AI-Assisted Lab Summaries: Progressive Disclosure Interfaces for Patient Health Portals

Hala Mukheimer

David R. Cheriton School of Computer Science, University of Waterloo, hala.mukheimer@uwaterloo.ca

1 INTRODUCTION

Nowadays, rapid access to data plays a major role in our daily life. We can find information on anything, anywhere, within seconds. However, accessing and interpreting personal medical information isn't as easy for non-expert users. Historically, patients faced a lot of difficulties to gain access to their medical records [1]. In response, governments and healthcare systems started pushing towards transparency in medicine [1, 2]. A key milestone was the American 21st Century Cures Act Final Rule, which mandates that patients be able to access their electronic medical records transparently [3]. While such policies solve data availability problems, they introduce an important user interface challenge: presenting raw clinical data without translation for non-experts risks the emergence of patient anxiety and misinterpretation.

Patient portals hold strong potential to improve health awareness, strengthen patient-provider communication, and encourage adherence to treatment [4]. In practice, however, modern portals benefit highly health-literate or tech-savvy users, while increasing distress and anxiety for others [5]. On the patient side, barriers include preference for direct human communication, confusion from accessing complex medical results, hard-to-navigate user interfaces, and other security concerns [5]. For healthcare providers, portals add administrative workloads and risk widening health disparities for vulnerable populations [5]. Researchers have explored patient experiences accessing abnormal test results through their portals [6]. Results showed that while patients preferred access to all types of abnormal test results, they still raised several concerns including not being able to interpret how the results really affect their health [6]. These findings highlight that patient experience with portals could be improved by the development of better interfaces to help non-expert patients understand and manage the results once received [6].

Recent advancements in Generative Artificial Intelligence (AI) and Large Language Models (LLMs) offer a way for bridging the understanding gap between users and the medical field by translating complex medical terms into plain language summary [7, 8]. However, simply adding conversational AI directly into patient interfaces introduces risks. First, presenting the AI's answer as a single block of explanation text can trigger automation-bias, where non-expert users trust inaccurate explanations, or get alarmed over dramatic language used in generation [9, 10]. Moreover, single text blocks fail when certain users prefer more depth in the information generated. Thus, to use Al translation safely, patient portals need interface designs that structure automated interpretations. Moreover, visual context and the ability to verify the evidence source are also required.

To address these challenges, we propose an interactive web-based portal extension that is designed to meet two core interaction principles: Progressive Disclosure and Visual Source-Grounding. Instead of overwhelming non-expert patients with dense information or ungrounded AI summaries, the proposed system displays the medical interpretation using three main interactive layers, a continuous visual spectrum bar, a short plain-language AI summary, and finally attribution markers that connect claims in the summary to actual clinical references.

We aim to answer three primary research questions through a user study as part of this research:

- *RQ1:* To what extent does a progressive disclosure interface reduce patient anxiety and cognitive workload compared to standard portal layouts?
- *RQ2:* To what extent does the interactive visual source-grounding improve users' understanding of diagnostic lab results?
- *RQ3:* How does source-grounded AI text affect users' verification behavior and trust calibration?

2 RELATED WORK

This section reviews existing literature across several key areas: first, patient portal adoption and equity barriers; second, usability failures and user experience (UX) dimensions in commercial health platforms; and finally, the informational and contextual support needs of patients managing diagnostic health data.

2.1 Patient Portal Adoption, Engagement, and Equity Barriers

While portal adoption is growing, current systems still favor tech-savvy users and leave the rest with access barriers. Government obligations have caused health systems to establish direct access to patient records in Electronic Health Record (EHR) deployment. However, the portals weren't designed for different age groups or background knowledge levels [14, 16]. Another issue is the engagement initiatives, where hospitals run campaigns like "Sign up for MyChart today!", which successfully gets people to create accounts. However, if the portal is confusing to use, patients log in once, get frustrated, and never open the account again [12, 13]. Moreover, elderly or low-income patients often face difficulties with using complex passwords, small fonts, and multi-factor authentication. They also worry that using electronic portals means face-to-face physician contact will be fully replaced [24]. While such portals promise patient empowerment, uploading raw clinical data to unassisted users widens the gap between tech-savvy, highly educated patients, and vulnerable populations [11, 15].

