EMBAIAnce

A Decision Support System incorporating a validated patient-specific, multi-scale Balance Hyper model towards early diagnostic **E**valuation and efficient **M**anagement plan formulation of **Balance** Disorders

EMBalance

Grant Agreement 610454

Overall Users' Evaluation Report

Deliverable D8.2

Project Title	EMBalance—A Decision Support System incorporating a validated patient-specific, multi-scale Balance Hypermodel towards early diagnostic **E**valuation and efficient **M**anagement plan formulation of **Balance** Disorders
Grant Number	610454
Project Type	Collaborative Project

Title of Deliverable	Overall Users' Evaluation
Subtitle of Deliverable	
Deliverable Number	D8.2
Dissemination Level	PU
Internal Rev. No.	V1
Contractual Delivery Date	M36
Actual Delivery Date	M38
Updated on	
Contributing WPs	WP8
Editor(s)	CITY
Author(s)	Ioannis Basdekis, George Spanoudakis and Maria Krotsiani
Reviewer(s)	Dimitris Kikidis, Doris Bamiou

Document History

Version	Type of Change	Author(s)
0.8	Table of content, 1stDraft	I. Basdekis (CITY)
0.87	Section on Correlation to D8.1 (methodology), questionnaire data (appendix c)	I. Basdekis (CITY)
0.89	Corrections, editions after 1st internal review	I. Basdekis (CITY)
0.91	Short description of UCL focus group	Laura Rammazzo (UCL)

Document Review

Date	Version	Reviewer	Comment
06/03/2017	0.88	Corrections, editions prior to internal review	D. Kikidis (UOA)
07/03/2017	0.90	Edits	DE Bamiou (UCL)

Executive Summary

This deliverable reports on the results of the second evaluation stage conducted with potential end-users (clinicians) of the EMBalance web platform. The platform has undergone two major evaluation phases to assess its usability and efficiency before coming to the current state (ver. 2.0). Preliminary results of the evaluation's first stage are reported in D8.1. The purpose of this document is to present the final outcomes of the usability testing activities performed during the last quarter (M32-35) of the EMBalance project.

User evaluation is of significant importance in order to assess the platform according to the end-users' needs and enhance the approval of the system from clinicians and the balance disorders-related healthcare community. To this purpose questionnaire results from EMBalance and other clinicians from the Health Care community, as well as feedback from three focus groups were collected, analysed and used to produce enhancements in both interaction and functionality elements (reported in D6.5). Methodologically, this stage continues the evaluation procedure that was previously designed and implemented, and as such reutilises the online questionnaire previously created (Appendix A) to collect information regarding user-friendliness and the efficiency of all the platform facilities. In addition, focus groups were conducted in a precise and methodologically adequate manner. These provided input for specific design refinements with special focus on shortening the overall data insertion time required for providing DSS suggested diagnosis(-es), a parameter regarded from all participants as the most influencing factor of success. There were three focus groups held in London, Athens and Freiburg (voice recordings).

This process revealed a small number of usability issues. As with the previous evaluation stage, the collected feedback:

- reaffirmed positive evidence regarding the usability of the platform

- confirmed that alterations introduced improved user experience (as reflected on increased approval ratings)

- corroborated the potential effect on clinical practice mainly as a tool offering education in clinical practice

- usability ratings were uniformly high, while the only generic concern raised had to do with time a doctor has to spend inputting several parameters (completion time)

Enhancements implemented in the final version 2.0 of the EMBalance platform (and reported in D6.5), were addressed successfully. It's worth mentioning that the vast majority of end-users acknowledge the potential value of the platform as an educational tool, which asserts that the implemented e-services are considered to be user-friendly even by non-experienced clinicians.

Contents

Table of figures

13

Chapter 1. Introduction

1.1 Background

The main aim of the EMBalance project was to develop a unified web-based platform, to be utilized by primary and secondary care physicians for the diagnostic evaluation, behaviour prediction and effective management planning of balance problems. Since this specific end-user target group operates in a highly stressful environment, user friendliness and usability of offered e-services should be achieved with the highest possible levels of acceptance in order for these to be taken into account on clinical practice.

To achieve these objectives, state-of-the-art methods incorporating Human Computer Interaction (HCI) techniques have been woven into the development process and influenced the design of the EMBalance web presentation and interaction elements (and indirectly the underlying functionalities). In order to assess the potential clinical service adoption of the EMBalance, the objective of the WP8 was to conduct and evaluate by means of various low-level assessments (of usability, efficiency, speed, accuracy), as well as to validate on a higher assessment level the proof of concept and potential value of the platform. The results of the continuous evaluation stages were constantly fed back to inform the design and development of several aspects towards the refinement of the platform.

In the context of WP8, the main objectives of the evaluation task were:

- to identify challenges and weak points in terms of interaction with offered e-services in real clinical practice and ways of addressing them, and

- to evaluate and improve the user-friendliness of the platform and reduce the workload of clinicians.

In summary, the evaluation was based on two main activities:

- Continuous user evaluation based on feedback from clinicians who were directly associated with the EMBalance project, and from focus groups (externals).

- A questionnaire-based evaluation involving both clinicians external to the project who provided feedback using it during the clinical trials (by invitation), and others willing to express their opinion (open version).
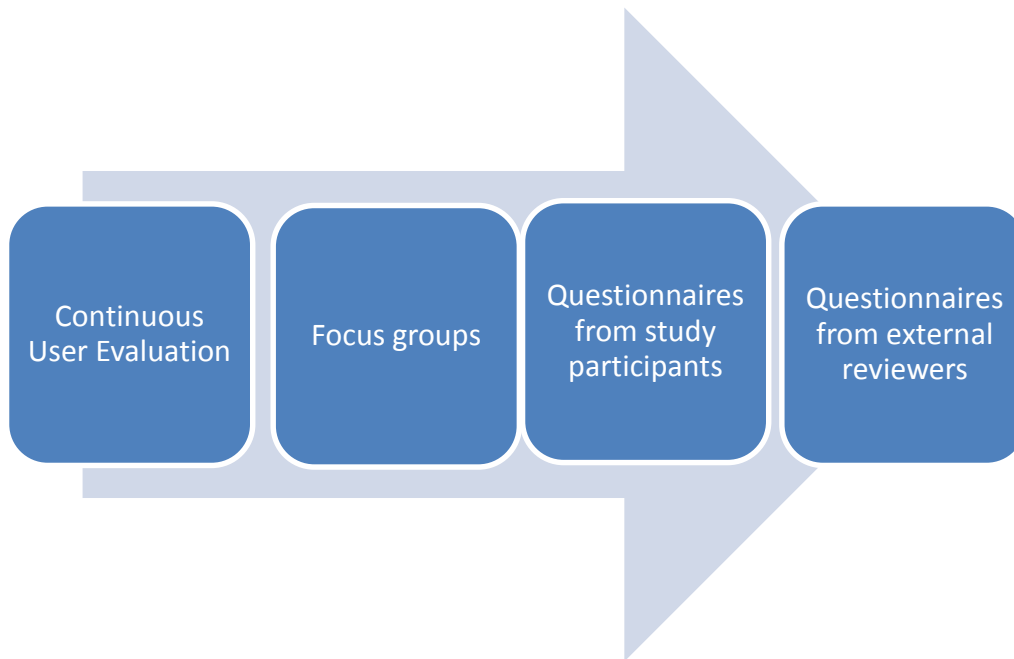
The outcome of this activity was twofold; the assessment from a usability point of view and as a consequence the introduction of specific design refinements with special focus on shortening the overall insertion time required for providing DSS suggested diagnosis(-es). In parallel, the final version of the platform incorporated models and parameters in a user friendly way to invoke DSS calculations in order to predict diagnosis in a more precise manner (Balance control, ICS Modelling tool). The latter is reported in detail in D6.5, while in this report we present and discuss the final outcome of the evaluation activity.

The majority of average satisfaction scores for all participants was positive indicating thus that participants were in general satisfied by the platform's design and the facilities provided. At the same time some new requirements were identified from pilot groups. Regarding completion time, which was highlighted as the most critical factor to be considered (and improved), specific alterations were implemented to simplify the flow of actions and improve even further the work load required by the clinicians. The following table (Table 1) presents alterations made (and

presented in D6.5) to accommodate users concerns and elements identified that required to be adapted.

Table 1: Changes in workflow and design in versions 1.5 and 2.0.

	Comments and actions	Version	Importance
	"An introductory home page is needed to inform end-users about implemented functionality".		Low
	A home page was implemented.	2.0	
	"After logging, instead of the search module, input form for creating a new patient considered as having higher priority, while search function should be a secondary option".		High
	At home page, the "create patient" form will appear. Viewing existing patients provides feedback and reassures of previously inserted records.	1.5	
	"Possibility to edit a visit previously entered".		High
	The end-user may edit/delete every previously inserted medical information (i.e., medical history, symptoms, clinical examinations and investigations, and clinical diagnosis-es) for a selected patient (within his/hers organization).	1.5	
	"Search button must be present at all times. It is essential having this functionality available in every view screen".		Medium

✔	Buttons will be placed on the top-left corner to indicate high priority task.	1.5	
👥	"A GP should be able to select any of the 12 diagnoses as clinical diagnosis, in addition to those diagnoses generated by DSS".		High
✔	End-user may select any from the 12 available diagnoses (suggested by DSS or not), or insert one (free text) of his own.	1.5	
👥	"Instead of downloading the explanatory video corresponding to a clinical examination parameter as a file, consider presenting the video in a separate window. Same behavior to be applied for text instructions".		Medium
✔	Explanatory videos were loaded from the EMBalance YouTube channel (to optimise speed) and appear in a JavaScript window. Same system behaviour implemented applied to explanatory instructions for clinical examinations.	1.5	
	Definitions for all examinations were added.	2.0	
👥	"Reduce the number of diseases. The ICD-10 taxonomy contains many values that confuse and delay process".		High
✔	Disease list updated to the following: Acoustic Neuroma, Atypical posterior benign paroxysmal positional vertigo, Atypical benign paroxysmal	2.0	

	positional vertigo, Benign neoplasm: Cranial nerves, Bilateral vestibular failure / dysfunction, BPPV, posterior, BPPV, lateral, BPPV, anterior, Central lesion, other, Cerebellar / Pontine lesion, Chronic subjective dizziness PPPD, CPA Acoustic Neuroma, Heart failure, unspecified, Meniere's, Migraine, unspecified, Multiple sclerosis, Neonatal hypertension, Psychological disorders, Stroke, not specified as haemorrhage or infarction, Typical benign paroxysmal positional vertigo, Typical posterior canal benign paroxysmal vertigo, Vestibular migraine, Vestibular neuronitis, Vestibular paroxysmia and Unilateral vestibular failure/weakness		
	"GP should be able to insert up to 9 vertigo triggers at the same window".		Medium
	Clinical examination form was altered accordingly.	1.5	
	"Reduce parameters to be inserted".		High
	Medical history parameters were prioritised (open vs toggle down sections).	2.0	
	Implement progressive disclosure where necessary to lower the user's cognitive load (show/hide associated form fields).	2.0	

			Medium
	Instructions page was implemented.	2.0	
	Changes in terminology and parameter values	1.5 & 2.0	High
	Information messages appear in view pages and forms.	2.0	High
	Integration of Balance Control Tool parameters and ICS Model tool parameters.	2.0	High

The remaining chapters of this evaluation document are structured as follows: Chapter 2 briefly elaborates on the evaluation method followed, while Chapter 3 presents final results of the usability assessment and corresponding analysis.

1.2 Correlation to D8.1

D8.1 described the rationale for the usability evaluation undertaken, and presented the preliminary results of each phase conducted for D8.1 (i.e., a) user-based evaluation from clinicians and b) preliminary results of the usability assessment via an online satisfaction questionnaire). User input and suggestions from members of the consortium (expert clinicians), as well as feedback from 10 clinicians (from UK and Greece) who were invited to participate during this first phase led to several enhancements and alternations of the platform (ver. 1.5). This preliminary analysis (presented in D8.1) indicated a positive attitude towards the scope and implemented functionalities, while the majority of satisfaction scores suggested that participants were in general terms satisfied by the platform's design and the facilities it provided, although at the same time most of them considered that the completion time (based on the usage scenario(s) followed prior to completion) could be improved further.

During the reported period, and while improvements in structuring, in emphasizing (or deemphasizing) elements and in introducing integration

with several models (to improve accuracy of DSS results) were undergoing, the same evaluation method via online questionnaire served as a further source for improvements and as a measure for overall usability and efficiency in comparison to the previous results (in D8.1). To this purpose, we present not only the satisfaction indicators derived from this, but also differences observed amongst participants groups of those who used extensively the latest version of e-services, vs. those who did not. To this end, we are presenting the results collected from community members who were invited (via email) to participate (Set A: 24 participants) during the second phase of the clinical trials, as well as those collected via a separate open version that made available to anyone willing to comment on the platform (Set B: 20 participants).

Chapter 2: Evaluation Method and Process

2.1 Usability challenges for the provision of a user-friendly web interface for the EMBalance platform

Nowadays, all e-services present usability challenges because of the need to serve differentiated target groups (specific vs generic), to provide access to a variety of devices (thus supporting basic vs complex interaction techniques), and to keep up with necessary adaptations that are required in order to support operation within diverse environments (traditional office vs mobile usage). In all circumstances, focusing on the characteristics of the target group and the environment in use implies an explicit design focus to address limitations derived from usage conditions.

Although generic usability challenges on how to develop usable Web applications are well documented (e.g., [Nielsen, 1999], [Spool, 2007]), designers are not informed when considering usability issues [Becker & Mottay, 2001]. Particularly in the context of an e-health web application such as the EMBalance web platform enabling activities to be performed by clinicians in stressful environments such as the ones this project aims for (e.g., hospitals, private clinics, etc.), usability improvements of the interaction elements targeted in the reduction of the information load, should be regarded as a major design consideration. Within this environment, end-users (clinicians) are in a hurry, can be confused by the platform sequence of actions (that may not resemble their own clinical practice) or by the terminology (use of basic ICT instead of medical terminology). These considerations are related directly to the satisfaction clinicians feel towards operating an e-health system, regardless of the quality of the produced outcome. As for the latter, general skepticism is reported on whether someone could trust e-health outcomes for decision making [Hesse & Shneiderman, 2007]. A generic assumption currently debated is that that most of the available systems were originally designed to manipulate data, and not to fit into a clinical environment and help clinicians' fulfil their tasks.

