# A critical assessment of the reliability and dependability of general practice (GP) websites - A case study of GPs around Bournemouth University

## **INTRODUCTION**

This study focuses on evaluating the trustworthiness and dependability of GP websites near Bournemouth University. Usability, which includes user satisfaction, ease of navigation, and learnability, is key to understanding how well users interact with these websites. GP websites are essential for patients to access information and make appointments, but the quality of information on these sites can vary. The study will assess five local GP websites to identify strengths and weaknesses in their content and functionality, ultimately providing recommendations for improvement to enhance patient experience and healthcare service delivery. The study is motivated by the increasing reliance on online health resources and concerns about the accuracy and reliability of information provided by GP websites.

**Aim and Objectives:**

The study aims to assess the reliability and dependability of general practice (GP) websites. The objectives are to:

* Identify usability criteria and user challenges in healthcare websites.
* Analyze user satisfaction with the design and usability of GP websites.
* Assess user engagement and satisfaction using usability techniques.
* Evaluate the simplicity and overall usability of GP websites.

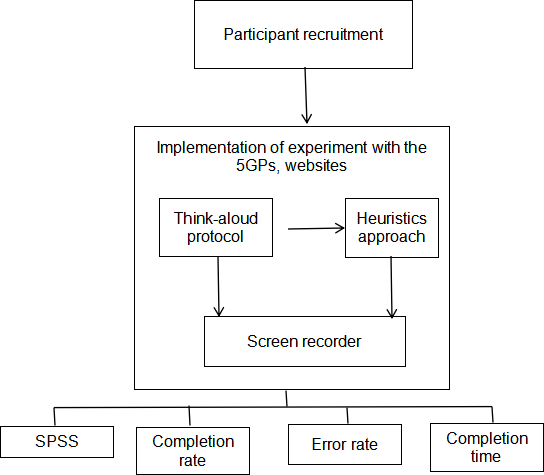
**Scope of the Study**

The study focuses on evaluating five GP websites in the Bournemouth University area. It will review each site’s accuracy, completeness, currency, and impartiality, providing insights into the strengths and weaknesses of these websites.

**Research Methods**

The first part of the method involves recruiting participants to partake in the usability testing. With this in mind, a total of 20 participants will be recruited to conduct the usability testing of the five GP websites around Bournemouth. A questionnaire containing various sections is used. The first part of the questionnaire, which is named pre-task, will focus on collecting general background information about the participants, like their age group, experience using GP websites, and gender. The second section of the questionnaire is the task scenario. In this part, the activity to be carried out by the participant is stated, and the tasks are a total of 8 tasks. Finally, the post-task section of the questionnaire provides insights on the usability of the GP websites based on the tasks performed by the participant. The method used consists of two approaches: thinkaloud protocol and heuristic approach. The participant session is recorded using a screen recorder application in order for the researcher to go through it for

any additional information when analysing the data collected. Figure 1 below shows the block diagram of the methods used.

 **Figure 1**

**Think-Aloud method**

The think-aloud method is a user-testing approach in which participants verbally express their ideas and behaviours while engaging with a product or system (Obrien and Wilson 2023). As users complete tasks, they express their impressions, decision-making processes, and any obstacles encountered in real time. This technique gives useful insights into users' preferences, thereby helping designers better understand how they perceive and interact with the interface. The think-aloud approach provides rich qualitative data by capturing users' ideas as they go through activities, which aids in identifying areas for improvement and directing design decisions. This method is common in usability testing and iterative development processes to improve user experiences.

**Heuristic Approach**

Heuristic assessment is a thorough evaluation of a product's user interface with the goal of identifying usability issues that may arise when consumers engage with it and developing solutions for them (Hussain 2023). The study uses three evaluators to examine the interface of each of the GP websites to detect usability flaws. Nielsen's heuristics were used for the user interface design testing. The heuristic comprises 10 concepts for determining the user design interface. Table 1 below provides the 10 Nielsen heuristics.

**Table 1**

|  |  |
| --- | --- |
| 1. Visibility of system status | 6. Recognition rather than recall |
| 2. Match between system and the real world | 7. Flexibility and efficiency of use |
| 3. User control and freedom | 8. Aesthetic and minimalist design |
| 4. Consistency and standards | 9. Help users recognize, diagnose, and recover from errors |
| 5. Error prevention | 10. Help and documentation |

**Participant Information**

This section of the questionnaire include the background information of the participant, such as age range and gender. Before conducting the experiment, the consent form is read out to the participant in which the participant answered verbally. The details about the study and why it is necessary will then be read to the participant.

**Pre-Test Questionnaire**

The pre-task section is verbally read to the participant. This aspect comprises of gathering information on user experience with e-health websites. The questions ranges from familiarity with e-health websites, purpose of the websites, and usefulness of the websites. This will make the participant to be more informed about the study goals.

**Tasks**

A task scenario has been developed for five GP websites, which are Talbot Medical Centre, Bournemouth University Medical Centre, Westbourne Medical Centre, The Village Surgery, and Haven Medical. This included standard tasks for the GP website, which accurately reflected their intended usage. Nielsen (1994) criteria for the beginning and end of assignments were taken into account while developing them. Easy exercises were chosen for the first and last tasks to make the user feel at ease and safe at the start, as well as proud once the evaluation was done. The first four tasks were easy, and the last four were a bit difficult. The task contains activities like booking an appointment, checking for information about common health practices, creating an account, submitting feedback, locating a list of services offered, exploring telehealth services, locating information about preventive care services , and locating emergency contact information.

**Post-Test questionnaire**

Once the task is completed, the post-task section is used to get feedback on the general usability of the GP websites. This part consists of 7 questions. The participant was asked to verbally answer the question as the researcher read it out. The question focuses on suggestions for improvement of the website and the general interface rating using the 5-point Likert scale method and the open-remark method.

**Pilot**

A pilot test is conducted before the main test to confirm the user-testing technique. This is a crucial phase since it allows you to rehearse the test while also identifying and refining any issues with the testing procedure, such as unrelated activities or deceptive questions. Three evaluator were selected to provide feedback and make any necessary changes prior to the actual test. Pilot test findings is examined, and used to improve the user testing questionnaire.

**Evaluation Procedure**

Each user-testing activity followed the same methodology. Data was collected using screen recorder software. The researcher introduces the purpose of the experiment and the participant’s right to resign at any time. The participant is also informed that the screen will be recorded using screen recording software throughout the session. The user is then asked to review and sign a consent form.Before beginning any website-related tasks, the user will be asked to visit the site for no more than 10 minutes. The pre-test questionnaire is read out by the researcher, and the participant answers accordingly. The participant is then given tasks to perform. The time estimate for each task was determined during the pilot stage. The researcher takes note of such things as the time taken to execute the task and any comments provided by the participant as each task is being conducted. Following completion of the activities for the inspected website, the user is provided a post-test questionnaire so as to provide feedback.

### Data Analysis

The data is analysed to determine which methods pinpoint the root of each usability problem. This analysis will be undertaken on an individual basis, and the results will be interpreted. The second part involved compiling a list of frequent usability concerns found by each technique using three expert evaluators. The problem identified is classified and analysed using a heuristic approach. The data collected from the participants testing is analysed using the completion rate, completion time, and error rate. Also, the SPSS package is used to further analyse the data.

## RESULT AND DISCUSSION

The performance of the heuristic and think-aloud protocols used in this study is presented. The think-aloud protocol further categorised participants based on their experience level.

### Evaluation metrics for think-aloud protocol

The study uses 20 participants. Each of the participants conducts an evaluation on the five selected GP websites. Participants level of experience is categorised into experience and no experience. To evaluate the performance of the websites based on the participant survey, three attributes were considered: completion rate (CR) in percentage, error rate (ER) in percentage, and completion time (CT) in seconds. The formulas for the three attributes are provided below.

Completion rate = × 100%

Error rate = × 100%

Completion time = Sec

### Performance of think-aloud method

From Section A of the questionnaire used for this study, the results indicate the participant's details, such as age group and gender, as shown in Tables 2 and 3, respectively. Many of the participants believed the websites were used to access medical information, and their level of experience was determined in Section B.

The findings from Section C, which involves the participants carrying out some tasks on the websites, are provided. A total of six tasks were conducted each for the five GP websites picked. Tables 4 to 8 present the performance of the participants for each GP website based on the evaluation metric used, while table 9 describes the average performance of the five GP websites using the metrics.

**Table 2: Age distribution of participants**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **AGE** | | | | | |
| Frequency | | | Percent | Valid Percent | Cumulative Percent |
| Valid | 15-20 years | 1 | 5.0 | 5.0 | 5.0 |
| 21-30 years | 13 | 65.0 | 65.0 | 70.0 |
| 31-40 years | 6 | 30.0 | 30.0 | 100.0 |
| Total | 20 | 100.0 | 100.0 |  |

The table indicates that 1 (5%) of the participants is between 15 and 20 years old, 13 (65%) are between 21 and 30 years old, and 6 (30%) are between 31 and 40 years old. The prevalent age range is 21 to 30, as shown in this study. Figure 4.1 shows the age distribution using a bar chart.

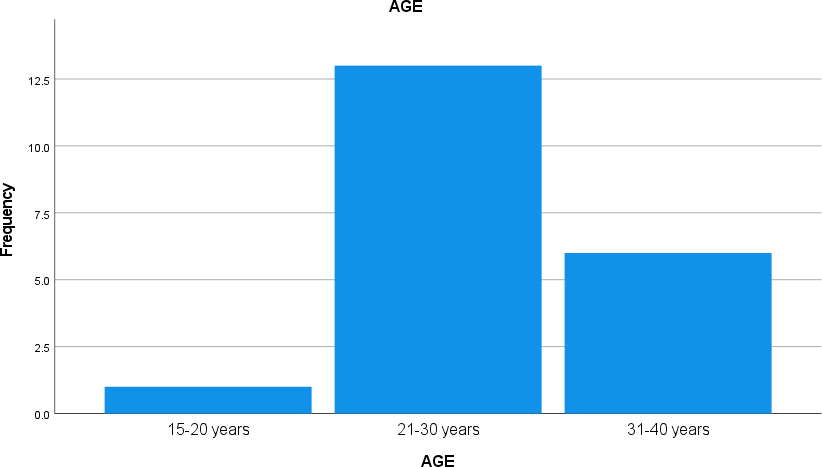
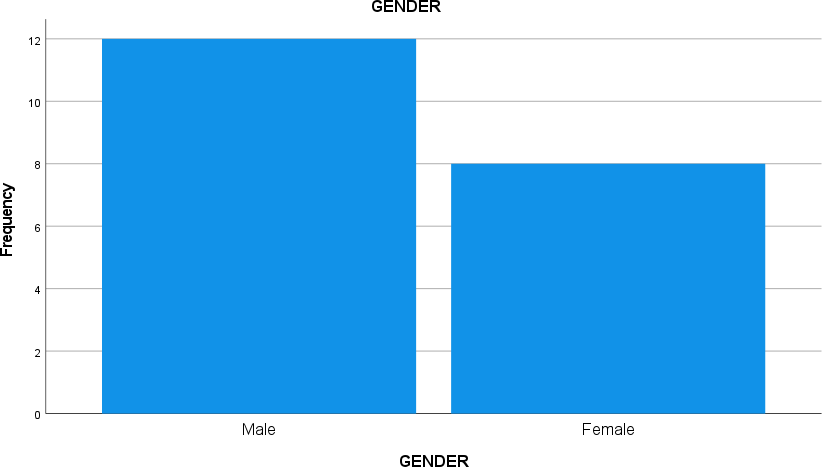


Figure 2: Age distributions of participants

**Table 3: Gender distribution of participants**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **GENDER** | | | | | |
| Frequency | | | Percent | Valid Percent | Cumulative Percent |
| Valid | Male | 12 | 60.0 | 60.0 | 60.0 |
| Female | 8 | 40.0 | 40.0 | 100.0 |
| Total | 20 | 100.0 | 100.0 |  |

The study demographic evaluation indicates that 12 (60%) of the participants are men, and the remaining 8 (40%) are women. This indicates that the dominant gender is male.

Gender distribution of perticipants

**Table 4: Performance of participants evaluation on Talbot medical centre based on the evaluation metrics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S/N** | **Tasks** | **Completion Rate (%)** | **Errors(%)** | **Completion Time**  **(in seconds)** |
| 1 | Create an account on the website. | 65 | 65 | 409 |
| 2 | Find information about common health issues. | 18 | 70 | 421 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 3 | Explore services rendered for mental health support or counselling. | 80 | 75 | 242 |
| 4 | Find information about the clinic's working hours. | 80 | 35 | 225 |
| 5 | Submit feedback. | 83 | 25 | 311 |
| 6 | Locate emergency contact information in the event of urgent  medical requirements. | 98 | 5 | 158 |

**Table 5: Performance of participants evaluation on Westbourne medical centre based on the evaluation metrics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S/N** | **Tasks** | **Completion Rate (%)** | **Errors(%)** | **Completion Time**  **(in seconds)** |
| 1 | Create an account on the website. | 95 | 0 | 325 |
| 2 | Find information about common health issues. | 50 | 100 | 438 |
| 3 | Explore services rendered for mental health support or counselling. | 98 | 5 | 148 |
| 4 | Find information about the clinic's working hours. | 95 | 10 | 198 |
| 5 | Submit feedback. | 85 | 30 | 138 |
| 6 | Locate emergency contact information in the event of urgent medical requirements. | 95 | 15 | 243 |

**Table 6: Performance of participants evaluation on Poole Town surgery based on the evaluation metrics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S/N** | **Tasks** | **Completion Rate (%)** | **Errors(%)** | **Completion Time**  **(in seconds)** |
| 1 | Create an account on the website. | 100 | 0 | 255 |
| 2 | Find information about common health issues. | 60 | 65 | 348 |
| 3 | Explore services renderedfor mental health support or counselling. | 80 | 10 | 169 |
| 4 | Find information about the clinic's working hours. | 75 | 50 | 287 |
| 5 | Submit feedback. | 68 | 65 | 274 |
| 6 | Locate emergency contact information in the event of urgent medical requirements. | 95 | 20 | 212 |

**Table 7: Performance of participants evaluation on The Village Surgery based on the evaluation metrics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S/N** | **Tasks** | **Completion Rate (%)** | **Errors(%)** | **Completion Time**  **(in seconds)** |
| 1 | Create an account on the website. | 93 | 15 | 288 |
| 2 | Find information about common health issues. | 45 | 90 | 404 |
| 3 | Explore services rendered for mental health support or counselling. | 43 | 40 | 294 |
| 4 | Find information about the clinic's working hours. | 93 | 15 | 127 |
| 5 | Submit feedback. | 75 | 50 | 191 |
| 6 | Locate emergency contact information in the event of urgent medical requirements. | 98 | 5 | 145 |

**Table 8: Performance of participants evaluation on Adams Practice based on the evaluation metrics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S/N** | **Tasks** | **Completion Rate (%)** | **Errors(%)** | **Completion Time**  **(in seconds)** |
| 1 | Create an account on the website. | 90 | 33 | 245 |
| 2 | Find information about common health issues. | 90 | 0 | 221 |
| 3 | Explore services rendered for  mental health support or counselling. | 78 | 15 | 251 |
| 4 | Find information about the clinic's working hours. | 98 | 10 | 170 |
| 5 | Submit feedback. | 85 | 20 | 152 |
| 6 | Locate emergency contact  information in the event of urgent medical requirements. | 58 | 40 | 315 |

**Table 9: Average of the GP websites based on the evaluation metrics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S/N | GP websites | Completion Rate (%) | Errors(%) | Completion  Time  (in seconds) |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 | Talbot Medical Centre | 71 | 46 | 294 |
| 2 | Westbourne Medical Centre | 86 | 27 | 248 |
| 3 | Poole Town Medical  Centre | 80 | 35 | 258 |
| 4 | The Village Surgery | 75 | 44 | 242 |
| 5 | Adams Practice | 82 | 20 | 226 |

From the analysis of participant findings in Table 9 above, Westbourne Medical Centre performed best with a completion rate of 86%, a 27% error rate, and a completion time of a completion time of 248 seconds. Adams Practice is the second-best-performing GP website, with a completion rate of 82%, an error rate of 20%, and a completion time of 226 seconds. The Village Surgery has a completion time of 75%, an error rate of 44%, and a completion time of 242 seconds. Lastly, Talbot Medical Centre has the worst performance, with a completion rate of 71%, an error rate of 46%, and a completion time of 294 seconds.

Section D of the questionnaire provides more insight into the participants feedback based on the task scenario carried out. Table 10 provides the statistics of the questions using SPSS package tools.

**Table 10: Statistics of post task questions**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| AGE | | | GENDR | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 |
| N | Valid | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| Missin g | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mean | | 2.25 | 1.40 | 3.55 | 2.80 | 3.75 | 3.75 | 4.05 | 3.90 | 4.10 |
| Median | | 2.00 | 1.00 | 4.00 | 3.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 |
| Mode | | 2 | 1 | 4 | 2a | 4 | 4 | 4 | 4 | 4 |
| Std. Deviation | | .550 | .503 | .945 | .894 | .550 | .851 | .510 | .641 | .718 |
| Variance | | .303 | .253 | .892 | .800 | .303 | .724 | .261 | .411 | .516 |
| Range | | 2 | 1 | 4 | 3 | 2 | 3 | 2 | 2 | 2 |

### Performance of heuristic evaluation

The heuristic evaluation was carried out on the five GP websites, indicating the common usability issues based on expert findings and issues identified during the participant sections of the think-aloud approach. This issue arises because some of Nielsen's 10 usability guidelines were not adhered to. The identified issues are presented in table 11, and table 12 provides the heuristic evaluation summary.

**Table 11: Usability issues identified for the GP websites**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| S/N | Issues | Talbot | Westbo urne | Poole Town | The Village Surgery | Adam Practic e | Description |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | Poor navigation |  | - | - |  | - | The users find it challenging to  locate the  information they are searching for on the website. |
| 2 | Complex interface design |  | - | - | - | - | Too many words and the irregularity of the design make it Uninteresting and difficult to locate important information. |
| 3 | Missing symptom checker option |  | - |  | - | - | Users could not find an option to check for disease symptoms, causes, and treatment on the  website. |
| 4 | Missing predicting search option. |  | - |  |  |  | The search box does not predict or make suggestions when searching for words. |
| 5 | Limited feedback mechanisms. |  | - | - |  | - | The website does not have a feedback option to lodge complaints. |
| 6 | Lack of integration with social media |  | - | - | - | - | The website does not have integration with  Social media platforms. |
| 7 | The footer is not well designed. | - |  | - | - | - | The footer does not have the proper information  required, and it is not well designed. |
| 8 | Language options are missing. | - |  | - | - | - | There is no option to select other languages. |
| 9 | Hidden content. | - | - |  | - | - | Important information is hidden behind multiple layers. |