2.2 Usability Failures, User Experience Dimensions, and Anxiety

Research has shown that when real patients are asked to perform basic tasks in a portal, like finding a lab result, they commit operational and navigation errors because of the complex UI hierarchy [17, 18]. Moreover, text-mining analysis of user reviews for portals on the app store showed that the biggest complaints weren't about the medical care, but about the hidden features, confusing navigation, and delayed data updates [19]. Large-scale national survey data shows that while patients like the idea of having online access, they still show immediate spikes in anxiety and distress the moment they encounter an uninterpreted result [20].

2.3 Complex Care Needs and the Need for Contextual Support

Standard portals treat every health record the same way, whether it's a routine cholesterol check or a complex chronic condition report. Hospitalized patients have different informational needs than someone sitting at home. Bedside patients need to know today's daily care plan, who their care team is, and what their next test is, things that traditional portals fail to highlight [23]. Another need is specialized condition portals. Patients with high-stakes conditions like stroke recovery or kidney dialysis need custom visualizations and tracking tools, not a generic document folder [21, 22]. Across all patient needs, whether it's routine lab reporting or complex disease management, there is an interface gap. Patients desperately need user-centered interface strategies, like easy contextual explanations and layered disclosure, to interpret medical data without getting overwhelmed [11, 17, 21, 22].

3 PROPOSED SOLUTION

To address the interaction challenges that were identified in the previous section, we propose a web-based portal extension which is designed to integrate easily into existing EHR patient portals. It is designed around two Human-Computer Interaction (HCI) principles: Progressive Disclosure [25] and Shneiderman's Visual Information-Seeking Mantra (*"Overview first, zoom and filter, then details-on-demand"*) [26]. This interactive UI system converts raw EHR data and AI-generated summaries into three structured interaction tiers as shown in Figure 1, a visual spectrum bar for immediate contextual orientation, a two-sentence plain-language AI summary for clear translation, and interactive source-grounding markers for evidence verification.

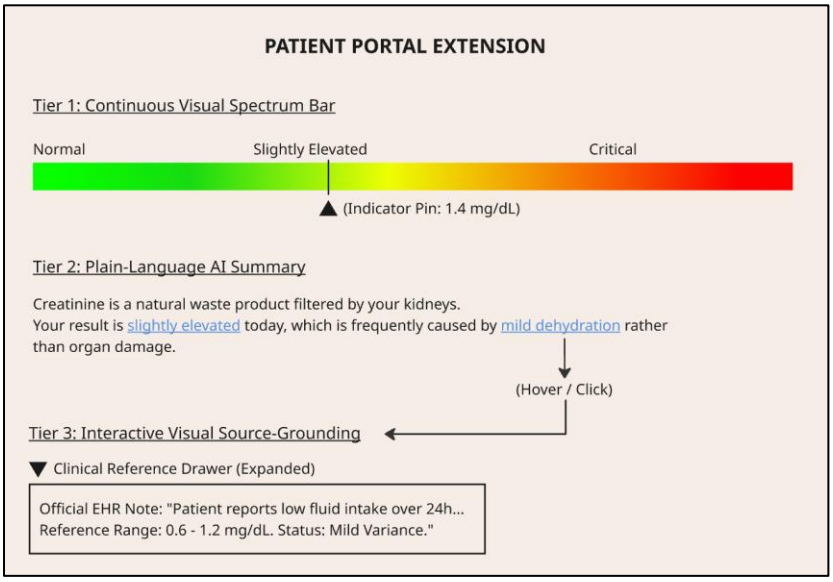

Figure 1: Architectural overview of the proposed 3-tier portal extension. The interface guides users from a visual status overview (Tier1) and a 2-sentence plain-language AI summary (Tier 2) to an interactive source-grounding drawer for verification (Tier 3).

3.1 Tier 1: Continuous Visual Spectrum Bar

The visual spectrum bar is a horizontal color gradient bar that represents the physiological range (e.g., Green = Normal, Yellow = Slightly Elevated, Red = Critical). It also has an indicator pin that displays where the patient's specific lab value falls along the continuous spectrum. This is designed to replace the binary red tags like "HIGH / LOW", because seeing a score sit just on the boundary of normal reassures patients that a slight variance is not an emergency. Furthermore, the human visual system processes color hue and spatial position faster than numerical data.

