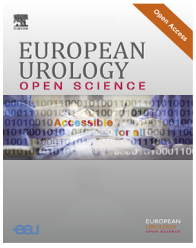
available at www.sciencedirect.com
journal homepage: www.eu-openscience.europeanurology.com



European Association of Urology

Prostate Cancer

Effectiveness of the Medical Chatbot PROSCA to Inform Patients About Prostate Cancer: Results of a Randomized Controlled Trial

Kilian Baumgärtner [a], *Michael Byczkowski* [b], *Tamara Schmid* [b], *Marc Muschko* [b], *Philipp Woessner* [b], *Axel Gerlach* [b], *David Bonekamp* [c], *Heinz-Peter Schlemmer* [c], *Markus Hohenfellner* [d], *Magdalena Görtz* [d,e,*]

[a] *Medical Faculty, Ruprecht-Karls University of Heidelberg, Heidelberg, Germany;* [b] *SAP SE, Walldorf, Germany;* [c] *Department of Radiology, German Cancer Research Center (DKFZ), Heidelberg, Germany;* [d] *Department of Urology, Heidelberg University Hospital, Heidelberg, Germany;* [e] *Junior Clinical Cooperation Unit 'Multiparametric Methods for Early Detection of Prostate Cancer', German Cancer Research Center (DKFZ), Heidelberg, Germany*

Article info

Article history:
Accepted August 30, 2024

Associate Editor:
Roderick van den Bergh

Keywords:
Artificial intelligence
Chatbot
Early detection
Large language model
Natural language processing
Prostate cancer
Randomized controlled trial

Abstract

Background and objective: Artificial intelligence (AI)-powered conversational agents are increasingly finding application in health care, as these can provide patient education at any time. However, their effectiveness in medical settings remains largely unexplored. This study aimed to assess the impact of the chatbot "PROState cancer Conversational Agent" (PROSCA), which was trained to provide validated support from diagnostic tests to treatment options for men facing prostate cancer (PC) diagnosis.
Methods: The chatbot PROSCA, developed by urologists at Heidelberg University Hospital and SAP SE, was evaluated through a randomized controlled trial (RCT). Patients were assigned to either the chatbot group, receiving additional access to PROSCA alongside standard information by urologists, or the control group (1:1), receiving standard information. A total of 112 men were included, of whom 103 gave feedback at study completion.
Key findings and limitations: Over time, patients' information needs decreased significantly more in the chatbot group than in the control group ($p = 0.035$). In the chatbot group, 43/54 men (79.6%) used PROSCA, and all of them found it easy to use. Of the men, 71.4% agreed that the chatbot improved their informedness about PC and 90.7% would like to use PROSCA again. Limitations are study sample size, single-center design, and specific clinical application.
Conclusions and clinical implications: With the introduction of the PROSCA chatbot, we created and evaluated an innovative, evidence-based AI health information tool as an additional source of information for PC. Our RCT results showed significant benefits of the chatbot in reducing patients' information needs and enhancing their understanding of PC. This easy-to-use AI tool provides accurate, timely, and accessible support, demonstrating its value in the PC diagnosis process. Future steps

* Corresponding author. Department of Urology, Heidelberg University Hospital, Im Neuenheimer Feld 420, Heidelberg 69120, Germany. Tel. +49 6221 568820; Fax: +49 6221 565188.
E-mail address: magdalena.goertz@med.uni-heidelberg.de (M. Görtz).

https://doi.org/10.1016/j.euros.2024.08.022

include further customization of the chatbot's responses and integration with the existing health care systems to maximize its impact on patient outcomes.

Patient summary: This study evaluated an artificial intelligence–powered chatbot—PROSCA, a digital tool designed to support men facing prostate cancer diagnosis by providing validated information from diagnosis to treatment. Results showed that patients who used the chatbot as an additional tool felt better informed than those who received standard information from urologists. The majority of users appreciated the ease of use of the chatbot and expressed a desire to use it again; this suggests that PROSCA could be a valuable resource to improve patient understanding in prostate cancer diagnosis.

1. Introduction

Prostate cancer (PC) is the second most frequent malignancy among men, accounting for 27% of all cancer diagnoses in males [1,2]. In low-income countries, PC was the number one cause of cancer death in males in 2022 [3]. Recent studies have shown that screening for prostate-specific antigen (PSA) has a long-term mortality benefit [4]. Patients who receive adequate and comprehensible information about their condition and treatment options report higher satisfaction levels [5]. However, there is evidence that the quality of PC information available online is moderate at best and often challenging to utilize [6].

Artificial intelligence (AI)-driven chatbots as validated informative conversational programs could contribute to the solution. These are part of a growing technological field of AI called conversational agents that can communicate with users in a human way [7]. Chatbots have potential for use in diagnostics [8], and are expected to benefit patients in disease management as well as physical, mental, and behavioral outcomes [9].

However, research is still in the early stages of gaining sufficient knowledge about the use of chatbots in medicine. Most studies have focused on mental health [10]. Randomized controlled trials (RCTs) are rare; many studies were simple proof–of-concept studies [10]. The effectiveness of chatbots, for example, concerning clinical endpoints, was investigated only in a few studies [11–13].

Recently, ChatGPT, a large language model (LLM) developed by OpenAI, gained increasing attention from urologists and has the potential to be a source of information for health issues [14–16]. A recently published study evaluated the quality of information of ChatGPT responses for urology patients compared with those of a urologist. However, it found poor-quality responses regarding urology cases, with only 52% of responses deemed appropriate [17].