**Table 12: Heuristic evaluation summary**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S/N** | **Issues** | **Severity rating** | **Ease of fix rating** | **Heuristic number** | **Recommendation** |
| 1 | Poor navigation | 3 | 2 | #1  #2  #3  #4 | Redesigning menu labels, grouping relevant items, and reorganising the navigation hierarchy. |
| 2 | Complex interface design | 3 | 2 | #1  #3  #4 | simplifying the layout, eliminating irrelevant features, and optimising whitespace to improve visual hierarchy. |
| 3 | Missing symptom checker option | 3 | 2 | #1  #2 | Integrating a third-party symptom checker API  into the website, such as |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | the NHS Health A-Z webpage. |
| 4 | Missing predicting search option. | 2 | 2 | #1  #2 | Creating a predictive search feature, which includes a frontend implementation that displays search recommendations as  users input data. |
| 5 | Limited feedback mechanisms. | 3 | 3 | #1  #2 | Implementing new feedback methods, such as frontend design and backend functionality for gathering and  Processing user comments. |
| 6 | Lack of integration with social media | 2 | 2 | #1  #2 | By adding new integration options to connect with social  networking APIs and services. |
| 7 | The footer is not well designed. | 2 | 3 | #1  #2 | A complete redesign of the footer, including a new layout, information organisation, and visual  style. |
| 8 | Language options are missing. | 3 | 2 | #1  #2 | by adding a language selection API to the website. |
| 9 | Hidden content. | 3 | 3 | #1  #2 | Reevaluating the website's design and content organisation to ensure that critical information is  prominently displayed and easily accessible |

***Note: Severity rating:*** *1=Not a usability issue; 2=Minor usability issue; 3=Major usability problem; 4= Severe Usability problem.*

***Ease of fix rating****: 0= extremely easy to fix; 1= can be fixed quickly & easily, 2= fixing may take moderate amount of time, 3= fixing take planning & significant amount of time, 4= requires a significant effort to fix*

**Discussion of results**

This project uses two approaches to evaluate the performance of five GP websites around Bournemouth. For the think-aloud protocol, a total of 20 participants were recruited. The questionnaire used is divided into four sections. Each participant performed a total of six tasks on the five selected GP websites in Section C. A screen recorder was used to record the screen sessions with participants. To get the performance analysis, three metrics were used. The metrics are completion rate, error rate, and completion time. For Section A, participant information was collected. male dominating with a tota of 12 (60%), and for age range, 21 to 30 years dominate with a total of 13 (65%). From the metrics, Westbourne had the best performance in terms of completion rate, with a CR value of 86%, followed by Adams practice with a CR value of 82%. Talbot has the worst performance in terms of completion rate. For error rate value, Adams practice has the best performance with a low error rate value of 20%, followed by Westbourne with 27%, and Talbot also has the worst performing value of 46%. Last but not least, for completion time, Adams practice produced the best performance with a low CT of 226 seconds, followed by the village surgery with 242 seconds. Talbot also had the worst performance with a CT value of 294 seconds. Furthermore, SPSS tools were used to determine the mean, median, mode, standard deviation, and variance for the post-task question in Section D.

The second method used for the study is a heuristic approach. Nielsen’s usability guidelines were followed. A total of 9 usability issues were identified with the aid of experts in website evaluation. The common issues were compared between the five GP websites, and a profound recommendation was suggested.

### Conclusion

The usability evaluation based on the reliability and dependability of general practice (GP) websites indicates the overall performance using the think-aloud and heuristic approaches. This study uses the selected technique to evaluate the five GP websites. The comparison between the GP websites was presented using both techniques. From the think-aloud approach, the evaluation metric result was that Westbourne had the best performance in terms of completion rate, with a CR value of 86%, followed by Adams practice with a CR value of 82%. Talbot has the worst performance in terms of completion rate. For error rate value, Adams practice has the best performance with a low error rate value of 20%, followed by Westbourne with 27%, and Talbot also has the worst performing value of 46%. Last but not least, for completion time, Adams practice produced the best performance with a low CT of 226 seconds, followed by the village surgery with 242 seconds. Talbot also had the worst performance, with a CT value of 294 seconds. The heuristic approach identified a total of nine issues; the issues were compared across the GP websites, and recommendations were provided. Furthermore, the severity, ease of use rating, and heuristic rules violated are also presented.