3.2 Tier 2: Plain Language AI Summary

The second part of this solution is the AI summary. It provides a two-sentence plain-language translation of the diagnostic result, which is generated by an LLM. However, there are certain content constraints that need to be followed. The first sentence should explain what the test measures in everyday language (e.g., *"Creatinine is a natural waste product filtered*

by your kidneys"). Then, the second sentence translates the result in a calm tone (e.g., *"Your result is slightly elevated today, which is frequently caused by temporary mild dehydration rather than organ damage"*). When we restrict the AI output to only these two parts, we will prevent the wall-of-text fatigue that leads to automation bias or panic from clinical terminology.

3.3 Tier 3: Interactive Visual Source-Grounding

The third and final part of our proposed solution is the verification. When the AI outputs the summary, specific claims like "mild dehydration" are designed to be interactive phrases. When the user hovers over or taps one of these phrases, a visual connection is triggered (e.g., highlighting). It points to the official lab report or a physician's note from an expandable clinical reference drawer. In this drawer, raw laboratory numbers and original clinical notes stay hidden. If a tech-savvy user or anyone who wants deep technical details wants to access it, they can simply do so with a click. Visual source-grounding encourages active verification, which allows non-expert users to calibrate their trust rather than blindly accepting or rejecting the AI's output. Table 1 below shows a summary of the 3-level solution.

Table 1: Summary of the 3-tiered interaction design, corresponding UI elements, and underlying HCI design rationales.

Interaction Layer	UI Element	Primary Function	HCI Principle / Design Rationale
Tier 1 (Overview)	Continuous Visual Spectrum Bar	Immediate visual status (Normal vs. Abnormal)	Pre-attentive Visual Processing; Mitigates Binary Flag Anxiety
Tier 2 (Translation)	2-Sentence Plain-Language AI Text	Simple, empathetic explanation of the result	Cognitive Load Reduction; Limits Wall-of-Text Automation Bias
Tier 3 (Verification)	Interactive Source Highlights & Drawer	Links AI text claims directly to raw EHR notes	Progressive Disclosure; Trust Calibration

4 STUDY DESIGN

This section describes the study design planned to evaluate the proposed 3-tiered portal extension and address our three research questions. To evaluate our system, we will combine a controlled lab experiment with a think-aloud protocol and post-study interviews. These HCI methodologies fit our research questions well. First, quantitative metrics allow us to measure whether our interface actually reduces anxiety while improving comprehension (RQ1, RQ2). Second, qualitative think-aloud data and interview responses help us understand why users choose to verify information and how they build trust with the AI (RQ3), which numbers alone cannot show.

4.1 Participants

For this study, our target audience are non-expert adults with no formal clinical background or medical training. We aim to recruit a sample size of N=60 participants in total, which will then be divided into three equal groups (n=20 per group) as the following:

- Group 1: Standard EHR Portal Layout (Raw values with red "HIGH/LOW" labels)
- Group 2: Unstructured AI Text Summary (Single block of text without source links)
- Group 3: The 3-Tiered Progressive Disclosure Interface (Proposed Solution)

As for the recruitment constraints, participants must be over the age of 18. They should also be screened for health literacy using Chew's Health Literacy Scale [27], so that we can balance demographics such as age, gender, and digital literacy evenly across the three groups. Furthermore, participants are obliged to complete a consent form that emphasizes voluntary participation, data confidentiality, and minimal risk.

4.2 Objectives

For the objectives, our focus aligns with our three research questions. The first objective is to measure if progressive disclosure reduces post-review anxiety (STAI-6) [29] and cognitive workload (NASA-TLX) [28] compared to standard portals with red "HIGH/LOW" labelling. The second objective is to assess how well the two-sentence summary and visual spectrum bar improve the right understanding of the results. The final objective is to evaluate whether interactive visual source-grounding encourages active verification of AI claims without causing either automation bias or complete distrust. In summary, we want to assess the following three things:

- Anxiety and Workload (RQ1)
- Diagnostic Comprehension (RQ2)
- Trust Calibration and Verification (RQ3)

4.3 Procedure

The procedure is divided into four main phases, a pre-study phase, a task execution phase, post-task questionnaires, and finally post-study interview and debrief. It is evaluated to need 45 minutes in total per participant in a single session.