A professional information service that supports patients from screening through diagnosis to treatment is highly desirable in the early detection of PC. A reliable source of information that is always accessible could increase patients' understanding throughout their diagnostic process and satisfy the high information need of the patients. This study investigated the effectiveness of the customized chatbot PROState cancer Conversational Agent (PROSCA) as an additional validated educational tool to the standard medical information in the context of PC diagnostics in an RCT.

2. Patients and methods

2.1. Development of the chatbot

An interprofessional team of the Urology Clinic of Heidelberg University Hospital, the German Cancer Research Center, and SAP SE developed the chatbot PROSCA (Fig. 1), which was evaluated in an exploratory RCT. The development was carried out using the web interface SAP Conversational AI (SAP CAI) [18], a low-code chatbot platform that allows users to utilize the variable AI or natural language processing (NLP) model for personalized purposes, as described previously [19]. In brief, the SAP CAI employs a hybrid approach of prebuilt NLP models and adaptable machine learning models to decode and interpret users' input. These models are mainly based on deep learning algorithms and are trained on extensive datasets to precisely extract information from text inputs. Contrary to LLMs, these custom machine learning models offer the flexibility of being trained with context-related data [20]. This facilitates the development of a chatbot customized to and validated for its particular field, capable of reacting to, for example, PC-specific terminology. The chatbot's operation field was defined to accompany patients during the diagnostic process of PC and answer questions, for example, about prostate diseases, diagnostic procedures, and treatment options for PC, as well as to support the patient from admission to the clinic to symptom checking at home.

The chatbot "PROSCA" is based on an expert system designed to provide responses predetermined by experts. Enrolled patients do not receive incorrect information from the chatbot, as the responses have been formulated and thoroughly verified by certified expert urologists. This procedure ensures the accuracy and reliability of the information provided to the patients. The chatbot's fixed responses were formulated and evaluated by two senior urologists (M.G. and M.H.) in accordance with published highest-quality scientific evidence including the European Association of Urology (EAU) guidelines [21]. Seven urologists and urology residents tested the chatbot with common questions of PC patients, and a prototype test was

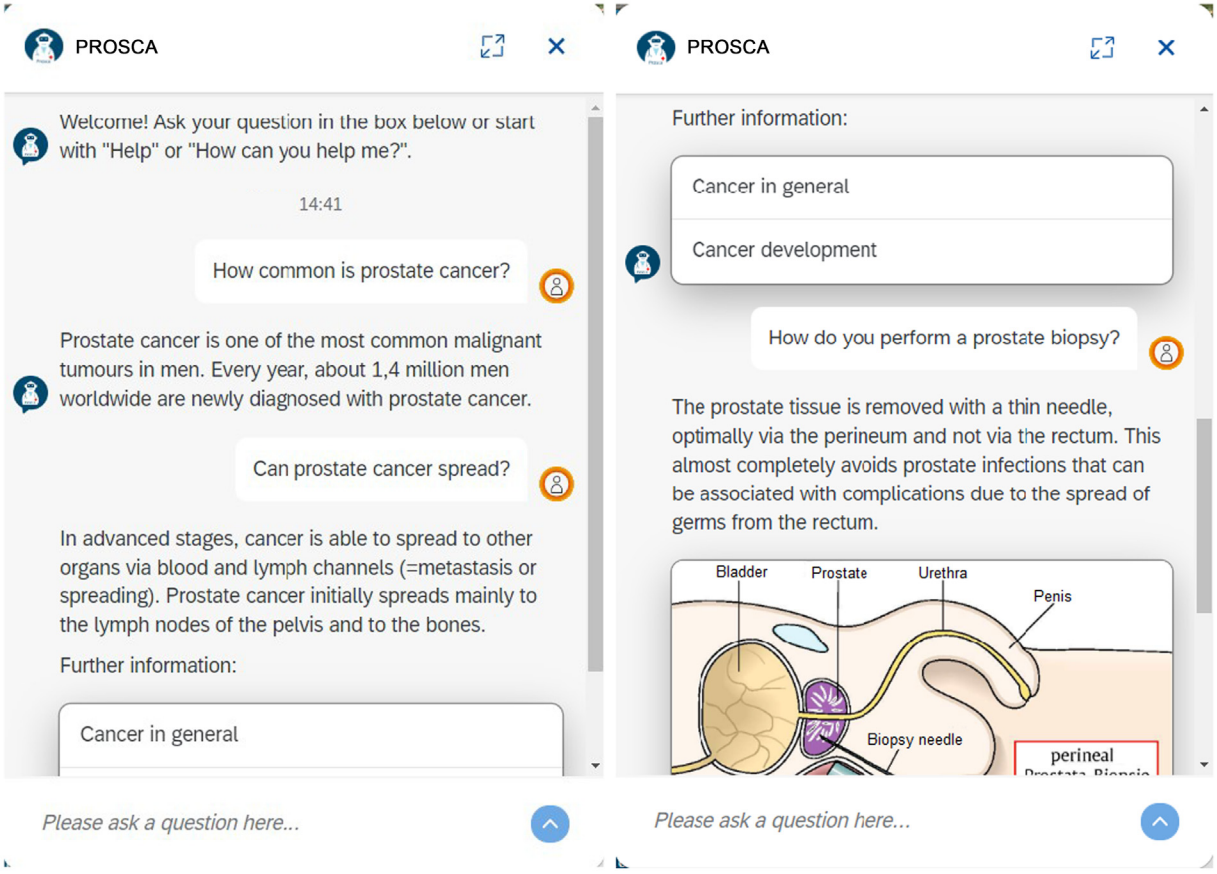


Fig. 1 – PROSCA's chat design and exemplary conversation. Chat interface design, exemplary text message input, and corresponding text message as well as image and button choice output are shown. The chatbot has been translated into English; the original chatbot pages can be found in Supplementary Figure 1. PROSCA = PROState cancer Conversational Agent.

conducted with ten patients. The chatbot uses a combination of free text input and predefined navigation fields to cater to users' different preferences, as reported in previous studies [10]. The chatbot provides its response via text message, images, web links, as well as e-mail and call forwarding.

