients respond quicker than females.
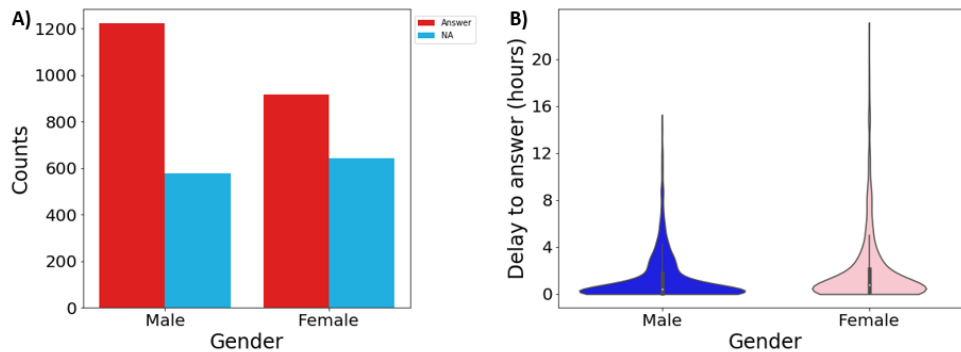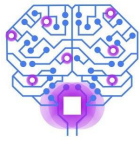

Figure 6 – A) Histogram of the answers given to the chatbot's service, as well as the absence of answers recorded for each patient, as a function of patient's gender. **B)** Violin plot of the delay to answer to the chatbot's service message as a function of patients' gender.

Fan and colleagues analyzed the interaction between 16.519 users with a chatbot used for detecting health problems. In this study, 54.80% of the interactions were

addressed to male users, and regarding the age, there was a larger interaction of the users between 20-39 years range (17). A small part of the elder patients (age greater than 60) also interacted with the chatbot. In the current study, it was possible to observe a higher interaction with the chatbot among the male patients, once this group responded quicker to the chatbot than females. However, the age range with the most significant interaction was found to be 40-60 years, indicating that they adhere better to the chatbot's technology than younger (age less than 40 years) and the elder (ages greater than 60 years) patients.

Figure 7 shows a scatter plot as a function of the day from the beginning to the end of the study, the number of answers recorded by the chatbot, as well as the absence of answers to the chatbot's question in each day of the study period. It can be seen that answers given by the patients to the chatbot's service decrease approximately linearly as the number of days increases, and, on the other hand, the number of no answers per patient per day assumes approximately a linear increase, regardless of some peaks in these curves, one of which correspond to New Year's Day (day 20). The calculated patient's adhesion prior to the treatment (black dotted line shown in Figure 7) was found to be 74% on average, whereas the patient's adhesion after the end of the sessions was found to be 62% on average. These results show that the patients exhibit good adhesion to the chatbot's service even after the end of the treatment, suggesting that the chatbot might be a powerful tool for monitoring the patients' pain.
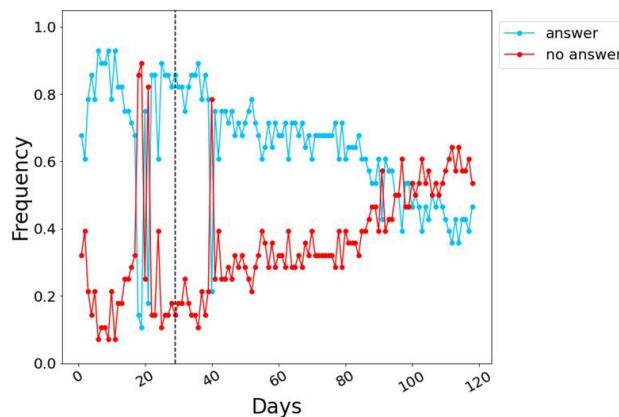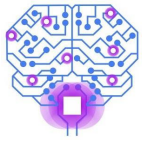

Figure 7 – Scatter plot of the number of responses collected by the chatbot, as well as the absence of answers to the chatbot, per day.

Conclusion

We have presented a data analysis of the use of a chatbot system as a tool to monitor the pain intensity of patients undergoing low back pain clinical conditions. The results suggest that female patients take longer to respond to the chatbot, and ages in the range of 30-60 years respond quicker to the chatbot, regardless of patient's gender. A strong correlation between the chatbot's vNRS and sessions' vNRS indicates that the pain intensity recorded by the chatbot is very reliable. Due to the small number of patients and problems that occurred in the execution of the study during the pandemic, the results presented here are preliminary regarding the delay to answer, and adhesion of the patients by gender, BMI, age, and phototype. Therefore, in future studies we intend to implement this follow-up with a larger number of patients, including people with different ages, widening the age distribution. Furthermore, we intend to test new methodologies to provide better interaction and adhesion between the user and the chatbot as well. Moreover, we want to apply supplementary questions to complement the database of the evolution of the pain of the patients. In this way, we can fine-tune the therapy according to the clinical condition of the patients.
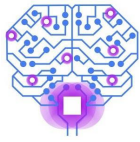
Acknowledgments

