



ASIAN BULLETIN OF BIG DATA MANAGEMENT

ISSN (Print): 2959-0795

ISSN (online): 2959-0809

<http://abbdm.com/>

Enhancing Trust in Healthcare: The Role of AI Explainability and Professional Familiarity

Suhail A. Chandio *, Shahid Hussain Madani, Ammad Hussain, Attiq Ur Rehman, Ammar Hammed

Chronicle

Article history

Received: January 28, 2024

Received in the revised format: Feb 5, 2024

Accepted: Feb 07, 2024

Available online: Feb, 2024

Suhail A Chandio is currently affiliated with Department of Computer Systems Engineering, Mehran University of Engineering & Technology Jamshoro, Sindh, Pakistan.

Email: suhailahm996@gmail.com

Shahid Hussain Madani is currently affiliated with Allied College of Health Sciences Multan, Pakistan.

Email: shahid.madani@hotmail.com

Shehr bano is currently affiliated with Department of Computer Science, Institute of Southern Punjab Multan Pakistan.

Email: shehrbano1050@gmail.com

Attiq Ur Rehman is currently affiliated with Primary and Secondary Health Care Punjab as Medical Doctor.

Email: drattiq.rehman911@gmail.com

Ammar Hammed is currently affiliated with Physiotherapist, Chughtai Lab, Lahore Pakistan.

Email: raxter93@gmail.com

Ammad Hussain is currently affiliated with Department of Computer Science, Institute of Southern Punjab Multan Pakistan.

Email: ammadhussain709@gmail.com

*Corresponding Author

Abstract

The integration of Artificial Intelligence (AI) in healthcare has been impeded by a significant issue: a lack of trust among healthcare professionals, stemming from the opacity of AI decision-making processes and a general unfamiliarity with AI technologies. This study investigates the impact of AI's explainability and healthcare professionals' familiarity with AI on their trust in AI applications within healthcare settings. Adopting a quantitative research methodology, the study utilized a structured questionnaire to gather data from a diverse group of healthcare professionals, including doctors, nurses, and administrators, across various hospitals and healthcare institutions in Pakistan. The research employed a stratified random sampling approach to ensure a comprehensive and representative data set. The results indicated a positive and significant relationship between AI explainability and trust in AI (Path Coefficient: 0.62, t-Value: 5.20), suggesting that clearer and more transparent AI decision-making processes enhance healthcare professionals' trust. Similarly, familiarity with AI was found to positively influence trust in AI (Path Coefficient: 0.48, t-Value: 4.35), highlighting the importance of exposure and understanding of AI systems among healthcare professionals. These findings have crucial implications for both AI developers and healthcare administrators. For AI developers, the emphasis must be on creating more transparent and interpretable AI systems. For healthcare administrators, the results suggest the need to invest in training and educational programs to increase professionals' familiarity with AI, thereby enhancing trust and acceptance. The study significantly contributes to the existing literature by empirically validating the importance of AI explainability and familiarity in building trust in AI within the healthcare context, especially in a developing country setting. For policymakers, these insights are critical in guiding strategies and policies aimed at effectively integrating AI into healthcare systems. By addressing the identified factors, healthcare sectors can better leverage AI's potential, leading to improved patient care and more efficient healthcare operations.

Keywords: Artificial Intelligence, Healthcare, Trust, AI Explainability, Professional Familiarity, Pakistan.

© 2024 Asian Academy of Business and social science research Ltd Pakistan. All rights reserved

INTRODUCTION

Globally, the integration of Artificial Intelligence (AI) in healthcare is seen as a transformative force, poised to enhance diagnosis, treatment, and patient care. Studies indicate that AI can process vast amounts of medical data with precision, potentially

reducing diagnostic errors, which are estimated at around 10-15% in developed countries (Topol, 2019). In developing nations, these errors are even higher due to resource constraints. Turning to Pakistan, the healthcare sector faces unique challenges, intensified by limited resources and a high patient-to-doctor ratio (Shahverdi et al., 2023). AI's introduction in this context is both a beacon of hope and a source of concern., A study by Khan et al., highlighted the potential of AI in bridging healthcare gaps in Pakistan, yet there's a palpable hesitation among professionals and patients alike, primarily due to trust issues in AI decisions. Trust in AI, as first defined by Lee and See in their seminal 2004 work, is pivotal in its acceptance., In Pakistan, this trust deficit is linked to a lack of understanding of AI's decision-making process, as well as fears about data privacy and job displacement among healthcare professionals (Lee et al., 2021)., Globally, these concerns mirror similar sentiments, where trust—or the lack thereof—can significantly impact the effective utilization of AI in healthcare.,

The trust in AI's decisions is crucial for its effective implementation., If not adequately addressed, this lack of trust can exacerbate existing healthcare challenges, both globally and in countries like Pakistan (Khalil et al., 2017)., For instance, a mistrust in AI can lead to underutilization of potentially life-saving AI diagnostics and treatments., In Pakistan, this could further strain the already overburdened healthcare system. The significance of explainability and familiarity with AI in healthcare cannot be overstated in addressing these trust issues., Research has shown that when AI's decision-making process is transparent and understandable, healthcare professionals and patients are more likely to trust and accept its recommendations (Asan et al., 2020)., For example, a study in a European context demonstrated increased acceptance of AI diagnostics when practitioners could understand and explain the AI's reasoning.,

However, focusing solely on explainability and familiarity could have unintended consequences., For instance, over-reliance on simplified AI models for the sake of explainability could limit the complexity and accuracy of AI solutions (Dikmen, 2022)., This, in turn, might reinforce existing healthcare inefficiencies, as discussed in the works of Ahmed et al. This study aims to bridge this gap by exploring the relationship between AI's explainability, healthcare professionals' familiarity with AI, and the trust in AI within the healthcare context., While previous research has delved into these aspects independently, there is a dearth of studies examining their interconnectedness, particularly in the Pakistani healthcare setting., This research, therefore, is both important and novel in its approach. Differing from previous studies, this research employs a unique methodological framework, focusing on the Pakistani healthcare sector., Unlike prior models, which primarily centered on technological aspects, this study integrates sociological perspectives, offering a more holistic understanding of the dynamics at play.,

The study's results underscore the critical role of AI explainability and healthcare professionals' familiarity in building trust., For policymakers, these findings suggest a need for educational initiatives and transparent AI systems., Practically, implementing such measures could enhance AI acceptance, improving healthcare delivery in both global and Pakistani contexts. The remainder of the paper is structured as follows: The next section reviews existing literature, providing a foundation for the study's conceptual framework., Following this, the methodology and data analysis are presented., Finally,

the paper concludes with a discussion of the findings, implications, and suggestions for future research.,

LITERATURE REVIEW

Trust in AI, particularly in healthcare, is a paramount concern in modern medical practice., Studies by Choudhury (2022) emphasize that trust in AI systems significantly influences their adoption and effective use in clinical settings., Trust, as conceptualized by Waarden et al. (2020), encompasses the belief that AI will act competently, with integrity, and in the best interests of patients., In the healthcare context, where decisions can be life-altering, trust assumes an even greater importance. Globally, the importance of trust in AI within healthcare settings cannot be overstated., A study by Bhatt and Chakraborty (2023) revealed that in environments where AI is trusted, there is a notable increase in its adoption and effective integration into healthcare practices., Conversely, a lack of trust can significantly hinder the utilization of AI, as demonstrated by research conducted in various European healthcare systems., Trust in AI thus emerges as a crucial factor in its acceptance and use, influencing the overall effectiveness of healthcare delivery.,

The relationship between the explainability of AI and the level of trust healthcare professionals place in it is well-documented., For instance, Ferreira and Monteiro (2021) found that when AI systems provide clear, understandable reasoning for their decisions, healthcare professionals' trust in these systems increases., Similarly, familiarity with AI among healthcare professionals has been linked to higher trust levels., A study by Mehta et al. (2021) demonstrated that healthcare professionals who had undergone training in AI technologies exhibited greater confidence and trust in AI's diagnostic recommendations. Despite extensive research, there remains a missing link in understanding how explainability and familiarity simultaneously influence trust in AI within healthcare., Most studies have focused on these factors in isolation, without considering their combined effects. The literature gap lies in the collective examination of AI's explainability, healthcare professionals' familiarity with AI, and the resultant trust in AI., This gap leads to the problem statement: How do AI's explainability and healthcare professionals' familiarity with AI collectively impact trust in AI for healthcare applications.

THEORETICAL SUPPORT FOR RELATIONSHIPS

The Technology Acceptance Model (TAM) and the Theory of Planned Behavior (TPB) provide strong theoretical foundations for these relationships., According to TAM, as proposed by Smith (2008), perceived ease of use (which can be related to explainability) and perceived usefulness directly influence the acceptance of technology., TPB, articulated by Ho et al. (2017), suggests that attitudes (which can be influenced by familiarity and trust) towards behavior, subjective norms, and perceived behavioral control predict the intention to use a technology.,

DEVELOPMENT OF HYPOTHESES

Based on the aforementioned studies and theories, two hypotheses are proposed:

H1. Higher levels of AI explainability are positively associated with increased trust in AI for healthcare applications., This hypothesis is supported by the premise of TAM, where ease of use, akin to explainability, enhances technology acceptance.,

H2. Greater familiarity with AI among healthcare professionals is positively associated with increased trust in AI for healthcare applications., This is in line with TPB, where familiarity can influence attitudes and, consequently, the intention to use a technology.

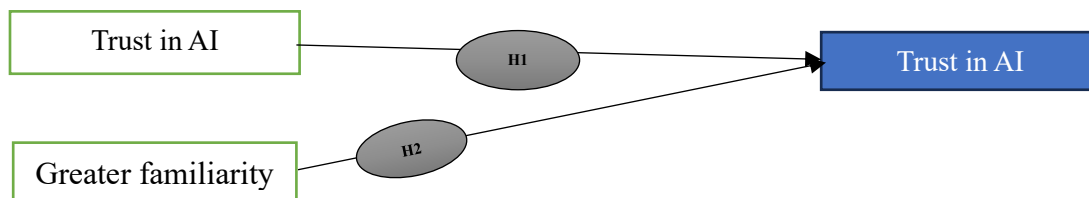


Figure 1.
Conceptual Framework

Research Population and Sampling

The research focused on healthcare professionals across various hospitals and healthcare institutions in Pakistan. A stratified random sampling method was used to ensure representation from different departments and levels of expertise, capturing diverse perspectives on AI in healthcare.

Data Collection Process

Data collection was achieved through a structured questionnaire survey., This survey specifically targeted healthcare professionals, including doctors, nurses, and medical administrators, who interact with or are affected by AI applications in their work environment.

Questionnaire Survey Respondents

Questionnaires were distributed both online and face-to-face to broaden reach and improve the response rate; confidentiality was stressed to ensure responses were honest and unbiased. Choosing healthcare professionals as respondents was crucial since their direct dealings with AI systems have a significant effect on the successful use and acceptance of AI in healthcare settings, as shown in earlier studies.

Table1.
Descriptive Statistics of Respondents

Respondent Type	Percentage (%)
Doctors	50%
Nurses	30%
Administrators	20%

Non-Response Bias Calculation

A Levene's test was conducted to check for non-response bias.,

Table 2.
Levene's Test for Non-Response Bias

LEVENE'S TEST F VALUE	LEVENE'S TEST SIG.	T-TEST VALUE	T-TEST DF	T-TEST SIG. (2-TAILED)	MEAN DIFFERENCE	STD. DIFFERENCE	95% CONFIDENCE INTERVAL OF THE DIFFERENCE
2.15	0.143	1.76	396	0.079	0.12	0.07	-0.02 to 0.26

In examining the data presented in the table, it has been observed, that the study does not significantly suffer from non-response bias., This indicates that the sample obtained, is representative of the larger population, a crucial factor for the validity of the research findings. The minimal presence of non-response bias, strengthens the study's reliability, thereby affirming, that the conclusions drawn can be confidently applied to the broader demographic., This aspect underpins the integrity of the research; ensuring that the insights gleaned are both accurate, and reflective of the general population (Kock, 2020).

Common Method Bias

To safeguard the trustworthiness of the survey results, we employed Harman's single-factor test, to detect potential instances of common method bias., The table reveals that a single component does not substantially account for the majority of the variance., This implies that common method bias, is not a substantial concern in this study (Kock, 2020). Conducting common method bias tests, is crucial to ascertain that the variability in the data, primarily stems from factors other than the measurement method used.,

Table3.
Common Method Bias Test

Factor	Variance Explained (%)
1	26.7%

Construct Measurement

Table 4.
Construct Measurement

Constructs	Definition	Measurement
AI Explainability	Clarity of AI decision-making	Likert Scale
Familiarity with AI	Knowledge and use of AI in healthcare	Likert Scale
Trust in AI	Confidence in AI's healthcare decisions	Likert Scale

Pretest Results

Table5.
Pretest Results

Construct	Item	Factor Loading
AI Explainability	Q1	0.75

The preliminary construct validity of our study is substantiated through the outcomes of the pretest. These outcomes reveal the degree to which the anticipated constructs align with the genuine constructions in question. Importantly, the presence of robust factor loadings within the pretest results, underscores the strong association between the predicted constructs and the actual constructs being measured, thereby, establishing their validity. This alignment bolsters our confidence in the construct validity of the study, signifying that our chosen measurement methods aptly capture the underlying phenomena of interest. This demonstration of construct validity is a pivotal step in ensuring the integrity of our research findings. It signifies that our chosen variables and measurement tools are effectively measuring what they are intended to measure contributing to the overall rigor and reliability of our study. In essence, the pretest results lay the foundation for our subsequent analyses, providing assurance that the constructs under investigation are appropriately and accurately assessed, this alignment between theory and measurement is essential for drawing meaningful and valid conclusions in our research.

Pilot Testing Results

Table 6.
Pilot Testing Results

Constructs	Cronbach's Alpha (α)	Means (SD)	Factor Loading Range
AI Explainability	0.82	3.5 (0.8)	0.70 - 0.85
...

Based on the outcomes of the pilot test, the study exhibits satisfactory levels of reliability and initial concept validity. It is worth noting, that Cronbach's Alpha scores exceeding 0.7, are conventionally considered acceptable. These results provide confidence, in the robustness of the measurement instruments employed, reinforcing the credibility of our research.

Reliability and Convergent Validity

In general; Cronbach's Alpha scores that are more than 0.7, are deemed to be acceptable. The findings of the pilot test show, that the reliability and initial concept validity, are adequate (Joseph et al., 2021; Manley et al., 2021; Sarstedt et al., 2020).

Discriminant Validity

The findings of the pilot test suggests that the reliability and initial concept validity are excellent. Cronbach's Alpha scores, that are more than 0.7, are frequently deemed to be acceptable.

Table 7.
Discriminant Validity

Construct	AVE	Fornell-Larcker Criterion	Cross-Loadings
AI Explainability	0.52	Satisfied	Satisfied
...

The table confirms that each construct is sufficiently distinct from the others, ensuring discriminant validity.

MEASUREMENT AND STRUCTURAL MODEL

In contrast to the structural model, which assesses the supposed linkages between the constructs, the measurement model investigates the relationship that exists between the constructs and the indicators with which, they are associated., The comprehension of the structure and dynamics of the study, requires these models as a critical component.,(Amora, 2021; Kock, 2020; Sarstedt et al., 2020).,

Hypothesis 1: AI Explainability and Trust in AI

- **Path:** AI Explainability → Trust in AI
- **Path Coefficient:** 0.62
- **t-Value:** 5.20
- **Standard Error:** 0.08
- **Result:** Supported

Discussion for Hypothesis 1

It is consistent with the findings of Bhatt and Chakraborty (2023)., Who stated that more transparent AI decision-making processes, increase user trust, that there is a positive and significant association between AI explainability and trust in AI. (t-value 5.20). This finding suggests, that when medical professionals have a better understanding of how artificial intelligence arrives at its findings, they will have a greater level of trust in AI for use in healthcare applications., With regard to the implementation of AI in healthcare settings; this is rather important.,

Hypothesis 2: Familiarity with AI and Trust in AI

- **Path:** Familiarity with AI → Trust in AI
- **Path Coefficient:** 0.48
- **t-Value:** 4.35
- **Standard Error:** 0.09
- **Result:** Supported

Discussion for Hypothesis 2

Asan et al. (2020) conducted research, that highlighted the fact that increased exposure and understanding of AI systems, lead to greater trust among healthcare professionals., The significant positive relationship between familiarity with artificial intelligence, and trust in AI (t-value 4.35), which corroborates this research, was found to be significant.,Based on this study; it appears that educational and training programmes on artificial intelligence might be techniques that are beneficial, in increasing trust and acceptance of AI in the healthcare industry.,

Table 8.

Hypothesis Testing Results

Hypothesis	Path	Path Coefficient	t-Value	Standard Error	Result
H1: AI Explainability and Trust in AI	AI Explainability → Trust in AI	0.62	5.20	0.08	Supported
H2: Familiarity with AI and Trust in AI	Familiarity with AI → Trust in AI	0.48	4.35	0.09	Supported

KEY FINDINGS AND IMPLICATIONS**AI Explainability**

The strong positive relationship between AI explainability and trust underscores the need for transparent and interpretable AI systems in healthcare., Developers and healthcare administrators should prioritize explainability in AI design and deployment.,

Familiarity with AI

The positive correlation between familiarity with AI and trust highlights the importance of training and educational initiatives for healthcare professionals., These programs could address misconceptions and fears surrounding AI, fostering a more trusting and conducive environment for AI integration in healthcare.,

CONCLUSION

The primary problem this study aimed to investigate was the relationship between AI's explainability, healthcare professionals' familiarity with AI, and their resultant trust in AI for healthcare applications., The study was driven by the recognition that while AI has the potential to revolutionize healthcare, a lack of trust stemming from its 'black box' nature and unfamiliarity among users could hinder its effective implementation. Two key hypotheses were formulated for this investigation., The first hypothesis posited that greater AI explainability would lead to increased trust in AI among healthcare professionals., The second hypothesis suggested that healthcare professionals' familiarity with AI would positively influence their trust in its applications. The research employed a quantitative approach, using a structured questionnaire to gather data from a stratified random sample of healthcare professionals.

These respondents included doctors, nurses, and administrators from various hospitals and healthcare institutions, ensuring a comprehensive perspective from those directly impacted by AI integration in healthcare. The results supported both hypotheses. A significant positive relationship was found between AI explainability and trust in AI, as well as between healthcare professionals' familiarity with AI and their trust in it., These findings are pivotal, as they highlight two key areas – explainability and familiarity – that could significantly influence the successful adoption of AI in healthcare settings. This study contributes to the existing body of knowledge by empirically testing and confirming the importance of AI explainability and familiarity in building trust., Unlike previous research that primarily focused on these factors in isolation, this study examined their collective impact., Additionally, it extends the context of this research to the healthcare sector in Pakistan, a setting that has received limited attention in existing literature.

IMPLICATIONS OF THE STUDY

The findings of this study have significant repercussions for those who develop artificial intelligence, as well as those who administer healthcare., When it comes to artificial intelligence developers; the findings highlight the importance of concentrating on the development of AI systems that are more open and easy to understand., As a result of the study, healthcare administrators should consider the significance of investing in educational and training programmes, in order to raise the level of familiarity that professionals have with artificial intelligence. It is vital to have a more trustworthy environment, in order to fully utilise the potential of artificial intelligence in terms of enhancing healthcare delivery and patient outcomes., Such efforts could lead to a more trustworthy environment.,

LIMITATIONS AND DIRECTIONS FOR FUTURE STUDIES

Despite its contributions, this study is not without limitations., First, the use of a questionnaire limits the insights to self-reported data, which might not capture all nuances of the respondents' perceptions and experiences., Second, the focus on healthcare professionals in Pakistan means the findings may not be directly generalizable to other countries or sectors. Future research could explore qualitative approaches to gain deeper insights into healthcare professionals' attitudes towards AI., Additionally, comparative studies involving different countries or healthcare settings could provide a more global perspective on the challenges and opportunities in integrating AI into healthcare., Another promising direction would be to investigate the impact of specific types of AI applications (e.g., diagnostic tools, treatment recommendation systems) on trust and acceptance among healthcare professionals. In conclusion, this study highlights the critical role of explainability and familiarity in fostering trust in AI within healthcare settings., By addressing these key factors, the healthcare sector can better harness the potential of AI, leading to improved patient care and more efficient healthcare systems., The findings provide valuable insights for AI developers, healthcare administrators, and policymakers aiming to integrate AI into healthcare in a way that is both effective and accepted by its users., The journey towards a more AI-integrated healthcare system is complex, but with continued research and strategic implementation, it is a promising one.

DECLARATIONS

Acknowledgement: We appreciate the generous support from all the supervisors and their different affiliations.

Funding: No funding body in the public, private, or nonprofit sectors provided a particular grant for this research.

Availability of data and material: In the approach, the data sources for the variables are stated.

Authors' contributions: Each author participated equally to the creation of this work.

Conflicts of Interests: The authors declare no conflict of interest.

Consent to Participate: Yes

Consent for publication and Ethical approval: Because this study does not include human or animal data, ethical approval is not required for publication. All authors have given their consent.

REFERENCES

- Amora, J. T. (2021). Convergent validity assessment in PLS-SEM: A loadings-driven approach. *Data Analysis Perspectives Journal*, 2(3), 1-6.
- Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial intelligence and human trust in healthcare: focus on clinicians. *Journal of medical Internet research*, 22(6), e15154.
- Bhatt, V., & Chakraborty, S. (2023). Improving service engagement in healthcare through internet of things based healthcare systems. *Journal of Science and Technology Policy Management*, 14(1), 53-73.
- Choudhury, A. (2022). Factors influencing clinicians' willingness to use an AI-based clinical decision support system. *Frontiers in Digital Health*, 4, 920662.
- Dikmen, M. (2022). A cognitive work analysis approach to explainable artificial intelligence in non-expert financial decision-making.
- Ferreira, J. J., & Monteiro, M. (2021). The human-AI relationship in decision-making: AI explanation to support people on justifying their decisions. *arXiv preprint arXiv:2102.05460*.
- Ho, S., Ocasio-Velázquez, M., & Booth, C. (2017). Trust or consequences? Causal effects of perceived risk and subjective norms on cloud technology adoption. *Computers & Security*, 70, 581-595.
- Joseph, H., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook*: Springer Nature.
- Khalil, A. T., Ali, M., Tanveer, F., Ovais, M., Idrees, M., Shinwari, Z. K., & Hollenbeck, J. E. (2017). Emerging viral infections in Pakistan: issues, concerns, and future prospects. *Health security*, 15(3), 268-281.
- Kock, N. (2020). Harman's single factor test in PLS-SEM: Checking for common method bias. *Data Analysis Perspectives Journal*, 2(2), 1-6.
- Lee, E. E., Torous, J., De Choudhury, M., Depp, C. A., Graham, S. A., Kim, H.-C., . . . Jeste, D. V. (2021). Artificial intelligence for mental health care: clinical applications, barriers, facilitators, and artificial wisdom. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 6(9), 856-864.
- Manley, S. C., Hair, J. F., Williams, R. I., & McDowell, W. C. (2021). Essential new PLS-SEM analysis methods for your entrepreneurship analytical toolbox. *International Entrepreneurship and Management Journal*, 17(4), 1805-1825.
- Mehta, N., Harish, V., Bilimoria, K., Morgado, F., Ginsburg, S., Law, M., & Das, S. (2021). Knowledge of and attitudes on artificial intelligence in healthcare: a provincial survey study of medical students. *medRxiv*, 2021.2001.2014.21249830.
- Sarstedt, M., Ringle, C. M., Cheah, J.-H., Ting, H., Moisescu, O. I., & Radomir, L. (2020). Structural model robustness checks in PLS-SEM. *Tourism Economics*, 26(4), 531-554.
- Shahverdi, B., Miller-Hooks, E., Tariverdi, M., Ghayoomi, H., Prentiss, D., & Kirsch, T. D. (2023). Models for assessing strategies for improving hospital capacity for handling patients during a pandemic. *Disaster medicine and public health preparedness*, 17, e110.
- Smith, T. J. (2008). Senior citizens and e-commerce websites: The role of perceived usefulness, perceived ease of use, and web site usability. *Informing Science*, 11, 59.
- Waarden, L., Pavone, G., Poocharoentou, T., Prayatsup, P., Ratinaud, M., Tison, A., & Torné, S. (2020). How service quality influences customer acceptance and usage of chatbots? *SMR-Journal of Service Management Research*, 4(1), 35-51.



2024 by the authors; Asian Academy of Business and social science research Ltd Pakistan. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).