These well-known challenges were taken into consideration from the early design stages of EMBalance web interface. The application of HCI

techniques throughout the development process of medical systems, such as in particular decision support systems, enhances the probability of achieving intended goals in terms of user satisfaction ([Marcilly et al, 2013], [Purves et al, 1999], and [Walton et al, 1997]). As mentioned in [Purves et al, 1999], the NHS-funded decision support system PRODIGY utilized several common HCI techniques (such as interviews, interaction design, information architecture, user testing and evaluation) with the intent that the outcomes of these techniques would lead to greater satisfaction and adoption of the system. Only after conducting user evaluations (and adhering with suggestions and feedback) it was concluded that general practitioners received the system positively.

With reference to D6.3 and D8.1, the involvement of clinicians – project partners - early in the design process ensured not only high probability of successfully delivering acceptable e-services, but also better resemblance of the EMBalance DSS supported diagnostic evaluation to that during standard clinical practice. During this early stage, different sketches were given to clinicians to observe the problems they were facing in performing basic tasks (the "think Aloud" technique [Erikson &Simon, 1985]. These discussions (conducted face to face in a dedicated meeting in Athens and digitally via skype in numerous consecutive calls) helped the designers materialise scenarios and abstract work flows into concrete interface blocks, while valuable comments provided early feedback on how to cope with usability issues that might arise at a later development stage. During this phase, clinicians provided valuable feedback for all stages of the development procedure. This feedback originated from consortium members, as well as from other clinicians working in the four clinical sites of the project, but not directly involved in the EMBalance project. To this purpose, several modifications of the platform's initial design– in a form of detailed design sketches – were presented in order to improve user experience. [Self, J. 2016] argues that sketching is considered as a critical part of conceptual design since ideas presented can be assessed early for their suitability, or act as a scaffold for reasoning between design problems and their potential solutions. As an aftermath of this process, simplification, interface consistency, effectiveness and resemblance of paper-based clinical practice (Figure 1) were considered to be the most crucial aspects to be taken into consideration and subject to repeated discussions with them. As we were dealing with an e-health system, effective usage was considered more important than efficiency since accuracy and easiness in inputting patient's health data are more important than marginal gains in speed [Whitney, 2009].

Consequently, user support was enhanced through the provision of consistent interaction mechanisms (e.g., input forms, presentation elements, colour coding, etc.) that attempted to replicate actual medical workflows and practice, while incorporating functional specifications of the EMBalance repository tool and the hardrepository constraints (reported om D3.1).

Figure 1 – Example of a paper imprinting of patient's medical history.

During the construction of several alternative designs of the platform, the aforementioned restrictions influenced every step of the design process (e.g., Figure 2). This exercise eventually led to a more acceptable overall design of EMBalance web platform that was completely different than the one presented in early drawings (D2.2).

Example: selection of date picker

Figure 2 – Example of date input element: actual clinical practice favors 3 different input fields instead of a most commonly utilized date picker for inputting a birthday, as in several cases patients do not recall such data during visits (exact dates or months).

Figure 3 presents only a fraction of the successive prototypes that were developed to elicit user feedback during the iterative design. This systematic work, which was considered to be the first evaluation stage, led eventually to the first implementation of the EMBalance platform version 1.0 with functional features presented in D6.3, and subsequently to the final version 2.0 presented in D6.5 (Figure 4).
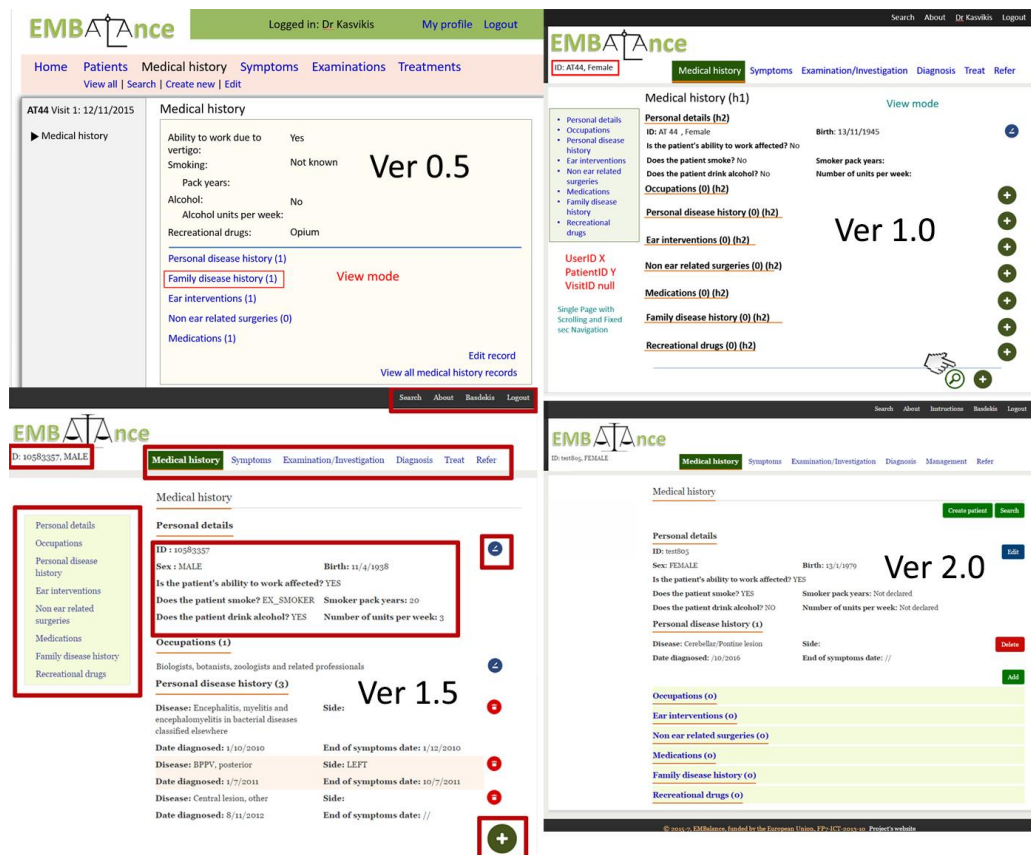
Figure 3 – Example of the progression of alternative design mock-ups presenting patient's medical history record: A) early sketch (version 0.5), B) final design (version 1.0), C) intermediate version 1.5, D) final version 2.0to increase.
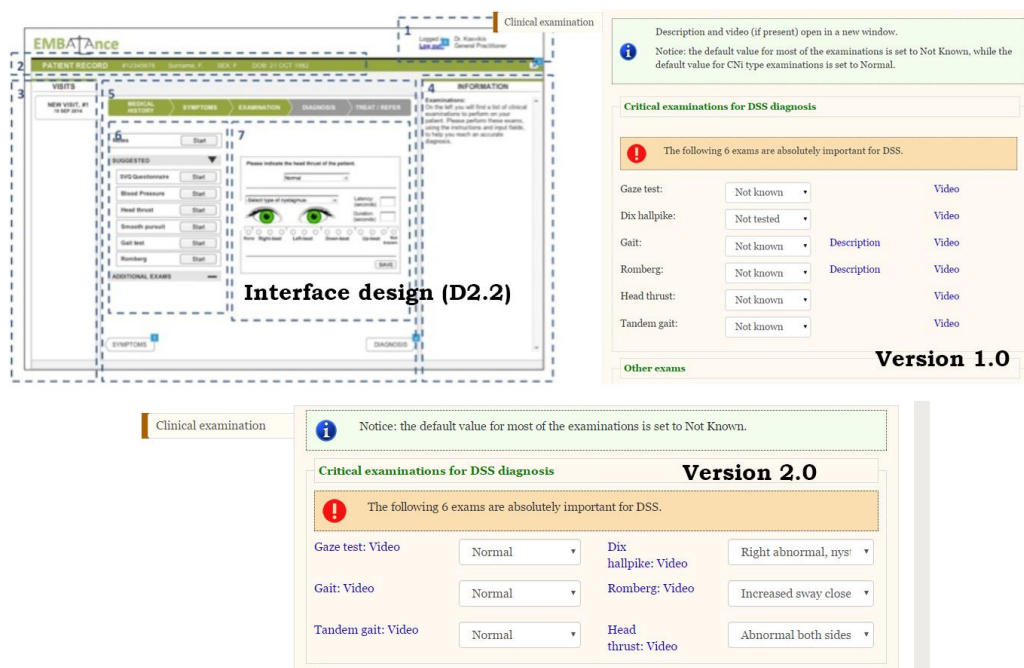
Figure 4 – Comparison of the early web interface design (D2.2), the one implemented in the alpha version of the EMBalance platform (version 1.0) and the final one available in version 2.0.

2.2 Usability and Efficiency Evaluation

██████ Methodology

The early feedback retrofitted the design choices and continuously improved user experience. The first operable version (version 1.0) of the EMBalance platform was delivered on January 9th 2016 and provided the basic functionality. Shortly after the introduction of this first operable version, feedback from clinicians who used this functionality indicated that targeted design changes were necessary in order to accommodate newly identified requirements, as well as to increase their overall user satisfaction. This input led to the successive version 1.5 that was made available on May 26th 2016, while the iterative design and (re-)development process continued. In order to evaluate the implemented e-services and to determine important factors that may improve user experience.

The general purpose of a formal usability evaluation is to diagnose whether end-users can learn and subsequently use efficiently a product tool to

achieve their goals, as well as to measure their satisfaction levels [Usability.gov, 2017]. In order to accomplish this purpose towards delivering the final version of the platform, evaluation activities conducted during the final stage of the project included comments collected during lab tests (Clinical trials: Phase B), data from the online questionnaire, and feedback from the focus groups held in London, Athens and Freiburg (with 3 participants respectively) (Figure 5).
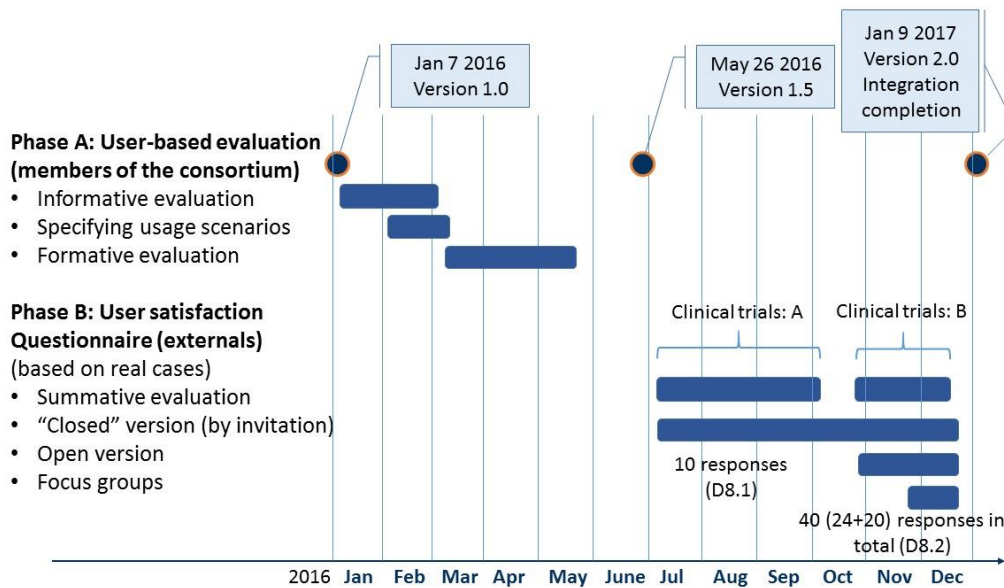
Figure 5 – Different testing phases provided critical feedback to identify usability problems, and determine end-user's satisfaction.

The online questionnaire utilized during phase A of clinical trials (D8.1) was composed with the intention to provide to (external) end-users the opportunity to express anonymously their thoughts regarding the usability of the facilities offered. In fact, this method was considered to be one of the most crucial aspects of the evaluation, since the results would come out from actual community members. This questionnaire assessed:

- end-users' overall satisfaction from their interaction with the platform, and

- actual usefulness of the functionality been introduced in clinical practice.

The usability section (i.e., closed ended questions 1 to 13) was based primarily on [Lund, 2001]. The reason for this was that a community member could declare that the platform did not satisfy him/her for a

specific purpose (e.g., as educational tool), or had a useful generic characteristic (e.g., navigation). In other words to focus on aspects that should be further improved in order to serve his/her needs. As [Lund, 2001] indicates, the first aim was to make the items as simply worded as possible, and as general as possible and constructed as five-point Likert rating scales (1: strongly disagree to 5: strongly agree). The questionnaire (its rationale presented in D8.1) included also close-ended questions (i.e., questions 14 to 20) aimed at evaluating the usefulness of the EMBalance functionality. Three open-ended questions could be used for other thoughts and comments. During the second stage of the clinical trials participators from different ages, genders and profession occupations (but relevant to the scope of the project) were invited to fill questionnaire.

In parallel, and during this final stage of the evaluation, an open version of the same questionnaire was made available for anyone willing to comment on the offered e-services. This version was disseminated through specific email lists (again relevant to the scope of the project), while an indicative message (Figure 6) on the web home page of the EMBalance urged visitors to use the EMBalance functionality (via a demo account credentials) and, at later stage, to complete the online questionnaire.

Requesting your feedback

We would kindly like to ask your feedback in regards to a decision support system about balance disorders. It is an easy to use platform and does not require previous education. Only think that requires attention is that the user should press the green button at the end of each screen after entering data.

You may use the following credentials:

Username: test
Password: p@ssword

After entering the platform, you can enter the details of a (real or imaginary) patient. Some fields with asterisk are required. The platform will suggest diagnosis and treatment.Please feel free to use the platform as long as you wish to have an opinion. More detailed instructions are also provided in case needed.

After taking some time to explore the platform and maybe enter the details of one or more patients (real or imaginary), you are kindly requested to fill in the specific online form (it will take less than 3 minutes).

 ℹ   EMBalance questionnaire

We will be more than happy to provide any information or clarification necessary. Thank you very much in advance.

Figure 6 – Home page: a message on how to use the EMBalance platform with a demo account.