2.2. *Study design and recruitment*

Patient recruitment was conducted prospectively from December 2021 to June 2022. Patients were recruited in this RCT when they first presented at the Urology Clinic of Heidelberg University Hospital due to a PC suspicion because of a PSA elevation or positive digital rectal examination. The study inclusion criteria were the regular use of a laptop, computer, smartphone, or tablet, and sufficient command of the German language. Following the inclusion of patients and the provision of written informed consent, randomization was conducted to guarantee an even distribution of the chatbot to selected patients.

The sample size for the RCT was determined in close collaboration with the Institute for Medical Biometrics at Heidelberg University Hospital through a thorough statistical calculation process. This process aimed to ensure that the study was powered adequately to detect a clinically meaningful effect. The key considerations included defining the primary outcome, estimating the expected difference between the intervention and control group, and approximating the dropout rate. During the planning phase of the study in 2020, previous studies and pilot data were rare. A minimum consent sample size of $n = 30$ was recommended. However, we aimed for a larger sample size to extend the approach beyond proof-of-concept attempts conducted by other studies at that time.

The primary endpoint of this study was to assess whether patients reported a benefit from the additional provision of information by the chatbot. The secondary objectives included evaluating the suitability and user friendliness of the chatbot for patients during a hospital stay.

After randomization, the control group received routine clinical care in terms of standard information from doctors and written materials, while the chatbot group had access to the chatbot PROSCA via a web page in addition to the standard education. The standard care for patients with suspected PC, which both study groups received, included several detailed patient education modules. The PC diagnosis itself took place in a urological university hospital with extensive experience in specialized oncological diagnosis and treatment. Throughout the study, various dedicated doctor-patient consultations occurred, including those for biopsy counseling, on the day of the biopsy procedure,

and for discussing the histology results and treatment strategies. In addition to the information provided by urologists, patients benefited from the expertise of other highly trained professionals, such as specialized nurses. All patients were offered a wide range of informational materials, including brochures on early detection of PC from the German Cancer Aid, a reference to the Cancer Information Service of the German Cancer Research Center, and the informed consent form for prostate biopsy.

A patient-focused questionnaire was handed to the patients on the day of enrolment and after they received the histological results of their prostate biopsy. The evaluation was primarily based on a four-point Likert scale ranging from "strongly agree" to "disagree". As there is currently no established questionnaire for chatbots, a questionnaire was adapted based on the MAUQ questionnaire for patient apps [22].

2.3. Ethical and organizational framework

Data were collected prospectively. The study was approved by the ethical committee of the University of Heidelberg (approval no. S-005/2021) and registered in the German Clinical Trials Register (DRKS00034484). All study participants provided written informed consent. Patients were informed that the use of the chatbot cannot replace a personal, medical consultation, especially in cases of acute complaints. The chatbot was accessed via a customized website and was hosted on Heidelberg University Hospital infrastructure.

2.4. Statistical analysis

Patient characteristics were evaluated with a basic descriptive analysis. For group comparison, two-sided *t* test and Pearson's chi-square test were used. Analyses were conducted using Microsoft Excel and IBM SPSS Statistics version 29.0.

3. Results

3.1. Recruitment of study patients

A total of 145 patients were screened for participation in the study, of whom 112 men were enrolled (Fig. 2). The most common reason for exclusion was a lack of technical equipment, followed by insufficient language comprehension capabilities. Randomization completed with 59 men in the chatbot group and 53 in the control group. As nine study participants were lost during follow-up, the feedback of 103 patients was included in the analysis. Of 54 study participants in the chatbot group, 43 used the chatbot and their feedback evaluated the usability and feasibility of the chatbot.

3.2. Characteristics of the study groups

Demographic data included age, device usage, and Gleason score of the prostate biopsy. The participants in both study arms (n = 103) had a similar age distribution (p = 0.944). The median age of the study cohort was 64 yr (interquartile range 59–71). PC was detected in 53 (51.5%) patients, with no significant difference between the study groups (p = 0.135) and Gleason scores ranging from 6 to 9.

Most study participants (n = 94; 91.0%) regularly used two or more different devices to access the Internet, most frequently a computer or laptop in combination with a smartphone. Men with an age of ≥75 yr (n = 13) behaved similarly to the entire group, with 11 (74.6%) using two or three devices.

3.3. Patients' information needs in the chatbot group versus the control group

The level of information need was observed through a four-point Likert scale ranging from "1 = very high" to "4 = very low" on both the day of enrolment (T_0) and in the final questionnaire (T_1), distributed after the histopathological review of the prostate biopsy.

At the beginning of the diagnostic process (T_0), 88/103 patients (85.4%) reported a high to very high information need, with no significant differences between the groups (p = 0.098). At T_1, the information need of the whole cohort decreased in 35/103 (34.0%) patients and increased in 26/103 (25.2%) patients; the rest of the patients indicated the same level of information need.

Remarkably, the patients in the chatbot group experienced a more frequent (40.8% vs 26.5%) and more often a higher (two categories; 9.3% vs 4.1%) decrease in information need than those in the control group (Table 1), whereas information need increased in 16.7% (vs 34.7% in the control group).

When comparing the change in information need as the difference between T_1 and T_0 in terms of "decreased/constant information need" versus "increased information need", the chatbot group had a significant benefit compared with the control group (p = 0.035).

3.4. Chatbot usage

During the study period, the chatbot facilitated 176 conversations, and on average, each user generated 7.1 inputs per conversation. Of 54 patients in the chatbot group, 43 (79.6%) stated that they used PROSCA at least once. PROSCA was used by the participants on a total of 115 d, with an average of 3.4 d and a maximum of 14.0 d. According to patients, PROSCA was used most frequently (22/43 patients, 51.2%) on the days following the biopsy. The most often triggered user intents were information on the purpose of the prostate biopsy, description of the procedure of the biopsy, procedure of radical prostatectomy, possible biopsy complications, and general information on PC treatment, which together accounted for 34.4% of all 1595 recognized user intents. Figure 3 shows all intents that were triggered more than 15 times.

3.5. Chatbot usability and feasibility

None of the chatbot users had previously used a medical chatbot in a clinical setting. Of 43 chatbot users, 36 (83.7%) stated not needing any help when using the chatbot and five users (11.6%) agreed partly, with two users (4.7%) requiring ongoing help from a technically experienced person. All users rated the chatbot as easy to use.

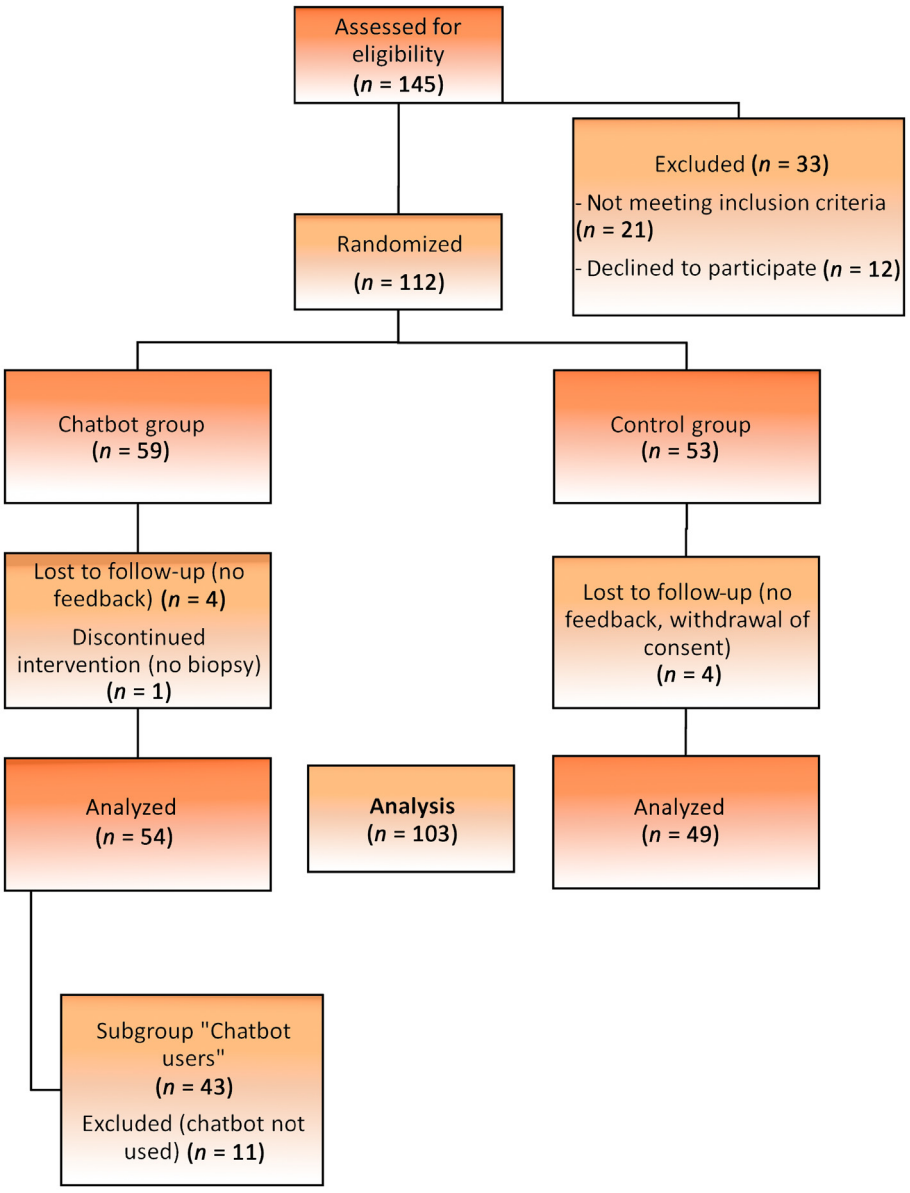


Fig. 2 – Study flowchart according to CONSORT. After inclusion criteria check, patients were randomized into the chatbot and control groups. The chatbot group was further categorized into actual chatbot users for a feasibility and usability analysis.

Table 1 – Comparison of information need and its change over time in study groups

Information need		Total group, *n* (%)	Chatbot group, *n* (%)	Control group, *n* (%)
At T_0	Very high	33 (32.0)	20 (37.0)	13 (26.5)
	High	55 (53.4)	29 (53.7)	26 (53.1)
	Low	15 (14.6)	5 (9.3)	10 (20.4)
	Very low	0 (0)	0 (0)	0 (0)
At T_1	Very high	33 (32.0)	15 (27.8)	18 (36.7)
	High	42 (40.8)	22 (40.7)	20 (40.8)
	Low	25 (24.3)	16 (29.6)	9 (18.4)
	Very low	3 (2.9)	1 (1.9)	2 (4.1)
Increased by one category (T_0 – T_1)		26 (25.2)	9 (16.7)	17 (34.7)
No change (T_0 – T_1)		42 (40.8)	23 (42.6)	19 (38.8)
Decreased by one category (T_0 – T_1)		28 (27.2)	17 (31.5)	11 (22.4)
Decreased by two categories (T_0 – T_1)		7 (6.8)	5 (9.3)	2 (4.1)

T_0 = day of enrolment; T_1 = day the final questionnaire was distributed.
Information need was measured with a four-pointe Likert scale ranging from "very high" to "very low" at the time of the patients' first clinical appointment for biopsy preparation (T_0) and 4 wk later (T_1), after the biopsy had been taken place and the patients had been informed of the histological findings.

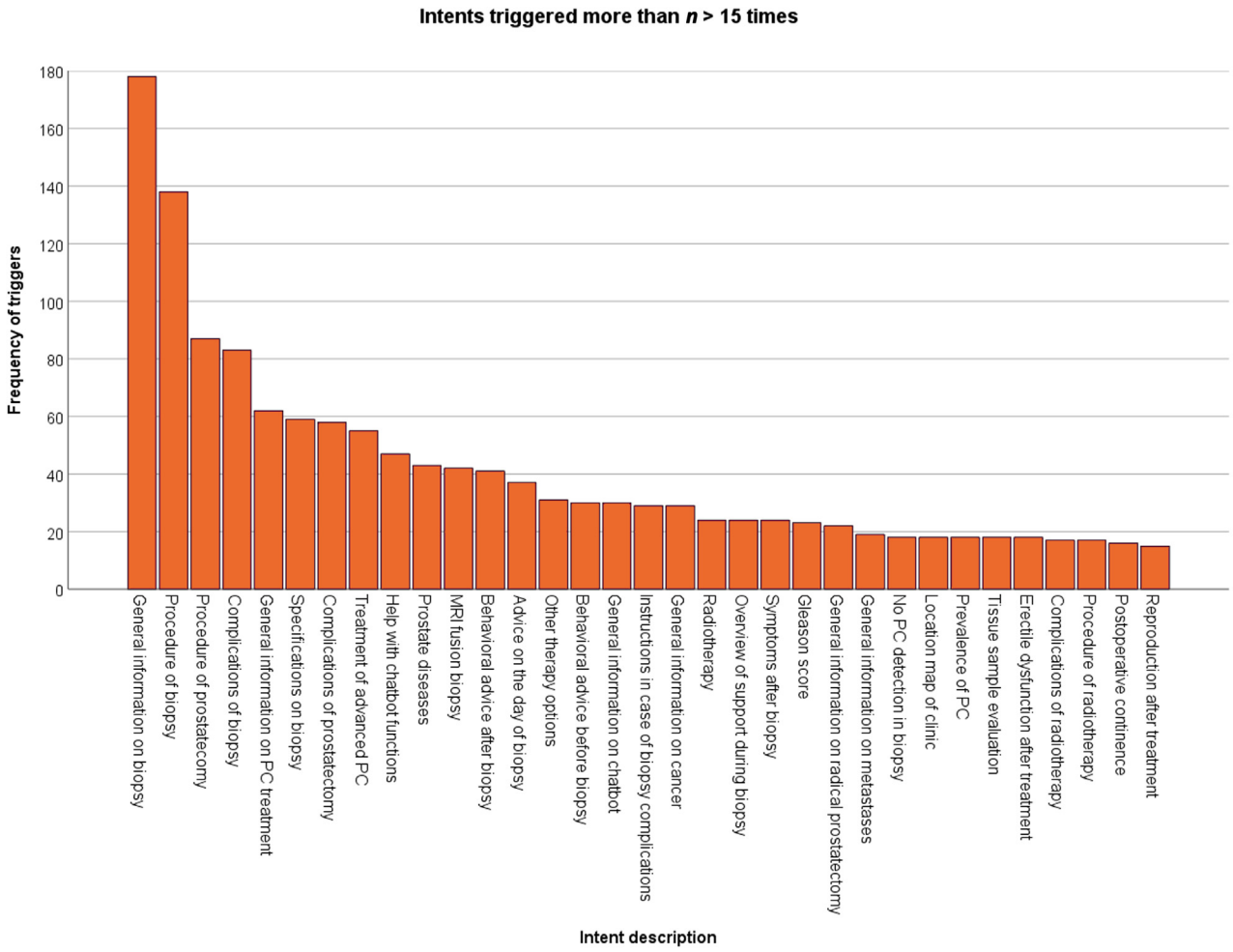


Fig. 3 – Frequency of intent recognition by PROSCA—frequency of intents that were triggered more than 15 times by the natural language processing model during the study period. The intents were triggered by a user's input and led to an action of the chatbot such as a text reply. The intent description used in this figure is a summary of the theme or category that the intent was created for. MRI = magnetic resonance imaging; PC = prostate cancer; PROSCA = PROState cancer Conversational Agent.

Of the chatbot users, 73.2% fully to partially agree that they gained substantial information regarding diagnostic tests and prostate biopsy, and 71.4% regarding PC disease and treatment. Of 43 chatbot users, 39 (90.7%) would like to use the chatbot again during another hospital visit (30.2% agreed fully, 32.6% agreed, and 27.9% agreed partially; Fig. 4A), and 41 (95.3%) recommend the general application of the chatbot in clinical routine (32.6% agreed fully, 32.6% agreed, and 30.2% agreed partially; Fig. 4B).

4. Discussion

In this study, we developed and evaluated a chatbot for early PC detection in a clinical setting in terms of an RCT with 103 patients. Patients universally reported the ease of use and comfort with handling of the chatbot PROSCA, with positive perceptions of usability and feasibility. A statistical analysis revealed a significant benefit, with the chatbot group exhibiting a more frequent reduced information need than the control group. Notably, the chatbot successfully improved information dissemination although the chatbot was used as an additional tool, and both groups received the same routine in-hospital education. The high acceptance and adherence of patients to the chatbot, its ease of use, and the high recommendation rate among PROSCA users support its suitability for the target group. The study results suggest that the additional provision of the chatbot can reduce patients' information need during the diagnostic process of PC compared with the standard provision of information. This highlights the potential for routine use of such medical chatbots for supporting patients during clinical procedures.

The study's clinical importance lies in its exploration of a chatbot's impact on patient information needs in the context of PC detection and addresses the issue of patient education in complex medical conditions. Meeting patients' information needs leads to higher levels of treatment adherence and lower anxiety [23,24]. This suggests that ongoing, tailored information provision is crucial for maintaining patient satisfaction. PROSCA has the ability to address the individualized information needs of patients

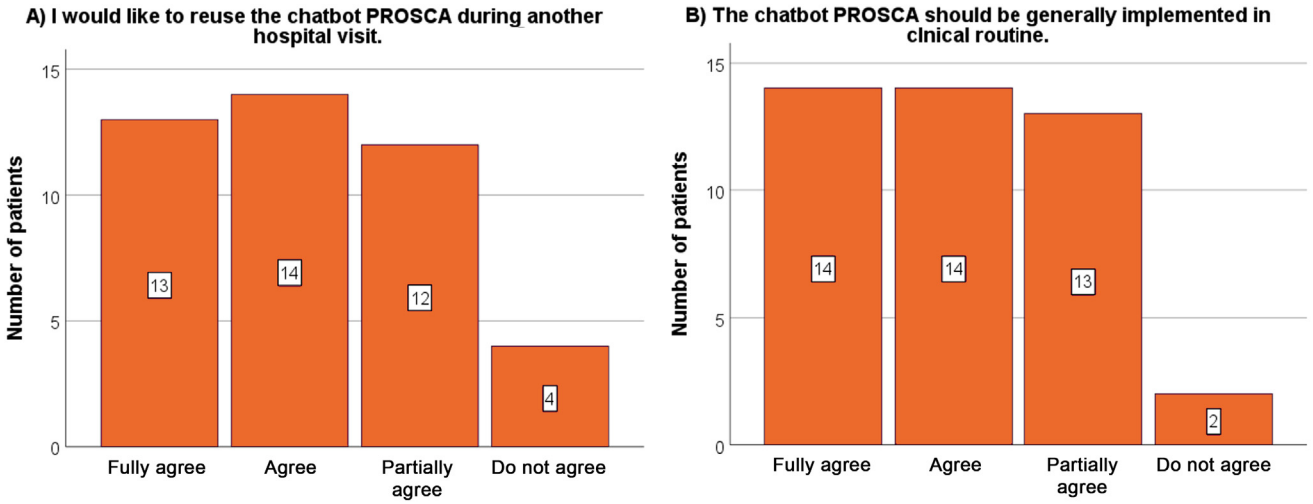


Fig. 4 – Evaluation of patient satisfaction with the chatbot—patients' evaluation regarding whether (A) they would like to reuse the chatbot PROSCA during another hospital visit and (B) they recommend the general implementation of the chatbot PROSCA in clinical routine. PROSCA = PROState cancer Conversational Agent.

facing PC diagnosis with an easy-to-learn application. Its role as an adaptive information resource that can be accessed at any time and place becomes crucial for satisfying patients' information needs throughout the diagnostic and treatment journey. PROSCA addresses relevant information needs that cancer patients require: information on diagnosis, course of disease, treatments, side effects, and psychological support [24]. Understanding the potential benefits of incorporating AI-based technology, such as chatbots, in health care can enhance patient engagement and decision-making, contributing to more effective and efficient care strategies. For the successful implementation of AI chatbots into clinical routine, patient's willingness and trust in AI applications are crucial. A recent study by Rodler et al [25] showed that patients confronted with diagnostic or therapeutic interventions for PC prefer AI-assisted urologists over urologists alone and AI alone. Consequently, AI would be best implemented in a medical professional–controlled setting as an additional tool. The supportive integration of PROSCA into the clinical setting without replacing the direct patient-doctor interaction in our RCT might be an interesting application for future AI chatbots.

The potential benefits of chatbots for patients and health care systems are in contrast with the lack of RCTs evaluating chatbots in health care, particularly in oncology. The chatbot Vik provided information on breast cancer, including its treatments, side effects, and strategies to improve quality of life. One prospective study and one RCT showed that Vik can increase study participant adherence and that the quality of information provided by Vik is comparable with that of a medical committee [12,26]. Among the available chatbot technologies, ChatGPT is poised to transform health care by providing patients and clinicians with greater access to medical information. While the use of a chatbot has the potential to enhance the ability to learn about complex topics in a time-efficient manner, the accuracy of the chatbot's output must be validated prior to its use [27]. One key problem with LLMs such as ChatGPT is their risk of hallucination, where the model generates apparently correct statements that turn out to be incorrect [28]. A recent study by Lombardo et al [29] aimed to analyze the quality of ChatGPT responses to PC-related queries compared with the EAU PC guidelines. ChatGPT had poor accuracy when answering questions about the EAU guidelines, with 26% being completely correct, 26% correct but inadequate, 24% correct but misleading, and 24% incorrect. Pan et al [30] characterized the quality of information and the presence of misinformation about prostate, skin, lung, breast, and colorectal cancers generated by four AI chatbots. They found that the chatbots generally provided accurate and high-quality information on cancer-related queries. However, the responses were not easily actionable, limiting their practical use for patient education and decision-making. A recent study by Hershenhouse et al [31] evaluated the quality of ChatGPT 3.5 outputs to PC-related questions from both physician and public perspectives. Both original and simplified responses were assessed for accuracy, completeness, and clarity, receiving high ratings for correctness and clarity from urology providers and the public. The authors concluded that while ChatGPT showed promise for patient education on PC, the technology's lack of reliability and controllability indicates that it is not yet suitable for delivering patient information. The chatbot PROSCA presented in this study with its guaranteed correctness of responses is a novel contribution to the literature with specific utility in the important context of early PC detection, distinguishing itself by urologist-validated responses and patient feedback investigated in the form of an RCT.

There are several dimensions that we can envision for extending the chatbot in future projects. Chatbots are cost effective to run and can automate repetitive tasks, allowing doctors to provide higher-quality, more personalized, and empathetic care to their patients. Next steps include the integration of the chatbot into existing health care systems to enable it to customize the responses even more to the patient's medical history and current situation. Understanding how chatbots can complement existing clinical

workflow is crucial for their successful adoption. This includes cost-benefit analyses to assess the economic implications of integrating chatbots into health care settings. Understanding the potential cost savings in human resources associated with improved patient outcomes will enable informed decisions about widespread adoption. In addition, research shall be extended to explore the use of chatbots in different clinical contexts, for example, other cancers or chronic diseases, providing disease-specific information needs.

Several limitations must be acknowledged. First, the study's sample size limits generalizability. It is important to exercise caution when generalizing the findings due to its single-center and specific clinical application.

Second, the exclusion of patients with insufficient language comprehension capabilities may introduce a selection bias into the study. This decision could result in a sample that is not fully representative of the broader patient population, limiting the generalizability of our findings and underscoring the need for further research that includes a more diverse patient cohort.

Third, we did not consider social determinants of health (SDOH) in our analysis. The study did not evaluate PROSCA's ability to tailor responses according to patient demographics related to SDOH, such as education, employment, economic background, and health care access. These factors can significantly influence patient outcomes and the appropriateness of clinical recommendations. Future studies should incorporate SDOH to better understand how these factors might impact the model's performance across diverse patient populations.

Fourth, the reliance on self-reported measures, including patient information needs, introduces a potential bias. An urgent need for reporting recommendations for the evaluation of the performance of chatbots was already reported [32].

Fifth, although the 4-wk observation period provided initial insights, it cannot capture the long-term effects of chatbot-supported interventions. Further longitudinal studies to assess the long-term effects of chatbots such as PROSCA on patient outcomes are recommended.

5. Conclusions

With the introduction of the PROSCA chatbot, we created and evaluated an innovative, evidence-based AI health information tool as an additional source of information for PC. Our RCT results showed significant benefits of the chatbot in reducing patients' information needs and enhancing their understanding of PC. This easy-to-use AI tool provides accurate, timely, and accessible support, demonstrating its value in the PC diagnosis process. Future steps include further customization of the chatbot's responses and integration with existing health care systems to maximize its impact on patient outcomes.

***Author contributions*:** Magdalena Görtz had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Görtz.
Acquisition of data: Baumgärtner.
Analysis and interpretation of data: Baumgärtner, Görtz.
Drafting of the manuscript: Baumgärtner, Görtz.
Critical revision of the manuscript for important intellectual content: Byczkowski, Schmid, Muschko, Woessner, Gerlach, Bonekamp, Schlemmer, Hohenfellner.
Statistical analysis: Baumgärtner, Görtz.
Obtaining funding: None.
Administrative, technical, or material support: Byczkowski, Schmid, Muschko, Woessner, Gerlach, Hohenfellner.
Supervision: Byczkowski, Görtz, Hohenfellner.
Other: None.

Financial disclosures: Magdalena Görtz certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: Michael Byczkowski, Tamara Schmid, Marc Muschko, Philipp Woessner, and Axel Gerlach are employees of SAP SE.

***Funding/Support and role of the sponsor*:** None.

***Acknowledgments*:** The training of the chatbot PROSCA was supported by SAP SE. SAP Conversational AI as a low-code chatbot building platform was provided free of charge during the execution of the randomized controlled trial.

***Ethics statement*:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and the 1964 Declaration of Helsinki and its later amendments. The study was approved by the ethical committee of the University of Heidelberg (approval no. S-005/2021). The consent obtained from study participants was in written form.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.euros.2024.08.022.

References

[1] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin 2022;72:7–33.
[2] Gandaglia G, Leni R, Bray F, et al. Epidemiology and prevention of prostate cancer. Eur Urol Oncol 2021;4:877–92.
[3] Cancer IAfRo. Cancer Today. https://gco.iarc.fr/today/en/dataviz/bars?types=1&mode=cancer&group_populations=1&sort_by=value0&sexes=1&populations=108_140_148_180_231_232_270_324_4_408_430_450_454_466_508_562_624_646_694_706_728_729_760_768_800_854_887.
[4] de Vos I, Meertens A, Hogenhout R, Remmers S, Roobol MJ. A detailed evaluation of the effect of prostate-specific antigen-based screening on morbidity and mortality of prostate cancer: 21-year follow-up results of the Rotterdam section of the European Randomised Study of Screening for Prostate Cancer. Eur Urol 2023;84:426–34.
[5] Goerling U, Faller H, Hornemann B, et al. Information needs in cancer patients across the disease trajectory. A prospective study. Patient Educ Couns 2020;103:120–6.

[6] Xu AJ, Myrie A, Taylor JI, et al. Instagram and prostate cancer: using validated instruments to assess the quality of information on social media. Prostate Cancer Prostatic Dis 2022;25:791–3.
[7] Denecke K, May R. Developing a technical-oriented taxonomy to define archetypes of conversational agents in health care: literature review and cluster analysis. J Med Internet Res 2023;25:e41583.
[8] Ghosh S, Bhatia S, Bhatia A. Quro: facilitating user symptom check using a personalised chatbot-oriented dialogue system. Stud Health Technol Inform 2018;252:51–6.
[9] Palanica A, Flaschner P, Thommandram A, Li M, Fossat Y. Physicians' perceptions of chatbots in health care: cross-sectional web-based survey. J Med Internet Res 2019;21:e12887.
[10] Milne-Ives M, de Cock C, Lim E, et al. The effectiveness of artificial intelligence conversational agents in health care: systematic review. J Med Internet Res 2020;22:e20346.
[11] Bickmore TW, Utami D, Matsuyama R, Paasche-Orlow MK. Improving access to online health information with conversational agents: a randomized controlled experiment. J Med Internet Res 2016;18:e1.
[12] Bibault J-E, Chaix B, Guillemassé A, et al. A chatbot versus physicians to provide information for patients with breast cancer: blind, randomized controlled noninferiority trial. J Med Internet Res 2019;21:e15787.
[13] Röbbelen A, Schmieding ML, Kopka M, Balzer F, Feufel MA. Interactive versus static decision support tools for COVID-19: randomized controlled trial. JMIR Public Health Surveill 2022;8:e33733.
[14] Eppler M, Ganjavi C, Ramacciotti LS, et al. Awareness and use of ChatGPT and large language models: a prospective cross-sectional global survey in urology. Eur Urol 2024;85:146–53.
[15] Kaufmann B, Busby D, Das CK, et al. Validation of a zero-shot learning natural language processing tool to facilitate data abstraction for urologic research. Eur Urol Focus 2024;10:279–87.
[16] Davis R, Eppler M, Ayo-Ajibola O, et al. Evaluating the effectiveness of artificial intelligence-powered large language models application in disseminating appropriate and readable health information in urology. J Urol 2023;210:688–94.
[17] Cocci A, Pezzoli M, Lo Re M, et al. Quality of information and appropriateness of ChatGPT outputs for urology patients. Prostate Cancer Prostatic Dis 2024;27:103–8.
[18] SAPConversationalAI. SAP Conversational AI. https://www.sap.com/products/artificial-intelligence/business-technology-platform.html.
[19] Görtz M, Baumgärtner K, Schmid T, et al. An artificial intelligence-based chatbot for prostate cancer education: design and patient evaluation study. Digit Health 2023;9:20552076231173304.
[20] SAPConversationalAI. Key features of SAP Conversational AI. https://help.sap.com/doc/5befc39ddee84fe681d565cadd98ce05/latest/en-US/KeyFeaturesOfSAPConversationalAI.pdf.
[21] Mottet N, van den Bergh RCN, Briers E, et al. EAU-EANM-ESTRO-ESUR-SIOG guidelines on prostate cancer—2020 update. Part 1: screening, diagnosis, and local treatment with curative intent. Eur Urol 2021;79:243–62.
[22] Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019;7:e11500.
[23] Lamers RE, Cuypers M, Husson O, et al. Patients are dissatisfied with information provision: perceived information provision and quality of life in prostate cancer patients. Psychooncology 2016;25:633–40.
[24] Faller H, Koch U, Brähler E, et al. Satisfaction with information and unmet information needs in men and women with cancer. J Cancer Surviv 2016;10:62–70.
[25] Rodler S, Kopliku R, Ulrich D, et al. Patients' trust in artificial intelligence–based decision-making for localized prostate cancer: results from a prospective trial. Eur Urol Focus, in press. https://doi.org/10.1016/j.euf.2023.10.020.
[26] Chaix B, Bibault J-E, Pienkowski A, et al. When chatbots meet patients: one-year prospective study of conversations between patients with breast cancer and a chatbot. JMIR Cancer 2019;5:e12856.
[27] Gupta R, Pedraza AM, Gorin MA, Tewari AK. Defining the role of large language models in urologic care and research. Eur Urol Oncol 2024;7:1–13.
[28] Ferber D, Kather JN. Large language models in uro-oncology. Eur Urol Oncol 2024;7:157–9.
[29] Lombardo R, Gallo G, Stira J, et al. Quality of information and appropriateness of Open AI outputs for prostate cancer. Prostate Cancer Prostatic Dis, in press. https://doi.org/10.1038/s41391-024-00789-0.
[30] Pan A, Musheyev D, Bockelman D, Loeb S, Kabarriti AE. Assessment of artificial intelligence chatbot responses to top searched queries about cancer. JAMA Oncol 2023;9:1437–40.
[31] Hershenhouse JS, Mokhtar D, Eppler MB, et al. Accuracy, readability, and understandability of large language models for prostate cancer information to the public. Prostate Cancer Prostatic Dis, in press. https://doi.org/10.1038/s41391-024-00826-y.
[32] Huo B, Cacciamani GE, Collins GS, McKechnie T, Lee Y, Guyatt G. Reporting standards for the use of large language model-linked chatbots for health advice. Nat Med 2023;29:2988.