References

1. Vasconcelos FH, Araújo GC de. Prevalence of chronic pain in Brazil: a descriptive study. Brazilian Journal Of Pain [Internet]. 2018 [citado 25 de julho de 2022];1(2). Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2595-31922018000200176&lng=en&nrm=iso&tlng=en

2. Cohen SP, Vase L, Hooten WM. Chronic pain: an update on burden, best practices, and new advances. The Lancet. maio de 2021;397(10289):2082–97.
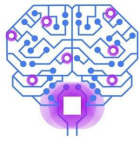
3. Dahlhamer J, Lucas J, Zelaya C, Nahin R, Mackey S, DeBar L, et al. Prevalence of Chronic Pain and High-Impact Chronic Pain Among Adults - United States, 2016. MMWR Morb Mortal Wkly Rep. 14 de setembro de 2018;67(36):1001–6.

4. Fayaz A, Croft P, Langford RM, Donaldson LJ, Jones GT. Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. BMJ Open. 20 de junho de 2016;6(6):e010364.

5. Aguiar DP, Souza CP de Q, Barbosa WJM, Santos-Júnior FFU, Oliveira AS de. Prevalence of chronic pain in Brazil: systematic review. Brazilian Journal Of Pain [Internet]. 2021 [citado 26 de julho de 2022]; Disponível em: http://www.scielo.br/scielo.php?script=sci_issues&pid=2595-3192&lng=en&nrm=isso

6. Hauser-Ulrich S, Künzli H, Meier-Peterhans D, Kowatsch T. A Smartphone-Based Health Care Chatbot to Promote Self-Management of Chronic Pain (SELMA): Pilot Randomized Controlled Trial. JMIR Mhealth Uhealth. 3 de abril de 2020;8(4):e15806.

7. Vugts MAP, Joosen MCW, van der Geer JE, Zedlitz AMEE, Vrijhoef HJM. The effectiveness of various computer-based interventions for patients with chronic pain or functional somatic syndromes: A systematic review and meta-analysis. Yang J, organizador. PLoS ONE. 16 de maio de 2018;13(5):e0196467.

8. Nicholl BI, Sandal LF, Stochkendahl MJ, McCallum M, Suresh N, Vasseljen O, et al. Digital Support Interventions for the Self-Management of Low Back Pain: A Systematic Review. J Med Internet Res. 21 de maio de 2017;19(5):e179.

9. Bickmore TW, Mitchell SE, Jack BW, Paasche-Orlow MK, Pfeifer LM, Odonnell J. Response to a Relational Agent by Hospital Patients with Depressive Symptoms. Interact Comput. 1º de julho de 2010;22(4):289–98.

10. Chung K, Park RC. Chatbot-based heathcare service with a knowledge base for cloud computing. Cluster Comput. janeiro de 2019;22(S1):1925–37.

11.   Goldenthal SB, Portney D, Steppe E, Ghani K, Ellimoottil C. Assessing the feasibility of a chatbot after ureteroscopy. Mhealth. 2019;5:8.

12.   Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. Archives of Dermatology. 1º de junho de 1988;124(6):869–71.

13.   Weltgesundheitsorganisation. The SuRF report 2: surveillance of chronic disease risk factors : country-level data and comparable estimates. Geneva: World Health Organization; 2005.

14.   de David CN, Deligne L de MC, da Silva RS, Malta DC, Duncan BB, Passos VM de A, et al. The burden of low back pain in Brazil: estimates from the Global Burden of Disease 2017 Study. Popul Health Metrics. setembro de 2020;18(S1):12.

15.   Bento TPF, Genebra CV dos S, Maciel NM, Cornelio GP, Simeão SFAP, Vitta A de. Low back pain and some associated factors: is there any difference between genders? Brazilian Journal of Physical Therapy. janeiro de 2020;24(1):79–87.

16.   Cateb GF, Amaral S, Gonçalves SCL, Oliveira IJR, Prates RO, Chagas BA, et al. Estudo piloto de validação de um chatbot de rastreamento, implementado para direcionar a teleassistência em COVID-19. Em: Anais Estendidos do XXI Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS Estendido 2021) [Internet]. Brasil: Sociedade Brasileira de Computação (SBC); 2021 [citado 2 de agosto de 2022]. p. 97–102.   Disponível   em: https://sol.sbc.org.br/index.php/sbcas_estendido/article/view/16108

17.   Fan X, Chao D, Zhang Z, Wang D, Li X, Tian F. Utilization of Self-Diagnosis Health Chatbots in Real-World Settings: Case Study. J Med Internet Res. 6 de janeiro de 2021;23(1):e19928.