4.3.1 Phase 1: Pre-Study (10 minutes)

During this phase, participants will provide informed consent and answer a demographic/health-literacy questionnaire. We will also conduct a quick tutorial and a task to practice the Concurrent Think-Aloud Protocol.

4.3.2 Phase 2: Task Execution (20 minutes)

We will present to each participant three clinical lab scenarios in random order (e.g., elevated creatinine, low hemoglobin, abnormal TSH). For each scenario, participants will review their assigned interface condition while speaking their thoughts aloud.

4.3.3 Phase 3: Post-Task Questionnaires (10 minutes)

For this phase, we will administer NASA-TLX, STAI-6, and a quick comprehension quiz after each scenario. We should also log system interactions for Group 3. These interactions include things like hover frequency over Tier 2 phrases, click counts on Tier 3 drawer, and time spent reading raw EHR notes.

4.3.4 Phase 4: Post-Study Interview (5 minutes)

During the final phase, we will conduct a semi-structured interview exploring how clear the summaries felt, trust in AI, and reliance on source links. We will also use a 7-point Likert system satisfaction scale.

4.4 Data Analysis

4.4.1 Quantitative Analysis

After quantitative metrics (NASA-TLX, STAI-6, and the comprehension scores) are collected, we will analyze them using One-Way Analysis of Variance (ANOVA) to determine overall statistical significance. If data distributions violate normality assumptions, we will use the non-parametric Kruskal-Wallis test instead. Post-hoc Tukey HSD tests ($\alpha = 0.05$) will also be conducted to evaluate pairwise differences between specific interface conditions where significant main effects are found. Moreover, for group 3, descriptive statistics like means, standard deviations, and dwell time should be calculated for verification interaction logs.

4.4.2 Qualitative Analysis

For qualitative analysis, we will transcribe think-aloud recordings and interview audio. Transcripts will then be evaluated using Thematic Analysis following the six-phase framework by Braun and Clarke [30]. Coding will focus on finding recurring themes related to trust calibration, automation bias, and verification triggers.

5 EXPECTED RESULTS

Based on our proposed 3-tiered progressive disclosure design, we anticipate quantitative and qualitative outcomes across our three research questions (RQ1–RQ3).

5.1 Reduction in Anxiety and Cognitive Workload (RQ1)

Participants using the proposed solution (Group 3) will report statistically lower post-review anxiety (STAI-6) and lower cognitive workload (NASA-TLX) compared to participants from groups 1 and 2. Replacing "HIGH/LOW" flags with a continuous spectrum bar gives immediate visual context, which prevents panic when a lab value is only slightly outside standard bounds.

5.2 Improvement in Diagnostic Comprehension (RQ2)

Participants in Group 3 will achieve higher accuracy on the comprehension quiz than Group 1, and comparable or superior accuracy to Group 2. Limiting the AI summary to a two-sentence structure eliminates the wall-of-text fatigue. This allows non-expert users to understand the test measures without being confused by clinical terms.

5.3 Calibrated Trust and Active Verification Behavior (RQ3)

Participants in Group 3 will show higher verification behavior without giving in to automation bias or completely untrusting the AI. Group 2 (unstructured AI text) participants will either accept the summary or be skeptic. We anticipate think-aloud transcripts will show that users feel empowered to verify specific claims against official EHR notes, which will lead to appropriate trust calibration.

REFERENCES

[1] Danielle Scheurer. 2015. Hospitalists Should Embrace Advances, Transparency in Health Record Technology. The Hospitalist (March 2015). Retrieved from https://www.the-hospitalist.org/hospitalist/article/121703/hospitalists-should-embrace-advances-transparency-health-record

[2] Bethany Bruno, Scott Steele, Justin Carbone, Katherine Schneider, Lori Posk, and Susannah L. Rose. 2022. Informed or anxious: patient preferences for release of test results of increasing sensitivity on electronic patient portals. Health and Technology 12, 1 (2022), 59–67. https://doi.org/10.1007/s12553-021-00628-5