Focus groups

In addition to the previous formal form of gathering user input, three focus groups were held in London, Athens and Freiburg, providing the

opportunity to the participants to thoughtfully answer questions at a deeper level. The methodology to conduct them based upon the following elements ([Marczak & Sewell, 2017] and [Simon, 1999]):

- Individuals to be invited had to be clinicians, experts in the field, in order to contribute to helping the research gain a greater understanding of the topic, subset of the indented target group,

- Focus group(s) may have 3-5 participants, as this was considered the optimal size to promote discussion,

- Invitations to participants should be sent at least 4 weeks before trials (phase B), and their participation should be verified,

- Facilitators should inform about the voice recording and ask for participants permission, and remind that all input is confidential,

- Facilitators should be assisted by a note taker, since it is important – in addition to voice recordings if applicable - to record the focus group feedback. Discussion should not be impeded by the facilitator delaying the discussion in order to take notes,

- In the beginning, facilitators would inform on the study purpose,

- Facilitator should present basic functionally implemented, prior to discussion, however one participant should be invited to interact to the system during discussion. This would also enable the facilitator to keep the group on task,

- One of the participants was given 4-5 task to perform, while others were watching (via projector) and discussing ("think aloud" technique),

- Five open-ended questions to promote discussion (similar to ones from the online questionnaire) were selected as considered of most importance. These were:

  o Q5 - Do you think the user-interface of the DSS was intuitive? (What did you like the most, and what do you think should be changed?)

  o Q6 - How easy it was to complete tasks using the system?

  o Q17- Was the DSS successful in providing the correct diagnosis?

- o Q9 - Do you think the EMBalance DSS could be used as routine practice in primary care clinics? (Why?)

- o Q10 - Do you think the EMBalance DSS could be used as an educational tool for healthcare professionals with interest in ENT medicine? (Why?)

Feedback from this stage was consistent to the one resulted from the analysis of questionnaire data, and presented in the following chapter (3.2).

Chapter 3. Results of the usability assessment

3.1 Introduction

This chapter presents the final outcome of the usability assessment conducted with potential end-users (clinicians) of the EMBalance web platform. The platform has undergone two major evaluation phases to assess its usability and efficiency before reaching its current state (ver. 2.0). Preliminary results of the first stage involved an internal part during which preliminary feedback from 10 clinical practitioners directly involved in the project was collected and reported in D8.1. This chapter presents the final outcomes of the usability testing activities performed during the last quarter (M32-35) of the EMBalance project, during which more than forty clinical practitioners, not related to the project, participated and stated their opinion about the offered e-services. Two different results are presented in order. The first one briefly elaborates on the outcome of the focus groups, while the second presents results derived from the questionnaire.

Questionnaire results from EMBalance and open community clinicians, as well as the feedback from three focus groups were collected, analysed and retrofitted enhancements in both interaction and functionality elements in version 2.0 of the platform (reported in D6.5). Methodologically, this stage continues upon the evaluation procedure designed and implemented previously, and as such reutilised the online questionnaire created (Appendix A) to collect information regarding user-friendliness and the usefulness of all the platform facilities. In addition, focus groups although considered rather a secondary method for evaluating usability, were conducted in a precise manner. These provided input for specific design refinements with special focus on shortening the overall insertion time required for providing DSS suggested diagnosis(-es).

3.2 Focus Groups Outcome

Nielsen, the guru of usability, states that "in interactive systems development, the proper role of focus groups is not to assess interaction styles or design usability, but to discover what users want from the

system" [Nielsen, J. 1997]. In the context of this statement, three focus groups were conducted in Athens, London and Freiburg to collect specific impressions from clinicians that had already utilised the EMBalance platform.

According to the specific protocol (briefly mentioned in 2.2), five questions were selected to evoke discussion. Expert clinicians, prior to discussion, tested the implemented features and provided their feedback in that given context. Although all groups expressed positive comments about the usefulness and feasibility for the final outcome, they raised concerns about the total time required to input health data during a visit which should be limited under 10 minutes. Overall:

end-users expressed in their inputs a positive attitude towards the scope and implemented functionality,

they were in general satisfied by the platform's design and the facilities provided,

completion time was highlighted by them as the most critical factor to be considered (and improved), and

suggested specific alterations to simplify the flow of actions (e.g., prioritisation) and to improve even further the work load that would be required by them.

In the case of University of Athens:

- Two Professors of ENT-Neurology and one experienced physician with more than 25 years in the field of Neurology were invited to comment

- Their overall comments were positive and did find the user interface intuitive. They considered tasks self-explanatory and thought that online instructions for the users would be enough and no face to face education would be necessary.

- All cases that they tried returned correct diagnoses and this feature increased their trust to the system. They had some concerns for adaptation in routine clinical practice, since system usage requires considerable time and level of knowledge to correctly populate cases in pending for many primary healthcare clinicians, however they considered this a valuable tool for specific and difficult cases.

- finally they strongly recommended to further integrate the system in order to develop an educational tool.

In the case of UCL:

- Two trainee GPs who participated in the EMBalance clinical trial were invited to the focus group.

- It was highlighted that one of the biggest strengths of the system is the learning material available through the platform. They found the process enriching and a great opportunity to learn more about ENT medicine. Furthermore, they were also asked whether they found the system difficult to use, and if they considered that 1 hour training session was enough to correctly apply the DSS in their daily clinic. Overall, they confirmed that 1 hour session was enough training, and although the first patient took slightly more time to input, usage became easier with some experience.

- Among the weaknesses of the system, it was mentioned that the EMBalance DSS required on average 20 minutes to correctly populate. However, during a standard visit, GPs have a 10 mins time-slot per patient. In their opinion, the system as it is requires too much time to be fully implemented in their routine practice. Nonetheless, they remarked that they found the system very useful for learning purposes.

3.3 Questionnaire Results

Results acquired through the usability questionnaire (mainly from closed version) are addressed in what follows.

█████Generic info

Participants were (Figure 7):

- Closed version: 24 clinicians (15 male, 9 female). In this version, participants in clinical trials were invited to participate by

- Open version: 20 clinicians (15 male, 5 female, 7 from Greece, 3 from England and 10 without declaring)

All participants that took part in phase B of clinical trials were also invited (via email) to participate in the questionnaire. Emails were sent to various

mail lists of otolaryngologists. Participants had been informed that responses should be given anonymously.
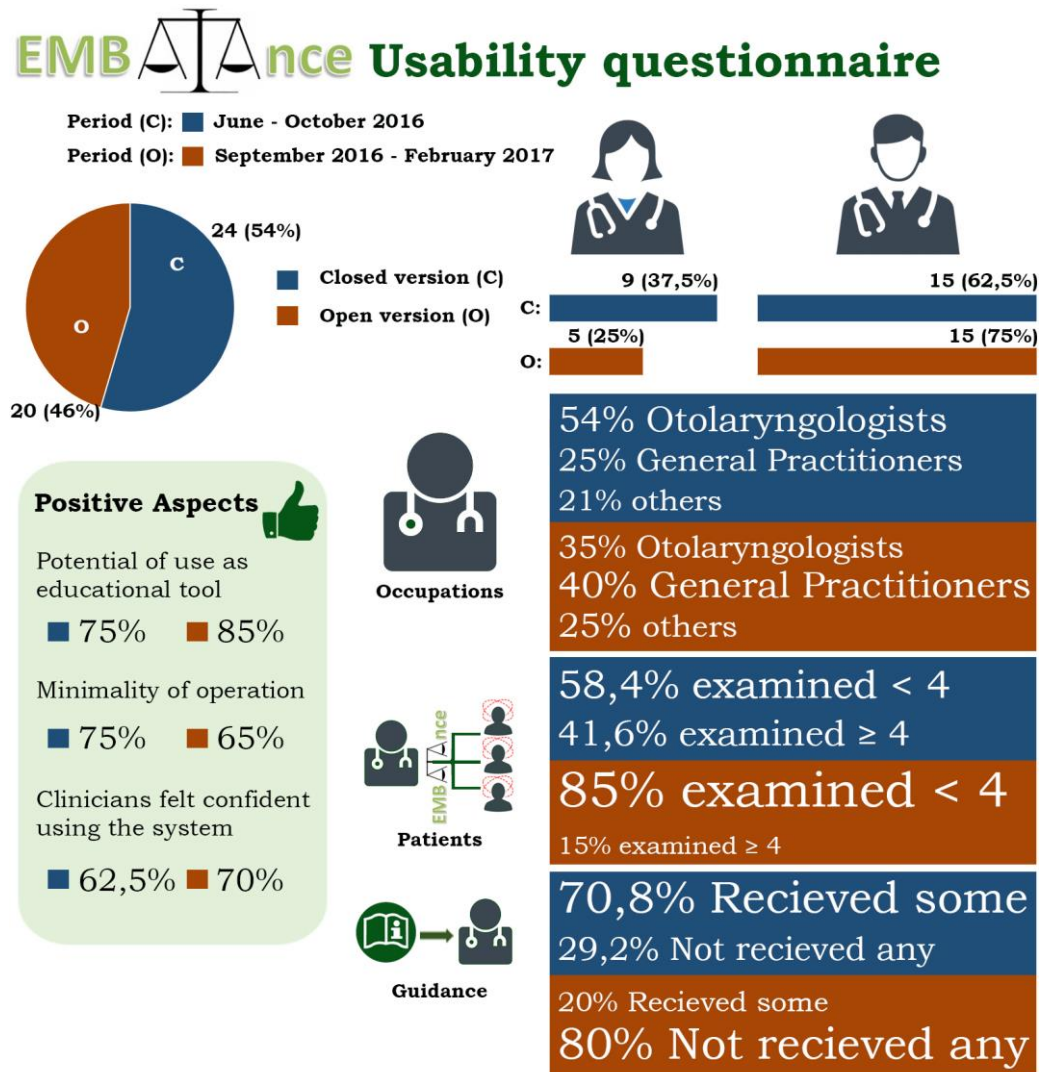
Figure 7 – Questionnaire's infographic.

3.4 Closed version

▮▮▮ General profiling questions

24 clinicians from UK, Greece and Germany participated during the reported period. Of those:

- 62,5% were male and 37,5% female (Figure 8) (previously reported 60% and 40% respectively).

- 54% were Otolaryngologists, 25% General Practitioners and 21% other categories (Figure 9) (previously reported 60% ENT doctors and 40% General Practitioners respectively).

- 58,4% examined less than 4 patients and 41,6% 4 or more with the use of the EMBalance platform (Figure 10) (previously reported 60% and 40% respectively).

- 70,8% received some guidance as to what EMBalance platform offers to the end user, while 29,2% did not (Figure 11) (previously reported 80% and 20% respectively).

Figure 8 – Participants' gender.
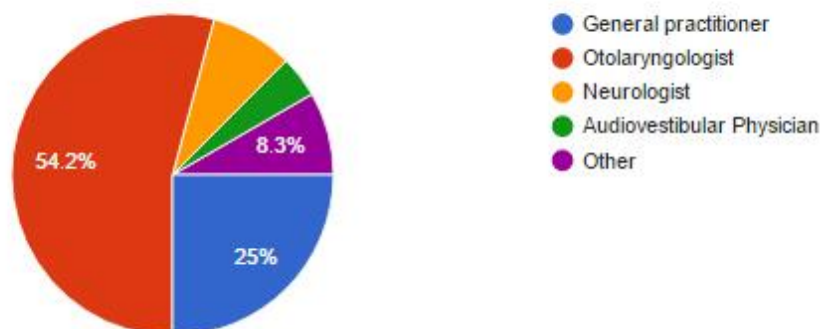
Figure 9 – Participants' specialty.

Disciplines distribution is representative of the medical specialties managing patients with balance disorders. Multiple disciplines ensure that results are representative of various points of views and backgrounds and consequently are of an increased value.

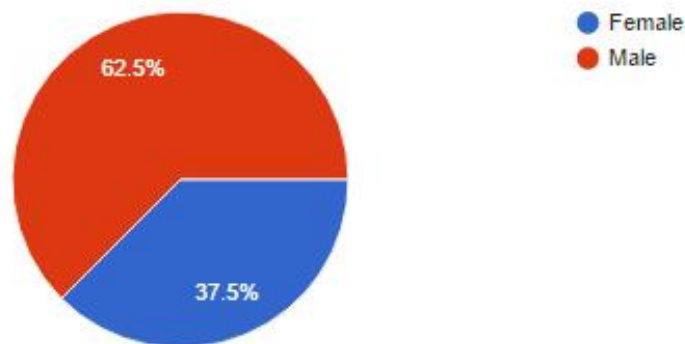How many patients have you seen using the EMBALANCE platform before completing this questionnaire?
(24 responses)
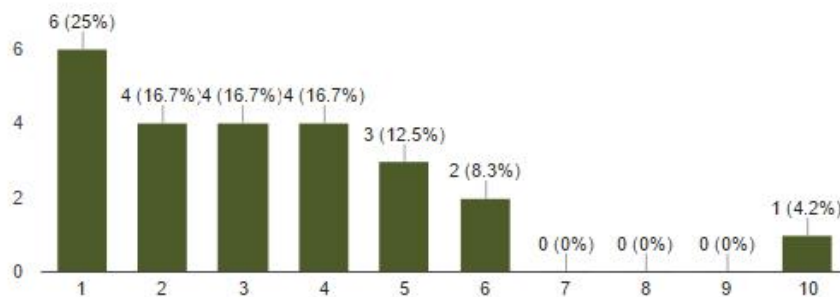
Figure 10 – Number of cases examined with the system.

The majority of participants stated that they had examined more than 3 patients with use of the DSS, which is considered satisfactory to express an opinion, since their initial impression is determinative in order to judge a decision support system and decide whether one should use it or not. Moreover, according to the study protocol, participants should not examine more than 6 patients, in order to avoid learning curve effect in the study results.

Have you received training or any assistance in the use of EMBalance platform?
(24 responses)
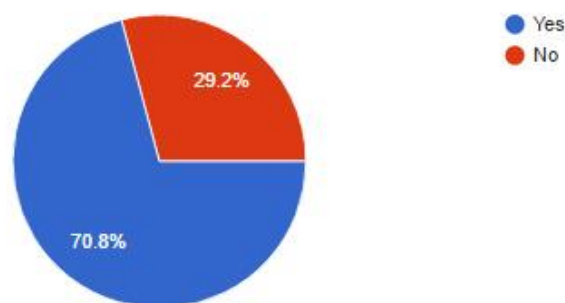
Figure 11 – Provision of guidance and instructions prior of completing the questionnaire.

█████Usability

Regarding the usability of the EMBALANCE platform, the following occurred:

Positive aspects:

- *Intuitiveness of user interface:* 50% of the participants considered that the interface is intuitive (i.e., rate 4 or more), while 29,2% of them expressed a neutral view regarding it (i.e., rate 3) and 20,9% found it not intuitive (i.e., rate 2 or less) (Figure 12) (previously reported 40%, 40% and 20% respectively).

The interface of the system was intuitive. (24 responses)
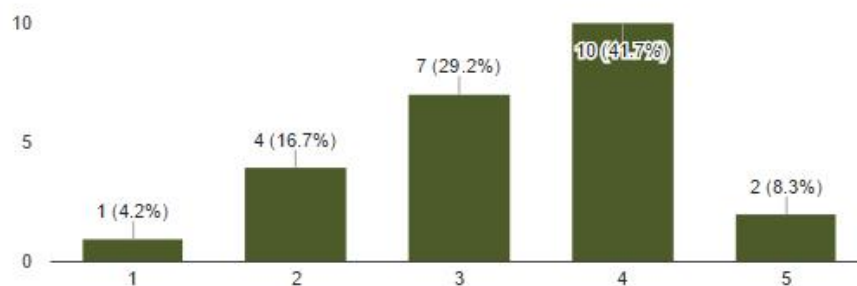
Figure 12 – Intuitiveness of user interface.

- *Ease of use:* Half of the participants (i.e., 50%) considered the system easy to use (i.e., rate 4 or more), 29,2% of them expressed a neutral view regarding it (i.e., rate 3) and 21,8% of them found it not easy (i.e., rate 2 or less) (Figure 13) (previously reported 50%, 20% and 30% respectively).

It was easy to complete tasks using the system. (24 responses)
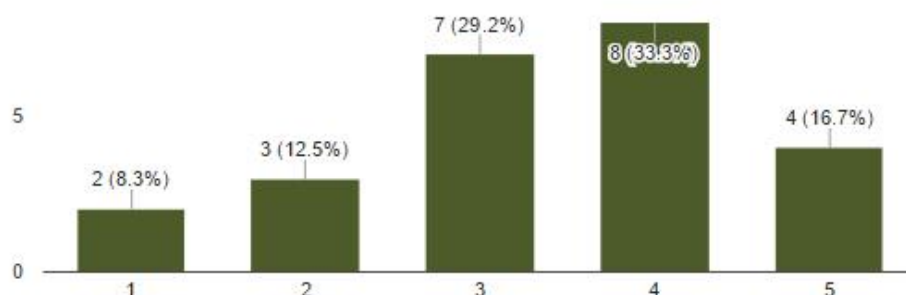
Figure 13 – Easiness of task completion.

- *Navigation mechanism:* 45,8% of the participants considered the navigation mechanisms of the platform to be effective (i.e., rate 4 or more), 20,8% of them expressed a neutral view regarding it (i.e., rate 3) and 33,3% of them found it not easy (i.e., rate 2 or less) (Figure 14) (previously reported 40%, 30% and 30% respectively).

Navigation through the different options of the system was effective.
(24 responses)
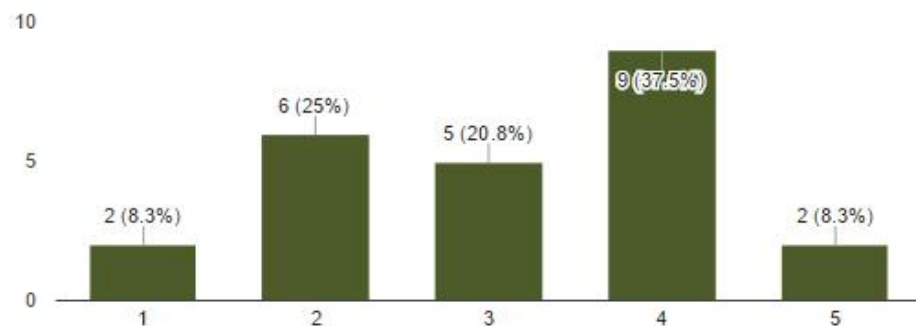
Figure 14 – User-friendly navigation.

- *User confidence:* 62,5% the participants felt confident in using the system (i.e., rate 4 or more), 29,2% of them expressed a neutral view regarding it (i.e., rate 3) and 8,3% of them found it not easy (i.e., rate 2 or less) (Figure 15) (previously reported 50%, 40% and 10% respectively).

I felt confident using the system. (24 responses)

Figure 15 – User confidence.

- *Potential of use in clinical practice:* 45,9% the participants indicated that they would use the EMBALANCE platform in clinical practice

(i.e., rate 4 or more), 25% of them expressed a neutral view regarding it (i.e., rate 3) and 29,1% of the participants expressed a less positive view (i.e., rate 2 or less) (Figure 16) (previously reported 50%, 20% and 30% respectively).

I could use this system frequently in normal clinical practice. (24 responses)
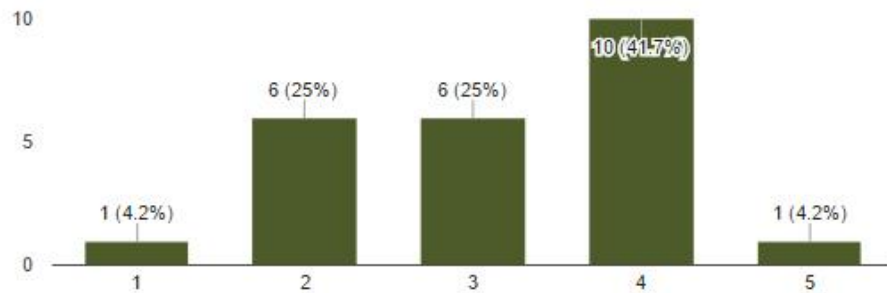
Figure 16 – Utilisation of the system as it is in clinical practice.

- *Potential of use as an educational tool*: 75% the participants indicated that they would use the EMBALANCE platform frequently as an educational tool (i.e., rate 4 or more), 16,7% of them expressed a neutral view regarding it (i.e., rate 3) and 8,3% of the participants expressed a less positive view (i.e., rate 2 or less) (Figure 17) (previously reported 50%, 30% and 20% respectively).

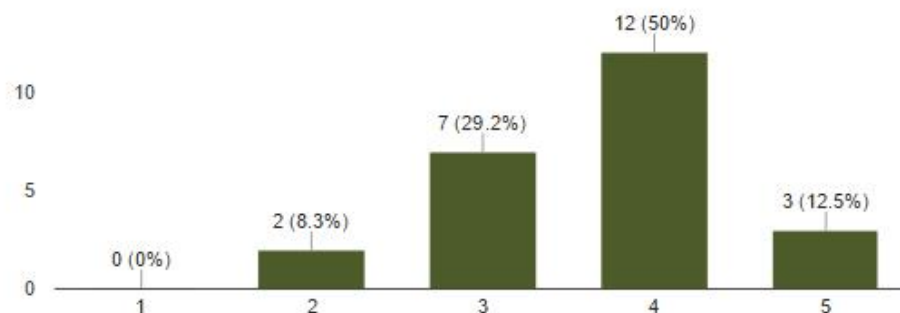**I could use this system frequently as an educational tool.** (24 responses)
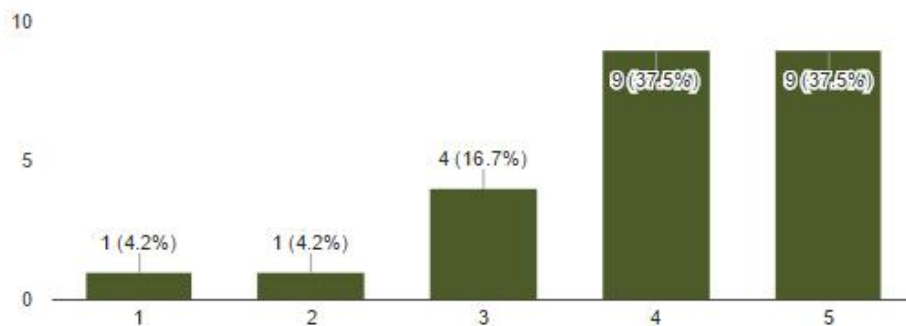
Figure 17 – Utilisation of the system as educational tool.

- *Effective requires technical advice:* the system was considered as self-explanatory as it was anticipated since 54,1% of the participants indicated that they would not need technical assistance to use it (i.e., rate 2 or less), 16,7% of them expressed a neutral

view regarding it (i.e., rate 3) and 29,2% of the participants expressed a less positive view (i.e., rate 4 or more) (Figure 18) (previously reported 50%, 16,7% and 33,3% respectively).

I would need technical advice in order to use the system effectively.
(24 responses)
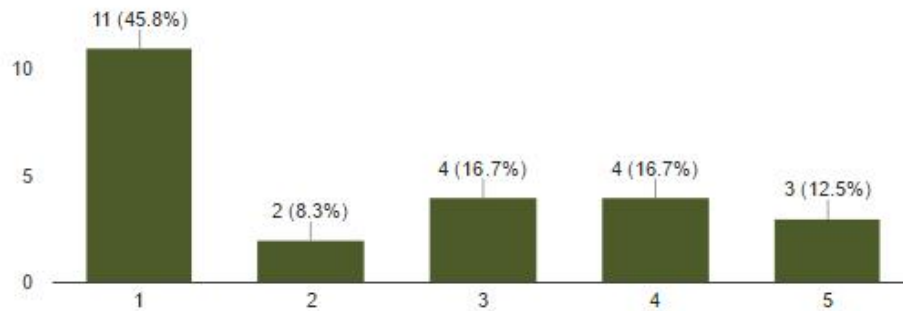
Figure 18 – Self-explanatory usage.

Neutral or negative aspects:

- *Guidance and instructions required for effective usage:* 50% of the participants indicated that they could operate the system without specific training in order to use it effectively (i.e., rate 2 or less), while 41,7% expressed opposite view (i.e., gave it a score of 4 or more) and 8,3% of them expressed a neutral view regarding it (i.e., rate 3) (Figure 19) (previously reported 50%, 50% and 0% respectively). This positive shift can be justified due to: a) introduction of help instructions, and b) the extensive usage for more than one patients by the majority of end users.

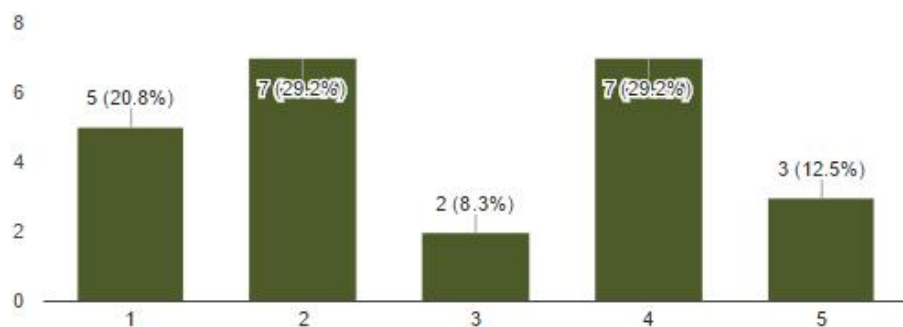**I would need training in order to use the system effectively.** (24 responses)

Figure 19 – Guidance and instructions required for effective usage.

40

- Completion time: Still 54,1% of the participants indicated the time required to complete a visit-task with the use of the system as not satisfactory when comparing to their everyday practice (i.e., rate 2 or less), while 37,6% expressed a positive view (i.e., rate 4 or more) and 8,3% of them expressed a neutral view regarding it (i.e., rate 3) (Figure 20) (previously reported 50%, 30% and 20% respectively).

I feel satisfied with the amount of time it took to complete a task in comparison to my previous working method.
(24 responses)
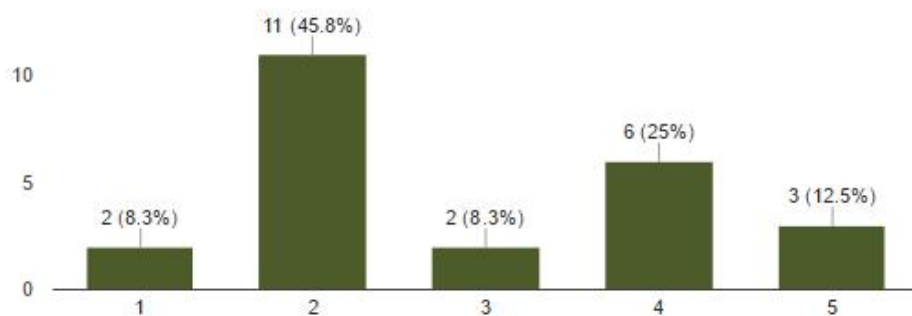
Figure 20 – How quickly a task can be completed in comparison to normal clinical practice.

Usefulness of EMBALANCE platform functionality

The questionnaire included also questions (i.e., questions 14 to 20) that aimed at evaluating the usefulness of the EMBALANCE functionality. The outcomes of our preliminary evaluation based on these questions are summarized below:

Positive aspects:

- *Enhancing effectiveness:* Half of the participants (50%) indicated that the system improved their effectiveness in the follow up of their patients visits (i.e., rate 4 or 5), 37,5% of the participants expressed a neutral view (i.e., rate 3) and another 12,5% expressed a less positive view (i.e., rate 2 or less) (Figure 21) (previously reported 60%, 20% and 20% respectively).

The use of the system has helped me to be more effective in seeing patients.
(24 responses)

Figure 21 – Effect of EMBalance on practitioner effectiveness.

- *Enhancing productivity: Again t*he majority of the participants (50%) indicated that the system improved their productivity (i.e., rate 4 or more),25% expressed a n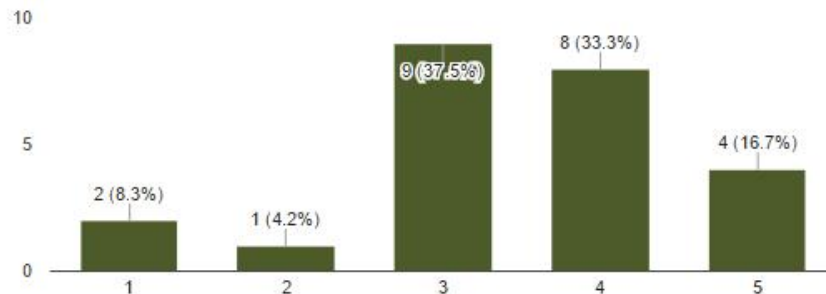eutral view(i.e., rate 3) and another 25% expressed a less positive view (i.e., rate 2 or less) (Figure 22) (previously reported 60%, 0% and 40% respectively).

The use of the system has helped me to be more productive. (24 responses)
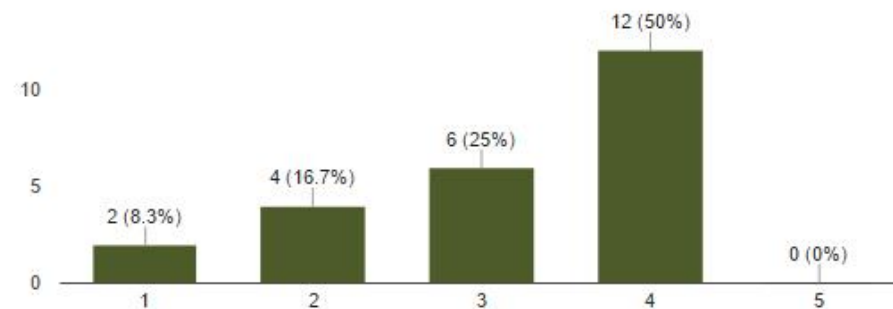
Figure 22 – Effect of EMBalance on practitioner productivity.

- *Successful use*: 45,8% indicated that they used the system successfully every time was required during trials or elsewhere (i.e., rate 4 or more), 41,7% of the participants expressed a neutral view (i.e., rate 3) and another 12,5% expressed a less positive view (i.e., rate 2 or less) (Figure 23) (previously reported 50%, 30% and 20% respectively).

Furthermore, 45,8% of the participants indicated that the system was successful in performing the task they had been asked to execute (i.e., rate 4 or more), 41,7% gave a neutral response (i.e., rate 3), and 12,5% expressed a less positive view (i.e., rate 2 or

less) (Figure 24) (previously reported 40%, 60% and 0% respectively).

I used it successfully every time. (24 responses)
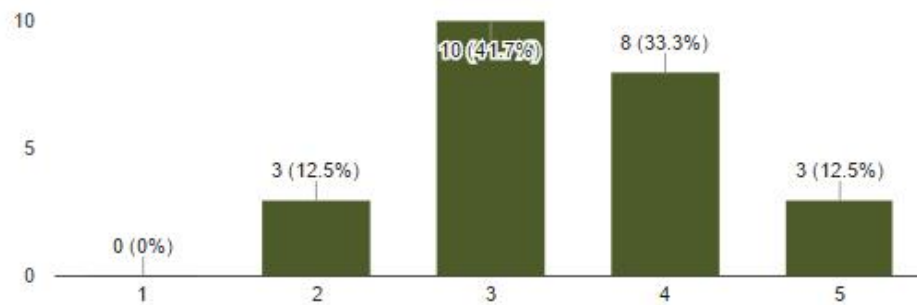
Figure 23 – Successful use of the platform.

It is successful in performing its intended task. (24 responses)
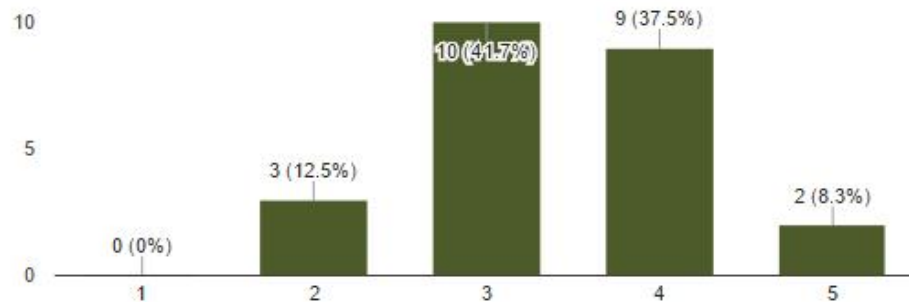
Figure 24 – Success of platform in supporting an intended task.

- *Operating according to expectations:* 54,1% of the participants indicated that the system operated according to their expectation (i.e., rate 4 or more), 25% of the participants expressed a neutral view in this respect (i.e., rate 3)and another 20,8% expressed a less positive view (i.e., rate 2 or less) (Figure 25) (previously reported 30%, 50% and 20% respectively).

It works the way I want it to work. (24 responses)
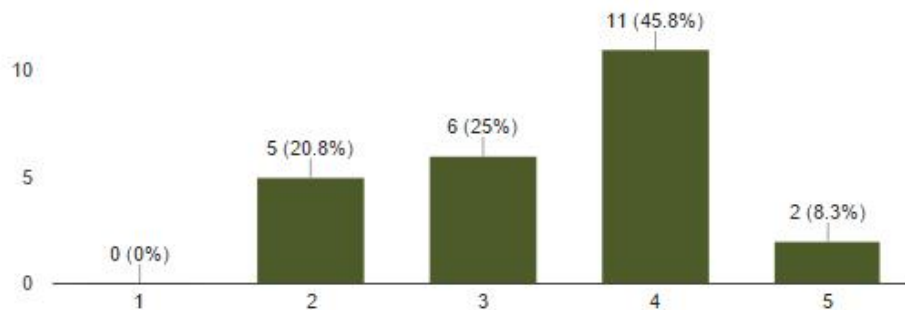
Figure 25 – Meeting end-user expectations.

- *Minimality of operation:* 45.9% of the participants indicated that the system operated without requiring any unnecessary or redundant steps (i.e., rate 4 or more), while 37,5% of the participants expressed a neutral view in this respect (i.e., rate 3) and another 16,6% expressed a less positive view (i.e., rate 2 or less) (Figure 26) (previously reported 70%, 30% and 0% respectively).

The system did not require taking unnecessary steps in order to accomplish what I wanted to do with it.
(24 responses)

Figure 26 – Minimality of operation.

- *Meeting clinical practitioners' expectations:* 58,4% of the participants indicated that the system meets their needs (i.e., rate 4 or more), while 25% of the participants expressed a neutral view in this respect (i.e., rate 3) and another 16,6% expressed a less positive view (i.e., rate 2 or less) (Figure 27) (previously reported 70%, 30% and 0% respectively).
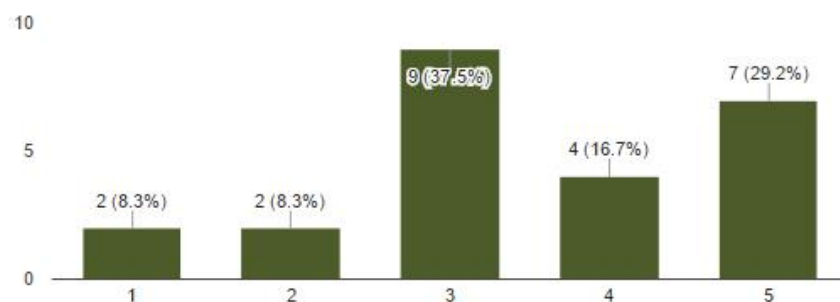
44

The system meets the needs of a clinical practitioner. (24 responses)
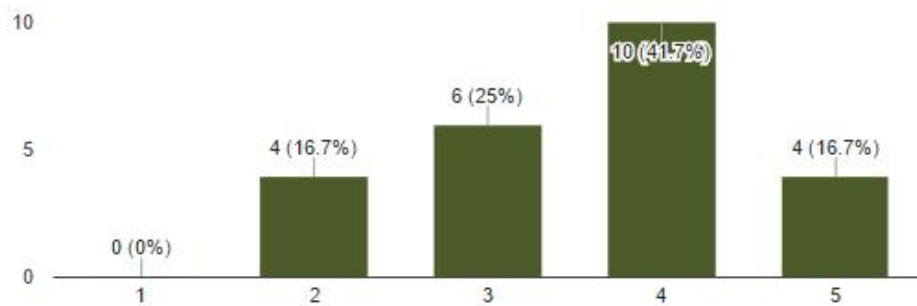
Figure 27 – Meeting clinical practitioners' expectations.

3.5 Open version

General profiling questions

This version was disseminated through specific email lists relevant to the scope of the project, and by an indicative message in the home page of EMBalance. As a result, 20 self-proclaimed clinicians participated. Of those:

- 75% were male and 25% female (Figure 28).

- Half of the participants declared their country of residence. Of them 70% were from Greece and 30% from England (Figure 29).

- The majority of them stated their age. Of them 35,7% were between 25-34, 28,6% between 35-44, 21,4% between 45-54 and 14,3% over 55 (Figure 30).

- 40% were General Practitioners, 35% Otolaryngologists, and 25% other categories (Figure 31).

- 85% examined less than 4 patients and 15% 4 or more with the use of the EMBalance platform (Figure 32).

- 80% received some guidance as to what EMBalance platform offers to the end user, while 20% did not (Figure 33).

45

What is your gender? (20 responses)
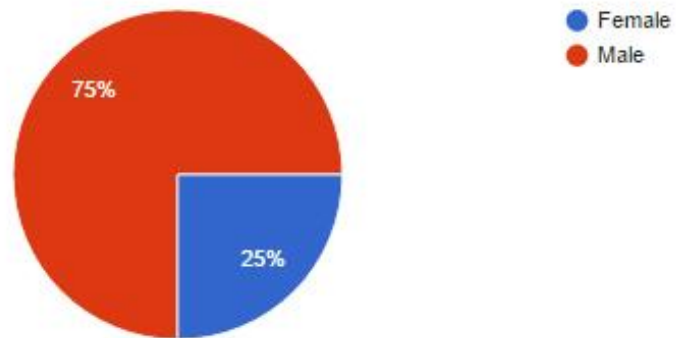
Figure 28 – Participants' gender.

What is your country of residence? (10 responses)
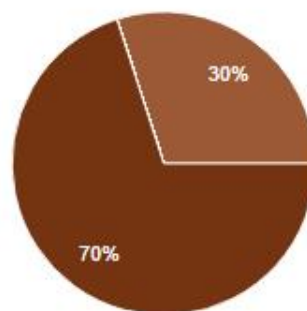
Figure 29 – Participants' country of residence.

Please select your age: (14 responses)
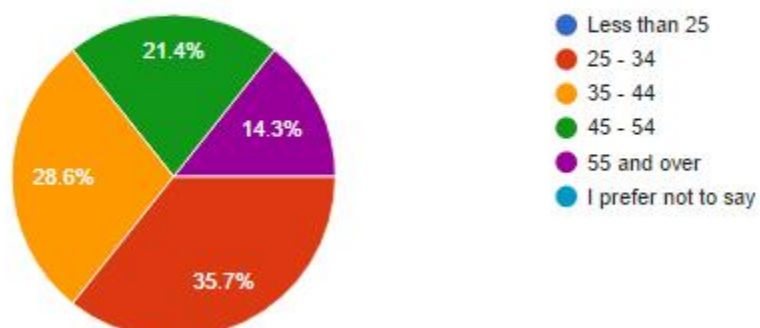
Figure 30 – Participants' age.

What is your professional occupation? (20 responses)
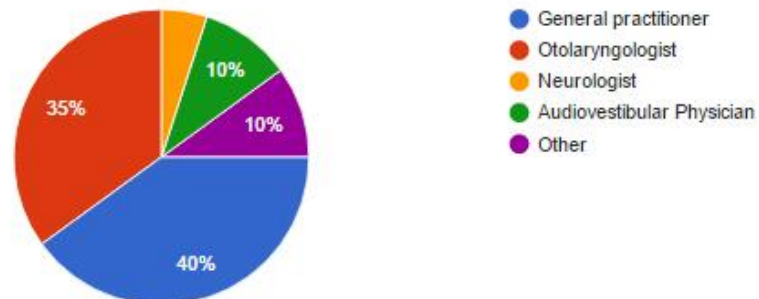
Figure 31 – Participants' specialty.

How many patients have you seen using the EMBALANCE platform before completing this questionnaire?
(20 responses)
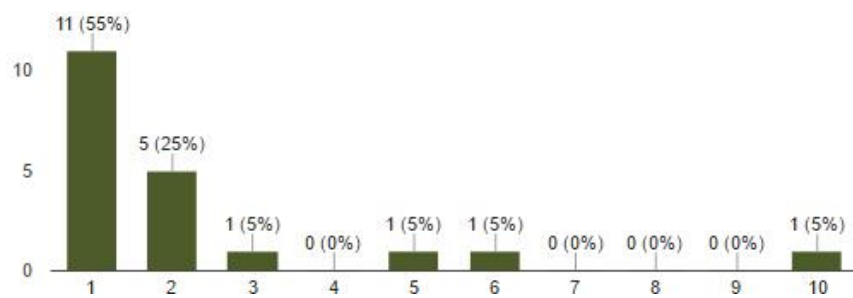
Figure 32 – Number of cases examined with the system.

Have you received training or any assistance in the use of EMBalance platform?
(24 responses)
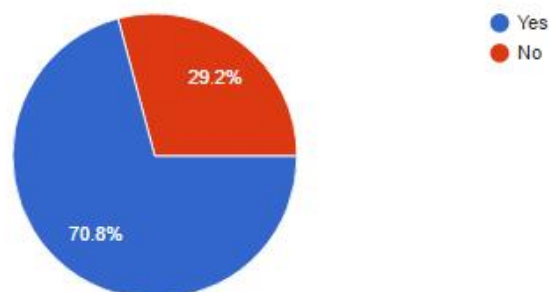
Figure 33 – Provision of guidance and instructions prior of completing the questionnaire.

47

████ Usability

Regarding the usability of the EMBALANCE platform, the following occurred:

Positive aspects:

- *Intuitiveness of user interface:* 60% of the participants considered that the interface is intuitive (i.e., rate 4 or more), while 35% of them expressed a neutral view regarding it (i.e., rate 3) and only 5% found it not intuitive (i.e., rate 2 or less) (Figure 34).
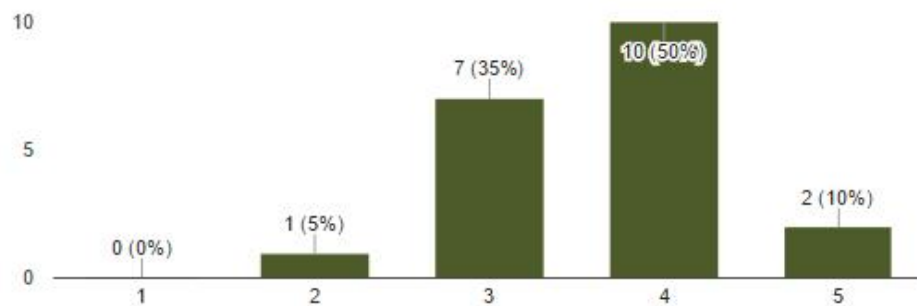
Figure 34 – Intuitiveness of user interface.

- *Ease of use:* 85% considered the system easy to use (i.e., rate 4 or more), while 15% of them expressed a neutral view regarding it (i.e., rate 3) (Figure 35).

Figure 35 – Easiness of task completion.

48

- *Navigation mechanism:* 85% of the participants considered the navigation mechanisms of the platform to be effective (i.e., rate 4 or more), while 15% of them expressed a neutral view regarding it (i.e., rate 3) (Figure 36).

Navigation through the different options of the system was effective.
(20 responses)
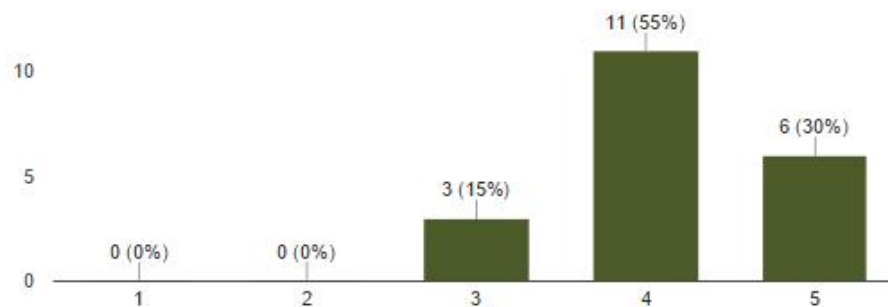
Figure 36 – User-friendly navigation.

- *User confidence:* 70% the participants felt confident in using the system (i.e., rate 4 or more), while 30% of them expressed a neutral view regarding it (i.e., rate 3) (Figure 37).

I felt confident using the system. (20 responses)
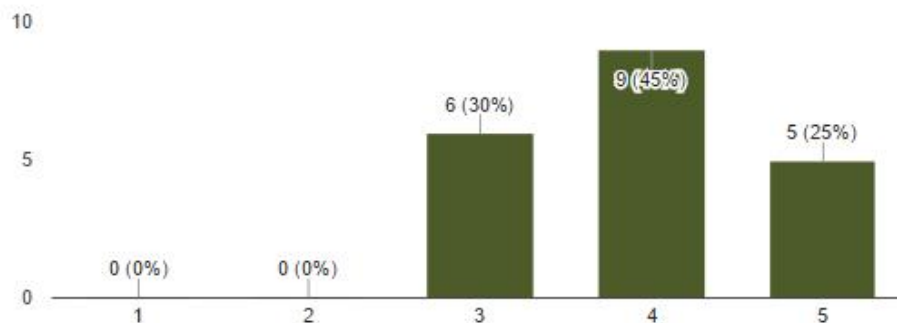
Figure 37 – User confidence.

- *Potential of use in clinical practice*: 45% the participants indicated that they would use the EMBALANCE platform in clinical practice (i.e., rate 4 or more), 35% of them expressed a neutral view regarding it (i.e., rate 3) and 20% of the participants expressed a less positive view (i.e., rate 2 or less) (Figure 38).

49
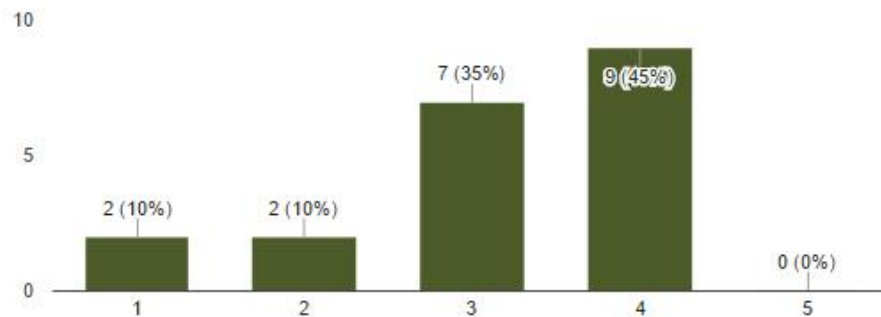
Figure 38 – Utilisation of the system as it is in clinical practice.

- *Potential of use as an educational tool*: 85% the participants indicated that they would use the EMBALANCE platform frequently as an educational tool (i.e., rate 4 or more), while 15% of them expressed a neutral view regarding it (i.e., rate 3) (i.e., rate 2 or less) (Figure 39).
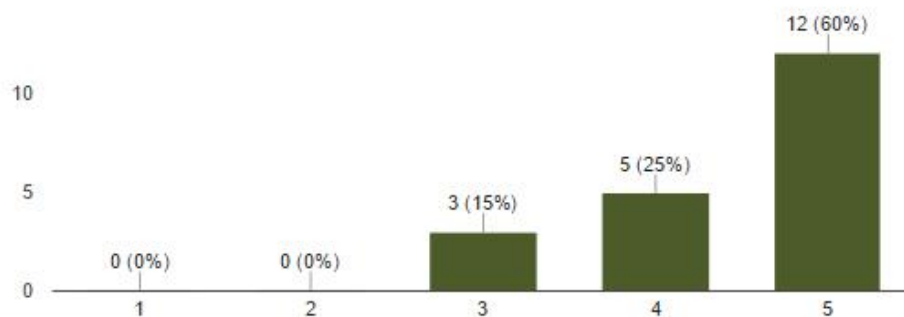
Figure 39 – Utilisation of the system as educational tool.

- *Effective requires technical advice:* the system was considered as self-explanatory as it was anticipated since 55% of the participants indicated that they would not need technical assistance to use it (i.e., rate 2 or less), 15% of them expressed a neutral view regarding it (i.e., rate 3) and 30% of the participants expressed a less positive view (i.e., rate 4 or more) (Figure 18).

50

I would need technical advice in order to use the system effectively.
(20 responses)

8 (40%)   6 (30%)

5

3 (15%)   3 (15%)

0 (0%)

0

1   2   3   4   5
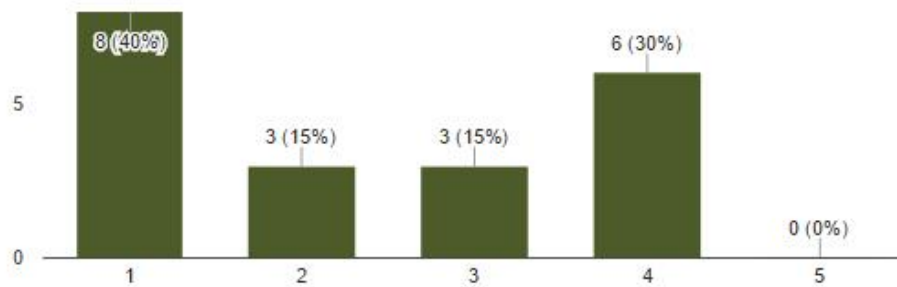
Figure 40 – Self-explanatory usage.

- *Guidance and instructions required for effective usage:* 55% of the participants indicated that they could operate the system without specific training in order to use it effectively (i.e., rate 2 or less), while 30% expressed opposite view (i.e., gave it a score of 4 or more) and 15% of them expressed a neutral view regarding it (i.e., rate 3) (Figure 41).

I would need training in order to use the system effectively. (20 responses)

8

6

7 (35%)

4 (20%)   4 (20%)

4

3 (15%)

2 (10%)

2

0

1   2   3   4   5
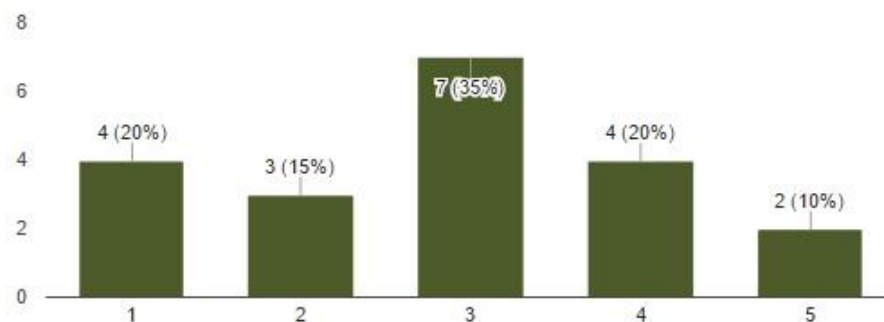
Figure 41 – Guidance and instructions required for effective usage.

- Completion time: Still 54,1% of the participants indicated the time required to complete a visit-task with the use of the system as not satisfactory when comparing to their everyday practice (i.e., rate 2 or less), while 37,6% expressed a positive view (i.e., rate 4 or more) and 8,3% of them expressed a neutral view regarding it (i.e., rate 3) (Figure 42) (previously reported 50%, 30% and 20% respectively).

51

I feel satisfied with the amount of time it took to complete a task in comparison to my previous working method.
(20 responses)
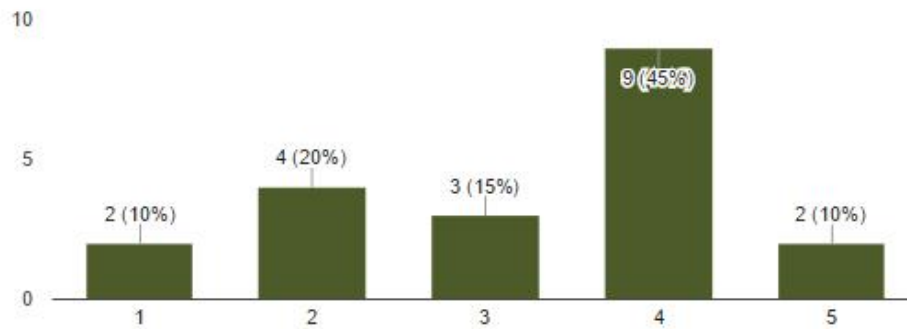
Figure 42 – How quickly a task can be completed in comparison to normal clinical practice.

▮▮▮ Usefulness of EMBALANCE platform functionality

The questionnaire included also questions (i.e., questions 14 to 20) that aimed at evaluating the usefulness of the EMBALANCE functionality. The outcomes of our preliminary evaluation based on these questions are summarized below:

Positive aspects:

- *Enhancing effectiveness:* Half of the participants (50%) indicated that the system improved their effectiveness in the follow up of their patients visits (i.e., rate 4 or 5), 40% of the participants expressed a neutral view (i.e., rate 3) and another 10% expressed a less positive view (i.e., rate 2 or less) (Figure 43).

The use of the system has helped me to be more effective in seeing patients.
(20 responses)
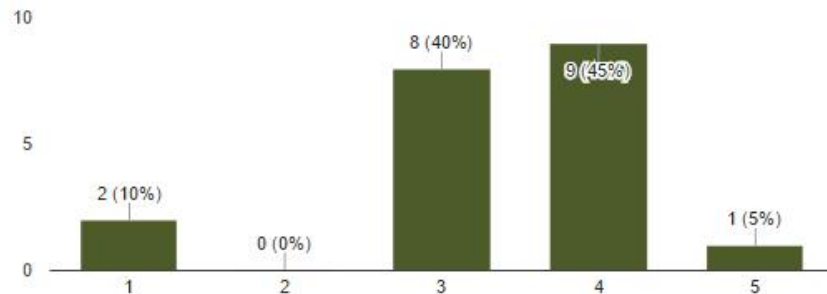
Figure 43 – Effect of EMBalance on practitioner effectiveness.

- *Enhancing productivity: Again* the majority of the participants (50%) indicated that the system improved their productivity (i.e., rate 4 or more), 40% expressed a neutral view (i.e., rate 3) and another 10% expressed a less positive view (i.e., rate 2 or less) (Figure 44).

The use of the system has helped me to be more productive. (20 responses)
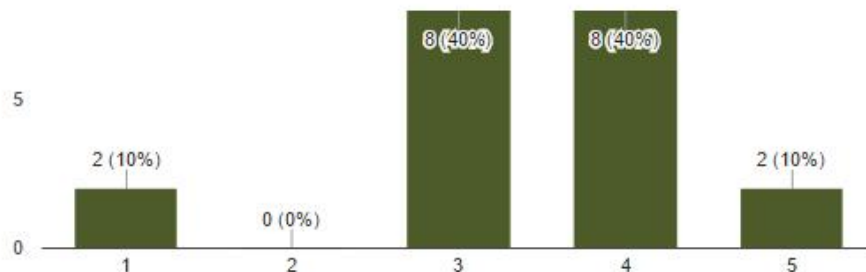
Figure 44 – Effect of EMBalance on practitioner productivity.

- *Successful use:* 50% indicated that they used the system successfully every time was required during trials or elsewhere (i.e., rate 4 or more), 40% of the participants expressed a neutral view (i.e., rate 3) and another 10% expressed a less positive view (i.e., rate 2 or less) (Figure 45).

Furthermore, 80% of the participants indicated that the system was successful in performing the task they had been asked to execute (i.e., rate 4 or more), while 20% gave a neutral response (i.e., rate 3) (Figure 46).

53

I used it successfully every time. (20 responses)
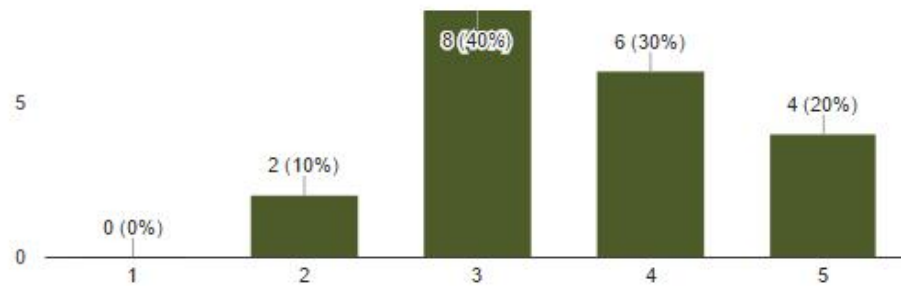
Figure 45 – Successful use of the platform.

It is successful in performing its intended task. (20 responses)
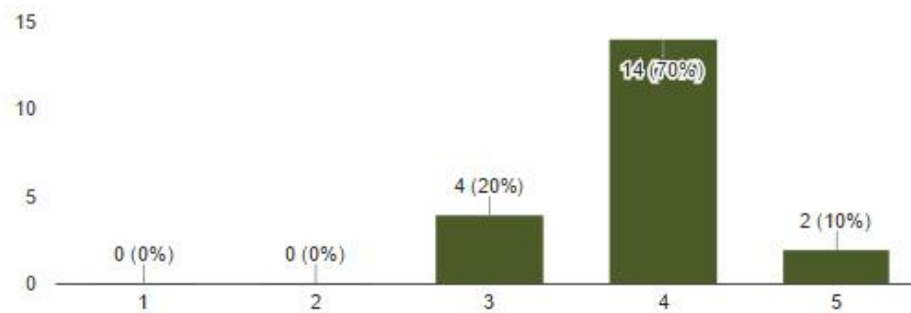
Figure 46 – Success of platform in supporting an intended task.

- *Operating according to expectations:* 70% of the participants indicated that the system operated according to their expectation (i.e., rate 4 or more), 25% of the participants expressed a neutral view in this respect (i.e., rate 3) and another 5% expressed a less positive view (i.e., rate 2 or less) (Figure 47).

It works the way I want it to work. (20 responses)
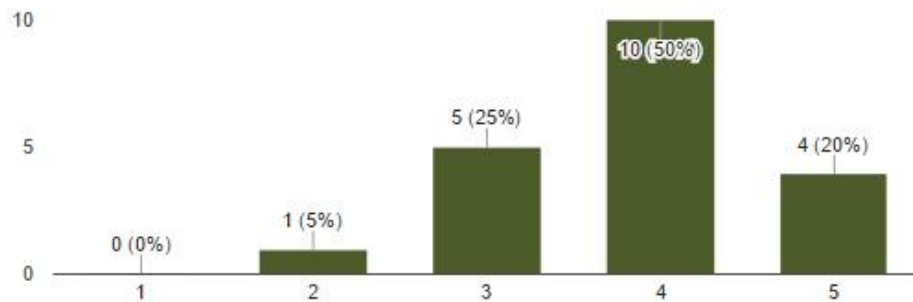
Figure 47 – Meeting end-user expectations.

- *Minimality of operation:* 65% of the participants indicated that the system operated without requiring any unnecessary or redundant steps (i.e., rate 4 or more), while 10% of the participants expressed a neutral view in this respect (i.e., rate 3) and another 28% expressed a less positive view (i.e., rate 2 or less) (Figure 48).

The system did not require taking unnecessary steps in order to accomplish what I wanted to do with it.
(20 responses)
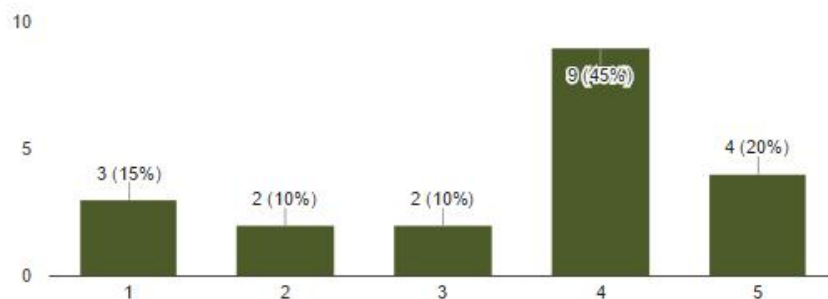
Figure 48 – Minimality of operation.

- *Meeting clinical practitioners' expectations:* 65% of the participants indicated that the system meets their needs (i.e., rate 4 or more), while 35% of the participants expressed a neutral view in this respect (i.e., rate 3) (Figure 49).

The system meets the needs of a clinical practitioner. (20 responses)
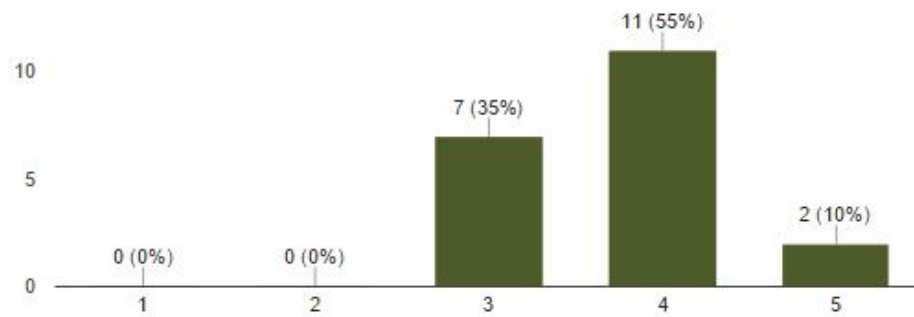
Figure 49 – Meeting clinical practitioners' expectations.

3.6 Comparison of Results Derived from Closed and Open Version of the Questionnaire

The following tables (Table 2, Table 3 and Table 4) present results from both the closed and open versions of the questionnaire (Raw data from both questionaries' are been presented in Appendix B: Questionnaire results). As aforementioned, 24 unique participants filled the "closed" version of the questionnaire, while 20 completed the "open" version (44 responses in total).

Table 2: Comparison of generic info. ▲Indicates increase (in comparison to the previous period reported), ▼decrease and ■stability.

| Generic info					
#	Question	Version			
1	Gender	Closed	▲62,5% male	▼37,5% female	
		Previous	60% male	40% female	
		Open	75% male	25% female	
2	Professional occupation	Closed	▲54% Otolaryngologists	▼25% General Practitioners	▲21% others
		Previous	50% Otolaryngologists	40% General Practitioners	10% others
		Open	35% Otolaryngologists	40% General Practitioners	25% others
3	Usage: # of patients examined	Closed	▼58,4% # < 4	▲41,6% # ≥ 4	
		Previous	60% # < 4	40% # ≥ 4	
		Open	85% # < 4	15% # ≥ 4	
4	Received training	Closed	▼70,8% Yes	▲29,2% No	
		Previous	80% Yes	20% No	
		Open	20% Yes	80% No	

Table 3: Comparison of usability section results. ▲Indicates increase (in comparison to the previous period reported), ▼decrease and ■stability.

Usability					
#	Question	Version	Positive	Neutral	Negative
5	Intuitiveness of user interface	Closed	▲50%	▼29,2%	■20,9%
		Previous	40%	40%	20%
		Open	60%	35%	5%
6	Ease of use	Closed	■50%	▲29,2%	▼21,8%

#	Question	Version	Positive	Neutral	Negative
		Previous	50%	20%	30%
		Open	85%	15%	0%
7	Navigation mechanism	Closed	▲45,8%	▼20,8%	▲33,3%
		Previous	40%	30%	30%
		Open	85%	15%	0%
8	User confidence	Closed	▲62,5%	▼29,2%	▼8,3%
		Previous	50%	40%	10%
		Open	70%	30%	0%
9	Potential of use in clinical practice	Closed	▼45,9%	▲25%	▼29,1%
		Previous	50%	20%	30%
		Open	45%	35%	20%
10	Potential of use as an educational tool	Closed	▲75%	▼16,7%	▼8,3%
		Previous	50%	30%	20%
		Open	85%	15%	0%
11	Requires technical advice	Closed	▲54,1%	■16,7%	▼29,2%
		Previous	50%	16,7%	33,3%
		Open	55%	15%	30%
12	Guidance for effective usage	Closed	■50%	▲8,3%	▼41,7%
		Previous	50%	0%	50%
		Open	35%	35%	30%
13	Completion time	Closed	▲37,6%	▼8,3%	▲54,1%
		Previous	30%	20%	50%
		Open	55%	15%	30%

Table 4: Comparison of usefulness section results. ▲Indicates increase (in comparison to the previous period reported), ▼decrease and ■stability.

Usefulness					
#	Question	Version	Positive	Neutral	Negative
14	Enhancing effectiveness	Closed	▲50%	▼29,2%	■20,9%
		Previous	40%	40%	20%
		Open	50%	40%	10%
15	Enhancing productivity	Closed	■50%	▲29,2%	▼21,8%
		Previous	50%	20%	30%
		Open	50%	40%	10%
16	Successful use (a)	Closed	▲45,8%	▼20,8%	▲33,3%
		Previous	40%	30%	30%
		Open	50%	40%	10%
17	Successful use (b)	Closed	▲62,5%	▼29,2%	▼8,3%
		Previous	50%	40%	10%
		Open	80%	20%	0%
18	Operating according to expectations	Closed	▼45,9%	▲25%	▼29,1%
		Previous	50%	20%	30%
		Open	70%	25%	5%
19		Closed	▲75%	▼16,7%	▼8,3%

	Minimality of operation	Previous	50%	30%	20%
		Open	65%	10%	25%
20	Meeting clinical practitioners' exceptions	Closed	▲54,1%	■16,7%	▼29,2%
		Previous	50%	16,7%	33,3%
		Open	65%	35%	0%

From this comparison the following findings can be derived:

- Participants of the open version had a different starting experience: most of them did not examine as many patients as their colleagues participated in the closed session, and they did not utilise any instruction (to be noted: portal's instructions page was implemented at a later stage).

- The majority of the participants present an almost identical criticism as for the usability. While they were stating the same positive attitude as for most of the aspects of the interaction (Q5, Q6, Q7, Q8, Q9, Q10, Q11, Q12), again both sets indicated that completion time (Q13) was a factor to be further examined (and improved).In fact, open version participants evaluate in a more positive manner these parameters.Concluding, comparison showed that both results sets demonstrated the same positive attitude towards the user friendliness of interaction elements.

- Again, the majority evaluate similarly the usefulness of e-services. Both sets, in all occasions expressed positive view (Q14 to Q20).

- Highest mark was given to Q9 Potential of use as an educational tool (75% and 85% respectively), followed Q8 User confidence (62,5% and 70%), and Q6 Ease of use (50% and 85%). As for usefulness, Q19 Minimality of operation considered as the strongest point (75% and 65% respectively).

- The worst marks were obtained for Q13 Completion time, which constitutes the only case in the closed version where more negative than positive marks were stated by the participants (54,1% negatives vs 37,6% positives), which is in line with suggestions derived from the focus groups. The opposite trend was presented in the open version (30% negatives vs 55% positives).

3.7 Assessing the Validity of Questionnaire Results

Since EMBalance questionnaire was constructed as five-point Likert rating scales (1: strongly disagree to 5: strongly agree), one of the best choices to analyse the quality of those Likert data is to use a Nonparametric tests,

such as the Mann-Whitney test[1] [Winter & Dodou, 2012]. Among several choices on how to generate distinct populations, two most "promising" ones that may generate differences was

- Test 1: whether the occupation (i.e., clinician type) may result significant differences

- Test 2: whether the extended use of the system (i.e., usage of EMBalance for more than two patients) affects the outcome

Both of these tests were conducted to the closed version since these results are considered having the highest validity.

To this purpose we defined:

Test 1

As Set A (sample size = 8): Participants having occupation: General practitioner or Neurologist

As Set B (sample size = 16): all other occupations

Test: whether two sample means are equal or not

Significance Level: 0,05

Two-tailed, Version: Closed (final)

Table 5: Mann-Whitney test: results

Mann-Whitney test 1		
#	**Question**	**Result**
5	Intuitiveness of user interface	The U-value is 49. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
6	Ease of use	The U-value is 62. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.

---

[1]For calculations the online Mann-Whitney U Test Calculator was used: http://www.socscistatistics.com/tests/mannwhitney/default2.aspx

7	Navigation mechanism	The U-value is 63.5. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
8	User confidence	The U-value is 53.5. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
9	Potential of use in clinical practice	The U-value is 60. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
10	Potential of use as an educational tool	The U-value is 48. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
11	Requires technical advice	The U-value is 55. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
12	Guidance for effective usage	The U-value is 49.5. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
13	Completion time	The U-value is 49.5. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
14	Enhancing effectiveness	The U-value is 45.5. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
15	Enhancing productivity	The U-value is 43. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
16	Successful use (a)	The U-value is 61. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
17	Successful use (b)	The U-value is 59. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
18	Operating according to expectations	The U-value is 49.5. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
19	Minimality of operation	The U-value is 55. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.
20	Meeting clinical practitioners' exceptions	The U-value is 40. The critical value of U at $p < .05$ is 31. Therefore, the result is not significant at $p < .05$.

Test 2

As Set A (sample size = 14): Extended use of EMBalance system: participants that examined # of patients > 2, and

As Set B (sample size = 10): participants that examined # of patients ≤ 2

Test: whether two sample means are equal or not

Significance Level: 0,05

Two-tailed, Version: Closed (final)

Table 6: Mann-Whitney test: results

Mann-Whitney test		
#	Question	Result
5	Intuitiveness of user interface	The U-value is 56. The critical value of U at p < .05 is 36. Therefore, the result is not significant at p < .05.
6	Ease of use	The U-value is 62. The critical value of U at p < .05 is 36. Therefore, the result is not significant at p < .05.
7	Navigation mechanism	The U-value is 47. The critical value of U at p < .05 is 36. Therefore, the result is not significant at p < .05.
8	User confidence	The U-value is 62.5. The critical value of U at p < .05 is 36. Therefore, the result is not significant at p < .05.
9	Potential of use in clinical practice	The U-value is 46. The critical value of U at p < .05 is 36. Therefore, the result is not significant at p < .05.
10	Potential of use as an educational tool	The U-value is 46.5. The critical value of U at p < .05 is 36. Therefore, the result is not significant at p < .05.
11	Requires technical advice	The U-value is 68.5. The critical value of U at p < .05 is 36. Therefore, the result is not significant at p < .05.
12	Guidance for effective usage	The U-value is 54.5. The critical value of U at p < .05 is 36. Therefore, the result is not significant at p < .05.
13	Completion time	The U-value is 51.5. The critical value of U at p < .05 is 36. Therefore, the result is not significant at p < .05.
14	Enhancing effectiveness	The U-value is 63.5. The critical value of U at p < .05 is 36. Therefore, the result is not significant at p < .05.
15	Enhancing productivity	The U-value is 64. The critical value of U at p < .05 is 36. Therefore, the result is not significant at p < .05.

16	Successful use (a)	The U-value is 69.5. The critical value of U at $p < .05$ is 36. Therefore, the result is not significant at $p < .05$.
17	Successful use (b)	The U-value is 60.5. The critical value of U at $p < .05$ is 36. Therefore, the result is not significant at $p < .05$.
18	Operating according to expectations	The U-value is 66.5. The critical value of U at $p < .05$ is 36. Therefore, the result is not significant at $p < .05$.
19	Minimality of operation	The U-value is 50.5. The critical value of U at $p < .05$ is 36. Therefore, the result is not significant at $p < .05$.
20	Meeting clinical practitioners' exceptions	The U-value is 63. The critical value of U at $p < .05$ is 36. Therefore, the result is not significant at $p < .05$.

The outcome of both of these tests provide evidence that the distributions of both populations are equal, thus selective distributions do not differentiate results.

Conclusions

This document reports on the results of usability evaluation activities in the context of WP8, and specifically of the usability assessment via an online satisfaction questionnaire. These results provided a better understanding of user behaviors of specific user group (clinicians) and input to support subsequent development, leading to the final version 2.0of the EMBalance platform.

Participants in general rated the e-services as very useful. The results of the questionnaire indicate clearly that user friendliness was of very good level, but at the same time appreciated from the system to provide suggestions much sooner in order to be utlised in medical practice. Almost all acknowledged its potential as an education tool. As with previous evaluation stage, the collected feedback:

- reaffirmed positive evidence regarding the usability of the platform,

- corroborated the potential effect on clinical practice mainly as a tool offering education in clinical practice

Enhancements suggested were implemented in the final version 2.0 of the EMBalance platform (and reported in D6.5). It's worth mentioning that the vast majority of end-users acknowledge the potential value of the platform as an educational tool, which asserts that implemented e-services considered user-friendly even to non-experienced clinicians.

Bibliography

[1] Spool, J. 2007. Five Usability Challenges of Web-Based Applications, online: https://articles.uie.com/usability_challenges_of_web_apps/

[2] Nielsen, J. 1999. Designing Web Usability: The Practice of Simplicity, New Riders, ISBN 156205810X.

[3] Becker, S.A. and Mottay, F.E. 2001. "A global perspective on Web site usability". IEEE Software, vol. 18, no. 1, pp. 54-61

[4] Lauterbach, C. 2010. Exploring the Usability of E-Health Websites, online: http://usabilitynews.org/exploring-the-usability-of-e-health-websites/

[5] Hesse, B.W., and Shneiderman, B. 2007. eHealth research for the user's perspective. American Journal of Preventative Medicine, 32(5); 97-103.

[6] Quesenbery, Whitney 2009. What do we mean by user friendly? Meeting expectations for e-health applications, a talk at Informatics for Consumer Health: Summit on Communication, Collaboration, and Quality, November 6, 2009, online: http://www.wqusability.com/articles/consumer-health-usability.html

[7] T. D. Erikson, & H. A. Simon, 1985. Protocol analysis: verbal reports as data. Cambridge, MIT Press.

[8] Marcilly, R., Beuscart-Zephir, MC., &Beuscart, R. Integrating Human Factors in an International Research Project: Lessons Learned from the PSIP Project, Studies in health technology and informatics 183 (2013): 162-7

[9] Purves, I. N., Sugden, B., Booth, N., &Sowerby, M., The PRODIGY project - the iterative development of the release one model. In Proceedings of the AMIA Symposium. American Medical Informatics Association, 359, 1999

[10]   Walton, R. T., Gierl, C., Yudkin, P., Mistry, H., Vessey, M. P., & Fox, J., Evaluation of computer support for prescribing

(CAPSULE) using simulated cases. Bmj, 315(7111), 791-795, 1997

[11]   Self, J. 2016. Why is Sketching (Still) Important (To Design)? To sketch is to design, online: http://www.core77.com/posts/52948/Why-is-Sketching-Still-Important-To-Design

[12]   Usability.gov, 2017. Usability Evaluation Basics, online: https://www.usability.gov/what-and-why/usability-evaluation.html

[13]   Lund, A.M. 2001. Measuring Usability with the USE Questionnaire. STC Usability SIG Newsletter, 8:2.

[14]   Joost C. F. de Winter and Dimitra Dodou, 2012. Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon, Department of BioMechanical Engineering, Delft University of Technology, Volume 15, Number 11, October, 2010, online: http://pareonline.net/getvn.asp?v=15&n=11

[15]   Nielsen, J. 1997. The Use and Misuse of Focus Groups, Topics: Research Methods, online: https://www.nngroup.com/articles/focus-groups/

[16]   Mary Marczak & Meg Sewell, 2017. Using focus groups for evaluation, online: https://cals.arizona.edu/sfcs/cyfernet/cyfar/focus.htm

[17]   Judith Sharken Simon, 1999. How to Conduct a Focus Group, online: https://www.tgci.com/podcasts/how-conduct-focus-group-judith-sharken-simon

Appendix A. EMBalance questionnaire

General section (* Required)

1. What is your gender? *{Female, Male}

2. What is your professional occupation? *{General practitioner, Otolaryngologist, Neurologist, Audiovestibular Physician, Other}

3. How many patients have you seen using the EMBALANCE platform before completing this questionnaire? * {1 to 10}

4. Have you received training or any assistance in the use of EMBalance platform? * {Yes, No}

Platform's Usability (1: Strongly disagree , 5 Strongly agree)

5. The interface of the system was intuitive. *

6. It was easy to complete tasks using the system. *

7. Navigation through the different options of the system was effective. *

8. I felt confident using the system. *

9. I could use this system frequently in normal clinical practice. *

10. I could use this system frequently as an educational tool. *

11. I would need technical advice in order to use the system effectively. *

12. I would need training in order to use the system effectively. *

13. I feel satisfied with the amount of time it took to complete a task in comparison to myprevious working method. *

Usefulness of the functionality been introduced (1: Strongly disagree , 5 Strongly agree)

14. The use of the system has helped me to be more effective in seeing patients.*

15. The use of the system has helped me to be more productive. *

16. I used it successfully every time. *

17. It is successful in performing its intended task. *

18. It works the way I want it to work. *

19. The system did not require taking unnecessary steps in order to accomplish what I wantedto do with it. *

20. The system meets the needs of a GP.

21. If you noticed unnecessary steps in the workflow, please list them below:

22. If you noticed inconsistencies in the user interface, please list them below:

23. Any other comments?

Thank you for your participation!

Figure 50 – The open version of the questionnaire (via Google forms)

Appendix B. Questionnaire results

Closed version results

Usability

Timestamp	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13
5/30/2016	M	Otolaryngologist	10	Yes	4	4	5	5	3	5	1	1	4
6/1/2016	M	General practitioner	1	Yes	3	4	4	3	4	4	2	4	2
6/2/2016	M	Otolaryngologist	3	Yes	3	2	2	3	3	3	3	2	3
6/2/2016	M	Otolaryngologist	5	Yes	2	3	1	3	1	1	3	5	2
6/6/2016	F	Otolaryngologist	2	Yes	3	2	3	2	2	2	4	4	1
6/6/2016	F	General practitioner	5	No	3	4	3	4	4	3	4	4	2
6/7/2016	M	Otolaryngologist	1	No	4	5	4	4	4	4	1	2	4
6/6/2016	F	General practitioner	6	Yes	4	3	2	3	4	4	4	4	3
6/7/2016	F	Otolaryngologist	2	Yes	2	1	3	4	2	3	1	1	2
6/7/2016	M	General practitioner	2	Yes	4	4	4	4	4	5	1	1	5
6/28/2016	M	Neurologist	4	Yes	5	4	4	5	3	4	1	1	2
6/28/2016	M	Audiovestibular Physician	3	Yes	3	2	3	3	2	4	1	4	2
6/28/2016	F	General practitioner	4	Yes	4	3	4	5	2	4	1	2	2
9/27/2016	M	General practitioner	6	Yes	2	3	2	3	2	4	3	3	2
9/26/2016	F	Otolaryngologist	4	Yes	4	4	2	4	5	5	4	4	4
9/28/2016	F	Otolaryngologist	3	Yes	3	3	2	3	3	4	5	5	2
10/3/2016	F	Otolaryngologist	3	No	4	4	4	4	4	5	1	2	4
10/9/2016	M	Otolaryngologist	4	Yes	4	5	4	4	2	4	2	5	2
10/9/2016	F	Otolaryngologist	1	No	5	5	4	4	4	5	3	3	4
10/27/2016	M	Neurologist	5	Yes	4	3	2	4	3	3	1	1	2
10/28/2016	M	Other	1	No	1	3	3	4	4	5	5	4	5
10/28/2016	M	Other	1	Yes	2	1	1	2	4	5	5	2	1
1/25/2017	M	Otolaryngologist	1	No	3	5	4	4	3	5	1	2	5
2/13/2017	M	Otolaryngologist	2	No	4	4	5	4	4	5	1	2	4

Usefulness

Timestamp	Q1	Q2	Q14	Q15	Q16	Q17	Q18	Q19	Q20
5/30/2016	M	Otolaryngologist	4	4	4	5	3	5	4
6/1/2016	M	General practitioner	4	4	3	4	3	4	4
6/2/2016	M	Otolaryngologist	3	2	2	3	2	3	3
6/2/2016	M	Otolaryngologist	1	1	5	3	3	3	4
6/6/2016	F	Otolaryngologist	2	2	2	4	3	2	3
6/6/2016	F	General practitioner	5	4	4	3	4	3	4
6/7/2016	M	Otolaryngologist	4	4	5	3	4	5	5
6/6/2016	F	General practitioner	4	4	3	3	2	4	5
6/7/2016	F	Otolaryngologist	3	2	3	3	3	4	3
6/7/2016	M	General practitioner	5	4	5	5	4	5	5
6/28/2016	M	Neurologist	3	3	3	4	4	3	4
6/28/2016	M	Audiovestibular Physician	4	2	3	3	4	1	2
6/28/2016	F	General practitioner	3	3	4	4	4	3	5
9/27/2016	M	General practitioner	4	4	4	3	2	3	2
9/26/2016	F	Otolaryngologist	5	4	3	3	5	3	3
9/28/2016	F	Otolaryngologist	3	3	3	3	4	3	4
10/3/2016	F	Otolaryngologist	3	3	4	4	4	5	4
10/9/2016	M	Otolaryngologist	3	4	4	4	4	2	2
10/9/2016	F	Otolaryngologist	5	4	3	4	4	5	4
10/27/2016	M	Neurologist	3	3	2	2	2	5	3
10/28/2016	M	Other	3	4	3	2	3	3	3
10/28/2016	M	Other	1	1	3	2	2	1	2
1/25/2017	M	Otolaryngologist	4	3	4	4	4	4	4
2/13/2017	M	Otolaryngologist	4	4	4	4	5	5	4

Open version results

Usability

Timestamp	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13
15/9/2016	M	General practitioner	1	No	3	4	4	4	3	4	1	1	4
15/9/2016	F	Other	1	No	2	3	3	3	1	3	2	3	1

Timestamp														
15/9/2016	M	Otolaryngologist	1	No	5	5	5	5	1	5	1	2	3
19/9/2016	F	Other	1	No	3	4	4	3	3	5	3	3	4
19/9/2016	M	General practitioner	5	Yes	3	3	4	3	3	4	3	4	3
4/10/2016	M	Neurologist	2	No	4	4	4	3	3	5	4	4	4
9/10/2016	M	Audiovestibular Physician	1	No	4	5	4	5	4	4	1	1	4
22/10/2016	M	Otolaryngologist	10	No	4	4	4	4	4	3	2	2	2
15/11/2016	M	General practitioner	2	Yes	4	4	3	5	4	5	3	4	3
15/11/2016	F	General practitioner	2	Yes	3	3	4	4	2	5	1	3	2
28/12/2016	M	General practitioner	2	No	4	5	4	4	4	5	4	3	4
28/12/2016	M	Audiovestibular Physician	3	No	3	4	5	4	4	5	4	5	4
28/12/2016	F	General practitioner	1	No	4	5	4	3	3	5	4	4	4
11/2/2017	M	Otolaryngologist	1	No	4	5	5	4	3	4	1	3	5
11/2/2017	F	General practitioner	2	No	4	5	5	5	4	5	1	1	4
11/2/2017	M	General practitioner	1	No	3	5	5	4	4	5	2	2	5
13/2/2017	M	Otolaryngologist	1	No	5	5	5	4	4	5	1	3	4
13/2/2017	M	Otolaryngologist	1	No	4	4	3	3	3	4	4	5	2
14/2/2017	M	Otolaryngologist	1	No	4	4	4	5	2	5	1	1	1
20/2/2017	M	Otolaryngologist	6	Yes	3	4	4	4	4	3	4	3	2

Usefulness

Timestamp	Q1	Q2	Q14	Q15	Q16	Q17	Q18	Q19	Q20
15/9/2016	M	General practitioner	3	4	3	4	3	1	4
15/9/2016	F	Other	1	1	3	3	2	3	3
15/9/2016	M	Otolaryngologist	3	5	3	5	5	5	5
19/9/2016	F	Other	3	3	3	3	4	4	3
19/9/2016	M	General practitioner	1	1	4	3	3	2	3
4/10/2016	M	Neurologist	4	3	4	4	5	1	4
9/10/2016	M	Audiovestibular Physician	4	4	5	5	5	4	5
22/10/2016	M	Otolaryngologist	4	4	4	4	4	4	4
15/11/2016	M	General practitioner	3	4	2	4	4	5	4
15/11/2016	F	General practitioner	5	5	4	4	4	2	3
28/12/2016	M	General practitioner	4	4	3	4	4	4	4

28/12/2016	M	Audiovestibular Physician	3	3	4	4	4	4	4
28/12/2016	F	General practitioner	3	3	3	4	4	4	3
11/2/2017	M	Otolaryngologist	4	3	3	4	4	1	4
11/2/2017	F	General practitioner	4	4	5	4	4	4	4
11/2/2017	M	General practitioner	4	4	4	4	5	5	4
13/2/2017	M	Otolaryngologist	4	3	5	4	4	5	4
13/2/2017	M	Otolaryngologist	3	3	2	4	3	4	3
14/2/2017	M	Otolaryngologist	4	3	5	3	3	4	3
20/2/2017	M	Otolaryngologist	3	4	3	4	3	3	4