[3] Office of the National Coordinator for Health Information Technology. 2020. About ONC's Cures Act Final Rule. Retrieved November 18, 2020 from https://www.healthit.gov/curesrule/overview/about-oncs-cures-act-final-rule

[4] Elettra Carini, Leonardo Villani, Angelo Maria Pezzullo, Andrea Gentili, Andrea Barbara, Walter Ricciardi, and Stefania Boccia. 2021. The Impact of Digital Patient Portals on Health Outcomes, System Efficiency, and Patient Attitudes: Updated Systematic Literature Review. J. Med. Internet Res. 23, 9 (September 2021), e26189. https://doi.org/10.2196/26189

[5] Monika J. Krause, Verena Lengston, and Simon Frazer. 2026. Patient portal pitfalls: When digital access creates more work and anxiety. Journal of Family Medicine (February 2026).

[6] Traber Davis Giardina, Varsha Modi, Danielle E. Parrish, and Hardeep Singh. 2015. The patient portal and abnormal test results: An exploratory study of patient experiences. Patient Exp. J. 2, 1 (2015), 148–154.

[7] Niklas Mannhardt, Elizabeth Bondi-Kelly, Barbara Lam, Chloe O'Connell, Hussein Mozannar, Mercy Asiedu, Alejandro Buendia, Tatiana Urman, Irbaz B. Riaz, Catherine E. Ricciardi, Monica Agrawal, Marzyeh Ghassemi, and David Sontag. 2024. Impact of Large Language Model Assistance on Patients Reading Clinical Notes: A Mixed-Methods Study. arXiv preprint arXiv:2401.09637 (2024). https://doi.org/10.48550/arXiv.2401.09637

[8] Karan Singhal, Shekoofeh Azizi, Tao Tu, S. Sara Mahdavi, Jason Wei, Hyung Won Chung, Nathan Scales, Ajay Tanwani, Heather Cole-Lewis, Stephen Pfohl, Perry Payne, Martin Seneviratne, Paul Gamble, Chris Kelly, Abubakr Babiker, Nathanael Schärli, Aakanksha Chowdhery, Philip Mansfield, Dina Demner-Fushman, Blaise Agüera y Arcas, Dale Webster, Greg S. Corrado, Yossi Matias, Katherine Chou, Juraj Gottweis, Nenad Tomasev, Yun

Liu, Alvin Rajkomar, Joelle Barral, Christopher Semturs, Alan Karthikesalingam, and Vivek Natarajan. 2023. Large language models encode clinical knowledge. Nature 620, 7972 (August 2023), 172–180. https://doi.org/10.1038/s41586-023-06291-2

[9] Giuseppe Romeo and Daniela Conti. 2025. Exploring automation bias in human–AI collaboration: a review and implications for explainable AI. AI & SOCIETY 41, 1 (July 2025), 259–278.

[10] Todd Kulesza, Simone Stumpf, Margaret Burnett, Sherry Yang, Irwin Kwan, and Weng-Keen Wong. 2013. Too much, too little, or just right? Ways explanations impact end users' mental models. In Proceedings of the 2013 IEEE Symposium on Visual Languages and Human Centric Computing (VL/HCC '13). IEEE, Los Alamitos, CA, 3–10. https://doi.org/10.1109/VLHCC.2013.6645235

[11] Jessica L. Baldwin, Hardeep Singh, Dean F. Sittig, and Traber Davis Giardina. 2017. Patient portals and health apps: Pitfalls, promises, and what one might learn from the other. Healthcare 5, 3 (2017), 81–85.

[12] Jamie Keiko Fujioka, Julia Bickford, Jennifer Gritke, Vess Stamenova, Trevor Jamieson, R. Sacha Bhatia, and Laura Desveaux. 2021. Implementation Strategies to Improve Engagement With a Multi-Institutional Patient Portal: Multimethod Study. Journal of Medical Internet Research 23, 10 (2021), e28924.

[13] Alexandra Ramsey, Erin Lanzo, Hattie Huston-Paterson, Kathy Tomaszewski, and Maria Trent. 2023. Increasing Patient Portal Usage: Preliminary Outcomes From the MyChart Genius Project. Journal of Adolescent Health 72, 2 (2023), S34–S35.

[14] Donald A. Redelmeier and Nicole C. Kraus. 2018. Patterns in Patient Access and Utilization of Online Medical Records: Analysis of MyChart. Journal of Medical Internet Research 20, 2 (2018), e43.

[15] Shelley Vanderhout, Shipra Taneja, Kamini Kalia, Walter P. Wodchis, and Terence Tang. 2022. Patient Experiences and Perspectives When MyChart is Introduced in a Large Community Hospital: Mixed Methods Study. JMIR Human Factors 9, 2 (2022), e34593.

[16] Melita Avdagovska, Tania Stafinski, Mark Ballermann, Devidas Menon, Karin Olson, and Pauline Paul. 2020. Tracing the Decisions That Shaped the Development of MyChart, an Electronic Patient Portal in Alberta, Canada: Historical Research Study. Journal of Medical Internet Research 22, 4 (2020), e16794.

[17] Po-Yin Yen, Daniel M. Walker, Jessica M. Garvey Smith, Michelle P. Zhou, Terri L. Menser, and Ann Scheck McAlearney. 2018. Usability evaluation of a commercial inpatient portal. International Journal of Medical Informatics 110 (2018), 10–18.

[18] Daniel M. Walker, Terri Menser, Po-Yin Yen, and Ann Scheck McAlearney. 2019. Optimizing the User Experience: Identifying Opportunities to Improve Use of an Inpatient Portal. Telemedicine and e-Health 25, 3 (2019), 203–211.

[19] Mohammad Al-Ramahi and Cherie Noteboom. 2020. Mining User-generated Content of Mobile Patient Portal: Dimensions of User Experience. ACM Transactions on Social Computing 3, 3, Article 15 (2020), 24 pages.

[20] Saija Simola, Iiris Hörhammer, Yuhui Xu, Annika Bärkås, Asbjørn Johansen Fagerlund, Josefin Hagström, Mari Holmroos, Maria Hägglund, Monika Alise Johansen, Bridget Kane, Anna Kharko, Isabella Scandurra, and Sari Kujala. 2023. Patients' Experiences of a National Patient Portal and Its Usability: Cross-Sectional Survey Study. Journal of Medical Internet Research 25 (2023), e45974.

[21] Zhiqiang Huo, Timothy Neate, David Wyatt, Sophie Rowland-Coomber, Martin Chapman, and Iain J. Marshall. 2024. Co-Designing a User-Centred Digital Portal to Support Health-Related Self-Management for Stroke Survivors. In Proceedings of the 2024 IEEE 12th International Conference on Healthcare Informatics (ICHI '24). IEEE, 1–10.

[22] Ramsay Meiklem, Karen Stevenson, Sabine Richarz, David B. Kingsmore, Matt-Mouley Bouamrane, Mark Dunlop, and Peter Thomson. 2021. Advanced Kidney Disease Patient Portal: Implementation and Evaluation with Haemodialysis Patients. In Human-Computer Interaction – INTERACT 2021 (Lecture Notes in Computer Science, Vol. 12933). Springer, 175–196.

[23] Shefali Haldar, Sonali R. Mishra, Maher Khelifi, Ari H. Pollack, and Wanda Pratt. 2019. Beyond the Patient Portal: Supporting Needs of Hospitalized Patients. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). ACM, 1–14.

[24] Celine Latulipe, Amy Gatto, Ha T. Nguyen, David P. Miller, Sara A. Quandt, Alain G. Bertoni, Alden Smith, and Thomas A. Arcury. 2015. Design Considerations for Patient Portal Adoption by Low-Income, Older Adults. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). ACM, 3859–3868.

[25] Nielsen, Jakob. 1994. *Usability Engineering*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[26] Shneiderman, Ben. 1996. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages (VL '96)*, IEEE Computer Society, 336–343. https://doi.org/10.1109/VL.1996.545307

[27] Chew, LD., Bradley, KA., and Boyko, EJ. 2004. Brief questions to identify patients with inadequate health literacy. *Family Medicine*, 36(8), 588–594.

[28] Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology* (Vol. 52, pp. 139-183). North-Holland.

[29] Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, 31(3), 301-306.

[30] Braun, V. and Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa