Full clinical study validation results

24/03/2023

WP6 - Multicentre clinical study

Project title	Patients-centred SurvivorShIp care plan after Cancer treatments based on Big Data and Artificial Intelligence technologies
Grant Agreement number	875406
Call and topic identifier	SC1-DTH-01-2019 - Big data and Artificial Intelligence for monitoring health status and quality of life after the cancer treatment
Funding schema	RIA
Coordinator	FUNDACION CENTRO TECNOLOXICO DE TELECOMUNICACIONS DE GALICIA (GRADIANT)
Website	projectpersist.com
Document keywords	Cancer survivors, mHealth, mClinician
Document Abstract	Clinical analysis the impact of PERSIST and the expected results of the clinical studies. Detailed evaluation of the individual co-creation workshops with clinicians. Feedback about possible application to other sectors.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

DOCUMENT	
Authors	Dina Bēma (UL), Inese Poļaka (UL), Krista Arcimoviča (UL), Anna Marija Leščinska (UL), Marcela Chavez (CHU), Sandra Martič (UKCM), Maja Molan (UKCM), Saša Nikolič (UKCM), Liliana Pires (RUBY); Patrick Duflot (CHU); Valérie Bleret (CHU); Ariadna Mato Montero (SERGAS), Beatriz Calderón Cruz (SERGAS), Gaetano Manzo (HESSO), José Aguayo Arjona (SERGAS), Tunç Cerit (EMODA), Laurence Seidel (CHU), Nathalie Maes (CHU), Liliana Pires (RUBY), Ignacio García Vega (GRADIANT), Alberto Sánchez (GRAD)
Internal reviewers	SERGAS, CHU
Work package	WP6
Task	T6.4
Nature	Report
Dissemination Level	Public

VERSION	DATE	CONTRIBUTOR	DESCRIPTION
0.1.	12.12.2022.	UL	Development of work-plan
0.2.	21.12.2022.	UL	Initial draft. Work assignments.
0.3.	31.01.2023.	HESSO	Additions from HESSO about trajectories.
0.4.	01.02.2023.	GRADIANT	Additions from GRADIANT about recurrence predictions.
0.5.	09.02.2023.	EMODA	Additions from EMODA about CVS, breast cancer and colon cancer recurrence.
0.6.	10.02.2023.	SERGAS	Additions from SERGAs
0.7.	17.02.2023.	UKCM	Additions from UKCM
0.8.	20.02.2023.	UL	Compilations of all additions.
0.9.	03.03.2023.	UL	Last review, additions.
1.0.	08.03.2023.	CHU	Review.
1.1.	09.03.2023.	SERGAS	Review
1.2.	10.03.2023.	UL	Corrections according to CHU and SERGAS review.
1.3.	21.03.2023.	UL	Corrections according to GRADIANT.
1.4.	22.03.2023.	GRADIANT	Final Review

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

DISCLAIMER

ACKNOWLEDGEMENT

This document is a deliverable of the PERSIST project. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 875406

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Table of contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Acronyms and abbreviations

ACRONYM	TITLE
AI	Artificial intelligence
API	Application Programming Interface
app	Application
AVG	Average
BMI	Body mass index
BRC	Breast cancer
C 50	Breast cancer diagnosis
C18 / C19	Colorectal cancer diagnosis
CASE-cancer	Communication and Attitudinal Self-Efficacy scale for cancer
CHU / CHUL	Centre Hospitalier Universitaire De Liège
COP	Cyclic Olefin Polymer
CRC	Colorectal cancer
CTC	Circulating tumour cells
CV	Coefficient of variation
DBP	Diastolic blood pressure
CDSS	Clinical decision support system
DSCF	Dwass-Steel-Critchlow-Fligner
ECOG	Eastern Cooperative Oncology Group
EHR	Electronic health records
F	Female
FHIR	Fast Healthcare Interoperability Resources
FIT	Faecal immunochemical test
FOBT	Faecal occult blood test
GAD7	Generalised Anxiety Disorder Assessment
GB	Gradient Boosting
GGT	Gamma Glytamyltransferase
HDL	High-density lipoprotein
HER2	Human epidermal growth factor receptor 2
HESSO	Haute Ecole Spécialisée de Suisse occidentale
HR	Heart rate
IQR	Interquartile range
LDL	Low density
m	Male
mBRC	Metastatic Breast Cancer

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

N	Number
NCCN	National Comprehensive Cancer Network
PAM	Patient activation measure
PHQ2	Patient Health Questionnaire-2
SBP	Systolic blood pressure
SD	Standard deviation
SERGAS	Complejo Hospitalario Universitario de Ourense (SERGAS)
Std	Standard deviation
STEP	Number of steps per day
SUS	System usability scale
TNM	Classification system of malignant tumours
TRL	Technology readiness level
UKCM	University Medical Centre Maribor
UL	University of Latvia
XGB	eXtreme Gradient Boost
y/o	Years old

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Executive Summary

This deliverable presents the results of the clinical study conducted and to analyse the impact of PERSIST in clinical settings. The study involved training patients and clinical personnel, collecting patient medical history data, and asking patients to complete questionnaires. In addition, individual consultations and co-creation workshops were conducted with clinicians across all four hospitals to obtain feedback on the PERSIST system, mClinician app, and web solutions to assess their potential. The report provides an overview of the results obtained from patients during the entire study, as well as the feedback received from clinicians through individual consultations, workshops, and questionnaires. The report evaluates the current readiness of the application and makes future recommendations for potential improvements and its use in other sectors.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Results of the PERSIST clinical trial

1. General information

Altogether 166 patients participated in the PERSIST clinical trial from four hospitals (see Table No.1). Among the recruited patients, 85 have had breast cancer and 81 colorectal cancer. The average age of the patients at the time of inclusion was 55 years old.

In total, there were 37 male and 129 female patients included in the study. The rarity of breast cancer in men and a slightly higher inclusion of women in the group of patients who presented colorectal cancer explain the greater inclusion of women and therefore the gender imbalance in inclusion. Altogether men showed lower interest in participating in the study than women (reported by clinicians).

CLINICAL PARTNER	RECRUITED PATIENTS	MEAN AGE	BREAST CANCER	COLORECTAL CANCER	MALE	FEMALE
UL	46	54	24	22	7	39
UKCM	40	57	20	20	11	29
CHU	41	55	21	20	7	34
SERGAS	39	56	20	19	12	27
TOTAL	**166**	**55**	**85**	**81**	**37**	**129**

Table 1 General description of patients

Altogether, (up to 31.10.2022.) 41 patients left the clinical study (see Table No. 2.). Mean age of patients that left were 54 that does not differ from other patients' mean age who stayed in the study. Therefore, we conclude that age is not an important factor for compliance in digital surveillance. Majority (31) were female and leaving in the breast cancer group was greater (24 vs 16). In total 27% males and 24% females left the study representing similar levels of leaving regarding the gender, meaning that gender was not important factor for compliance in digital surveillance

CLINICAL PARTNER	WITHDRAW	MEAN AGE	BREAST CANCER	COLORECTAL CANCER	MALE	FEMALE
UL	16	53	10	6	4	12
UKCM	4	57	2	2	0	4
CHU	16	52	8	7	5	11
SERGAS	5	55	4	1	1	4
TOTAL	**41**	**54**	**24**	**16**	**10**	**31**

Table 2 General description of patients who have left the study

A comprehensive analysis of the reasons why patients left the study highlighted that the most frequently cited factors were related to personal circumstances and technical issues, including smart-bracelet malfunctions and other technical problems (see Table 3). It is also important to note that "participation takes too much time" was a common reason given by

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

some patients for leaving the study. However, despite these challenges, most of the patients remained engaged in the study and provided valuable insights that will inform future research in this field. Patient cooperation in the study was stimulated by additional counselling by phone on the use of technology and other mitigation strategies, such as specialist lectures in workshops.

Reasons for leaving	Times mentioned
Personal life situation	11
Device malfunction, technical problems	10
Participation takes too much time	9
Does not like the system in general	7
Complaints about app	6
Induces stress, anxiety	6
Not specified	4
Reminds of cancer	3
No need for follow-up	2
Light at night from bracelet	2
Tired of participating	2
Patient died	2
Recurrence	1

Table 3. Summary of reasons patients mention upon leaving study

The clinical trial was extended until the end of December 2022 to allow for additional data collection and updates to the mClinician app. The hospitals and ethics committees were notified of this extension, and patients were contacted accordingly. Even after the initial end date of October 31st, 2022, the majority of patients remained committed to the study, with all patients from SERGAS and UKCM continuing to participate due to necessary blood withdrawal for CTC (circulating tumour cell) tests. Additionally, a significant number of patients from CHU and UL also chose to continue their participation, with 5 patients from CHU and 12 patients from UL remaining involved until the study's finalisation. Overall, this extension allowed for a more comprehensive and robust analysis of the study's results, and we are grateful for the ongoing commitment of our participants. In addition, the patients from UKCM and UL who attended the last PERSIST workshop in February 2023 expressed their enthusiasm and willingness to continue participating in future projects based on results gained in the project.

➔ Conclusions:

Majority (75,3%) of cancer survivors included in the study are motivated to participate in health surveillance activities by using digital wearable devices daily for a prolonged period. Adjustments to the possible timescale for the study length should be made, and the daily

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

workload of tasks related to technologies should be evaluated. Possible participation termination of 25% should be included in the initial project planning phase. Changes in personal life situations seem to be one of the key factors initiating participants to leave the study. This could be avoided if the patients were questioned more detailed about their nearest future plans before recruitment and warned of possible time-consuming participation. Appropriate device selection and testing, simplification of the technologies to make them easy-to-use before distributing them within a larger group of patients, could be among the solutions for the difficulty in using the devices and complaints about the technology. Also, evaluating digital literacy using standardised questionnaires at the recruitment could help.

2. Time of use and frequency of use of the mHealthApp

Note: To assess the usability of the mHealthApp, specifically user comfort and ease of management as outlined in the clinical protocol, we collected information from the mClinician app in the form of an Excel file. This file shows the frequency of use of the mHealthApp, including synchronisation time, measurements, diaries, and emotions. A 100% score in synchronisation, diaries and emotions indicates that the patient synced and uploaded at least one measurement each day while they were included in the study. Smart-band usage indicates the usage of the smart band and data gathered every hour (if a patient wears the bracelet all day long, hourly data points should be synchronised on the server).

- ✓ Patients from UL

The usage statistics for the mHealthApp were collected from 31 patients who were active until 31.10.2022. The data retrieved from the mClinician app (see Table No.4) showed that synchronisation was successful nearly 90% of the time. However, 30,61% of smart-band usage (24 hours) and 10,82% of emotion data were obtained, and only 1,61% of time diaries were recorded.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	Sync (%)	Smart band usage(%)	Diaries (%)	Emotions (%)
1	100,00	30,00	0,00	9,50
2	100,00	98,00	0,00	16,00
3	100,00	10,00	4,00	5,00
4	100,00	100,00	0,00	8,60
5	100,00	24,00	2,50	12,00
6	100,00	2,00	2,00	2,60
7	100,00	34,00	2,00	19,20
8	100,00	7,00	0,00	4,20
9	100,00	48,00	6,50	17,70
10	100,00	100,00	0,00	10,70
11	50,00	17,00	1,00	15,40
12	100,00	19,00	3,00	15,80
13	100,00	81,00	5,00	10,00
14	100,00	2,00	0,00	6,60
15	100,00	2,00	0,00	5,40
16	100,00	34,00	2,00	9,80
17	100,00	41,00	2,00	19,00
18	100,00	5,00	0,00	3,50
19	50,00	15,00	0,00	3,70
20	100,00	35,00	1,00	12,80
21	33,30	1,00	0,00	14,00
22	100,00	19,00	2,00	12,90
23	100,00	3,00	2,50	12,40
24	100,00	63,00	0,00	20,30
25	100,00	6,00	2,00	4,10
26	100,00	8,00	4,00	20,50
27	50,00	34,00	3,00	12,80
28	4,00	0,00	0,00	1,00
29	100,00	66,00	5,50	10,60
30	100,00	31,00	0,00	6,10
31	100,00	14,00	0,00	13,30
	89,91	**30,61**	**1,61**	**10,82**

Table 4. UL patients' mHealth usage statistics imported from mClinician

- ✓ Patients from CHU

The usage statistics for the mHealthApp were obtained from 26 patients who were active until 31.10.2022. The data retrieved from the mClinician app (see Table No.5) indicated that synchronisation was successful approximately 69,47% of the time. However, 23,88% of smart-band usage (24 hours) and 12,88% of emotion data were obtained, and only 2,86% of diaries were recorded.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	Sync (%)	Smart-band usage (%)	Diaries (%)	Emotions (%)
1	100,00	46,00	5,50	23,00
2	50,00	36,00	4,50	18,40
3	100,00	24,00	2,50	11,40
4	50,00	100,00	0,00	11,30
5	100,00	40,00	2,00	17,30
6	1,00	0,00	6,00	6,30
7	50,00	3,00	2,00	16,50
8	100,00	31,00	0,00	4,40
9	33,30	14,00	4,00	18,80
10	100,00	26,00	3,50	8,10
11	100,00	28,00	6,00	18,20
12	50,00	33,00	3,50	19,20
13	100,00	41,00	4,50	23,30
14	33,30	6,00	1,00	3,40
15	13,50	1,00	1,00	2,90
16	100,00	31,00	5,80	16,30
17	50,00	4,00	2,00	6,50
18	50,00	1,00	0,00	4,60
19	100,00	27,00	1,00	12,30
20	50,00	17,00	3,50	19,00
21	100,00	15,00	1,00	12,20
22	100,00	15,00	2,50	4,50
23	100,00	24,00	4,50	19,00
24	25,00	28,00	5,00	18,80
25	100,00	1,00	3,00	2,30
26	50,00	29,00	0,00	17,00
	69,47	**23,88**	**2,86**	**12,88**

Table 5. CHU patients' mHealth usage statistics imported from mClinician

✓ Patients from UKCM

The usage statistics for the mHealthApp were collected from 39 patients. The data retrieved from the mClinician app (see Table No. 6) showed that it was successful around 64,83% of the time. However, 19,65% of smart-band usage (24 hours) and 9,74% of emotion data were obtained, and only 2,35% of diaries were recorded.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	Sync (%)	Smart-band usage (%)	Diaries (%)	Emotions (%)
1	100,00	24,00	1,00	2,20
2	100,00	8,00	0,00	16,70
3	100,00	13,00	1,00	6,30
4	100,00	32,00	5,50	22,30
5	50,00	0,00	3,00	11,90
6	50,00	32,00	9,00	21,30
7	100,00	100,00	4,50	19,80
8	100,00	34,00	3,30	15,20
9	50,00	16,00	0,00	6,00
10	100,00	37,00	2,00	6,90
11	100,00	41,00	1,00	8,50
12	50,00	40,00	0,00	9,50
13	100,00	3,00	0,00	1,50
14	50,00	24,00	3,00	5,80
15	1,00	0,00	0,00	1,00
16	16,70	19,00	3,00	6,80
17	4,00	0,00	0,00	2,00
18	7,00	19,00	0,00	2,00
19	100,00	44,00	4,00	10,70
20	100,00	13,00	5,00	12,40
21	50,00	2,00	4,00	6,40
22	100,00	18,00	1,00	3,20
23	50,00	12,00	1,00	4,00
24	86,00	10,00	0,00	4,00
25	50,00	5,00	0,00	2,80
26	100,00	9,00	2,30	3,40
27	100,00	34,00	5,50	11,50
28	50,00	13,00	3,00	16,30
29	100,00	26,00	3,00	13,60
30	33,30	33,00	0,00	7,40
31	100,00	3,00	2,50	12,00
32	20,00	19,00	4,00	5,40
33	25,00	7,00	5,00	19,50
34	33,30	32,00	9,00	18,30
35	2,00	0,00	0,00	1,00
36	100,00	15,00	0,00	11,80
37	50,00	17,00	2,00	20,30
38	50,00	3,00	4,00	21,50
39	50,00	6,00	0,00	8,80
	64,83	**19,56**	**2,35**	**9,74**

Table 6. UKCM patients' mHealth usage statistics imported from mClinician

✓ Patients from SERGAS

The usage statistics for the mHealthApp were collected from 33 patients who participated until 31.10.2022. The data retrieved from the mClinician app (see Table No. 7) showed that

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

it was successful around 20,10% of the time. However, 22,27% of smart-band usage (24 hours) and 15,6% of emotion data were obtained, and only 3,95% of diaries were recorded.

	Sync (%)	Smart-band usage (%)	Diaries (%)	Emotions (%)
1	50	18,00	4,50	14,80
2	100	41,00	0	22
3	50	35,00	5,50	26,30
4	100	8,00	0	5,70
5	50	7,00	4,50	22
6	50	51,00	3,30	19
7	100	0,00	0	5,30
8	100	71,00	0	22,50
9	50	31,00	2,80	21,30
10	50	7,00	0	16,50
11	100	15,00	0	24,30
12	100	32,00	4,70	19
13	100	31,00	4,50	23,50
14	100	100,00	5	23
15	100	37,00	3,50	24
16	40	4,00	0	13,70
17	3,50	2,00	5	17,80
18	50	3,00	3,50	18,40
19	50	32,00	4	20,30
20	12,30	2,00	1	4,90
21	33,30	1,00	2	11,80
22	10	2,00	4,50	14,20
23	25	32,00	4	24,30
24	50	29,00	0	19,50
25	100	30,00	3,50	18
26	33,30	9,00	2	6,80
27	50	27,00	3,50	22,30
28	11	13,00	0	4
29	100	21,00	4	18,80
30	8,30	11,00	0	4,60
31	33,30	14,00	0	17,80
32	33,30	18,00	3,50	9,10
33	3,50	1,00	3,50	5,40
	20,10	**22,27**	**3,95**	**15,60**

Table 7. SERGAS patients' mHealth usage statistics imported from mClinician

➔ Conclusions:

The data collected suggests that there is room for improvement in the usability of the technologies studied. A study evaluating the use of Publicly Available Physical Activity Mobile Apps have revealed that the engagement is affected by multiple factors and is highly personalised [1]. This means that the quality of the app, such as its usability, accuracy,

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

quality of production, and scientific evidence-base, is just one of many factors that determine how patients interact with it.

The smart band usage measurements consider 24 h in a row, meaning that if a patient uses it 24h per day the usage is accounted for 100%. On the other hand if the smart band is not used 24h the system is not counting, and as many patients might take the band for several hours (sleeping time for example) then the overall number does not represent the overall usage in time. Additionally, some patients may have experienced issues with the data transfer process, leading to incomplete data being received into the server.

Despite these challenges, the high rate in almost all hospitals suggests that patients found the app easy to use and manage on a daily basis. Further analysis of the data needs to be considered in future related clinical studies, but it is important to keep in mind that the percentage of measurements gathered from the smart-band may not accurately reflect patients' everyday activity.

The low percentage for diary data (~1,6-4%) can be explained by infrequent diary recording, as patients were asked to do this every three days, resulting in a maximum expected data of 30% in this category. Patients noted in workshops and consultations that they found it difficult to record a video diary. As a result, they were instructed again on using this function in the app and allowed to record diaries less frequently, meaning that the expected data was decreased.

The data about emotions represent the patients' willingness to report emotions in the app. They reflected on their emotions between 9-16% of the time. Section 7 contains a detailed analysis of the reported patient data.

In conclusion, the usage of new technologies by cancer patients can be a limiting factor in today's society, as seen in the challenges faced during the study. However, we foresee that the adoption of these technologies will increase as they become more user-friendly and accessible. The data gathered from these technologies, such as the publicly available self-monitoring mobile apps or smart-bracelets, can be valuable for big data platforms like PERSIST. With further analysis of the potential large data points that can be collected with new technologies, we can gain insights into the daily activities and emotional states of cancer patients, which can ultimately improve their care and outcomes. Overall, the use of new technologies has the potential to enhance the healthcare experience for cancer patients, and we look forward to seeing continued advancements in this field.

3. Perceived self-efficacy of patients (CASE cancer questionnaire)

Patient capacity to interact with healthcare professionals, their understanding of treatment regimens and options, and their ability to participate in healthcare decisions seem to play an important role in the development of anxiety and depression [2,3]. Although it has not been proven that a patient's psychological attitude towards cancer affects cancer survival

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

and recurrence, it may influence patient communication with healthcare professionals and adherence to treatments and follow-up [4,3]. Therefore, in order to measure patient self-efficacy, patients were asked to complete the CASE-cancer questionnaire.

The CASE-cancer questionnaire is a 12-item questionnaire developed by Wolf et al. (2005) with the objective of measuring cancer patients' capacity for productive communication and maintaining a positive attitude towards cancer. Three factors were suggested for analysing patient responses, namely: 1) understanding and participating in care, 2) maintaining a positive attitude, and 3) seeking and obtaining information. Each factor ranges on a scale from 1 to 16, where a higher value indicates a more positive result.

In the study, the CASE-cancer questionnaire was used to measure patient self-efficacy as a primary endpoint and to understand if the PERSIST system had any effect on patient self-efficacy. Patients completed the questionnaire at three time points during the project: at recruitment (baseline data), before the virtual assistant was launched into the mHealth app (07.06.2022), and at the end of the study (24-31.10.2022).*

A total of 75 questionnaires were analysed, and descriptive statistics were calculated for each score factor (Table 8). No statistically significant differences in scores between the recruitment and last follow-up were found in any of the three factors using the Wilcoxon test (Table 9). However, patients in the study showed a high level of understanding and participation in their care, with no score below 9 in Factor 1: Understand & Participate in care. This suggests that patients in the study had a good understanding of their treatment regimens and options and their ability to participate in healthcare decisions. The scores for Factor 2: Maintain positive attitude ranged from 4 to 16, indicating some difficulty in maintaining a positive attitude for some patients. However, all patients seemed willing to stay informed about their disease to some extent, according to the scores obtained for Factor 3: Seek & obtain information. The result shows that the patients participating in PERSIST are people capable of understanding, managing, and obtaining information about their illness.

**To comply with the provisions of the protocol approved by the Research Ethics Committee of Pontevedra-Vigo-Ourense and with the Spanish Organic Law of 3/2018, of December 5, on the Protection of Personal Data and guarantee of digital rights and in the provisions of the General Data Protection Regulation (EU Regulation 2016/679 of the European Parliament and of the Council, of April 27, 2016) SERGAS confirms that in the month of May 2023, after the official review with the European Commission, will eliminate the personal, pseudonymized and anonymized data of all the patients included in the PERSIST study, saving only the results obtained from the statistical analyses.*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	Factor 1: Understand & Participate in care		**Factor 2: Maintain positive attitude**		**Factor 3: Seek & obtain information**	
	Score at recruitment	Score at last follow-up	Score at recruitment	Score at last follow-up	Score at recruitment	Score at last follow-up
N	75	75	75	75	75	75
Mean	13.73	13.75	13.28	13.17	13.81	13.55
Median	14	14	14	14	15	14
Std. Deviation	1,905	2,014	2,299	2,435	2,312	2,207
Minimum	9	9	6	4	7	8
Maximum	16	16	16	16	16	16
Percentiles 25	12	12	12	12	12	12
50	14	14	14	14	15	14
70	16	15	15	15	16	16

Table 8. Descriptive statistics of Factors 1, 2 and 3

It is important to highlight that finding statistically significant differences in the responses related to self-efficacy was not expected in this project, as the patients were testing an application under development. Therefore, it was not expected that the mHealth app would have a tangible effect on the extent to which the patients perceive themselves as able to communicate with their doctor, maintain a positive attitude, or seek information about their disease.

The CASE-cancer questionnaire will be maintained as a primary endpoint in a future phase of development of the PERSIST system, where the functionality of a fully developed application will be evaluated and statistically significant differences are expected.

Factor	**Scores at recruitment (N=75)**	**Scores at the last follow up (N=75)**	**P Value**
Understand & participate in care (Median [IQR])	14[12-16]	14[12-16]	0,985
Maintain positive attitude (Median [IQR])	14[12-16]	14[12-16]	0,661
Seek & obtain information (Median [IQR])	15[12-16]	14[12-16]	0,255

Table 9. Comparison of the median scores of the three factors at recruitment vs at the last follow up. P values were calculated with the Wilcoxon test.

➔ Conclusions:

The results of the study indicate that patients who participated in the PERSIST study showed a high level of understanding and active involvement in their care, with no score below 9 in Factor 1: Understand & Participate in care. This suggests that patients in the study had a good understanding of their treatment options and were motivated to take an active role in managing their illness. Despite some patients reporting difficulty in maintaining a positive attitude towards their cancer (Factor 2), the majority of patients still

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

scored positively in this factor, indicating that they were able to maintain a positive outlook. Additionally, all patients in the study demonstrated a willingness to seek and obtain information about their disease (Factor 3), highlighting the importance of self-management in cancer care. This suggests that patients were motivated to learn about their illness and take an active role in their care.

Although some patients did report difficulties in maintaining a positive attitude, the scores for this factor still ranged from 4 to 16, indicating that most patients were able to maintain a positive outlook. These findings suggest that the PERSIST system could potentially help patients maintain their positive attitude in the future by providing support and encouragement

Overall, the results of the CASE-cancer questionnaire suggest that the patients participating in the PERSIST study had a positive attitude towards managing their illness and were actively involved in their care. While the study did not find statistically significant differences in the scores of patients' perceived self-efficacy, the analysis of the data suggest that the PERSIST results may be useful in supporting the management of cancer patients and provide a promising baseline for future development.

4. Activation levels of patients (PAM questionnaire)

In this study, the PAM-13 questionnaire was used as a secondary endpoint to measure patient knowledge, skills, and confidence in self-management. PAM-13 is a shortened version of the PAM-22 questionnaire that has been proven to be as effective [5]. As cancer patients, participants in PERSIST should acquire an efficient role in self-management, which requires a high level of knowledge, skills, and confidence measured by the PAM-13 questionnaire. To provide high-quality cancer care, it is essential to support patients in their role as self-managers. This support is precisely one of the potential functions of a fully developed mHealth app.

During the study, the PAM-13 questionnaire was provided to patients to complete at three timepoints: at recruitment (baseline data), before the virtual assistant was launched into the mHealth app (07.06.2022.), and at the end of the study (24.-31.10.2022), along with the aforementioned CASE-cancer questionnaire.

The result analysis from the start until the end of the study was performed using the corresponding PAM Score. The PAM Score is an interval-level scale from 0-100 that correlates with one of the four levels of patient activation. PAM levels 1 and 2 indicate lower patient activation, while PAM levels 3 and 4 indicate higher patient activation (https://www.insigniahealth.com/products/pam/pamsurvey):

Level 1: Disengaged and overwhelmed, individuals are passive and lack confidence. Healthcare knowledge is low, goal orientation is weak, and adherence is poor.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Level 2: Becoming aware but still struggling, individuals have some health-care knowledge, but large gaps remain. They believe health is largely out of their control but can set simple goals.

Level 3: Taking action and gaining control, individuals have the key facts and are building self-management skills. They strive for best practice behaviours and are goal-oriented.

Level 4: Maintaining behaviours and pushing further, individuals have adopted new behaviours but may struggle at times of stress or change. Maintaining a healthy lifestyle is a key focus.

A total of 78 questionnaires were analysed, and the descriptive statistics of the final scores obtained at recruitment and the last follow-up are presented in Table 10. The mean score of self-management level of the PERSIST participating patients at baseline was 3, and this level was maintained throughout the entire project with a slight increase in the mean score from 65,10 to 65,75 (Table 10). These results are not surprising as patients who volunteer to participate in these types of studies are expected to have a certain level of self-management skills. Thus, the PERSIST participants started the project with a good level of self-management skills, and this disposition was maintained throughout the entire project, as they continued to gain control and build their self-management skills.

	Score at recruitment	Score at last follow-up
N	78	78
Mean	65,10	65,71
Median	63,10	63,10
Std. Deviation	14,605	16,063
Minimum	38	37
Maximum	100	100
Percentiles 25	53,20	52,65
50	63,10	63,10
70	75,00	77,70

Table 10. Descriptive statistics of PAM questionnaire scores at recruitment and at the last follow-up. P value calculated with the Wilcoxon test.

Tables 11 and 12 display the evolution of patients in each of the four activation levels during the study. As shown in Table 11 and Table 12, most patients reported having level 3 or 4 of activation at both recruitment and last follow-up (42,3% and 32,1% respectively), with a small increase in the number of patients reporting level 4 activation at the follow-up (from 32,1% to 35,9%).

As was the case with the CASE questionnaire analysis, no statistically significant difference was found in the percentage of patients at each level between recruitment and last follow-

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

up (Table 12). It is necessary to consider that finding statistically significant differences in the levels of self-management was not an objective of the project, and that the PAM-13 questionnaire was used as a tool to monitor this patient skill. Nonetheless, despite not having seen statistically significant differences in the activation levels of the participants throughout the project, these results show that the mHealth app still under development may be able to support patients in improving their self-management skills, which is encouraging when it comes to continuing the development of the PERSIST system.

I			PAM levels at last follow-up				Total
			1	2	3	4	
PAM levels at recruitment	1	Count	4	0	1	0	**5**
		% of Total	5,1%	0%	1,3%	0%	**6,.4%**
	2	Count	1	7	6	1	**15**
		% of Total	1,3%	90%	7,7%	1,3%	**19,2%**
	3	Count	1	8	14	10	**33**
		% of Total	1,3%	10,3%	17,9%	12.8%	**42,3%**
	4	Count	0	1	7	17	**25**
		% of Total	0%	1.3%	9.0%	21.8%	**32,1%**
Total		Count	**6**	**16**	**28**	**28**	**78**
		% of Total	**7.7%**	**20.5%**	**35.9%**	**35.9%**	**100,0%**

Table 11. Level PAM pre vs post Cross tabulation

Level	Recruitment (N=75)	Last follow-up (N=75)	P value
Level 1 n (%)	5 (6,4)	6 (7,7)	1,000
Level 2 n (%)	15 (19,2)	16 (20,5)	1,000
Level 3 n (%)	33 (42,3)	28 (35,9)	0,486
Level 4 n (%)	25 (32,1)	28 (35,9)	0,648

Table 12. Comparison of the percentage of patients in each level at the recruitment vs at the last follow-up. P values have been calculated with McNemar Test

➔ Conclusions:

The results of the PAM-13 questionnaire in the study suggest that the participating cancer patients had a good level of self-management skills at baseline, which was maintained throughout the study. Most patients reported having level 3 or 4 of activation at both recruitment and last follow-up, indicating that they were taking action and gaining control over their condition. Although no statistically significant differences were found in the activation levels of the participants throughout the project, the results suggest that the mHealth app under development may be able to support patients in improving their self-management skills, which is an encouraging finding for the development of the PERSIST system. Overall, these results suggest that the PERSIST results may contribute to supporting cancer patients in their role as self-managers, which is essential for high-quality cancer care.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

5. User acceptance of mHealthApp (SUS) from patients

During the development of the mHealthApp, we engaged in a co-creation phase that included user testing, where end users provided direct feedback and recommendations for improving usability. To evaluate usability, we implemented the System Usability Scale (SUS), which is a widely used questionnaire that measures user perception of system convenience and necessary skills. However, it should be noted that some authors have argued that it may not capture all unique aspects of mobile apps [6].

The SUS consists of 10 statements with five response options for respondents, ranging from "strongly agree" to "strongly disagree". We administered the questionnaire to patients in their native language, including Latvian, Russian, Slovenian, Spanish, or French.

Participants completed the survey in three rounds: at the beginning of 2022 (January-April), around two months after the virtual agents were presented in the App (07.06.2022), and at the end of the clinical study (last week of October 2022). As the questionnaires were available at all times in the mHealthApp, some participants chose to complete these questionnaires more than once. A statistician from the team analysed 27 from patients who responded at all the three time points.

For each patient, the SUS score group was calculated. Figure 1 shows the sum score of the 10 questions. According to the definition of the system usability level (Table 13), at the beginning of 2022 (Figure 1, Table 14), most of the patients thought that the system was "experiencing usability issues" (frequency of 10) and "acceptable to good" (frequency of 10). This could be related to the patients' previous experience with technology in general, including the use of various types of applications and the possibility to adapt to the mHealthApp, which was in the development process.

Level	Definition
<=50	Not easy to use
50-70	Experiencing usability issues
70-85	Acceptable to good
>85	Excellent usability

Table 13 The definition of system usability level

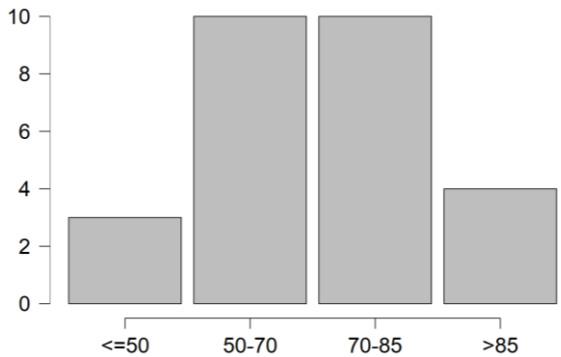

Figure 1 The sum score of the points acquired in all 10 questions in the beginning

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Score group first	Frequency	Percent	Valid Percent
<=50	3	11,111	11,111
50-70	10	37,037	37,037
70-85	10	37,037	37,037
>85	4	14,815	14,815
Missing	0	0,000	
Total	27	100,000	

Table 14 Frequencies for Score group first

During the study, the percentage of participants who rated the system as having "excellent usability" increased from 14% to 33% (Figure 2, Table 15). This could be attributed to the ongoing upgrades made to the mHealth app in collaboration with technical partners.

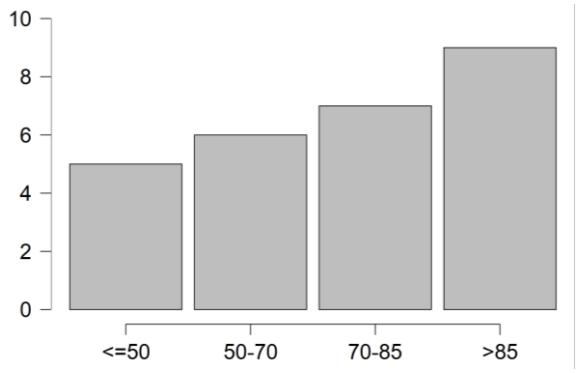

Figure 2 The score group in the middle of the study

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Score group mid	Frequency	Percent	Valid Percent
<=50	5	18,519	18,519
50-70	6	22,222	22,222
70-85	7	25,926	25,926
>85	9	33,333	33,333
Missing	0	0,000	
Total	27	100,000	

Table 15 Frequencies for Score group mid

At the end of the study, the most popular score group for the system was "Experiencing usability issues" (Figure 3, Table 16), which could be explained by negative feedback from patients about the virtual agent introduction. The virtual agent was repeating the same phrases in the app, but later, the option to exclude it was added. Nevertheless, 44,44% of patients evaluated the usability as good or excellent, combining the answers "Acceptable to good" and "Excellent usability" together.

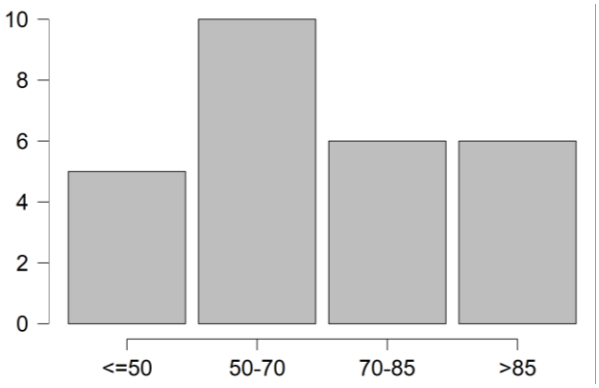

Figure 3 The score group at the end of the study

Score group mid	Frequency	Percent	Valid Percent
<=50	5	18,519	18,519
50-70	10	37,037	37,037
70-85	6	22,222	22,222
>85	6	22,222	22,222
Missing	0	0,000	
Total	27	100,000	

Table 16 Frequencies for Score group final

➔ Conclusions:

The use of user testing and the SUS questionnaire proved to be effective tools in identifying usability issues with the mHealthApp and making improvements that benefited the end user. As the study progressed, there was a notable increase in the number of participants

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

who perceived the system as having excellent usability, which could be attributed to the constant upgrades made to the app in collaboration with technical partners. Despite some negative feedback about the virtual agent introduction, 44,44% of patients still evaluated the usability of the app as good or excellent. These findings highlight the importance of user testing and continuous improvement to enhance the usability and user acceptance of mHealth apps.

The individual analysis of each statement in the SUS questionnaire can be found below, where participants were able to provide feedback on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) for all the statements in the survey.

- ✓ I think that I would like to use this system frequently.

The majority of the participants are neutral or have a slightly positive attitude towards using the system frequently (Table 17). The mean scores for the statement are between 3,444 and 3,519 across the three assessment periods. The median is also 3,0 for all three periods. The standard deviation decreases over time, indicating that the participants' opinions became more consistent over time.

However, it is important to note that the majority of the responses fall in the 4th category, which suggests that the participants agree with the statement to some extent. Only a small percentage of participants strongly disagree or strongly agree with the statement. Therefore, it may be necessary to conduct further research to explore the reasons behind the participants' attitudes towards using the system frequently.

	First	Mid	Final
Mean	3,444	3,481	3,519
Median	3,000	3,000	3,000
Std, Deviation	1,311	1,122	0,975
Minimum	1,000	1,000	1,000
Maximum	5,000	5,000	5,000
25th percentile	3,000	3,000	3,000
50th percentile	3,000	3,000	3,000
75th percentile	5,000	4,500	4,000

Table 17 Descriptive statistics on SUS 1st statement

Friedman One-Way Repeated Measure Analysis of Variance by Ranks shows that there are no statistically significant differences between any two time point answers (p=0,758; Conover's post-hoc comparisons: First-Mid p=0,599, First-Final p=0,833, Mid-Final p=0,462) and there is no significant difference between the answers of each centre (p=0,206).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

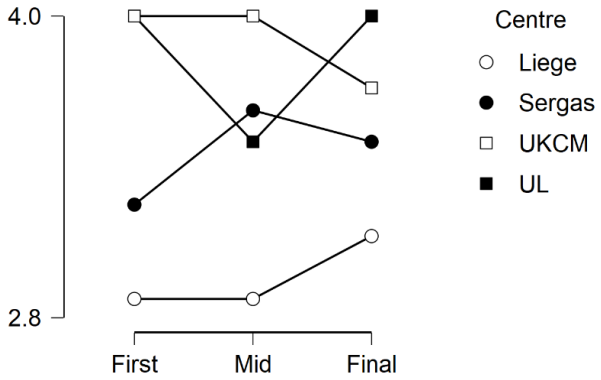

Figure 4: SUS 1st statement replies by the centres at all three time points.

➔ Conclusions:

Overall, while there were some differences in mean scores between the different centres (Figure 4), these differences were not statistically significant, indicating that there was overall consistency in participants' attitudes towards using the system across all centres and at all time points.

The majority of the participants have a neutral to slightly positive attitude towards using the system frequently. This suggests that the PERSIST system has potential to be useful to a wide range of users, as most participants did not strongly object to the idea of using it frequently.

Additionally, the fact that the standard deviation decreased over time suggests that the participants became more consistent in their opinions and potentially more familiar with the system. This could indicate that with more exposure to the system, users may become more comfortable and confident in using it, leading to higher levels of satisfaction and engagement.

Overall, these findings suggest that the system has potential to be well-received by users and could be further developed and improved upon based on user feedback

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

✓ I found the system unnecessarily complex.

	First	Mid	Final
Mean	2,148	2,037	2,000
Median	2,000	1,000	2,000
Std, Deviation	1,134	1,285	1,109
Minimum	1,000	1,000	1,000
Maximum	5,000	5,000	4,000
25th percentile	1,000	1,000	1,000
50th percentile	2,000	1,000	2,000
75th percentile	3,000	3,000	3,000

Table 18 Descriptive statistics on SUS 2nd statement

Friedman One-Way Repeated Measure Analysis of Variance by Ranks shows that there are no statistically significant differences between any two time points answers (p=0,758; Conover's post-hoc comparisons: First-Mid p=0,726, First-Final p=0,540, Mid-Final p=0,337) and there is no significant difference between the answers of each centre (p=0,383).

Participants did not find the system unnecessarily complex, as the mean scores for the SUS 2nd statement are relatively low across all time points and there were no statistically significant differences between the answers at different time points or between the different centres (Table 18). Specifically, the mean scores for the statement "I found the system unnecessarily complex" were 2,148 at the beginning of the study, 2,037 in the middle, and 2,000 at the end, all of which indicate a slightly negative attitude towards the complexity of the system. However, these differences were not statistically significant, and there was overall consistency in participants' attitudes towards the complexity of the system across all centres and at all time points.

Therefore, it can be concluded that the system was not perceived as unnecessarily complex by the majority of participants, and there were no significant changes in participants' attitudes towards the complexity of the system over time or between the different centres.

Nevertheless, apart from UKCM patients, several patients from the other 3 centres do not think that the system is unnecessarily complex (Figure 5). Thus, if the answers of UKCM, which show a different opinion from other three centres (that the system is complex), are excluded, a significant difference (p=0,013) would have been revealed between the answers gathered at the beginning of 2022 and the end of the study.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

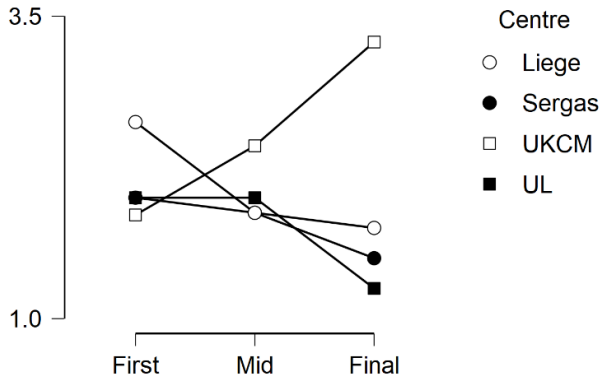

Figure 5: SUS 2nd statement replies by the centres at all three time points.

➔ Conclusions:

The participants did not find the system unnecessarily complex, as indicated by the relatively low mean scores for the SUS 2nd statement across all time points. This suggests that the system was designed in a user-friendly way and was not a source of frustration or confusion for the participants.

Additionally, the lack of significant differences between the answers at different time points or between the different centres indicates that the user-friendliness of the system was consistent across all groups and did not deteriorate over time. This is a positive outcome as it suggests that the system was able to maintain its user-friendliness throughout the study period, which can lead to increased user satisfaction and improved usability.

✓ I thought the system was easy to use.

	First	Mid	Final
Mean	4,148	4,037	3,889
Median	4,000	4,000	4,000
Std, Deviation	0,907	1,091	0,974
Minimum	2,000	1,000	2,000
Maximum	5,000	5,000	5,000
25th percentile	3,500	3,000	3,000
50th percentile	4,000	4,000	4,000
75th percentile	5,000	5,000	5,000

Table 19 Descriptive statistics on SUS 3rd statement

Friedman One-Way Repeated Measure Analysis of Variance by Ranks shows that there are no statistically significant differences between any two time points answers (p=0,504; Conover's post-hoc comparisons: First-Mid p=0,776, First-Final p=0,258, Mid-Final p=0,395) there is no significant difference between the answers of each centre (p=0,957).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

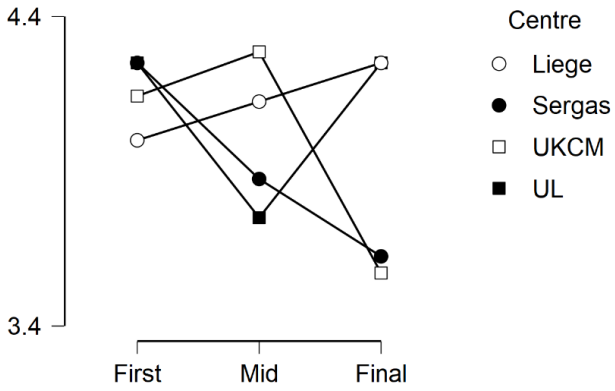

Figure 6: SUS 3rd statement replies by the centres at all three time points.

Patients from UL and CHU during the three check-ups have given slightly more points meaning that at the end of the study they thought that the system was easier to use (Figure 6). No statistically significant differences between any two time points could be found.

➔ Conclusions:

The participants found the system easy to use. The high mean scores for the SUS 3rd statement across all time points suggest that the system was designed in a user-friendly way and was not a source of frustration or difficulty for the participants.

The lack of significant differences between the answers at different time points or between the different centres indicates that the ease of use of the system was consistent across all groups and did not deteriorate over time. This is a positive outcome as it suggests that the system was able to maintain its user-friendliness throughout the study period, which can lead to increased user satisfaction and improved usability. Overall, the results of this statement are positive, as they indicate that the system was well-designed and easy to use for the participants.

✓ I think that I would need the support of a technical person to be able to use this system.

	First	Mid	Final
Mean	2,037	1,667	2,148
Median	2,000	1,000	2,000
Std, Deviation	1,126	1,074	1,231
Minimum	1,000	1,000	1,000
Maximum	5,000	4,000	5,000
25th percentile	1,000	1,000	1,000
50th percentile	2,000	1,000	2,000
75th percentile	3,000	2,000	3,000

Table 20 Descriptive statistics on SUS 4th statement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Friedman One-Way Repeated Measure Analysis of Variance by Ranks shows that there are no statistically significant differences between any any two time points answers except for mid and final (p=0,071; Conover's post-hoc comparisons: First-Mid p=0,076, First-Final p=0,705, Mid-Final p=0,033) and centre is not a statistically significant impact factor (p=0,933).

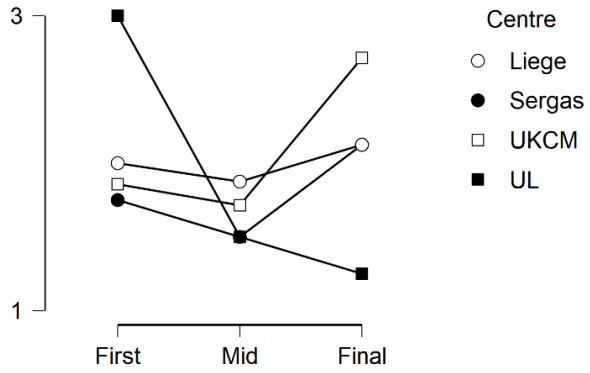

Figure 7: SUS 4th statement replies by the centres at all three time points.

➔ Conclusions:

The main conclusion from this is that users found the system difficult to use without the support of a technical person, as evidenced by the higher scores on the fourth statement of the SUS survey (Table 20). However, there was a significant improvement in the final survey compared to the mid survey, indicating that perhaps the technical support provided was helpful. The lack of significant differences between centres suggests (Figure 7) that the difficulty in using the system was a universal experience rather than a centre-specific issue.

✓ I found the various functions in this system were well integrated.

	First	Mid	Final
Mean	3,667	3,926	3,704
Median	3,000	4,000	4,000
Std, Deviation	0,961	0,997	0,953
Minimum	2,000	2,000	1,000
Maximum	5,000	5,000	5,000
25th percentile	3,000	3,000	3,000
50th percentile	3,000	4,000	4,000
75th percentile	4,500	5,000	4,000

Table 21 Descriptive statistics on SUS 5th statement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Friedman One-Way Repeated Measure Analysis of Variance by Ranks shows that there are no statistically significant differences between any any two time points answers (p=0,167; Conover's post-hoc comparisons: First-Mid p=0,077, First-Final p=0,765, Mid-Final p=0,139) and there is no significant difference between the answers of each centre (p=0,442).

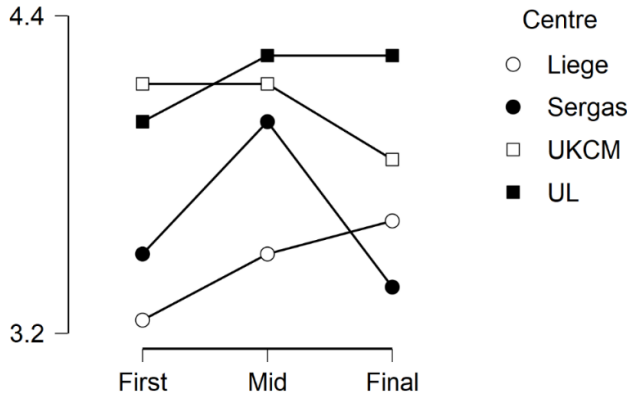

Figure 8: SUS 5th statement replies by the centres at all three time points.

➔ Conclusions:

Participants generally had a neutral to positive perception of the integration of various functions in the system (Table 20). This is indicated by the mean score for the statement being above 3 (which is the midpoint of the Likert scale), and the median score being 4 in the Mid and Final assessments. Additionally, there were no statistically significant differences between any two time points answers, suggesting that participants' perceptions of the system remained consistent over time. Finally, there were no significant differences between the answers of each centre (Figure8), indicating that participants from different centres had similar perceptions of the system's integration.

✓ I thought there was too much inconsistency in this system.

	First	Mid	Final
Mean	2,519	2,185	2,259
Median	3,000	2,000	2,000
Std, Deviation	1,014	1,145	1,095
Minimum	1,000	1,000	1,000
Maximum	5,000	4,000	4,000
25th percentile	2,000	1,000	1,000
50th percentile	3,000	2,000	2,000
75th percentile	3,000	3,000	3,000

Table 21 Descriptive statistics on SUS 6h statement

Friedman One-Way Repeated Measure Analysis of Variance by Ranks shows that there are no statistically significant differences between any any two time points answers

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

(p=0,363; Conover's post-hoc comparisons: First-Mid p=0,219, First-Final p=0,219, Mid-Final p>0,999) and there is no significant difference between the answers of each centre (p=0,856).

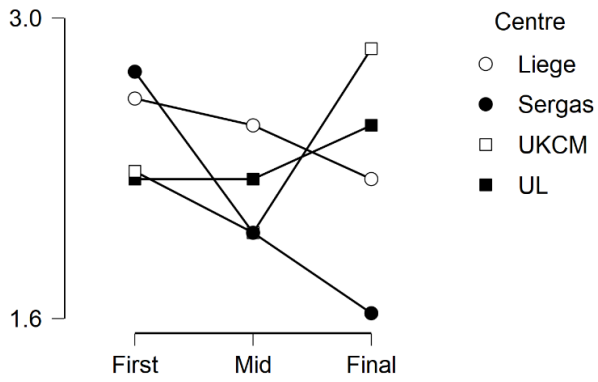

Figure9: SUS 6th statement replies by the centres at all three time points.

➔ Conclusions:

The mean score for the statement is below 3 (which is the midpoint of the Likert scale), indicating that participants generally disagreed with the statement (Table 21). However, the median score for the statement decreased from 3 to 2 between the First and Mid assessments, suggesting that some participants may have become more critical of the system's consistency over time. However, there were no statistically significant differences between any two time points answers or between the answers of each centre (Figure 9), indicating that any changes in participants' perceptions of the system's consistency over time were not significant.

✓ I would imagine that most people would learn to use this system very quickly.

	First	Mid	Final
Mean	4,222	4,111	3,926
Median	5,000	4,000	4,000
Std, Deviation	0,892	0,934	1,072
Minimum	3,000	2,000	1,000
Maximum	5,000	5,000	5,000
25th percentile	3,000	3,000	3,000
50th percentile	5,000	4,000	4,000
75th percentile	5,000	5,000	5,000

Table 22 Descriptive statistics on SUS 7th statement

Friedman One-Way Repeated Measure Analysis of Variance by Ranks shows that there are no statistically significant differences between any any two time points answers

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

(p=0,526; Conover's post-hoc comparisons: First-Mid p=0,566, First-Final p=0,253, Mid-Final p=0,566) and there is no significant difference between the answers of each centre (p=0,874).

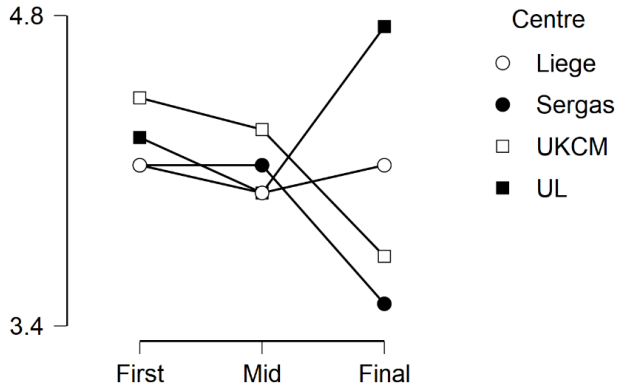

Figure 10: SUS 7th statement replies by the centres at all three time points.

➔ Conclusions:

The participants generally agreed with the statement that most people would learn to use the system very quickly (Table 22). This is indicated by the mean score for the statement being above 4 (which is closer to "strongly agree" on the Likert scale), and the median score being 4 in the Mid and Final assessments.

Furthermore, there were no statistically significant differences between any two time points answers or between the answers of each centre (Figure 10), indicating that participants' perceptions of the ease of learning the system were consistent over time and across different centres. Therefore, it can be inferred that participants found the system easy to learn and that it would not require a significant amount of training for most people.

✓ I found the system very cumbersome/awkward to use.

	First	Mid	Final
Mean	1,852	1,852	1,889
Median	2,000	1,000	2,000
Std, Deviation	0,949	1,027	0,934
Minimum	1,000	1,000	1,000
Maximum	4,000	4,000	4,000
25th percentile	1,000	1,000	1,000
50th percentile	2,000	1,000	2,000
75th percentile	3,000	3,000	3,000

Table 23: Descriptive statistics on SUS 8th statement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Friedman One-Way Repeated Measure Analysis of Variance by Ranks shows that there are no statistically significant differences between any any two time points answers (p=0,948; Conover's post-hoc comparisons: First-Mid p>0,999, First-Final p=0,774, Mid-Final p=0,774) and there is no significant difference between the answers of each centre (p=0,916).

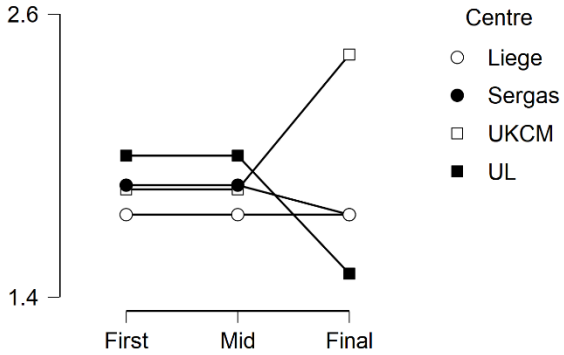

Figure 11: SUS 8th statement replies by the centres at all three time points.

➔ Conclusions:

No significant differences were found between any time points or between the centres (Figure 10) in relation to the 8th statement on the SUS survey. This suggests that overall, participants did not find the system to be too cumbersome or awkward to use. A positive conclusion that can be drawn is that the system was relatively easy to use for participants and did not cause significant frustration or difficulty.

✓ I felt very confident using the system.

	First	Mid	Final
Mean	3,889	4,148	3,741
Median	4,000	4,000	4,000
Std, Deviation	0,892	0,949	0,984
Minimum	2,000	2,000	2,000
Maximum	5,000	5,000	5,000
25th percentile	3,000	3,000	3,000
50th percentile	4,000	4,000	4,000
75th percentile	5,000	5,000	5,000

Table 24: Descriptive statistics on SUS 9th statement

Friedman One-Way Repeated Measure Analysis of Variance by Ranks (Table 24) shows that there are no statistically significant differences except between mid and final time points answers (p=0,138; Conover's post-hoc comparisons: First-Mid p=486, First-Final

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

p=0,199, Mid-Final p=0,050) (see table No. 25, Figure 12.). And there is no significant difference between the answers of each hospital patient (p=0,864).

SUS answers	Centre	Mean	SD	N
First	Liege	3,625	0,916	8
	Sergas	4,125	0,641	8
	UKCM	3,857	1,215	7
	UL	4,000	0,816	4
Mid	Liege	4,125	0,991	8
	Sergas	4,125	0,835	8
	UKCM	4,429	1,134	7
	UL	3,750	0,957	4
Final	Liege	3,500	1,309	8
	Sergas	3,625	0,916	8
	UKCM	4,000	0,816	7
	UL	4,000	0,816	4

Table 25: Patient points by centre

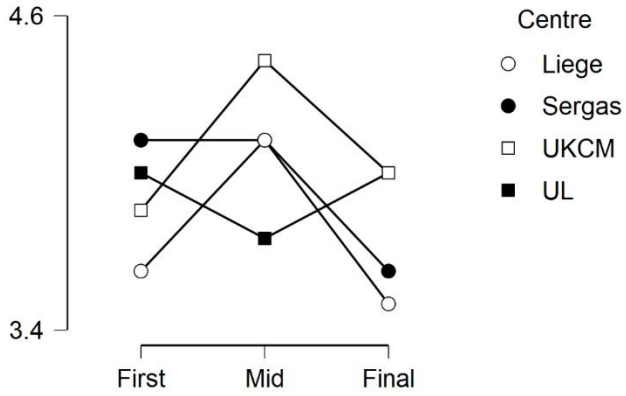

Figure 12: SUS 9th statement replies by the centres at all three time points.

From the table and Figure 12, we can see that the highest mean score was achieved by UKCM at the Mid time point, while the lowest mean score was achieved by Liege at the Final time point. The standard deviation also varies across centres and time points, indicating differences in patient responses.

➔ Conclusions:

Initially, the participants felt confident using the system, which is reflected in the mean and median scores for the first time point. Although the final scores showed a slight decrease in confidence, the mean and median scores were still relatively high, indicating that overall, the participants still felt confident using the system. The lack of significant differences between the answers of each hospital patient further supports this positive conclusion.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

✓ I needed to learn a lot of things before I could get going with this system.

	First	Mid	Final
Mean	2,222	2,074	2,148
Median	1,000	1,000	2,000
Std, Deviation	1,450	1,385	1,199
Minimum	1,000	1,000	1,000
Maximum	5,000	5,000	5,000
25th percentile	1,000	1,000	1,000
50th percentile	1,000	1,000	2,000
75th percentile	3,000	3,000	3,000

Table 26: Descriptive statistics on SUS 10th statement

Friedman One-Way Repeated Measure Analysis of Variance by Ranks (Table 26) shows that there are no statistically significant differences between any any two time points answers (p=0,684; Conover's post-hoc comparisons: First-Mid p=0,417, First-Final p=0,477, Mid-Final p=0,919) and there is no significant difference between the answers of each centre (p=0,208).

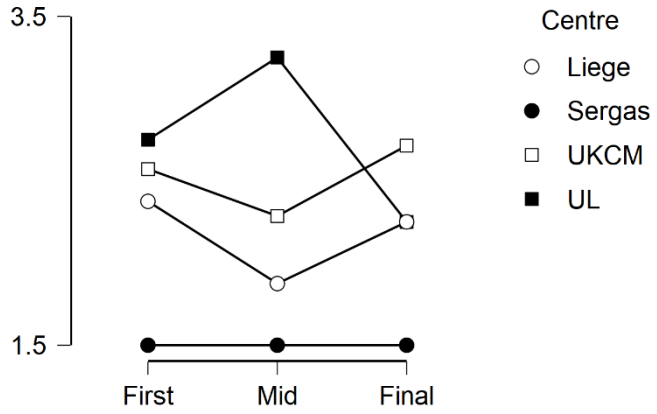

Figure 13: SUS 10th statement replies by the centres at all three time points

➔ Conclusions:

Despite the participants needing to learn a lot before they could use the system, there was no significant difference in their responses between the first, mid, and final time points. This suggests that the learning curve for using the system was not too steep and that participants were able to adapt and improve their understanding of the system over time. Additionally, there was no significant difference in responses between each centre (Figure 13), indicating that the system was equally accessible and usable across different locations.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

6. Blood pressure/heart rate and steps

All the data regarding patient blood pressure, heart rate and step measurement were analysed by CHU statisticians below.

Context and data sets

The enrolled patients were equipped with a smartwatch that allowed for the daily measurement of four parameters: the number of steps taken, heart rate (HR), and systolic and diastolic blood pressure (SBP and DBP, respectively).

The objective of this study is to describe the large amount of data collected, to analyse the evolution of the patients, and to examine the relationships between the four parameters. The data on the number of steps per day, as well as the mean, minimum, and maximum values per day of HR, SBP, and DBP for each patient from each hospital, were combined into a single file.

After removing observations where the number of steps was less than 10 per day, there were still 38,482 lines of data. However, for each line, measurements of all four parameters on the same day were not always available.

Statistical methods

Continuous variables were described using mean and standard deviation (±SD) or median and interquartile range (Q1 – Q3) as appropriate. Categorical variables were described using frequency tables (number and percent).

The comparisons of continuous variables between the hospitals were performed by the Kruskal-Wallis test and the post-hoc test of Dwass-Steel-Critchlow-Fligner (DSCF) was used for pairwise comparisons. Chi-square test was used for categorical variables.

We reported for each patient, the mean and the coefficient of variation (CV) per day of the 4 parameters. The CV is expressed in %, the higher it is, the greater the dispersion around the mean.

To study the participants' evolution, only those who had at least 10 days of measurements of the number of steps, mean HR and mean SBP/DB were considered. The slopes of linear evolution of each parameter of each patient were calculated and classified into 3 groups: “Significant positive slope”, “Significant negative slope”, “Non significant slope”. Slopes of the parameters were compared using Spearman correlation analysis. Slopes categories were compared using Fisher's Exact Test.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 875406

Missing data were not replaced, and calculations were always done on the maximum number of data available. Results were considered significant at the 5% critical level (p<0,05). Data analysis was carried out using SAS (version 9.4) and R (version 4.1.0) was used for the figures. Results are summarised in this report.

Descriptive statistics

➔ **Cancer survivor patients**

Table 27 describes the age and gender of the patients with at least one measurement, globally and by hospital. The patients were on average 55,0 ± 10,3 years old (range: 33-75 years old). There were 113 (73,4%) women and 41 (26,6%) men. Since the age and gender were comparable in the four hospitals (p=0,84 and p=0,31, respectively), a comparison between the participating countries allowed us to assess homogeneity.

		All	LIEGE	SERGAS	UKCM	UL	P-value
N		154	40	32	39	43	
Age (years)		55,0 ± 10,3	54,6 ± 11,0	54,8 ± 10,8	56,3 ± 8,4	54,5 ± 11,1	0,84
Gender	F	113 (73,4)	28 (70,0)	21 (65,6)	28 (71,8)	36 (83,7)	0,31
	M	41 (26,6)	12 (30,0)	11 (34,4)	11 (28,2)	7 (16,3)	

Table 27: Age and gender of participants for blood pressure/steps analysis (N=154, F=female; M= male)

➔ **Number of steps.**

A total of 34 001 measurements of the number of steps per day (STEP) were considered over all patients. The deciles of all these measurements were calculated. The overall first decile (= very low walking activity) was 2 452 steps/day, the overall median was 9 002 steps/day, and the overall last decile 9 (= very high walking activity) was 19 422 steps/day. Each of the 34 001 measurements was classified into one of the 10 categories defined by the deciles. The proportion of measurements above the overall median, or below the overall first decile, or above the overall last decile, was then possible to calculate for each patient. Thus, each participant could be compared to the others.

For each patient, the following information was gathered: Patient ID; Date of the first step measurement; Date of the last step measurement; Time between the first and last step measurement (days); Number of step measurements; Compliance of step measurements (%); Mean, standard deviation (SD), Coefficient of variation (CV), Minimum, 1st quartile (Q1), Median, 3rd quartile (Q3), and maximum of the number of steps; Number and proportion (%) of measurements lower than the overall 1st decile; Number and proportion (%) lower than the overall median; Number and proportion (%) higher than the overall median; Number and proportion (%) of measurements of the patient higher than the overall 9th decile.

As reported in Table 32, comparisons between the hospitals showed significant differences for the number of step measurements (p=0,042), the compliance (p=0,013), the median

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

number of steps (p=0,0002), the proportion of measurements lower than the overall 1st decile (p=0,0002), and the proportion of measurements higher than the overall median (p=0,0012). In general, participants from CHUL (Belgium) walked less than the others and had a lower compliance. This result is aligned with the observation of Althoff et al [7], who reported that Belgian citizens walked less than Spanish ones (average 5000 vs 5936 steps). In the frame of PERSIST, reasons such as low walkability of the environment of the Belgian participants, as well as the rainy weather in both winter and summer times in the Wallonia region, can be behind the lowest physical activity of CHUL participants among the four hospitals.

	All	CHU	SERGAS	UKCM	UL	P-value
N	156	40	35	40	41	
Time between first and last STEP measurement (days)	383 (98,5 - 519)	226 (60,5 – 478)	467 (109 – 521)	455 (267 – 526)	376 (71 – 525)	0,22
Number of STEP measurements	161 (29 – 388)	68,5 (14,5 - 282)	239 (39,5 – 446)	272 (101 – 415)	124 (20 – 353)	0,042
Compliance STEP measurements (%)	70,8 (41,5 - 86,7)	54,0 (25,5 - 77, 5)	85,3 (53,3 - 91,4)	70,8 (48,3 - 89,5)	74,2 (39,4 - 82,9)	0,013
Median number of steps/day	7278 (4842 - 11194)	5386 (3809 – 7530)	8697 (6102 – 11986)	9251 (6217 – 14305)	6257 (5064 – 9249)	0,0002
% measurements lower than overall 1st decile (<2452 steps/day)	12,5 (5,44 - 27,7)	23,9 (14,9 - 45,0)	6,41 (2,19 - 21,1)	7,44 (3,38 - 18,2)	15,5 (7 – 25)	0,0002
% measurements higher than overall median (>9002 steps/day)	38,5 (13,6 - 67,1)	22,8 (9,39 - 40,4)	46,3 (26,4 - 80,1)	53,1 (27,5 - 76,8)	29,6 (12,7 - 54,3)	0,0012
% measurements higher than overall 9th decile (>19422 steps/day)	0,73 (0 - 6,14)	0 (0 - 4,26)	3,24 (0 - 11,1)	0,91 (0 - 20,8)	1,37 (0 - 4,26)	0,17

Table 32. Number of steps/day vs. hospitals (N=156 patients with at least one measurement). Results are expressed as median (IQR), IQR=interquartile range Q1-Q3; P-value is the Kruskal-Wallis test

In addition, Eurostat [8] shows that a high percentage of the adult general population in Spain and Slovenia (> 67%) does sufficient physical activity. This trend is also followed by the Spanish and Slovenian PERSIST survivors who were the top 2 walkers. On the contrary, Eurostat shows that the percentage of the counterpart population decreases tremendously in Belgium (39%) and in Latvia (around 13%) [8]. Belgian cancer survivors again follow the Eurostat's report, i.e Belgian participants walked around 40% less than the Spanish and Slovenian survivors (median steps 5385 vs almost 8696 and 9250 respectively). Interestingly, the statistics of the Latvian participants do not follow their general population since even though they walked less than the Spanish and Slovenian ones these differences are not statistically significant (median steps 6257 vs almost 8696

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

and 9250 respectively). Thus, it seems that the pedometer-based walking of PERSIST had already from the beginning of the study a much higher positive impact in the Latvian cancer survivors than in the Belgian ones. This is all the more true since most (60%) of Latvian participants (with registration of the 4 parameters) did not improve their walking behaviour throughout the study and only a few increased their walking activity (14%). The other 26% decreased their step counts (Table 33).

Even though the Belgian patients walked less than the rest of the participants, the smart bracelet seems to have encouraged them more than their counterparts to walk as the study progressed. Indeed, 41% of Belgian patients with step measurements increased their physical activity vs 14% (Latvia); 23% (Spanish) and 29% (Slovenian) (Table 33).

Slope	N	UL (Latvian)	CHUL (Belgian)	Sergas (Spanish)	UKCM (Slovenian)
	131 (%)	35 (%)	32 (%)	30 (%)	34 (%)
Significant positive slope	35 (26,7)	5 (14)	13 (41)	7 (23)	10 (29)
Significant negative slope	28 (21,4)	9 (26)	3 (9)	9 (30)	7 (21)
Non-Significant slope	68 (51,9)	21(60)	16 (50)	14 (47)	17 (50)

Table 33. Walking behaviour among the patients who had registrations of the 4 parameters

The median compliance, (i.e. the percentage of days the patient walked during his participation), was 70,8%. More than 50% of the patients had a median number of steps per day higher than 7278. More than 50% of the patients had 12,5% of their measurements below the overall first decile, 38,5% above the overall median and 0,73% above the overall last decile

Unexpectedly, not only compliance increased with age (r=0,17, p=0,037), (Table 34) but the proportion of measurements lower than overall 1st decile decreased with age (r=-0.16, p=0,046). These are encouraging observations given that the literature reports that older adults are not meeting current physical activity recommendations [9, 10]. Analysing all the participants together, the compliance was higher in males (males 74,7% vs. females 67,1%, p=0.048, table 33) despite that they are older than female participants (61,7 vs 52,6 y/o). This difference can be explained by a higher number of female participants enrolled in the breast cancer survivor group, of which the literature reports that one in two suffered from joint pain. [11]. This important result prompted us to propose in a future version of the PERSIST solution other low-impact aerobic exercises alleviating specific pains of breast cancer survivors such as swimming or biking for those who do not like to walk [12].

Interestingly, analysis of the compliance of colorectal cancer survivors alone revealed that females walked more days than men (62,12% vs 47,02% compliance) during the time they participated in the study. This can be explained by the fact that the mean age of women (60 y/o) was lower than men (66 yo) (data from colorectal cancer CHUL participants). Although other factors can explain this difference, again more personalised physical activities in line with certain CRC-related discomforts have to be proposed in a future solution version.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	Age (years)		Gender		
	R	P-value	Female	Male	P-value
Time between first and last STEP measurement (days)	0,047	0,57	416 (178-523)	289 (61-512)	0,15
Number of STEP measurements	0,086	0,29	204 (42-405)	132 (21-400)	0,50
Compliance STEP measurements (%)	0,17	0,037	67,1 (39,3-81,4)	74, 7 (46,9-91,2)	0,048
Median number of steps/day	0,097	0,24	7479 (5178-11033)	6372 (3625-11075)	0,13
% measurements lower than overall 1st decile (<2452 steps/day)	-0,16	0,046	11,6 (6,02-22,5)	17,1 (4,39-48,2)	0,38
% measurements higher than overall median (>9002 steps/day)	0,028	0,73	40,4 (14,3-63,1)	34,5 (12,7-65,0)	0,38
% measurements higher than overall 9th decile (>19422 steps/day)	-0,049	0,55	0,91 (0-7,09)	0,41 (0-5,25)	0,47

Table 34. Number of steps/day vs. age and gender (N=156 patients with at least one measurement). The comparison with age is done by the Spearman correlation (r and p-value) and the comparison between gender is done by the Kruskal-Wallis test (median and IQR for each gender and p-value).

Although, there is no complete agreement on how many steps a day are optimal. As a useful guide, an adult achieving 10 000 or more daily steps is categorised as highly active, over 5000 but less than 10 000 as moderately active, and 5000 steps or below as inactive. The above results were confirmed with the data of participants showing at least 10 measurements of step with or without the 3 other parameters (Table 35). The Spanish and Slovenian participants had an important proportion of participants doing high physical activity. It seems that most of them either had this behaviour before the enrolment of the study or PERSIST helped them to put it into practice from the very beginning of the study without modifying it until the end of it (table 34). The same is true for moderate activity behaviour among the Latvian participants. Concerning the Belgian participants, although most of them never reached the high activity category as the Spanish and Slovenian participants, half of the Belgian survivors still acquired a moderate activity behaviour (table 35) which seems to be the result of the PERSIST solution (high number of Belgian participants with a positive slope, Table 34).

	UL (Latvian participants)	CHUL (Belgian participants)	Sergas (Spanish participants)	UKCM (Slovenian participants)
N (total 154)	36	40	36	42
highly active	8 (22%)	4 (10%)	16 (44%)	17 (40%)
moderate active	22 (61%)	22 (55%)	13 (36%)	18 (43%)
inactive	6 (17 %)	14 (35%)	6 (17%)	5 (12%)

Table 35. Participants with at least 10 measurements of steps (either with or without the other 3 parameters).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Heart rate.

Heart rate measurements were given more than once a day by patients. For the analysis, they were summarised per patient per day: mean, minimum and maximum HR/day.

The statistical characteristics of HR per day were listed. For each patient, the following information was gathered: Patient ID; Date of first HR measurement; Date of last HR measurement; Time between first and last HR measurement (days); Number of HR measurements; Mean, standard deviation (SD), Coefficient of variation (CV), Minimum, 1st quartile (Q1), Median, 3rd quartile (Q3) and maximum of the HR mean (respectively HR min and HR max).

The median number of days of HR measurements per patient was 118 (min=0, max=481), 156 patients had at least one day of HR measurement, and they were followed between 1 and 538 days. As reported in Table 36 comparison between the hospitals showed no significant differences in HR between hospitals.

	All	**CHU**	**SERGAS**	**UKCM**	**UL**	**P-value**
N	156	39	35	40	42	
Time between first and last HR measurement (days)	369 (113-518)	208 (54-505)	480 (174-513)	483 (274-525)	315 (71-492)	0,19
Median HR mean/day	79,1 (75,2-83,8)	79,3 (75,6-83,7)	80,4 (76,4-84,3)	77,7 (74,9-84,4)	79,3 (75,1-84,1)	0,58
Median HR min /day	57 (52-62)	56 (52-62)	57 (53-62)	57 (51,5-62)	56 (52-63)	0,96
Median HR max/day	110 (108-114)	109 (107-112)	111 (110-114)	110 (106-114)	110 (109-113)	0,20

Table 36. Heart rate/day vs. hospitals (N=156 patients with at least one measurement). Results are expressed as median (IQR), IQR=interquartile range Q1-Q3; P-value is the Kruskal-Wallis test

As reported in Table 37, no significant association was observed with age and gender.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	Age (years)		Gender		
	R	P-value	Female	Male	P-value
Time between first and last HR measurement (days)	0,076	0,35	386 (169-522)	344 (71,0-513)	0,50
Median HR mean/day	-0,001	0,99	79,3 (75,2-83,5)	78,3 (75,0-83,9)	0,80
Median HR min /day	0,028	0,73	57 (52-62)	55 (51,5-61,0)	0,28
Median HR max/day	-0,079	0,33	110 (108-113)	110 (108-116)	0,21

Table 37. Heart rate/day vs. age and gender (N=156 patients with at least one measurement). The comparison with age is done by the Spearman correlation (r and p-value) and the comparison between gender is done by the Kruskal-Wallis test (median and IQR for each gender and p-value)

Blood pressure

Systolic (SBP) and diastolic (DBP) blood pressure measurements were given more than once a day by patients. For the analysis, they were summarised per patient per day: mean, minimum and maximum SBP/day and mean, minimum and maximum DBP/day.

The statistical characteristics of SBP and DBP blood were gathered for each patient. The following information was gathered: Patient ID; Date of first BP measurement; Date of last BP measurement; Time between first and last BP measurement (days); Number of BP measurements; Mean, standard deviation (SD), Coefficient of variation (CV), Minimum, 1st quartile (Q1), Median, 3rd quartile (Q3) and maximum of the SBP mean (respectively SBP min, SBP max, DBP mean, DBP min, DBP max).

The median number of days of blood pressure measurements per patient is 117 (min=0, max=533), 161 patients have at least one day of blood pressure measurement, and they were followed between 1 and 571 days.

As reported in Table 38, comparison between the hospitals showed no significant differences in BP between hospitals.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	All	CHU	SERGAS	UKCM	UL	P-value
N	160	40	36	40	44	
Time between first and last BP measurement (days)	376 (110-523)	227 (63-488)	469 (213-526)	469 (224-523)	355 (66,5-524)	0,42
Median SBP mean/day	116 (112-120)	116 (110-120)	118 (114-122)	115 (111-120)	116 (110-120)	0,21
Median SBP min /day	113 (107-117)	113 (105-117)	114 (111-120)	113 (107-117)	112 (106-116)	0,11
Median SBP max/day	120 (116-126)	121 (116-126)	123 (116-129)	119 (114-123)	120 (116-126)	0,27
Median DBP mean/day	75 (72,5-78)	75,5 (72,0-78)	75,8 (73,5-79,5)	74,5 (72-77,7)	74,8 (72-77,8)	0,27
Median DBP min /day	72 (70-75)	72,5 (69-75)	73,5 (72-76,8)	72 (70-75)	72 (69-74,5)	0,073
Median DBP max/day	78 (74-82)	79 (75-82,3)	78,5 (75-84,8)	77 (74-79)	77,5 (75-81,5)	0,28

Table 38. Blood pressure/day vs. hospitals (N=156 patients with at least one measurement). Results are expressed as median (IQR), IQR=interquartile range Q1-Q3; P-value is the Kruskal-Wallis test

As reported in Table 39, no significant association was observed with age and gender.

	Age (years)	Gender			
	R	P-value	Female	Male	P-value
Time between first and last BP measurement (days)	0,010	0,90	432 (209-528)	243 (104-509)	0,072
Median SBP mean/day	-0,057	0,49	117 (113-120)	116 (110-120)	0,21
Median SBP min /day	-0,052	0,52	113 (108-117)	110 (104-117)	0,14
Median SBP max/day	-0,063	0,44	120 (116-124)	120 (114-127)	0,98
Median DBP mean/day	-0,013	0,88	75 (73-78)	75,4 (71,5-78)	0,55
Median DBP min /day	-0,027	0,74	72 (70-75)	71,5 (69-75)	0,43
Median DBP max/day	-0,034	0,68	78 (75-81)	79 (74-83)	0,83

Table 39. Blood pressure/day vs. age and gender (N=156 patients with at least one measurement). The comparison with age is done by the Spearman correlation (r and p-value) and the comparison between gender is done by the Kruskal-Wallis test (median and IQR for each gender and p-value).

Individual results (the number of steps, HR mean, SBP mean and DBP)

We reported for each patient, the mean and the coefficient of variation (CV) of all his measurements for the number of steps, the HR mean per day, the SBP mean per day and

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

the DBP mean per day. The characteristics of these 162 means and CV were calculated. As a reminder, the coefficient of variation is the ratio of the standard deviation to the mean and it is expressed in %. The higher it is, the greater the dispersion around the mean.

The mean (± SD) number of steps per day was 8412 ± 4566 and on average, the CV was 63,4 ± 28,1 %. The variation of the number of steps per day within a patient was quite huge.

If we consider the HR mean (± SD) per day, the overall mean was 79,8 ± 6,6 beats/min and the CV 9,2 ± 3,4%. For SBP mean and DBP mean, overall means were respectively 117 ± 6,9 mmHg and 76 ± 4,0 mmHg and overall CV were respectively 9,4 ± 3,3 % and 8,0 ± 2,8%. The variation of the HR and SBP/DBP measurements per day within a patient were quite small.

The individual coefficients of variation (CV) of the number of steps/day, mean HR/day, mean SBP and DBP /days are represented in Figure 14.

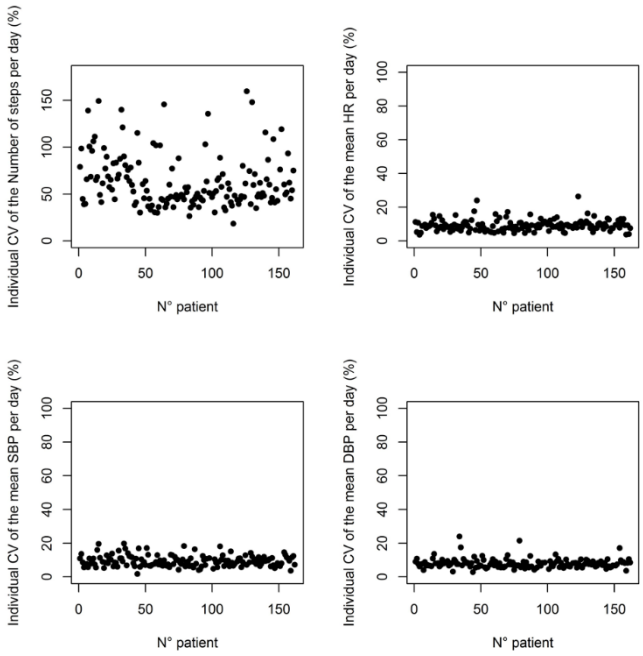

Figure 14 Individual coefficients of variation (CV) of the number of steps/day, mean HR/day, mean SBP and DBP /days.

Evolution with respect to time since first date

To study the evolution, we considered the 131 patients for whom there were at least 10 days of measurements of the number of steps, mean HR **and** mean SBP/DBP. This corresponds to 38009 days of measurements. The same proportion of patients was discarded from the analyses in the 4 hospitals (LIEGE: 20% UKCM: 16,7%, UL: 19,0%, SERGAS: 20,5%, p=0.98).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

The evolution of the 4 parameters (number of steps, mean HR per day, mean SBP per day and mean DBP per day) was plotted for 4 patients (see Figure 15, Figure 16, Figure 17 and Figure 18).

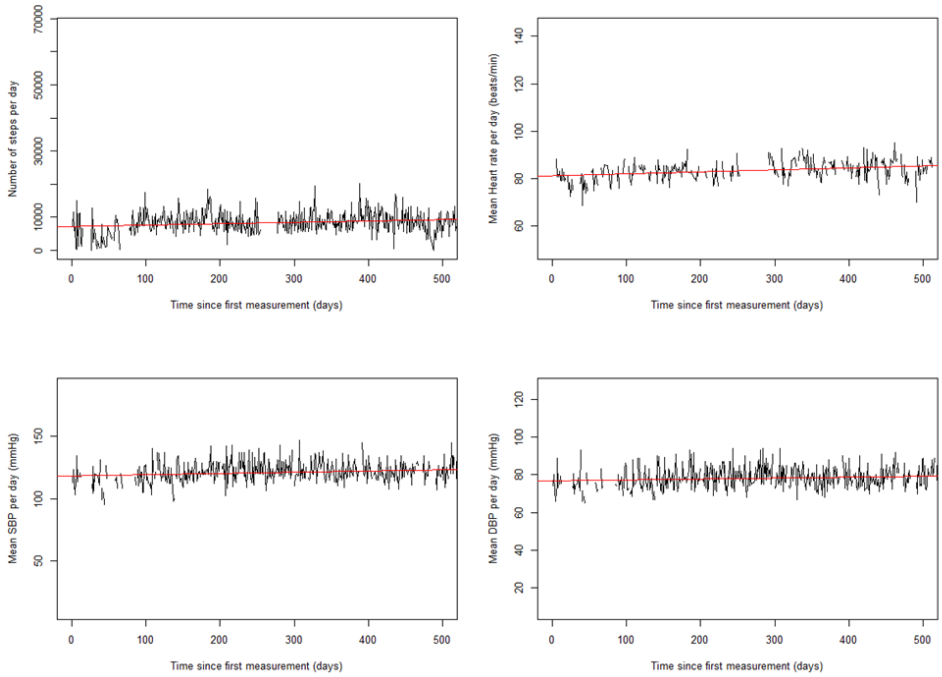

Figure 15 Evolution of number of steps/day, mean HR/day, mean SBP and DBP /days – PATIENT/LIEGE.

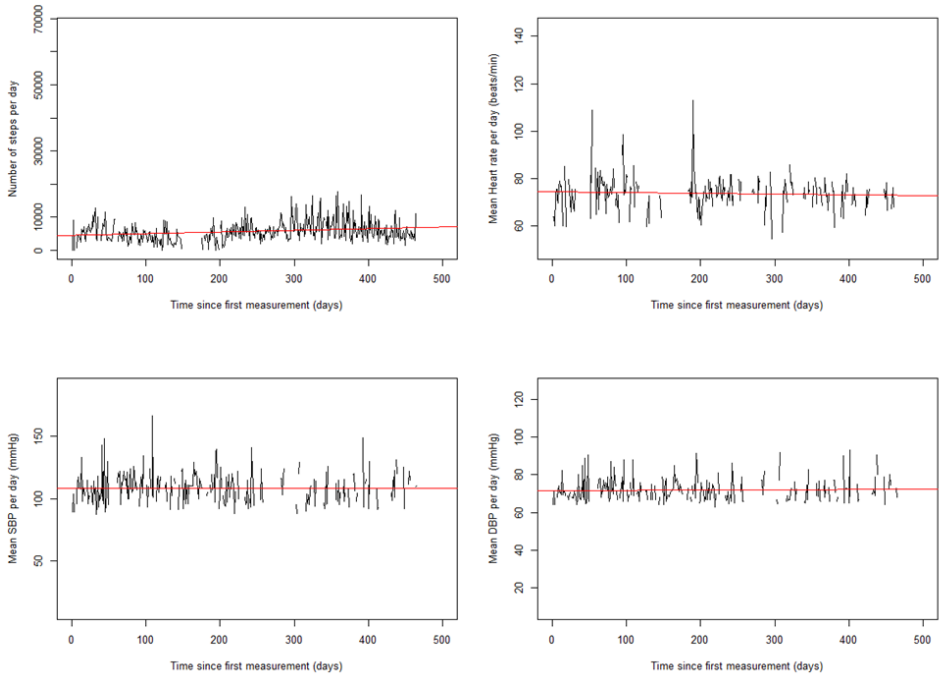

Figure 16 Evolution of number of steps/day, mean HR/day, mean SBP and DBP /days – PATIENT/UL.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

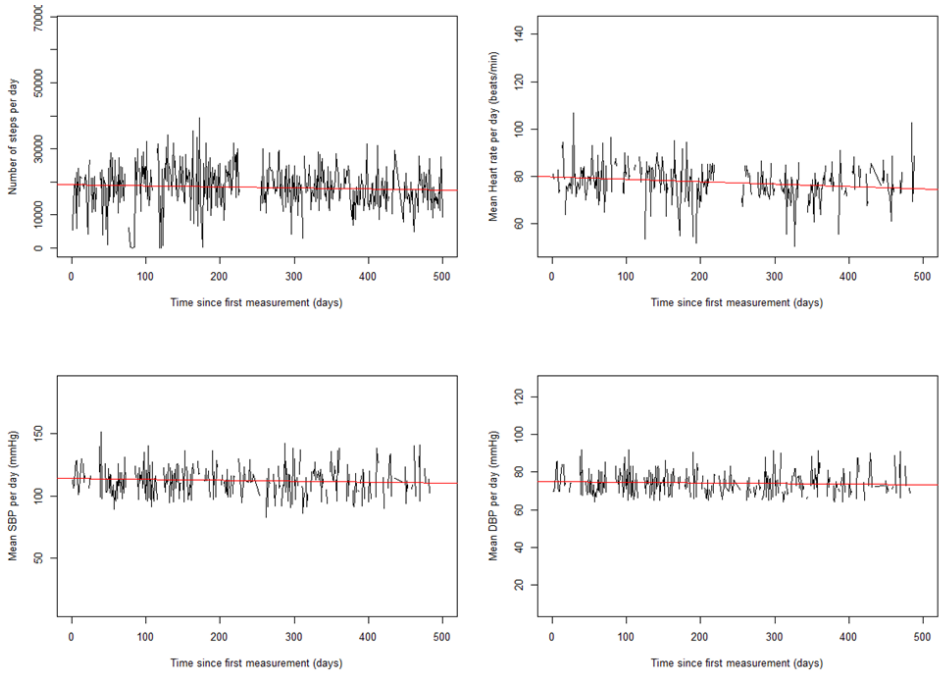

Figure 17 Evolution of number of steps/day, mean HR/day, mean SBP and DBP /days – PATIENT/UKCM.

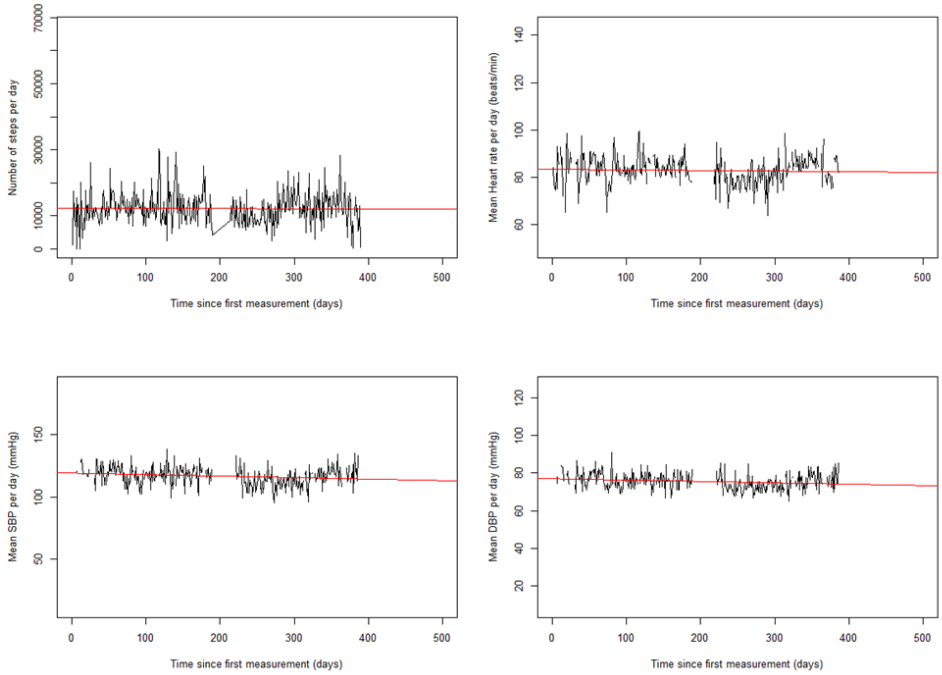

Figure 18 Evolution of number of steps/day, mean HR/day, mean SBP and DBP /days – PATIENT/SERGAS.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

For each of the 131 patients with at least 10 days of measurements of number of steps, HR and BP, slopes were calculated. The correlation between the slopes of the number of steps, of HR, SBP and DBP was calculated by the Spearman correlation (SeeTable 40). No association was observed between the slope of the number of steps and the slope of HR, SBP and DBP. Positive correlations were observed between HR, SBP and DBP slopes.

Variable X	Variable Y	r	P-value
Slope number of steps/day	Slope HR mean/day	0,11	0,91
Slope number of steps/day	Slope SBP mean/day	0,017	0,84
Slope number of steps/day	Slope DBP mean/day	-0,020	0,82
Slope HR mean/day	Slope SBP mean/day	0,38	<0,0001
Slope HR mean/day	Slope DBP mean/day	0,37	<0,0001
Slope SBP mean/day	Slope DBP mean/day	0,75	<0,0001

Table 40: Association between the evolutions of the number of steps, mean heart rate, mean Systolic and diastolic blood pressure (N=129 patients with at least 10 days of measurement). Spearman correlation coefficients between slopes (R) and p-values (p)

The slopes were then classified into 3 groups: “Significant positive slope”, “Significant negative slope”, “Non significant slope”.

The slope of the number of steps was significantly positive for 35 (26,7%) patients, while the slope of HR and SBP/DBP were significantly positive for respectively 20 (15,3%), 21 (16,0%) and 20 (15,3%) patients.

When testing the association between the classifications of the slopes, no significant association was observed between the number of steps and respectively HR, SBP and DBP (p=0,084, p=0,33, p=0,74).

Regular exercise can lower both HR and SBP and DBP. However, since walking in the context of PERSIST was considered mild (everyday activity), these decreases were obviously not observed.

Moreover, even if the volume of activity increases, the physiological parameters may not change. To achieve this objective, regular aerobic physical activity, gradually adapted in volume and intensity must be put into practice.

Thus, what is most important in the context of health promotion, is to improve the level of physical fitness rather than the level of Blood Pressure.

Nevertheless, measures of HR, SDP and DBP were important in PERSIST for rather medical reasons. Implemented alarms can notify the physician when a patient remotely exhibits values above their normal physiological parameters for more than 2-3 days.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 875406

➔ Conclusions:

The initial testing of the mHealth app has yielded promising results, indicating its potential to remotely monitor patients' physical activity and cardiovascular health, including heart rate and blood pressure. To encourage physical activity, the app can include notifications to remind patients to meet their daily physical activity goals. Additionally, if a patient's blood pressure or heart rate remains elevated over time, the app can suggest that they reach out to their doctor for further guidance.

7. Patient emotion wheel data from mHealth

The app uses emotions from the Plutchik Model or Wheel, created by psychologist Robert Plutchik. Patients can choose from different emotions, including the 8 basic emotions: joy, trust, fear, surprise, sadness, anticipation, anger, and disgust. The Emotion Wheel was created to help organise complex emotions so that people can gain clarity, identify, and label their emotions. Psychosocial distress is an important issue for cancer patients. Therefore, to evaluate the psychological state of patients, the emotion wheel was introduced into the mHealth app.

Patients were instructed to fill out the emotion's questionnaire in the mHealth app each day, and a reminder notification is sent to them. The Emotion Wheel was adapted from a previous version following patients' feedback (see Deliverable D6.2).

Emotion data analysis was done by UKCM statisticians and can be seen below:

Emotion data for all centres

The data does not include emotion measurements if:

- ➔ the number of measurements was less than 10 (looking at total N of measurements),
- ➔ the value was 0 (No value)
- ➔ no data. It could mean it was a mistake or data not transferred.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Patient mood	N	%	Average	Min. value	Max. value
Joy	12487	40,6%	93,4	2,03	100
good-bad	6973	22,7%	29,9	1,03	100
Fear	1993	6,5%	82,9	0,34	100
Optimism	1470	4,8%	100,0	100	100
Anger	1414	4,6%	84,5	5,04	100
Disgust	1206	3,9%	90,4	0,16	100
Trust	1102	3,6%	88,4	0,32	100
rested-tired	983	3,2%	55,5	1,58	100
Love	841	2,7%	100,0	100	100
Anticipation	734	2,4%	78,5	0,21	100
Sadness	393	1,3%	80,6	0,61	100
Disapproval	377	1,2%	100,0	100	100
Submission	374	1,2%	100,0	100	100
Surprise	150	0,5%	82,1	1,47	100
Remorse	136	0,4%	100,0	100	100
Awe	74	0,2%	100,0	100	100
Aggressiveness	38	0,1%	100,0	100	100
no pain-pain	16	0,1%	71,2	37,2	100
Contempt	14	0,0%	100,0	100	100
sum	30775	100,0			

Table 41. Emotion data for all centres in all period

Table 41 shows that the favourite emotion marked by patients was joy (40,6%) and afterwards they chose to mark how they feel on a scale from bad to good (22,7%). Contempt was the least chosen emotion with only 14 times selected in the entire study.

Conclusions: The results obtained show that the majority of patients had been in a good emotional state in the project period.

The data analysis from the app shows that patients most commonly chose joy as their favourite emotion, which could indicate a positive outlook on life and their cancer treatment. The low selection of negative emotions such as sadness and fear could also suggest that the app is helping patients to focus on positive emotions and manage negative ones. Overall, the use of the Emotion Wheel in the mHealth app appears to be a positive approach in improving the psychological well-being of cancer patients.

Patient mood analysis for UKCM

The data does not include emotion measurements if:

- ➔ the number of measurements was less than 10 (looking at total N of measurements),
- ➔ the value was 0 (no value)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

- ➔ patient left the trial (4 patients, less than 10 entries by patient) or
- ➔ no data. It could mean it was a mistake or data not transferred.

Number of patients=34 (9 males, 23 females)

Men = 10 (27,3%)

Female = 24 (72,7 %)

Average age (mean) = 56,2

Median = 56

Age: Men (average) = 61,8

Female (average) = 53,9

Patient mood	N	%	Average	Min. value	Max.value
Joy	3396	43,3%	95,1	2,03	100
Good-Bad	2569	32,8%	32,2	1,03	100
Optimism	449	5,7%	100,0	100	100
Fear	433	5,5%	81,3	14,8	100
Anger	363	4,6%	92,0	6	100
Love	183	2,3%	100,0	100	100
Anticipation	135	1,7%	80,8	16,8	100
rested-tired	91	1,2%	56,4	4,4	94,8
Trust	69	0,9%	80,3	0,32	100
Sadness	60	0,8%	64,5	1,25	100
Surprise	26	0,3%	83,5	1,47	100
Submission	19	0,2%	100,0	100	100
Disapproval	17	0,2%	100,0	100	100
Remorse	8	0,1%	100,0	100	100
Disgust	8	0,1%	100,0	100	100
Awe	5	0,1%	100,0	100	100
no pain-pain	4	0,1%	68,7	37,2	83,1
Aggressiveness	1	0,0%	100,0	100	100
sum	7836				

Table 42. Patient mood by number of measurements, min, max and average values.

The most mood measurements obtained are for Joy (43,3 %) and the one ranging from Good to -Bad (32,8 %). Other emotions have only a few measurements.

The highest intensity values were for Optimism, Love, Submission, Disapproval and also Remorse, Disgust, Awe. Nevertheless, these last three emotions have not many measurements and therefore the results can be questionable.

Statistical calculations were only performed with the emotions of JOY and GOOD-BAD given that they showed the highest numbers of measurements for those two emotions.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Conclusions: UKCM patients have been in a positive emotional state even above the joint average of all 4 centres.

Overall, analysis gives valuable information that could help healthcare providers to better understand patient mood and potentially improve patient experience.

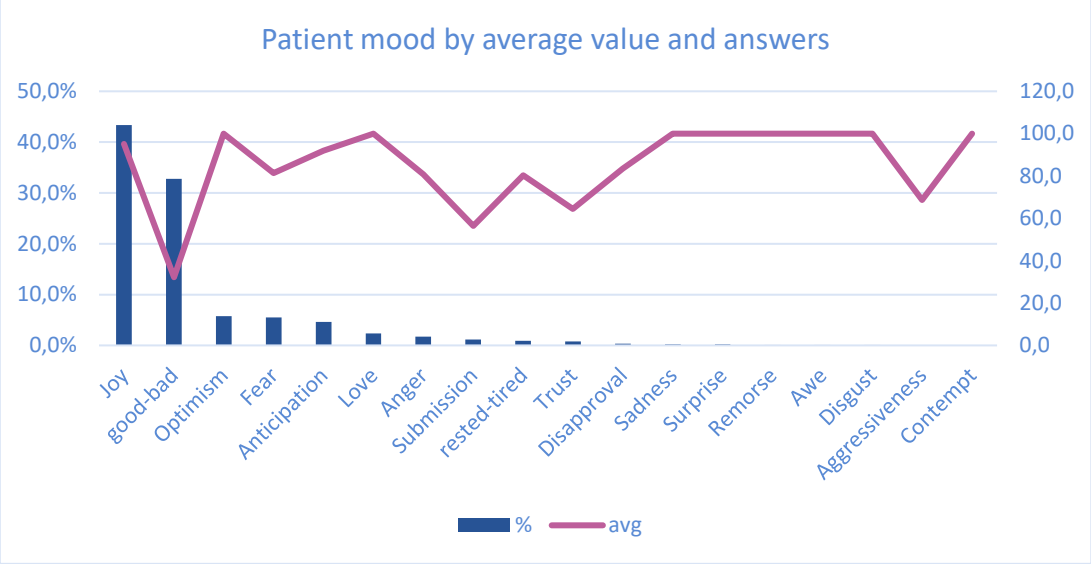

Figure 19: Patient mood by average value and answers (Y axis – percentage of patients; x – emotion intensity).

Patient mood	M (N)	M (%)	F (N)	F (%)	M (avg)	F (avg)
Joy	505	38,0%	2891	44,4%	92,4	95,5
good-bad	581	43,7%	1988	30,6%	25,5	34,2
Optimism	77	5,8%	372	5,7%	100,0	100,0
Fear	10	0,8%	423	6,5%	78,3	81,4
Anger	61	4,6%	302	4,6%	96,9	91,0
Love	18	1,4%	165	2,5%	100,0	100,0
Anticipation	28	2,1%	107	1,6%	86,4	79,3
rested-tired	19	1,4%	72	1,1%	36,1	61,8
Trust	18	1,4%	51	0,8%	86,2	78,3
Sadness		0,0%	60	0,9%		64,5
Surprise	2	0,2%	24	0,4%	47,8	86,4
Submission	3	0,2%	16	0,2%	100,0	100,0
Disapproval	2	0,2%	15	0,2%	100,0	100,0
Remorse	1	0,1%	7	0,1%	100,0	100,0
Disgust	3	0,2%	5	0,1%	100,0	100,0
Awe		0,0%	5	0,1%		100,0
no pain-pain		0,0%	4	0,1%		68,7
Aggressiveness	1	0,1%		0,0%	100,0	
Sum	1329	100,0 (17,0%)	6507	100,0 (83,0 %)		

Table 43: Patient mood by number of measurements and average values. By GENDER.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Table 43 shows that most women declare high values of Joy (average 95,5). Although men also report a very high value of joy (average 92,4), it remains lower than that of women. Interestingly, this difference is statistically significant ($p<0,01$) and can be explained somehow by a higher number of female (83%) than male (17%) participants .

When looking at the outcome of Good-Bad emotions, men's values are much lower than women's values (25,5 towards 34,2), this difference is also statistically significant ($p<0,05$).

When looking at the outcome of the intensity value of the Good-Bad emotions, men's values are much lower than women's values (25,5 versus 34,2), this difference is also statistically significant ($p<0,05$).

Conclusion: Women show slightly higher values of Joy than men, yet there is a bigger difference in Good-Bad emotions outcomes. Women have more complex emotional states than men. The emotions that were selected only by females are sadness, awe and no pain-pain.

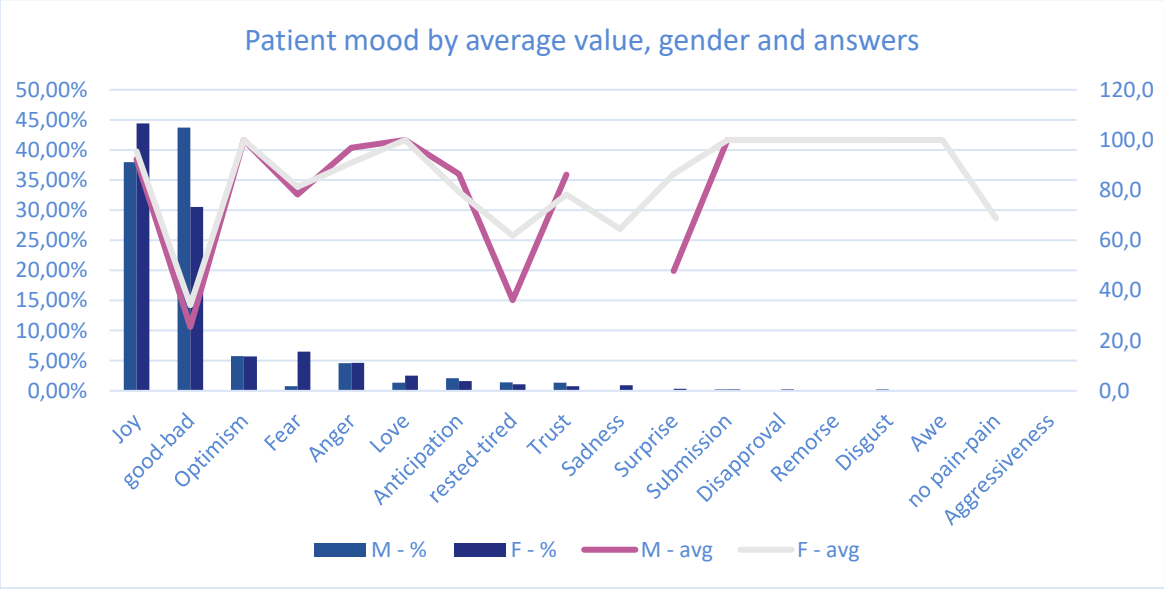

Figure 20: Patient mood by average value, gender and answers. (Y axis – percentage of patients; x – emotion intensity)

JOY	gender	N	Mean	Std. Deviation	Std. Error Mean
value quantity	1 male	505	92,362537	18,6452117	,8297010
	2 female	2891	95,539865	13,7380378	,2555057

Table 44. T-test calculation between gender and mood JOY.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

JOY - GENDER		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_ quantity	Equal variances assumed	79,452	,000	-4,521	3394	,000	-3,1773281	,7027830	-4,5552488	-1,7994073
	Equal variances Not assumed			-3,660	603,178	,000	-3,1773281	,8681515	-4,8822948	-1,4723613

Table 45. t-test for Equality of Means

There is a statistical significance (p<0,01) between gender and values of emotion JOY

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of gender	Independent samples Mann Whitney U test	**,004**	Reject the null hypothesis

Table 46. Mann Whitney test.

GOOD-BAD	gender	N	Mean	Std. Deviation	Std. Error Mean
value quantity	1 male	581	25,482135	17,8511556	,7405907
	2 female	1988	34,150916	24,3342585	,5457703

Table 47. T-test calculation between gender and mood GOOD-BAD

GOOD – BAD / GENDER		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value – quantity	Equal variances assumed	119,727	,000	-7,982	2567	,000	-8,6687802	1,0861054	-10,7985119	-6,5390484
	Equal variances not assumed			-9,423	1271,565	**,000**	-8,6687802	,9199673	-10,4736009	-6,8639594

Table 48. t-Test

There is a statistical significance (p<0,01) between gender and intensity values of emotion Good-Bad.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of gender	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 49: Mann Whitney test

Mood	2021			2022		
	N	**%**	**AVG**	**N**	**%**	**AVG**
Joy	1577	43,9%	89,4	1819	42,9%	100,0
good-bad	610	17,0%	29,4	1959	46,2%	33,1
Optimism	449	12,5%	100,0		0,0%	
Fear	280	7,8%	71,1	153	3,6%	100,0
Anger	148	4,1%	80,3	215	5,1%	100,0
Love	183	5,1%	100,0		0,0%	
Anticipation	124	3,5%	79,1	11	0,3%	100,0
rested-tired	27	0,8%	59,1	64	1,5%	55,3
Trust	68	1,9%	80,1	1	0,0%	100,0
Sadness	51	1,4%	58,2	9	0,2%	100,0
Surprise	23	0,6%	81,3	3	0,1%	100,0
Submission	19	0,5%	100,0		0,0%	
Disapproval	17	0,5%	100,0		0,0%	
Remorse	8	0,2%	100,0		0,0%	
Disgust	2	0,1%	100,0	6	0,1%	100,0
Awe	5	0,1%	100,0		0,0%	
no pain-pain	2	0,1%	82,5	2	0,0%	54,8
Aggressiveness	1	0,0%	100,0		0,0%	
sum	**3594**	**100,0%**		**4242**	**100,0%**	

Table 50. ***Patient mood by number of measurements and average values-by year****.*

The same result is observed when the data is broken down by year Thus, most measurements for both years come from the emotions of Joy (43,9 % in 2021 and 42,9 % in 2022) and Good-Bad (17,0 % in 2021 and 46,2 % in 2022) emotions. There is also a statistically significant difference ($p<0,01$) between the years for the two emotions. ($p<0,01$).

Note: It is possible that some emotions were not marked by participants, leading to their absence from the statistical analysis in 2022. However, the exclusion of certain emotions from the measurement process of the emotion wheel also contributed to their non-inclusion in the analysis.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

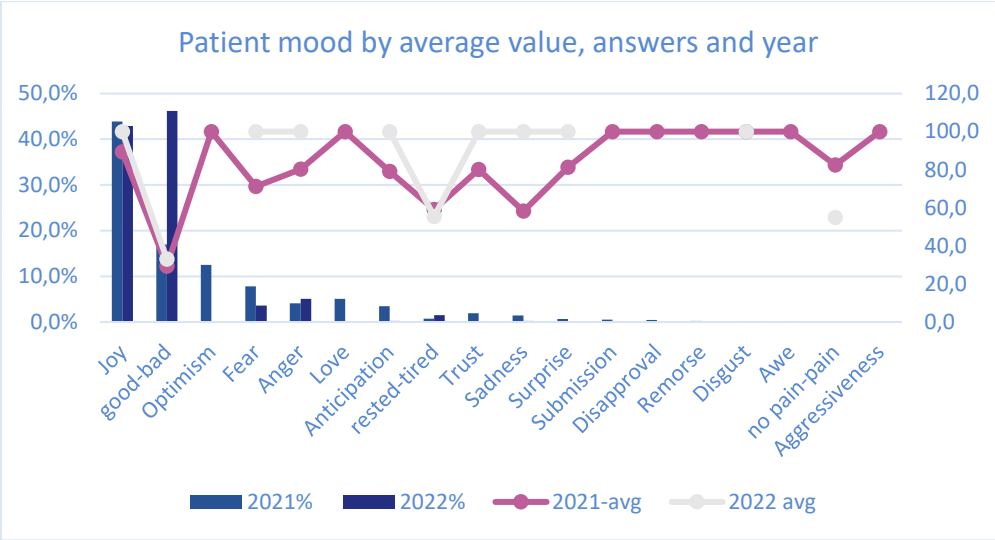

Figure 21: Patient mood by average intensity value, answers (emotions) and years.

JOY	year	N	Mean	Std. Deviation	Std. Error Mean
value quantity	2021	1577	89,377825	19,9886738	,5033478
	2022	1819	100,000000	0E-7	0E-7

Table 51. T-test calculation between data year and mood JOY

JOY - YEAR		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value – quantity	Equal variances assumed	3097,703	,000	-22,665	3394	**,000**	-10,6221750	,4686599	-11,5410593	-9,7032908
	Equal variances not assumed			-21,103	1576,000	**,000**	-10,6221750	,5033478	-11,6094767	-9,6348733

Table 52. t-test.

There is a statistically significant difference (p<0,01) between data year and values of emotion JOY.

Null hypothesis	**test**	sig	**decision**
The distribution of value quantity is the same across categories of year	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 53 Mann Whitney.

GOOD-BAD	year	N	Mean	Std. Deviation	Std. Error Mean
value quantity	2021	610	29,402086	23,0257737	,9322863
	2022	1959	33,058636	23,3344369	,5272057

Table 54: T-test calculation between data year and mood GOOD-BAD

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

GOOD-BAD / YEAR		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value – quantity	Equal variances assumed	5,872	,015	-3,390	2567	,001	-3,6565503	1,0785471	-5,7714610	-1,5416397
	Equal variances not assumed			-3,414	1028,080	,001	-3,6565503	1,0710293	-5,7582033	-1,5548973

Table 55: t-Test.

There is a statistically significant difference (p<0,01) between both years regarding the intensity values of the GOOD-BAD emotion.

Null hypothesis	test	sig	decision
The distribution of value quantity is the same across categories of year	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 56: Mann Whitney test.

mood	41-55 age			56-73 age		
	N	**%**	**avg**	**N**	**%**	**avg**
Joy	2125	53,9%	93,7	1271	32,7%	97,4
good-bad	961	24,4%	31,7	1608	41,3%	32,5
Optimism	219	5,5%	100,0	230	5,9%	100,0
Fear	141	3,6%	88,2	292	7,5%	78,0
Anger	99	2,5%	87,1	264	6,8%	93,8
Love	151	3,8%	100,0	32	0,8%	100,0
Anticipation	50	1,3%	77,1	85	2,2%	82,9
rested-tired	59	1,5%	63,9	32	0,8%	42,7
Trust	33	0,8%	72,6	36	0,9%	87,4
Sadness	56	1,4%	66,4	4	0,1%	37,8
Surprise	16	0,4%	79,6	10	0,3%	89,6
Submission	12	0,3%	100,0	7	0,2%	100,0
Disapproval	11	0,3%	100,0	6	0,2%	100,0
Remorse	1	0,0%	100,0	7	0,2%	100,0
Disgust	4	0,1%	100,0	4	0,1%	100,0
Awe	5	0,1%	100,0		0,0%	
no pain-pain	3	0,1%	63,9	1	0,0%	83,1
Aggressiveness		0,0%		1	0,0%	100,0
Sum	**3946**	**100,0%**		**3890**	**100,0%**	

Table 57. Patient mood by number of measurements and average values. By AGE GROUP

The median age was 56 and the average age was 56,2. Patients were divided into two age groups; 41-55 y/o and 56-73 y/o. Once again, the results show that most measurements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

correspond to the emotions of Joy and Good-Bad. The Joy emotion has more measurements in younger patients than in older patients (53,9 % against 32,7 %). There is also a statistically significant difference ($p<0,01$) between the intensity of Joy values in the two age groups.

As for Good-Bad emotions, more measurements are observed in older than in younger patients (41,3 % against 24,4 %). But the intensity values for Good-Bad mood are almost the same in both groups. Statistically significant difference ($p<0,05$) is observed between age groups (Mann Whitney test for nonparametric distribution). No significant difference is observed when comparing the intensity value of this mood.

Conclusion: Younger group showed significantly higher measurements of Joy then older group, as there were more measurements in younger patients' group.

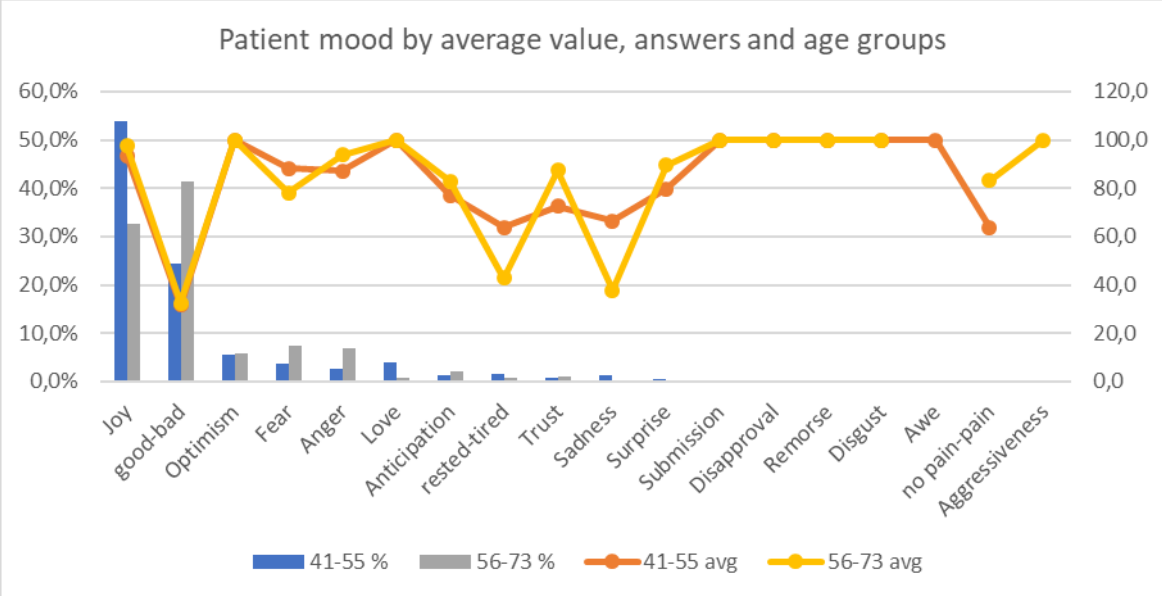

Figure 22: Patient mood by average value,% of answers and age groups.

JOY	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	41-55	2125	93,653147	15,8639358	,3441371
	56-73	1271	97,431858	11,8753917	,3331004

Table 58. T-test calculation between age groups and mood JOY

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

JOY – AGE GROUPS		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value – quantity	Equal variances assumed	,000	-7,349	3394	**,000**	-3,7787113	,5141790	-4,7868432	-2,7705794	,000
	Equal variances not assumed		-7,890	3228,631	**,000**	-3,7787113	,4789428	-4,7177740	-2,8396486	

Table 59. t-Test

There is a statistically significant difference (p<0,01) between age groups and intensity values of the Joy emotion.

Null hypothesis	**test**	sig	**decision**
The distribution of value quantity is the same across categories of age group	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 60: Mann Whitney test

GOOD-BAD	year	N	Mean	Std. Deviation	Std. Error Mean
value quantity	41-55	961	31,740123	27,9270084	,9008712
	56-73	1608	32,459504	20,0527221	,5000694

Table 61. T-test calculation between age groups and mood GOOD-BAD

Null hypothesis	**test**	sig	**decision**
The distribution of value_quantity is the same across categories of age group	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 62.Mann Whitney test

Patient moods by number of measurements, age and average value for each patient – are gathered in separate files.

Patient mood analysis for CHU

The data does not include emotion measurements if:

- ➔ the number of measurements was less than 10 (looking at total N of measurements),
- ➔ the value was 0 (no value),
- ➔ no data, It could mean it was a mistake or data not transferred.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Patients who left the study almost at the end were not removed for the analysis.

Number of patients = 38 (10m, 28f)

Men = 10 (26,3 %)

Female = 28 (73,7 %)

Average age (mean) = 55,2

Median = 53,5

Age: Men (average) = 63,2

Female (average) = 52,3

Patient mood	N	%	Average	Min. value	Max. value
Joy	2478	37,8%	87,8	2,85	100
good-bad	1439	21,9%	34,9	1,18	100
Anger	590	9,0%	77,6	5,04	100
Fear	414	6,3%	79,4	1,22	100
Trust	340	5,2%	87,6	9,43	100
Disgust	213	3,2%	89,4	0,22	100
Optimism	201	3,1%	100,0	100	100
Disapproval	166	2,5%	100,0	100	100
Anticipation	161	2,5%	74,6	0,21	100
Sadness	158	2,4%	85,4	0,61	100
rested-tired	137	2,1%	58,5	2,18	100
Love	120	1,8%	100,0	100	100
Submission	60	0,9%	100,0	100	100
Remorse	46	0,7%	100,0	100	100
Surprise	17	0,3%	82,0	37,9	100
Aggressiveness	13	0,2%	100,0	100	100
Contempt	8	0,1%	100,0	100	100
no pain-pain	2	0,0%	75,0	52,1	97,8
Awe	1	0,0%	100,0	100	100
sum	**6564**				

Table 63. Patient mood by number of measurements, min, max and average values.

The most mood measurements are for joy (37,8 %), good-bad (21,9 %) and anger (9,0 %). Other emotions have only a few measurements.

The highest measures values have Optimism, Disapproval, Love, Submission, Remorse, Aggressiveness, Contempt and Awe, but also the last two have not so many measurements (2 and 1) so the results can be questionable.

The two highest numbers of measurements are JOY and GOOD-BAD emotions so statistical calculations are done only for those two emotions.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

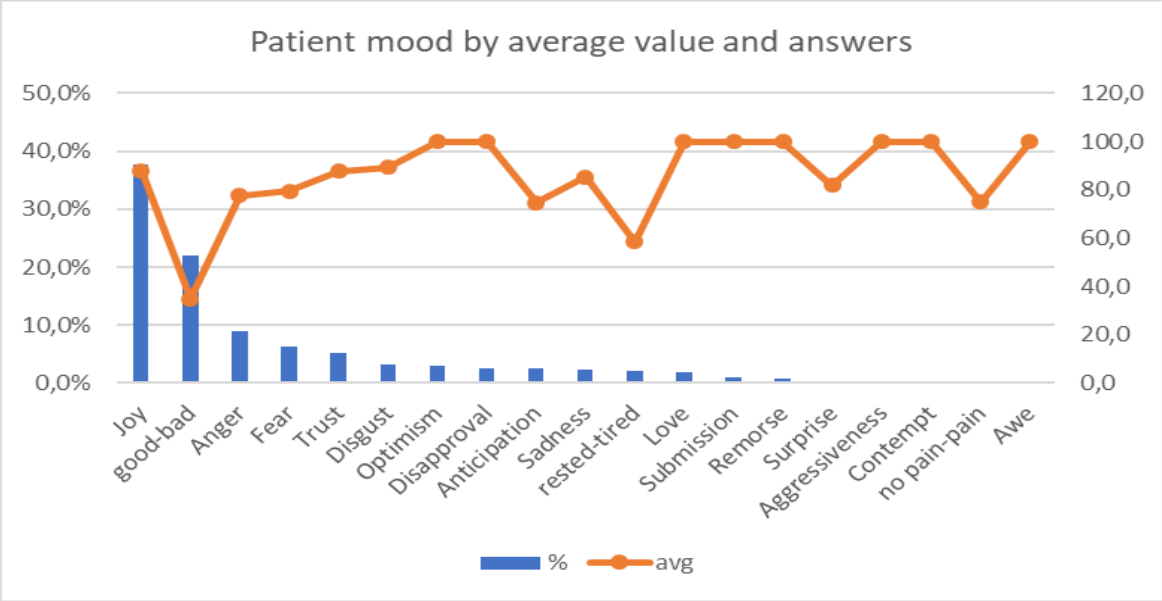

Figure 24: Patient mood by average of intensity value and answers (emotions).

Patient mood	M (N)	M (%)	F (N)	F (%)	M (avg)	F (avg)
Joy	1282	59,3%	1191	27,2%	92,0	83,3
Good-Bad	374	17,3%	1064	24,3%	21,1	39,7
Anger	89	4,1%	493	11,3%	78,8	77,0
Fear	63	2,9%	349	8,0%	74,6	80,1
Trust	157	7,3%	183	4,2%	86,1	88,9
Disgust	33	1,5%	175	4,0%	87,6	89,5
Optimism	9	0,4%	192	4,4%	100,0	100,0
Disapproval	1	0,0%	165	3,8%	100,0	100,0
Anticipation	17	0,8%	144	3,3%	64,1	75,8
Sadness	30	1,4%	127	2,9%	94,1	83,2
rested-tired	24	1,1%	112	2,6%	48,6	60,7
Love	63	2,9%	57	1,3%	100,0	100,0
Submission	7	0,3%	53	1,2%	100,0	100,0
Remorse	2	0,1%	44	1,0%	100,0	100,0
Surprise	9	0,4%	4	0,1%	84,9	57,2
Aggressiveness	1	0,0%	12	0,3%	100,0	100,0
Contempt		0,0%	8	0,2%		100,0
no pain-pain	1	0,0%	1	0,0%	97,8	52,1
Awe		0,0%	1	0,0%		100,0
Sum	**2162**	**100,0 (32,9%)**	**4375**	**100,0 (66,7 %)**		

Table 64. Patient mood by number of measurements and average values. By GENDER

Most men (59,3%) selected high intensity values of Joy emotion (average 92,0). Although fewer women (27.2%) reported the emotion Joy, their intensity values were also high (average 83.3). Interestingly, this intensity value difference is statistically significant ($p<0,01$) and can be explained somehow by a higher number of female (66,7%) than male (32,9%) participants.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Looking at the outcome of the intensity value of the Good-Bad emotion, men's values are much lower than women's values (21,1 versus 39,7), this difference is also statistically significant (p<0,01), but it should be taken into consideration that compared to men, more women selected this emotion (1064 measurements of women against 374 of men).

Conclusions: The results show that in the case of CHU more men selected Joy emotion. On the other hand, the Good-Bad value is much lower in the male population.

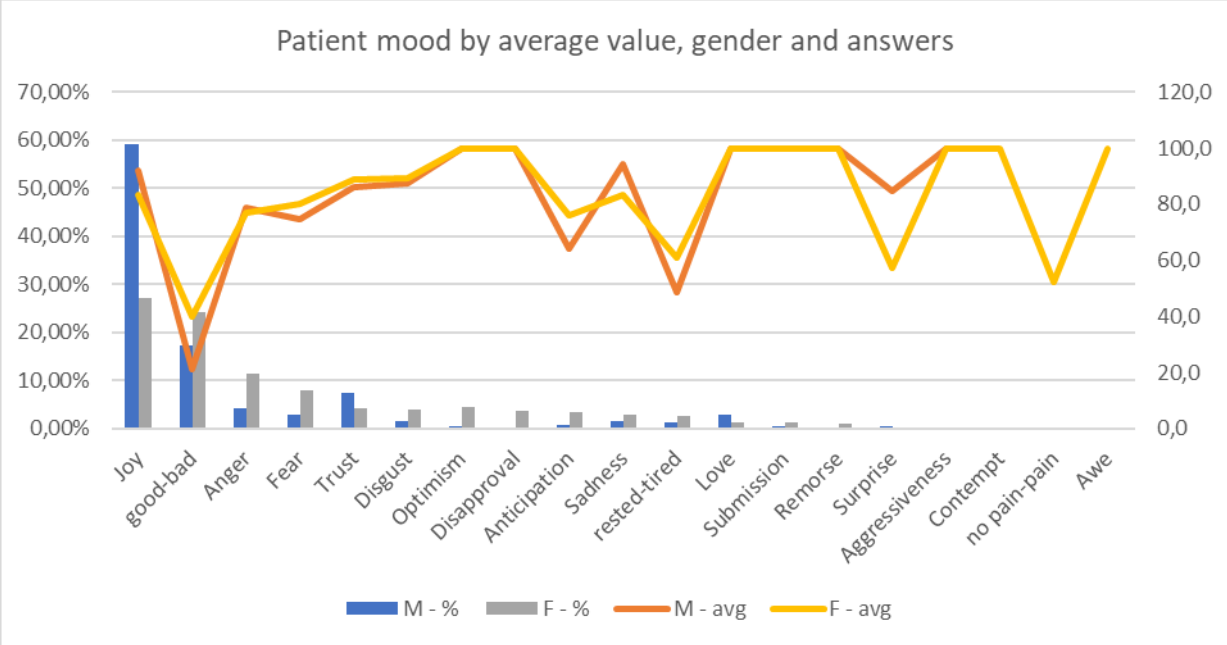

Figure 25: Patient mood by average value, gender and answers.

JOY	gender	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	F	1282	91,995615	17,1086025	,4778268
	M	1191	83,318887	24,8809562	,7209600

Table 65 T-test calculation between gender and mood JOY

JOY - GENDER		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	380,003	,000	10,165	2471	**,000**	8,6767274	,8536060	7,0028704	10,3505844
	Equal variances not assumed			10,032	2090,362	**,000**	8,6767274	,8649288	6,9805160	10,3729388

Table 66 t-Test

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

There is a statistical significance (p<0,01) between gender and values of emotion JOY

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of gender	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 67: Mann Whitney test

GOOD-BAD	gender	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	F	374	21,115135	7,9686530	,4120492
	M	1064	39,689061	20,9583305	,6425188

Table 68 T-test calculation between gender and mood GOOD-BAD

GOOD – BAD / GENDER		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	313,872	,000	-16,716	1436	**,000**	-18,5739267	1,1111270	-20,7535326	-16,3943207
	Equal variances not assumed			-24,334	1428,549	**,000**	-18,5739267	,7632922	-20,0712205	-17,0766328

Table 69 t-Test

There is a statistical significance (p<0,01) between gender and values of emotion GOOD-BAD.

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of gender	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 70: Mann Whitney test

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Mood	2021			2022		
	N	**%**	**AVG**	**N**	**%**	**AVG**
Joy	1647	40,1%	81,7	831	33,9%	100,0
Good-Bad	492	12,0%	35,2	947	38,6%	34,7
Anger	387	9,4%	65,9	203	8,3%	100,0
Fear	227	5,5%	62,3	187	7,6%	100,0
Trust	239	5,8%	82,4	101	4,1%	100,0
Disgust	110	2,7%	79,5	103	4,2%	100,0
Optimism	201	4,9%	100,0		0,0%	
Disapproval	166	4,0%	100,0		0,0%	
Anticipation	161	3,9%	74,6		0,0%	
Sadness	146	3,6%	84,2	12	0,5%	100,0
rested-tired	75	1,8%	61,0	62	2,5%	55,5
Love	120	2,9%	100,0		0,0%	
Submission	60	1,5%	100,0		0,0%	
Remorse	46	1,1%	100,0		0,0%	
Surprise	12	0,3%	74,4	5	0,2%	100,0
Aggressiveness	13	0,3%	100,0		0,0%	
Contempt	8	0,2%	100,0		0,0%	
no pain-pain	1	0,0%	52,1	1	0,0%	97,8
Awe	1	0,0%	100,0		0,0%	
sum	**4112**	**100,0%**		**2452**	**100,0%**	

Table 71: Patient mood by number of measurements and average values-by DATA YEAR

The same result is observed when the data is broken down by year. Thus, most measurements for both years come from the emotions of Joy (40,1 % in 2021 and 33,9 % in 2022) and Good-bad (12,0 % in 2021 and 38,6 % in 2022) emotions. There is also a statistically significant difference ($p<0,01$) between the years for the two emotions ($p>0,05$). In total, 62,6 % of all measurements happened in 2021 and only 37,4 % measurements in 2022.

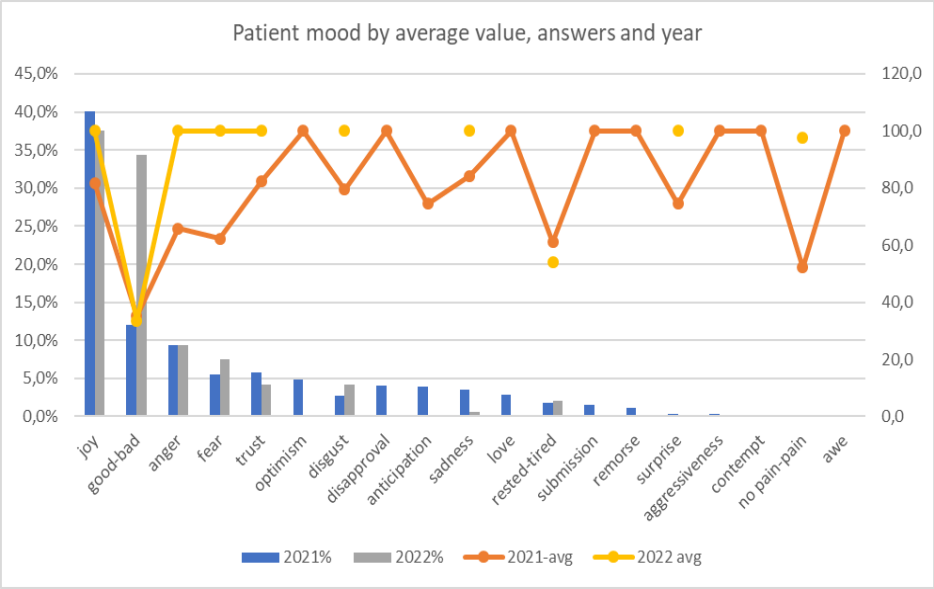

Figure 26: Patient mood by average value, answers and year.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

JOY	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	2021	1647	81,7068	24,3260517	,5994112
	2022	831	100,0000	0E-7	0E-7

Table 73 T-test calculation between data year and mood JOY

JOY - YEAR		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	3117,897	,000	-21,676	2476	**,000**	-18,2931557	,8439458	-19,9480680	-16,6382433
	Equal variances not assumed			-30,519	1646,000	**,000**	-18,2931557	,5994112	-19,4688445	-17,1174669

Table 74 t-Test

There is a statistically significant difference (p<0,01) between data year and values of emotion JOY.

GOOD-BAD	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	2021	492	35,183270	17,6485381	,7956575
	2022	947	34,728536	21,4300691	,6963837

Table 75: T-test calculation between data year and mood GOOD-BAD

GOOD-BAD / YEAR		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	15,254	,000	,405	1437	**,686**	,4547339	1,1235808	-1,7493004	2,6587682
	Equal variances not assumed			,430	1173,846	**,667**	,4547339	1,0573652	-1,6198028	2,5292706

Table 76: t-Test

There is no statistically significant difference (p>0,05) between data year and values of emotion GOOD-BAD.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

mood	23-54 age			55-75 age		
	N	**%**	**avg**	**N**	**%**	**avg**
Joy	634	21,3%	78,9	1839	51,7%	90,9
good-bad	680	22,8%	41,5	758	21,3%	28,9
Anger	445	14,9%	77,1	137	3,9%	77,9
Fear	321	10,8%	83,8	91	2,6%	63,1
Trust	122	4,1%	84,4	218	6,1%	89,4
Disgust	153	5,1%	93,4	55	1,5%	77,4
Optimism	150	5,0%	100,0	51	1,4%	100,0
Disapproval	105	3,5%	100,0	61	1,7%	100,0
Anticipation	21	0,7%	56,0	140	3,9%	77,4
Sadness	106	3,6%	85,0	51	1,4%	85,9
rested-tired	98	3,3%	60,2	38	1,1%	54,6
Love	55	1,8%	100,0	65	1,8%	100,0
Submission	48	1,6%	100,0	12	0,3%	100,0
Remorse	32	1,1%	100,0	14	0,4%	100,0
Surprise	6	0,2%	71,5	7	0,2%	80,6
Aggressiveness	1	0,0%	100,0	12	0,3%	100,0
Contempt	3	0,1%	100,0	5	0,1%	100,0
no pain-pain		0,0%		2	0,1%	75,0
Awe		0,0%		1	0,0%	100,0
Sum	**2980**	**100,0%**		**3557**	**100,0%**	

Table 77: Patient mood by number of measurements and average values. By AGE GROUP

The median age was 53,5 and average age was 55,2. Patients were divided into two age groups; 23-54 y/o and 55-75 y/o. Once again, the results show that most measurements correspond to the emotions of Joy and Good-Bad emotions. The Joy emotion has been selected more frequently by older (51,7 %) than by younger (21,3%) patients). In addition, more joy emotion measurements have been observed in the older (1839) than in the younger population (634). There is also a statistically significant difference ($p<0,01$) between the intensity of Joy values in the two age groups.

The values of measurement in Good-Bad mood are very different (41,5 against 28,9) and there is a statistical significance ($p<0,01$) between age groups.

Conclusion: In this case the older group showed significantly higher measurements of Joy then the younger group.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

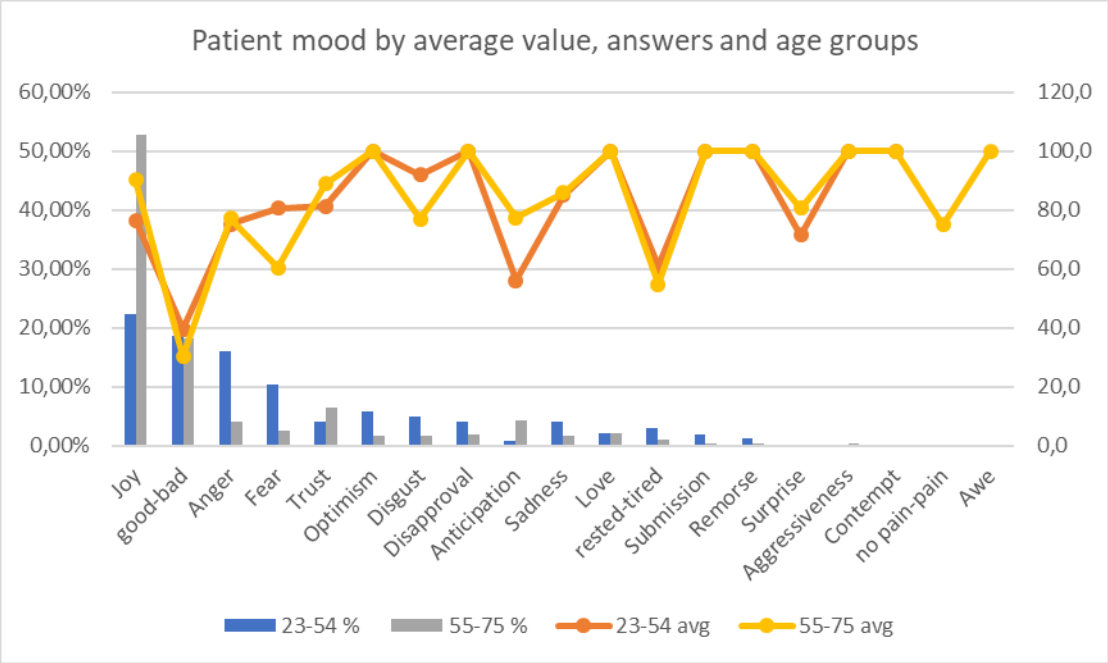

Figure 27: Patient mood by average value, answers (emotions) and age groups.

JOY	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	23-54	634	78,786389	27,0267053	1,00733676
	55-75	1839	90,899153	18,4786081	,4309019

Table 78: T-test calculation between age groups and mood JOY

JOY – AGE GROUPS		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	405,170	,000	-12,430	2471	**,000**	-12,02 27640	,9672 703	-13,91 95079	-10,12 60200
	Equal variances not assumed			-10,395	854,904	**,000**	-12,02 27640	1,156 6306	-14,29 29666	-9,752 5613

Table 79: t-Test

There is a statistical significance (p<0,01) between age groups and values of emotion JOY.

Null hypothesis	**test**	sig	**decision**
The distribution of value quantity is the same across categories of age groups	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 80: Mann Whitney test

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

GOOD-BAD	year	N	Mean	Std. Deviation	Std. Error Mean
value quantity	23-54	680	41,488950	24,3280510	,9329382
	55-75	758	28,909941	12,9776188	,4713683

Table 81:T-test calculation between age groups and mood GOOD-BAD

GOOD-BAD / AGE GROUP		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value – quantity	Equal variances assumed	296,249	,000	12,404	1436	**,000**	12,579008 9	1,0141203	10,585693 0	14,568234 8
	Equal variances not assumed			12,034	1012,832	**,000**	12,579008 9	1,0452568	10,527887 3	14,630130 5

Table 82 t-Test

There is a statistical significance (p<0,01) between age groups and values of emotion GOOD – BAD.

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of age groups	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 83: Mann Whitney test

Patient moods by number of measurements, age, collaboration and average value for each patient – are gathered in separate files.

Patient mood analysis for UL

The data does not include emotion measurements if:

- ➔ the number of measurements was less than 10 (looking at total N of measurements), (5 patients)
- ➔ the value was 0 (no value),
- ➔ no data, it could mean it was a mistake or data not transferred.

We did not remove the patients if they left, because some left almost at the end.

Number of patients=38

Men = 5 (13,2 %)

Female = 33 (86,8 %)

Average age (mean) = 53,1

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Median = 52,5

Age: Men (average) = 62,2

Female (average) = 51,8

Patient mood	N	%	Average	Min. value	Max. value
Joy	1891	27,3%	93,6	4,14	100
good-bad	1738	25,1%	28,4	1,24	100
Optimism	711	10,3%	100,0	100	100
Fear	592	8,6%	91,7	0,34	100
Anticipation	411	5,9%	80,4	1,11	100
Love	333	4,8%	100,0	100	100
Anger	252	3,6%	88,8	20,5	100
Submission	184	2,7%	100,0	100	100
rested-tired	183	2,6%	66,2	1,67	100
Trust	145	2,1%	80,6	7,44	100
Disapproval	117	1,7%	100,0	100	100
Sadness	81	1,2%	80,4	12,5	100
Surprise	76	1,1%	79,0	10,2	100
Remorse	65	0,9%	100,0	100	100
Awe	59	0,9%	100,0	100	100
Disgust	52	0,8%	90,4	0,16	100
Aggressiveness	22	0,3%	100,0	100	100
no pain-pain	5	0,1%	71,7	58	84,7
Contempt	5	0,1%	100,0	100	100
sum	**6922**				

Table 84: Patient mood by number of measurements, min, max and average values.

The most mood measurements are for joy (27,3 %), the one ranging from good-bad (25,1 %) and optimism (10,3 %). Other emotions have only a few measurements.

The highest intensity values were for Optimism, Love, Submission, Disapproval, Remorse, Awe, Aggressiveness and Contempt. Nevertheless, these last two emotions have not so many measurements (22 and 5) and therefore the results can be questionable.

Statistical calculations were only performed with the emotions of JOY and GOOD-BAD given that they showed the highest numbers of measurements for those two emotions.

Conclusions: The results obtained show that the majority of patients had been in a good emotional state in the project period.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

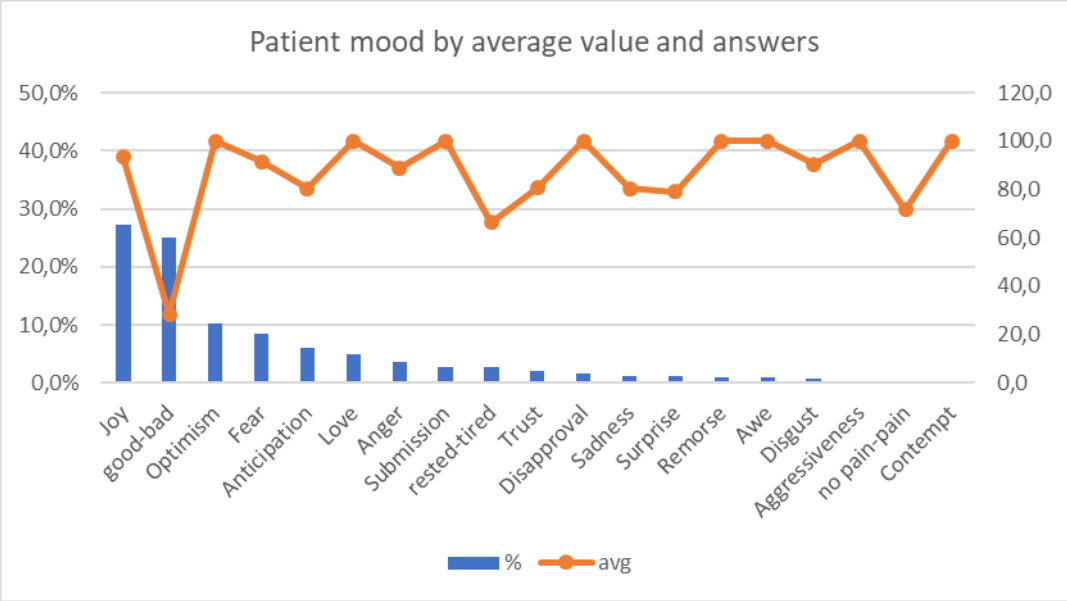

Figure 28: Patient mood by average value and answers.

Patient mood	M (N)	M (%)	F (N)	F (%)	M (avg)	F (avg)
Joy	535	43,5%	1356	23,8%	95,3	93,0
good-bad	358	29,1%	1380	24,2%	27,1	28,8
Optimism	46	3,7%	665	11,7%	100,0	100,0
Fear	67	5,4%	525	9,2%	94,3	91,3
Anticipation	74	6,0%	337	5,9%	79,5	80,6
Love	46	3,7%	287	5,0%	100,0	100,0
Anger	13	1,1%	239	4,2%	87,9	88,9
Submission	23	1,9%	161	2,8%	100,0	100,0
rested-tired	1	0,1%	182	3,2%	32,0	66,4
Trust	14	1,1%	131	2,3%	75,1	81,2
Disapproval	7	0,6%	110	1,9%	100,0	100,0
Sadness	12	1,0%	69	1,2%	88,0	79,0
Surprise	6	0,5%	70	1,2%	71,6	79,7
Remorse	10	0,8%	55	1,0%	100,0	100,0
Awe	8	0,7%	51	0,9%	100,0	100,0
Disgust	5	0,4%	47	0,8%	80,0	91,5
Aggressiveness	5	0,4%	17	0,3%	100,0	100,0
Contempt		0,0%	5	0,1%		100,0
no pain-pain		0,0%	5	0,1%		71,7
Sum	**1230**	**100,0 (17,7%)**	**5692**	**100,0 (82,2 %)**		

Table 85 Patient mood by number of measurements and average values. By GENDER

Table 82 shows that most men declare high values of Joy (average 95,3), women are also reporting high levels (average 93,0), and this difference is not statistically significant (p>0,05) but it also could be explained by a higher number of female measurements (82,2 %) than male (17,7 %).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

When looking at the outcome of Good-Bad emotions men`s values are almost the same as women`s values (27,1 against 28,8), this difference is also statistical significant (p<0,05), but it should be taken into consideration that more women measured this emotion (1091 measurements for women against 358 men).

Conclusions: The Good-Bad and Joy emotions are almost the same in male and female population.

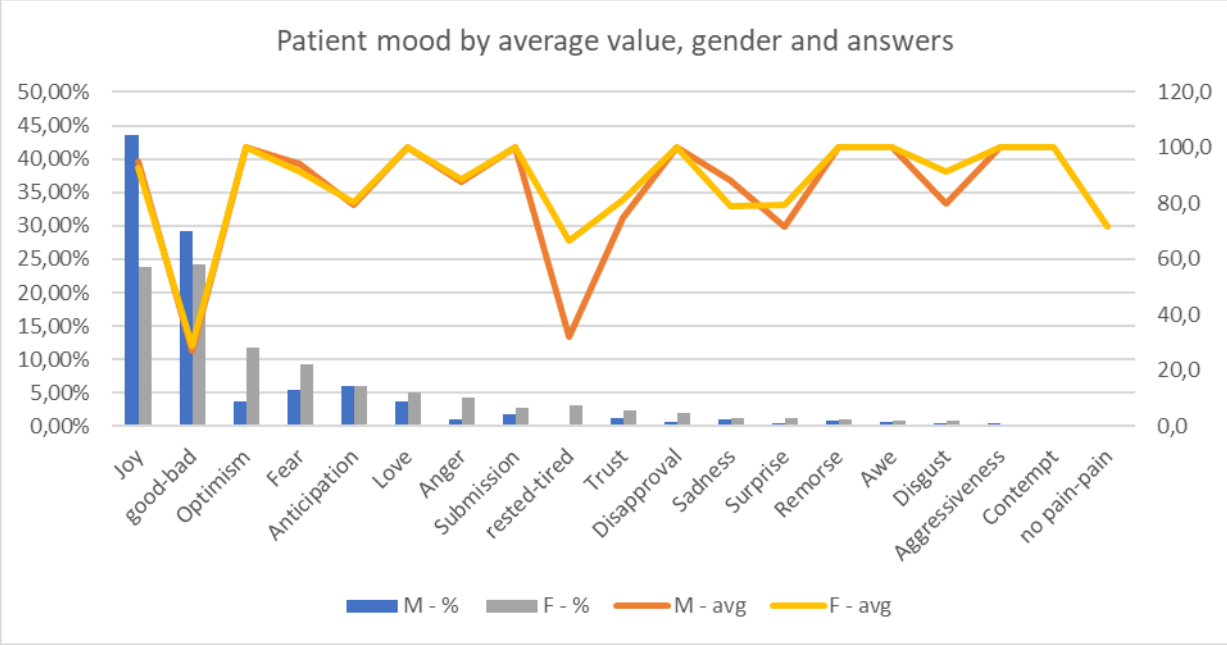

Figure 29: Patient mood by average value, gender and answers.

JOY	gender	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	M	535	95,328233	12,2919497	,5314273
	F	1356	92,968044	16,4010997	,4453927

Table 86 T-test calculation between gender and mood JOY

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of gender	Independent samples Mann Whitney U test	**,216**	Retain the null hypothesis

Table 87: Mann Whitney test

Good-bad	gender	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	M	535	95,328233	12,2919497	,5314273
	F	1356	92,968044	16,4010997	,4453927

Table 88: T-test calculation between gender and mood GOOD-BAD

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of gender	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 89: Mann Whitney test

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Mood	2021			2022		
	N	**%**	**AVG**	**N**	**%**	**AVG**
Joy	1193	28,3%	90,1	698	25,8%	99,8
good-bad	533	12,6%	27,9	1205	44,5%	28,7
Optimism	702	16,7%	100,0	9	0,3%	100,0
Fear	256	6,1%	80,7	336	12,4%	100,0
Anticipation	298	7,1%	73,5	113	4,2%	98,5
Love	331	7,9%	100,0	2	0,1%	100,0
Anger	125	3,0%	77,5	127	4,7%	100,0
Submission	184	4,4%	100,0		0,0%	
rested-tired	73	1,7%	65,9	110	4,1%	66,4
Trust	106	2,5%	74,2	39	1,4%	98,2
Disapproval	116	2,8%	100,0	1	0,0%	100,0
Sadness	72	1,7%	78,6	9	0,3%	94,6
Surprise	55	1,3%	71,0	21	0,8%	100,0
Remorse	65	1,5%	100,0		0,0%	
Awe	56	1,3%	100,0	3	0,1%	100,0
Disgust	22	0,5%	77,4	30	1,1%	100,0
Aggressiveness	22	0,5%	100,0		0,0%	
Contempt	5	0,1%	100,0		0,0%	
no pain-pain	1	0,0%	58,0	4	0,1%	75,2
sum	**4215**	**100,0%**		**2707**	**100,0%**	

Table 90: Patient mood by number of measurements and average values-by DATA YEAR

The same result is observed when the data is broken down by year. Thus, most measurements for both years come from emotions of Joy (28,3 % in 2021 and 25,8 % in 2022) and Good-bad (12,6 % in 2021 and 44,5 % in 2022) emotions. There is also a statistically significant difference ($p<0,01$) between the years for the two emotions ($p<0,01$). In total, 60,9 % of all measurements happened in 2021 and only 39,1 % measurements in 2022.

Some emotions have no measurement in 2022 because they were not measured any more or data was not submitted

Conclusion: Majority of patients have reported emotions of Joy and Good-Bad emotions.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

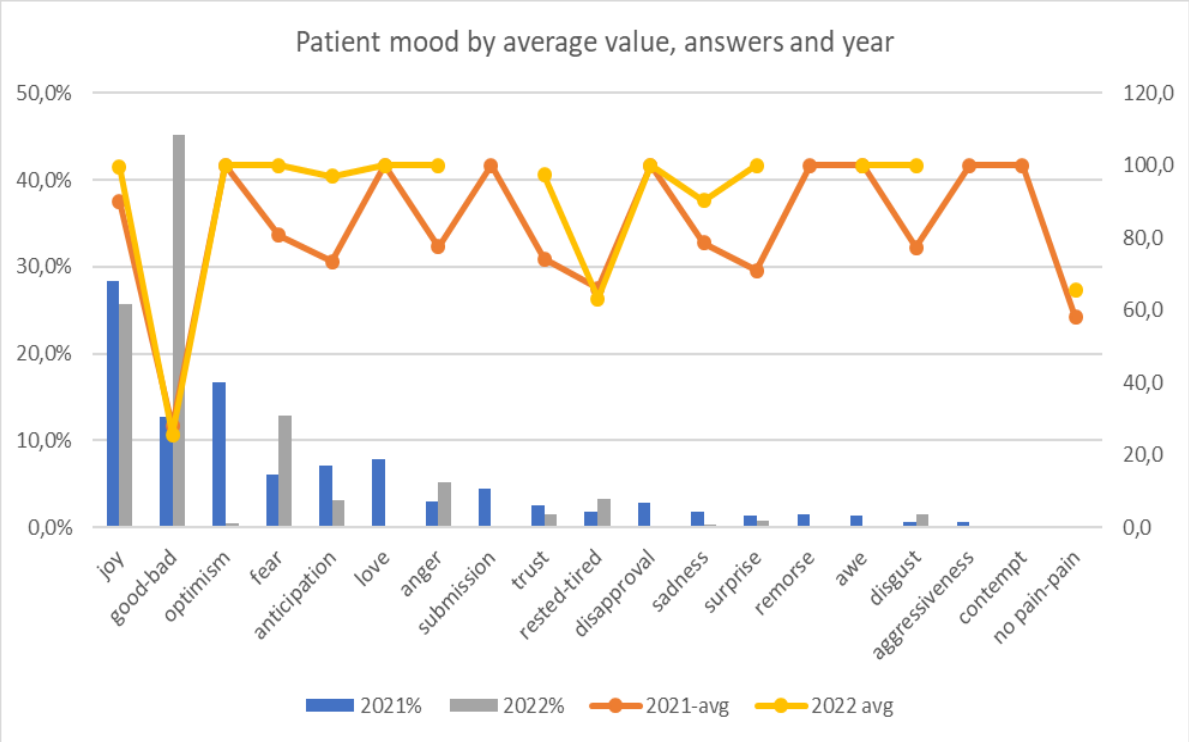

Figure 30: Patient mood by average value, answers (emotions) and year.

JOY	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	2021	1193	90,054289	18,2825845	,5293189
	2022	698	99,757171	3,2633481	,1235195

Table 91: T-test calculation between data year and mood JOY

JOY - YEAR		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	908,395	,000	-13,891	1889	,000	-9,7028822	,6984989	-11,072792 7	-8,3329716
	Equal variances not assumed			-17,851	1318,668	,000	-9,7028822	,5435398	-10,769179 3	-8,6365850

Table 92: t_test

There is a statistical significance (p<0,01) between data year and values of emotion JOY

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of year	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 93: Mann Whitney test

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

GOOD-BAD	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	2021	533	27,889605	18,6453717	,8076207
	2022	1205	28,677046	18,9140562	,5448678

Table 94 T-test calculation between data year and mood GOOD-BAD

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of year	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 95 Mann Whitney test

mood	34-53 age			54-74 age		
	N	**%**	**avg**	**N**	**%**	**avg**
Joy	638	21,4%	88,7	1253	31,8%	96,2
good-bad	780	26,2%	32,8	958	24,3%	24,9
Optimism	241	8,1%	100,0	470	11,9%	100,0
Fear	279	9,4%	87,6	313	7,9%	95,3
Anticipation	177	5,9%	68,4	234	5,9%	89,4
Love	164	5,5%	100,0	169	4,3%	100,0
Anger	180	6,0%	87,4	72	1,8%	92,5
Submission	53	1,8%	100,0	131	3,3%	100,0
rested-tired	108	3,6%	64,7	75	1,9%	68,4
Trust	70	2,4%	70,5	75	1,9%	90,1
Disapproval	71	2,4%	100,0	46	1,2%	100,0
Sadness	56	1,9%	81,0	25	0,6%	79,1
Surprise	22	0,7%	51,4	54	1,4%	90,3
Remorse	41	1,4%	100,0	24	0,6%	100,0
Awe	29	1,0%	100,0	30	0,8%	100,0
Disgust	47	1,6%	91,5	5	0,1%	80,0
Aggressiveness	12	0,4%	100,0	10	0,3%	100,0
Contempt	5	0,2%	100,0		0,0%	
no pain-pain	3	0,1%	63,1	2	0,1%	84,7
Sum	**2976**	**100,0%**		**3946**	**100,0%**	

Table 96: Patient mood by number of measurements and average values. By AGE GROUP

The median age was 52,5 and the average age was 53,1. Patients were divided into two age groups: 34-53 y/o and the 54-74 y/o. Once again, the results show that most measurements correspond to the emotions of Joy and Good-bad. The Joy emotion has more measurements in older patients (31,8 % against 21,4 % in younger patients) and also more joy emotion measurement in the older population (1253 against 638) . There is also a statistically significant difference ($p<0,01$) between the intensity of Joy values in two age groups.

As for Good-Bad emotions a slightly higher number (2 %) of measurements in younger patients than older patients was observed. The intensity values of measurement in good-bad mood are similar for both groups and there is a statistically significant difference ($p<0,01$) between age groups.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Conclusion: Also, in this case the older group showed significantly higher measurements of Joy then the younger group.

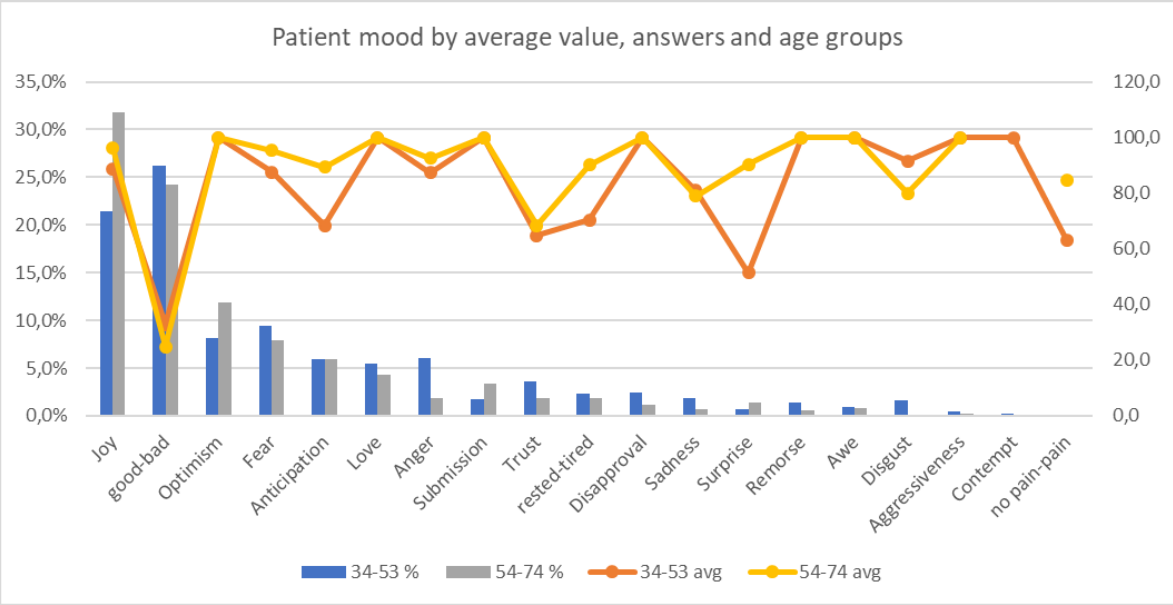

Figure 31: Patient mood by average value, gender and age groups.

JOY	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	34-53	638	88,691348	20,5492350	,8135520
	54-74	1253	96,153386	11,1188692	,3141124

Table 97: T-test calculation between age groups and mood JOY

JOY – AGE GROUPS		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	370,653	,000	-10,243	1889	**,000**	-7,4620377	,7284660	-8,8907202	-6,0333553
	Equal variances not assumed			-8,557	831,672	**,000**	-7,4620377	,8720857	-9,1737854	-5,7502900

Table 98: tTest

There is a statistical significance (p<0,01) between age groups and values of emotion JOY

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of age groups	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 99: Mann Whitney test

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

GOOD-BAD	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	34-53	780	32,811295	17,2308833	,6169646
	54-74	958	24,872850	19,3282557	,6244676

Table 100: T-test calculation between age groups and mood GOOD-BAD

GOOD-BAD / AGE GROUP		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value – quantity	Equal variances assumed	,054	,816	8,938	1736	**,000**	7,9384449	,8881898	6,1964104	9,6804795
	Equal variances not assumed			9,043	1721,765	**,000**	7,9384449	,8778412	6,2166976	9,6601923

Table 101: tTest

There is a statistical significance (p<0,01) between age groups and values of emotion GOOD – BAD.

Null hypothesis	test	sig	decision
The distribution of value quantity is the same across categories of age groups	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 102: Mann Whitney test

Patient moods by number of measurements, age, collaboration, and average value for each patient – are gathered in separate files.

Patient mood analysis for SERGAS

The data does not include emotion measurements if:

- ➔ count was less than 10 (looking at total N of measurements), (1 patient)
- ➔ the value was 0 (no value),
- ➔ no data, it could mean it was a mistake or data not transferred.

Number of patients = 35

Men = 11 (31,5 %)

Female = 20 (57,1 %)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

4 unknown (11,4 %)

Average age (mean) = 55,8

Median = 55

Age: Men (average) = 63,1

Female (average) = 51,8

Patient mood	N	%	Average	Min. value	Max. value
Joy	4722	50,0%	95,1	6,05	100
good-bad	1227	13,0%	21,6	1,03	100
Disgust	933	9,9%	90,6	1,47	100
rested-tired	572	6,1%	51,1	1,58	100
Fear	554	5,9%	77,4	5,2	100
Trust	548	5,8%	91,9	3,16	100
Anger	209	2,2%	85,5	16,5	100
Love	205	2,2%	100,0	100	100
Submission	111	1,2%	100,0	100	100
Optimism	109	1,2%	100,0	100	100
Sadness	94	1,0%	83,2	23,8	100
Disapproval	77	0,8%	100,0	100	100
Surprise	31	0,3%	88,8	6,07	100
Anticipation	27	0,3%	62,6	22,5	100
Remorse	17	0,2%	100,0	100	100
Awe	9	0,1%	100,0	100	100
no pain-pain	5	0,1%	71,1	62	100
Aggressiveness	2	0,0%	100,0	100	100
Contempt	1	0,0%	100,0	100	100
sum	**9453**				

Table 103 Patient mood by number of measurements, min, max and average values.

The most mood measurements are for Joy (50,0 %) and the one ranging from Good-Bad (13,0 %). Other emotions have only a few measurements.

The highest intensity values were for Love, Submission, Optimism, Disapproval, Remorse, Awe, Aggressiveness and Contempt. Nevertheless, these last two emotions have not many measurements (2 and 1) and therefore the results can be questionable.

Statistical calculations were only performed with the emotions of JOY and GOOD-BAD given that they showed the highest numbers of measurements for those two emotions.

Conclusion: The results obtained show that the majority of patients had been in a good emotional state in the project period stating emotion.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

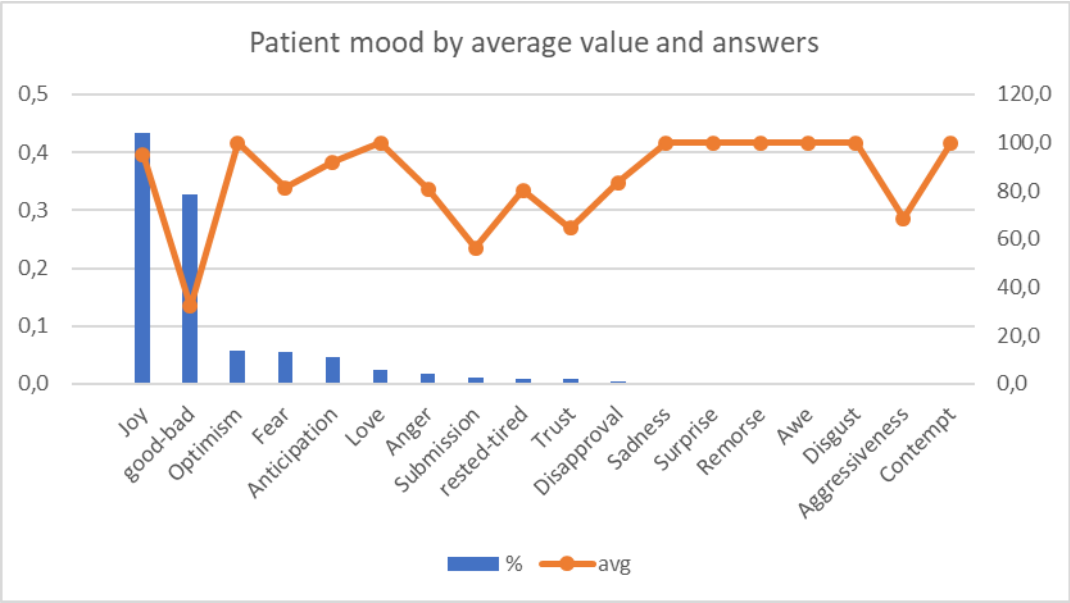

Figure 32: Patient mood by average value and answers.

Patient mood	M (N)	M (%)	F (N)	F (%)	M (avg)	F (avg)	None (N)	None (%)	None (Avg)
Joy	2682	70,4%	1778	35,7%	98,5	90,1	262	39,5%	93,7
good-bad	219	5,7%	1004	20,2%	28,3	20,2	4	0,6%	9,3
Disgust	509	13,4%	404	8,1%	90,5	90,4	20	3,0%	95,5
rested-tired	8	0,2%	563	11,3%	52,7	51,0	1	0,2%	84,3
Fear	152	4,0%	281	5,6%	43,6	86,5	121	18,2%	98,8
Trust	88	2,3%	250	5,0%	94,2	86,0	210	31,6%	97,9
Anger	12	0,3%	185	3,7%	87,6	86,0	12	1,8%	76,3
Love	55	1,4%	139	2,8%	100,0	100,0	11	1,7%	100
Submission	11	0,3%	99	2,0%	100,0	100,0	1	0,2%	100
Optimism	29	0,8%	75	1,5%	100,0	100,0	5	0,8%	100
Sadness	15	0,4%	70	1,4%	75,7	82,7	9	1,4%	99,8
Disapproval	19	0,5%	56	1,1%	100,0	100,0	2	0,3%	100
Surprise	3	0,1%	26	0,5%	71,8	89,9	2	0,3%	100
Anticipation	4	0,1%	21	0,4%	63,1	63,2	2	0,3%	55,2
Remorse	1	0,0%	15	0,3%	100,0	100,0	1	0,2%	100
Awe	3	0,1%	6	0,1%	100,0	100,0		0,0%	
no pain-pain		0,0%	5	0,1%		71,1		0,0%	
Aggressiveness		0,0%	1	0,0%		100,0	1	0,2%	100
Contempt	1	0,0%		0,0%	100,0			0,0%	
Sum	**3811**	**100,0 (40,3%)**	**4978**	**100,0 (52,7 %)**			**664 (7%)**		

Table 104 Patient mood by number of measurements and average values. By GENDER

Table 104 shows that most men declare high values of Joy (average 98,5). Although, women are also high (average 90,1), and this difference is statistically significant ($p<0,01$) but it also could be the result that there are most measurements with 70,4 % men and 35,7 % women.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

When looking at the outcome of the Good-Bad scale, men`s values are higher than women`s values (28,3 against 20,2). This difference is also statistically significant (p<0,05), but it should be taken into consideration that more women measured and chose this emotion (1004 measurements of women against 219 men).

Conclusion: Men show higher values of Joy than women, and differences in Good-Bad emotions outcomes.

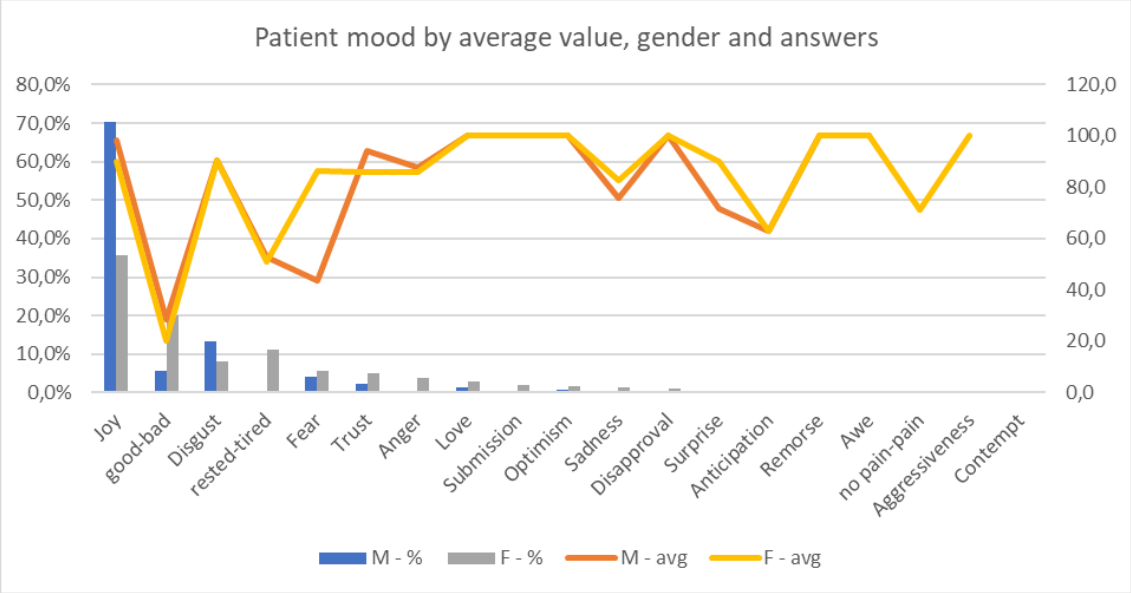

Figure 33: Patient mood by average value, gender and age groups.

JOY	gender	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	M	2682	98,470822	7,8075556	,1507599
	F	1778	90,110937	20,2006844	,4790714

Table 105 T-test calculation between gender and mood JOY

JOY - gender		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	1521,958	,000	19,362	4458	,000	8,3598844	,4317642	7,5134124	9,2063564
	Equal variances not assumed			16,645	2132,521	,000	8,3598844	,5022329	7,3749670	9,3448019

Table 106 t-test

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of gender	Independent samples Mann Whitney U test	**,216**	Retain the null hypothesis

Table 107 Mann Whitney test

Good-bad	gender	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	M	219	28,277803	21,8999419	1,4798609
	F	1004	20,202982	20,6850810	,6528154

Table 108 T-test calculation between gender and mood GOOD-BAD

GOOD-BAD / YEAR		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	12,174	,001	5,179	1221	**,000**	8,0748208	1,5592631	5,0156888	11,133952 8
	Equal variances not assumed			4,992	308,561	**,000**	8,0748208	1,6174536	4,8921865	11,257455 0

Table 109 t-test

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of gender	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 110 Mann Whitney test

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Mood	2021			2022		
	N	%	AVG	N	%	AVG
Joy	2590	50,1%	91,0	2132	49,8%	100,0
good-bad	466	9,0%	21,9	761	17,8%	21,5
Disgust	439	8,5%	80,0	494	11,5%	100,0
rested-tired	204	3,9%	55,8	368	8,6%	48,5
Fear	318	6,1%	60,6	236	5,5%	100,0
Trust	405	7,8%	89,0	143	3,3%	100,0
Anger	108	2,1%	72,0	101	2,4%	100,0
Love	205	4,0%	100,0		0,0%	
Submission	111	2,1%	100,0		0,0%	
Optimism	109	2,1%	100,0		0,0%	
Sadness	68	1,3%	76,8	26	0,6%	100,0
Disapproval	77	1,5%	100,0		0,0%	
Surprise	14	0,3%	75,2	17	0,4%	100,0
Anticipation	25	0,5%	59,6	2	0,0%	100,0
Remorse	17	0,3%	100,0		0,0%	
Awe	9	0,2%	100,0		0,0%	
no pain-pain	3	0,1%	75,1	2	0,0%	65,1
Aggressiveness	2	0,0%	100,0		0,0%	
Contempt	1	0,0%	100,0		0,0%	
sum	**5171**	**100,0%**		**4282**	**100,0%**	

Table 111 Patient mood by number of measurements and average values-by DATA YEAR

The same result is observed when the data is broken down by year. Thus most measurements for both years come from the emotions of Joy (50,1 % in 2021 and 49,8 % in 2022), Good-bad (9,0 % in 2021 and 17,8 % in 2022) emotions. There is also a statistically significant difference ($p<0,01$) between the years and joy. The average values in Good-Bad emotions are the same and therefore, there is no statistically significant difference between years and Good-Bad emotion.

In total, 54,7 % of all measurements happened in 2021 and only 45,3 % measurements in 2022.

Some emotions have no measurement in 2022 because they were not measured any more or data was not submitted.

Conclusion: The data shows the patients were less interested in measuring the emotions in the second year.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

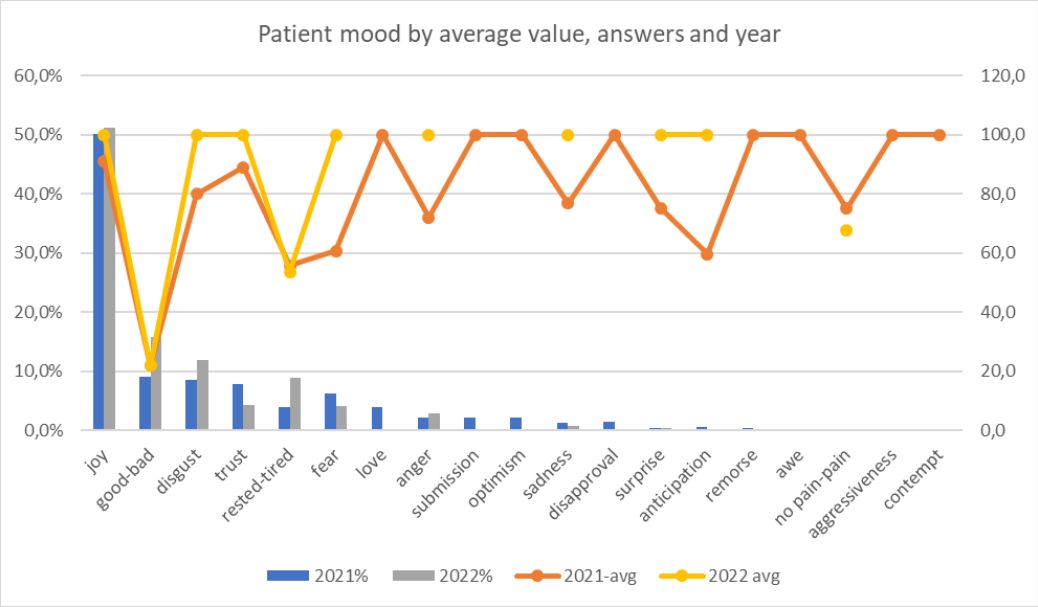

Figure 34: Patient mood by average value, answers and year.

JOY	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	2021	2590	90,994179	18,9072620	,3715171
	2022	2132	100,000000	0E-7	0E-7

Table 112 T-test calculation between data year and mood JOY

JOY - YEAR		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	2310,944	,000	-21,993	4720	,000	-9,0058211	,4094899	-9,8086124	-8,2030298
	Equal variances not assumed			-24,241	2589,000	,000	-9,0058211	,3715171	-9,7343217	-8,2773205

Table 113 T-test

There is a statistically significant difference between (p<0,01) the data year and values of emotion JOY

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of year	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 114 Mann Whitney test

GOOD-BAD	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	2021	466	21,857482	18,5007949	,8570328
	2022	761	21,456200	22,5642248	,8179525

Table 115 T-test calculation between data year and mood GOOD-BAD

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

GOOD-BAD / YEAR		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	18,084	,000	,323	1225	,747	,4012818	1,2419635	-2,0353294	2,8378930
	Equal variances not assumed			,339	1126,212	,735	,4012818	1,1847158	-1,9232166	2,7257802

Table 116 t-Test

There is no statistical significance (p>0,05) between data year and values of emotion GOOD-BAD.

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of year	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 117 Mann Whitney test

mood	44-54 age			55-71 age			none		
	N	%	avg	N	%	avg	N	%	avg
Joy	960	32,4%	86,8	3500	60,1%	97,4	262	39,5%	93,7
good-bad	229	7,7%	27,6	994	17,1%	20,3	4	0,6%	9,3
Disgust	305	10,3%	91,5	608	10,4%	90,0	20	3,0%	95,5
rested-tired	546	18,4%	50,7	25	0,4%	59,0	1	0,2%	84,3
Fear	238	8,0%	92,2	195	3,3%	46,1	121	18,2%	98,8
Trust	202	6,8%	85,0	136	2,3%	92,9	210	31,6%	97,9
Anger	157	5,3%	86,4	40	0,7%	85,0	12	1,8%	76,3
Love	74	2,5%	100,0	120	2,1%	100,0	11	1,7%	100
Submission	35	1,2%	100,0	75	1,3%	100,0	1	0,2%	100
Optimism	49	1,7%	100,0	55	0,9%	100,0	5	0,8%	100
Sadness	66	2,2%	83,4	19	0,3%	74,5	9	1,4%	99,8
Disapproval	43	1,5%	100,0	32	0,5%	100,0	2	0,3%	100
Surprise	25	0,8%	93,2	4	0,1%	55,4	2	0,3%	100
Anticipation	10	0,3%	80,2	15	0,3%	51,8	2	0,3%	55,2
Remorse	14	0,5%	100,0	2	0,0%	100,0	1	0,2%	100
Awe	6	0,2%	100,0	3	0,1%	100,0		0,0%	
no pain-pain	4	0,1%	63,8	1	0,0%	100,0		0,0%	
Aggressiveness	1	0,0%	100,0		0,0%		1	0,2%	100
Contempt		0,0%		1	0,0%	100,0		0,0%	
Sum	**2964**	**100,0%**		**5825**	**100,0%**				

Table 118 Patient mood by number of measurements and average values. By AGE GROUP

The median for age was 55 and average age was 55,8. Patients were divided into two age groups: 44-54 y/o and 55-71 y/o. Once again, the results show that most measurements correspond to the emotions of Joy and Good-Bad. The Joy emotions has more

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

measurements in older patients (60,1 % against 32,4 % in younger patients) and also more Joy emotion measurement in older population (3500 against 960) . There is also a statistically significant difference (p<0,01) between the intensity Joy values in the two age groups.

As for good-bad emotions there are a little bit more (7 %) measurements in younger patients than older patients. The values of measurement in good-bad mood are similar for both groups and there is a statistically significant difference (p<0,01) between age groups.

Conclusion: Younger group showed significantly higher measurements of Joy then older group. There were also more measurements in the younger group.

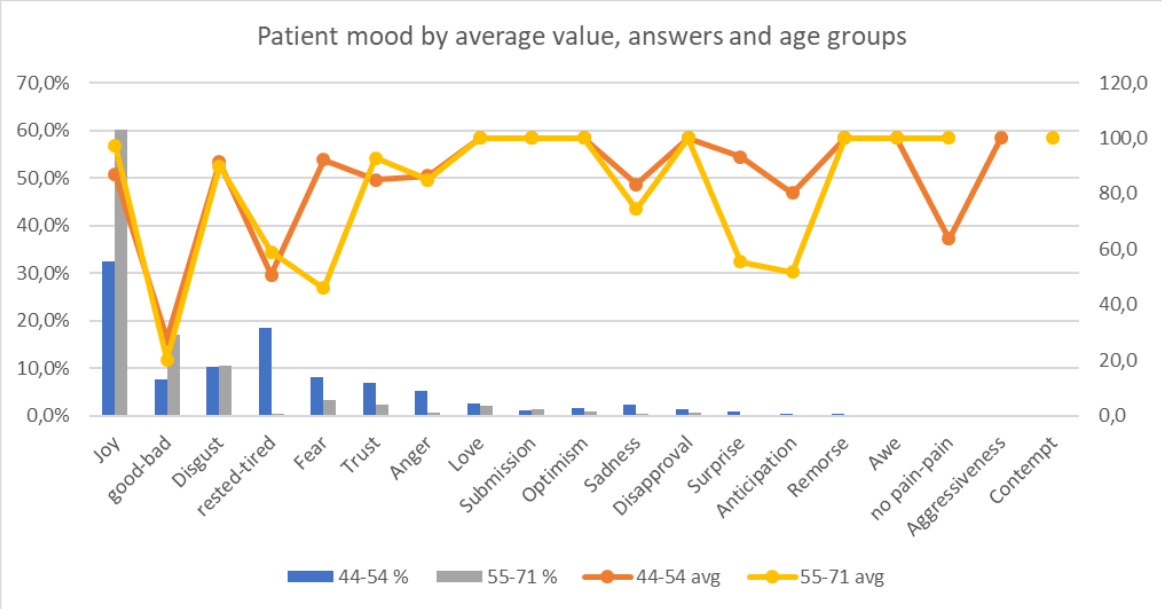

Figure 35: Patient mood by average value, gender and age groups.

JOY	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	44-54	960	86,793908	23,0906904	,7452488
	55-71	3500	97,426811	10,2385651	,1730633

Table 119 T-test calculation between age groups and mood JOY

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

JOY – AGE GROUPS		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	1306,090	,000	-20,795	4458	,000	-10,6329028	,5113321	-11,6353675	-9,6304380
	Equal variances not assumed			-13,898	1064,373	,000	-10,6329028	,7650796	-12,1341383	-9,1316673

Table 120 T-Test

There is a statistical significance (p<0,01) between age groups and values of emotion JOY.

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of age groups	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 121: Mann Whitney test

GOOD-BAD	year	N	Mean	Std. Deviation	Std. Error Mean
value_quantity	44-54	229	27,576312	21,1743337	1,3992394
	55-71	994	20,283358	20,8893270	,6625692

Table 122 T-test calculation between age groups and mood GOOD-BAD

GOOD-BAD / AGE GROUP		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
value_quantity	Equal variances assumed	1,915	,167	4,751	1221	,000	7,2929540	1,5351043	4,2812194	10,3046887
	Equal variances not assumed			4,711	337,809	,000	7,2929540	1,5481824	4,2476617	10,3382463

Table 123 t-test

There is a statistically significant difference (p<0,01) between age groups and values of emotion GOOD – BAD.

Null hypothesis	test	sig	decision
The distribution of value_quantity is the same across categories of age groups	Independent samples Mann Whitney U test	**,000**	Reject the null hypothesis

Table 124 Mann Whitney test

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Patient moods by number of measurements, age, collaboration and average value for each patient – are gathered in separate files.

Conclusions: altogether patients tend to report positive emotions (most of the measurements were for Joy). As later reported by patients in the workshop, this could be the case because that is one of rare positive emotions from which they can choose. On the other hand, patients have marked this as a great opportunity to get aware of their emotional state at least once a day.

Altogether these data about patient emotions helps clinician to gain overview of their mental health and in case of many negative emotions reported, recommend to contact psychologist.

8. General feedback from patients (PERSIST block ABC)

Part A: feedback about the project

To gather general feedback from patients, patient surveys were conducted at three different time points using an app-based questionnaire. The aim was to understand patients' experience of participating in the study and to identify and share their most important insights. In total, 32 participants from different healthcare institutions (6 from CHU, 8 from SERGAS, 14 from UKCM, and 4 from UL) were included in the analysis, which was carried out by statisticians from UL. The surveys were conducted at the beginning of the questionnaire introduction in the app, after the introduction of the virtual agent in the app, and at the end of the study period in October 2022

- ✓ **How do you rate your experience with participation in the PERSIST project (in general)?**

Table 125 shows that there were no statistically significant differences between any two time points, as determined by the Friedman One-Way Repeated Measure Analysis of Variance by Ranks (p=0,585). Conover's post-hoc pairwise comparisons also revealed no significant differences between the initial and mid points (p=0,391), the initial and final points (p=0,346), or the mid and final points (p=0,931).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	Init_experience	R_experience 5	V3rd_experience
Mean	7,406	7,750	7,688
Median	8,000	8,000	8,000
Std. Deviation	1,643	1,704	1,533
Minimum	4,000	5,000	4,000
Maximum	10,000	10,000	10,000
25th percentile	6,000	6,750	7,000
50th percentile	8,000	8,000	8,000
75th percentile	8,000	9,000	9,000

Table 125 Descriptive Statistics of 1st question

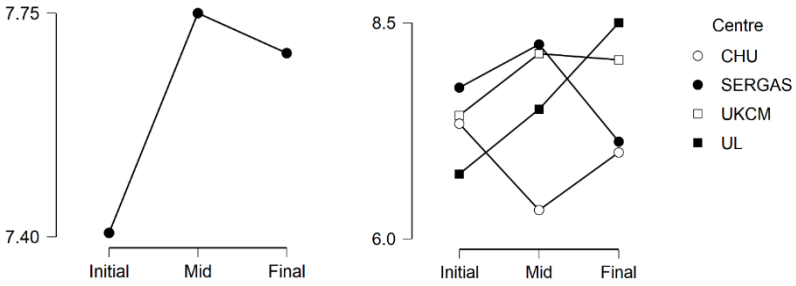

Figure 36: Descriptive plot.

➔ Conclusions:

Patient feedback suggests that they rated their participation in the PERSIST study positively, with slight improvements in their ratings over time. This indicates that patients found the project valuable and well-received. Furthermore, the lack of statistically significant differences between any two time points suggests that patient experiences remained stable throughout the project. However, patients from SERGAS and UKCM reported slightly lower evaluations, suggesting that there may be areas for improvement in those locations.

Overall, these results suggest that the PERSIST project successfully engaged patients and provided a positive experience for them.

✓ **Are the instructions and explanations about the project from personnel understandable to you?**

Table 126 shows that there were no statistically significant differences between any two time points, as determined by the Friedman One-Way Repeated Measure Analysis of Variance by Ranks (p=0,833). Conover's post-hoc pairwise comparisons also showed no significant differences between the initial and mid points (p=0,866), the initial and final points (p=0,672), or the mid and final points (p=0,554).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	Init_personnel_expl	R_personnel_expl6	V3rd_personnel_expl
Mean	8,531	8,531	8,469
Median	9,000	8,500	8,000
Std. Deviation	1,665	1,164	1,244
Minimum	2,000	6,000	6,000
Maximum	10,000	10,000	10,000
25th percentile	8,000	8,000	8,000
50th percentile	9,000	8,500	8,000
75th percentile	10,000	9,250	10,000

Table 126: Descriptive Statistics of 2nd question from PERSIST block ABC

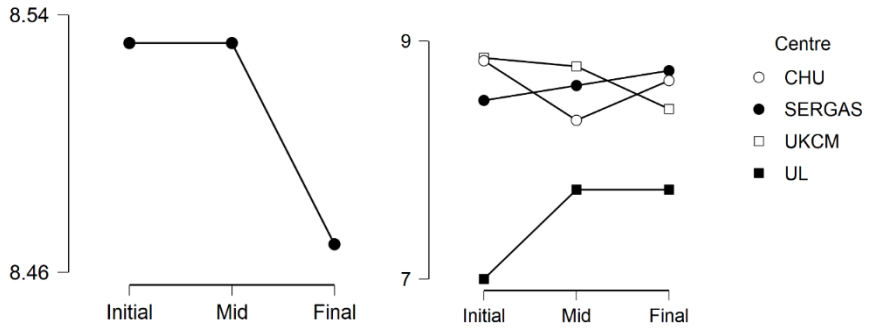

Figure 37: Descriptives plot of PERSIST block ABC 2nd question.

➔ Conclusions:

Patients found the instructions and explanations as understandable and rated their participation as great. Moreover, there were no statistically significant differences between any two time points, and patients from all centres reported slight improvements in the quality of the explanations over time. These findings suggest that the project was successful in engaging patients and providing them with clear instructions and explanations throughout the study period.

✓ **How does the participation in the PERSIST project make you feel?**

The results presented in Table 127 show no statistically significant differences between any two time points, as determined by the Friedman One-Way Repeated Measure Analysis of Variance by Ranks (p=0,502). Conover's post-hoc pairwise comparisons also revealed no significant differences between the initial and mid points (p=0,554), the initial and final points (p=0,238), or the mid and final points (p=0,554).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	Init_feel	R_feel4	V3rd_feel
Mean	8,125	8,188	8,063
Median	8,000	8,000	8,000
Std. Deviation	1,862	1,554	1,684
Minimum	4,000	5,000	5,000
Maximum	10,000	10,000	10,000
25th percentile	7,750	7,000	6,750
50th percentile	8,000	8,000	8,000
75th percentile	10,000	10,000	10,000

Table 127: Descriptive Statistics of 3rd question from PERSIST block ABC

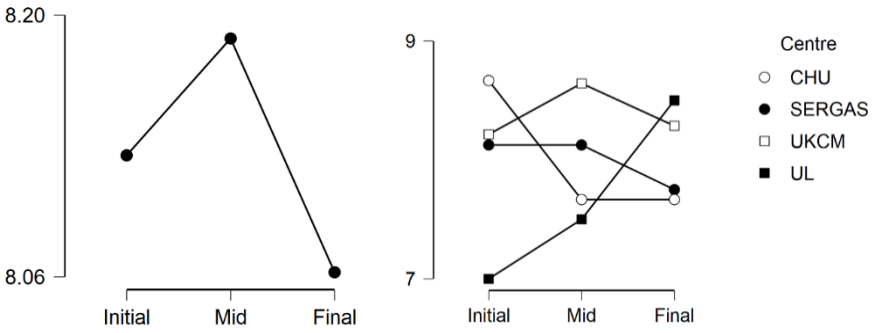

Figure 38: Descriptives plot of PERSIST block ABC 3rd question.

➔ Conclusions:

Overall, patients rate their participation in the PERSIST project as great, with an average score of 8 or higher across all time points. Patients' ratings of their participation and the quality of personnel explanations remained consistently high over time, indicating that the project was effective and provided clear guidance throughout. Furthermore, the absence of statistically significant differences between any two time points suggests that the results are reliable and consistent.

Part B: feedback about mHealth

Twenty participants responded to the survey at three different time points, with 4 participants from CHU, 4 from SERGAS, and 12 from UKCM. However, none of the participants were from UL as none of them replied in 3 time points to this questionnaire.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

- ✓ **How do you rate the emotion wheel/detection in the app?** From 1 (bad, confusing) to 10 (super, interesting)

Table 128 shows there were no statistically significant differences between any two time points, Friedman One-Way Repeated Measure Analysis of Variance by Ranks p=0,390 . Additionally, Conover's post-hoc pairwise comparisons indicate that there were no significant differences between the initial-mid time points (p>0,999), but there were slight differences between the initial-final and mid-final time points (p=0,235).

	1st	Middle	Last
Mean	6,500	6,350	6,850
Median	7,000	7,500	8,000
Std. Deviation	2,395	2,681	2,207
Minimum	2,000	1,000	2,000
Maximum	10,000	10,000	10,000
25th percentile	5,000	4,000	5,750
50th percentile	7,000	7,500	8,000
75th percentile	8,000	8,000	8,000

Table 128: Descriptive Statistics of par B 1st question "How do you rate the emotion wheel/detection in the app"

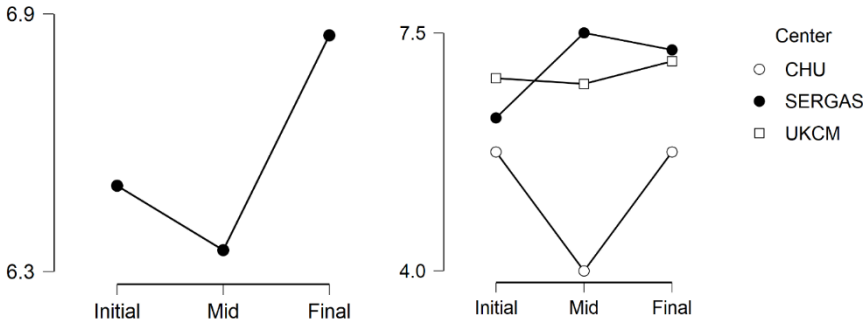

Figure 39: Descriptives plot of PERSIST block B 1st

➔ Conclusions:

The emotion wheel/detection in the app was generally well-received by the patients, with a mean score of 6.57 across all time points (1st, middle, and last). The median score was higher (7.5), indicating that most participants rated the app positively. The low standard deviation suggests that participants' ratings were consistent

However, the lack of statistically significant differences between any two time points suggests that there were not meaningful changes in participants' ratings over time. Therefore, it is unclear whether the effectiveness or usability of the feature improved or declined during the study period.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

While the data suggests that the emotion wheel/detection feature was well-received, further research may be needed to assess its long-term performance.

- ✓ **How do you rate your experience with questionnaires in the app?** From 1 (bad) to 10 (excellent)

Table 129 shows there were no statistically significant differences between any two time points, Friedman One-Way Repeated Measure Analysis of Variance by Ranks p=0,779. Additionally, Conover's post-hoc pairwise comparisons indicate that there were no significant differences between the p-values: initial-mid p=0,490, initial-final p=0,843, mid-final p=0,622.

	First	Middle	Last
Mean	7,600	7,250	7,600
Median	8,000	8,000	8,000
Std. Deviation	1,635	2,023	1,789
Minimum	5,000	2,000	4,000
Maximum	10,000	10,000	10,000
25th percentile	6,000	6,750	6,000
50th percentile	8,000	8,000	8,000
75th percentile	8,250	8,000	9,000

Table 129: Descriptive statistics of"How do you rate your experience with questionnaires in the app" question.

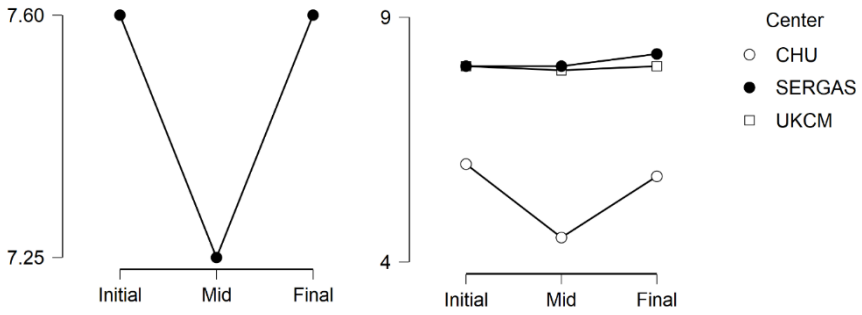

Figure 40: Descriptives plot of PERSIST block B 2nd question

- ➔ Conclusions:

Participants generally had a positive experience with questionnaires in the app, with a mean score of 7.48 across all time points (first, middle, and last). The median score was 8, which indicates that most participants rated their experience as "good" or "excellent." The standard deviation was also relatively low, suggesting that there was little variability in participants' ratings.

Additionally, the lack of statistically significant differences between any two time points suggests that patients' perceptions of their experience with questionnaires did not change

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

significantly over time. This implies that the app maintained a consistent level of usability and effectiveness throughout the study period.

Overall, the data suggests that participants had a positive experience with questionnaires in the app, which is a favourable outcome. However, it's important to note that this is only one aspect of the app's performance, and more research may be needed to fully evaluate its effectiveness and user-friendliness.

- ✓ **How do you rate your experience with diary recording?** From 1 (bad, confusing) to 10 (super, interesting)

Table 130 indicates that there were no statistically significant differences observed between any two time points, as revealed by Friedman One-Way Repeated Measure Analysis of Variance by Ranks (p=0,581). Conover's post-hoc pairwise comparisons also showed no significant differences between the initial-mid (p=0,304), initial-final (p=0,512), and mid-final (p=0,707) time points.

	First	Middle	Last
Mean	6,650	7,000	7,000
Median	7,000	8,000	8,000
Std. Deviation	2,455	2,753	2,695
Minimum	1,000	1,000	1,000
Maximum	10,000	10,000	10,000
25th percentile	5,750	6,750	6,000
50th percentile	7,000	8,000	8,000
75th percentile	8,000	9,000	9,000

Table 130: Descriptive statistics of" How do you rate your experience with diary recording" question.

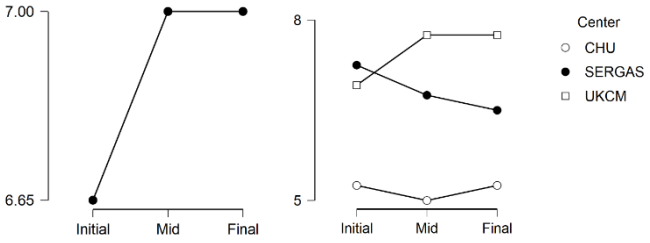

Figure 41: Descriptives plot of PERSIST block B 3rd question

- ➔ Conclusions:

Participants who replied to this question generally had a positive experience with diary recording in the app, with a mean rating of 7 out of 10. The median rating was also 8 out of 10 for both the middle and last time points, indicating that the positive experience for

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

these patients was consistent over time. Additionally, there were no statistically significant differences between any two time points, indicating that the positive experience with diary recording for these patients was stable throughout the study.

- ✓ **How do you rate your experience with the mHealth app?** From 1 (really bad) to 10 (excellent)

The data in Table 131 shows that there were no statistically significant differences between any two time points in terms of participants' ratings of the app's ease of use. The results of the Friedman One-Way Repeated Measure Analysis of Variance by Ranks suggest that the p-value was not significant at 0,279, indicating that any observed differences in the ratings were likely due to chance.

Furthermore, the Conover's post-hoc pairwise comparisons revealed that there were no significant differences between the initial and middle time points (p=0,891), but there were significant differences between the initial and final time points (p=0,138) and the mid and final time points (p=0,176).

	First	Middle	Last
Mean	7,600	7,350	7,900
Median	7,500	8,000	8,000
Std. Deviation	1,667	1,899	1,553
Minimum	5,000	3,000	5,000
Maximum	10,000	10,000	10,000
25th percentile	6,000	6,000	7,000
50th percentile	7,500	8,000	8,000
75th percentile	9,000	8,250	9,000

Table 131: Descriptive statistics of"How do you rate your experience with the mHealth app."

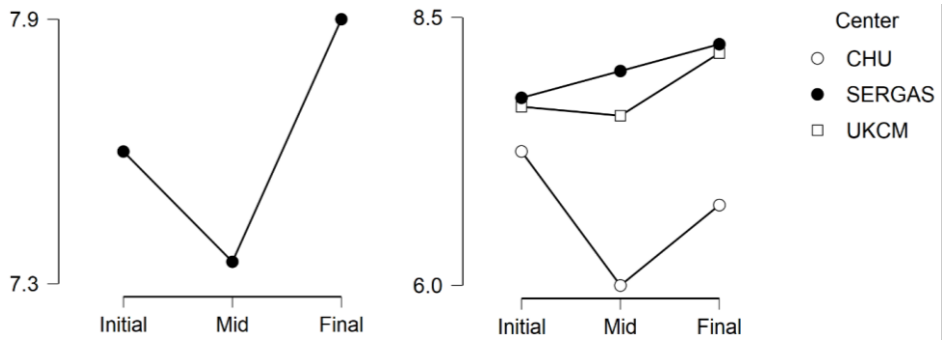

Figure 42: Descriptives plot of PERSIST block B 4th question

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

➔ Conclusions:

Overall, patients rate their experience with the mHealth app as good and the ratings slightly increase over time. Additionally, there were no statistically significant differences between any two time points, indicating that the app was consistently well-received by patients throughout the study. The fact that patients from all centres gave more points in the middle test is also a positive indication that the app was consistently useful to patients across different locations. Finally, while CHU patients gave the lowest ratings, there were still no statistically significant differences detected, suggesting that the app was generally well-received regardless of the specific centre.

✓ **Are the instructions and explanations about mHealth app usage understandable?** From 1 (completely confusing) to 10 (completely clear)

Table 132 indicates that there were no statistically significant differences between any two time points (p=0,109, Friedman One-Way Repeated Measure Analysis of Variance by Ranks). The post-hoc pairwise comparisons by Conover's test show that there was no significant difference between the initial and middle time points (p=0,910), but there was a trend towards a difference between the initial and final time points (p=0,078) and between the middle and final time points (p=0,062). However, these differences were not statistically significant.

	First	Middle	Last
Mean	8,600	8,600	8,250
Median	9,000	9,000	8,000
Std. Deviation	1,314	1,273	1,333
Minimum	5,000	6,000	6,000
Maximum	10,000	10,000	10,000
25th percentile	8,000	8,000	7,750
50th percentile	9,000	9,000	8,000
75th percentile	9,250	10,000	9,000

Table 132: Descriptive statistics of" Are the instructions and explanations about mHealth app usage understandable?"

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

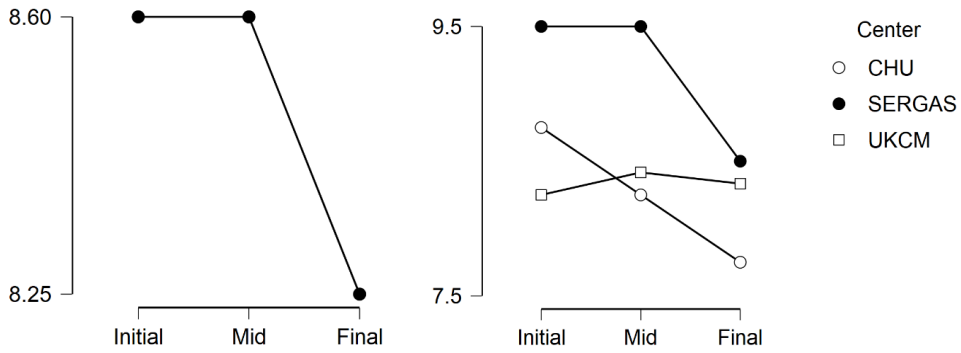

Figure 43: Descriptives plot of PERSIST block B 5th question

➔ Conclusions:

On average, patients rate the instructions and explanations about the mHealth app usage as clear (mean score of 8,483), with no statistically significant differences between any two time points. While patients from all centres (except UL, who was not included in the analyses as none of the patients had replied to the question in 3 given time points) reported a slight decrease in understanding over time, this was not statistically significant. Therefore, it can be concluded that the majority of patients found the instructions and explanations about the mHealth app usage to be understandable throughout the study period.

✓ **Do you follow up your gathered data in the mHealth app?** From 1 (no at all) to 10 (all the time)

As can be seen in Table 133 there are no statistically significant differences between any two time points (Friedman One-Way Repeated Measure Analysis of Variance by Ranks p=0,395; Conover's post-hoc pairwise comparisons: initial-mid p=0,704, initial-final p=0,189, mid-final p=0,345).

	First	Middle	Last
Mean	7,350	6,800	6,900
Median	8,000	7,500	8,000
Std. Deviation	2,889	2,783	2,532
Minimum	1,000	1,000	2,000
Maximum	10,000	10,000	10,000
25th percentile	5,750	5,500	5,750
50th percentile	8,000	7,500	8,000
75th percentile	10,000	8,250	8,250

Table 133: Descriptive statistics of "Do you follow up your gathered data in the mHealth app?"

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

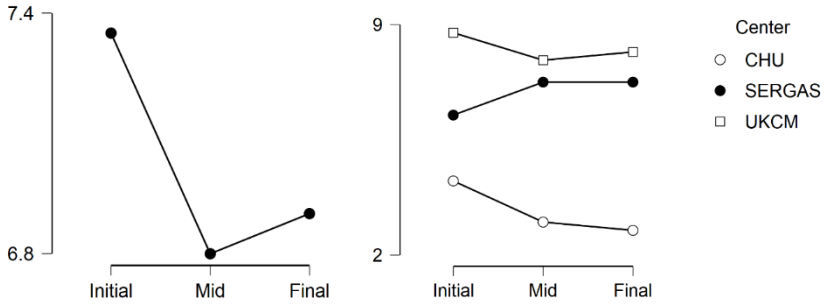

Figure 44: Descriptives plot of PERSIST block B 6th question

➔ Conclusions:

On average, patients tend to follow up on their gathered data in the mHealth app with a mean score of around 7 out of 10. However, there is a trend of patients doing this less as time goes on, although this trend is not statistically significant. Additionally, patients from CHU tend to follow up on their data less compared to patients from other centres.

✓ **Does the mHealth app affect your behaviour?** From 1 (no at all) to 10 (I modify my behaviour after looking at the data)

As can be seen in Table 134 there are no statistically significant differences between any two time points (Friedman One-Way Repeated Measure Analysis of Variance by Ranks p=0,755; Conover's post-hoc pairwise comparisons: initial-mid p=0,707, initial-final p=0,454, mid-final p=0,707).

	First	Middle	Last
Mean	5,500	5,750	6,150
Median	5,000	6,000	6,000
Std. Deviation	3,052	2,693	2,978
Minimum	1,000	1,000	1,000
Maximum	10,000	10,000	10,000
25th percentile	3,750	4,000	4,000
50th percentile	5,000	6,000	6,000
75th percentile	7,250	8,000	8,000

Table 134: Descriptive statistics of "Does the mHealth app affect your behaviour?"

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

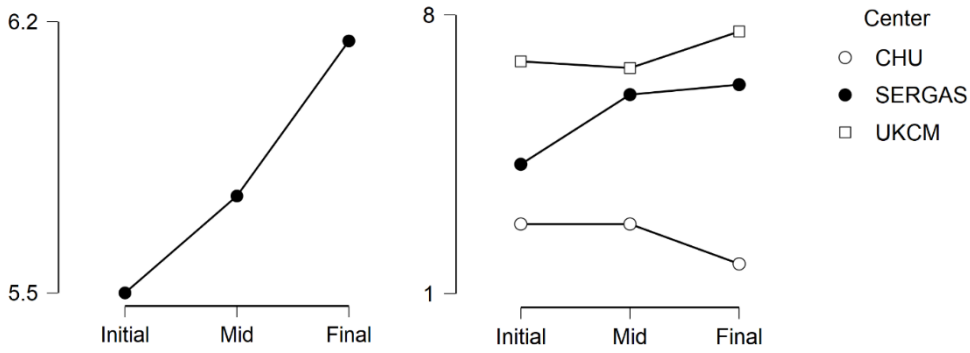

Figure 45: Descriptives plot of PERSIST block B 7th question.

➔ Conclusions:

On average, patients tend to follow their gathered data in the mHealth app (with a mean score of around 7 out of 10). This suggests that patients are engaged with the app and willing to monitor their health using it. Additionally, the slight increase in mean score from the first time point to the last suggests that patients may become more engaged with the app over time.

✓ **If you could change one or more things about the mHealth app, it would be** (name it/write it down)

Figure 46 compares patients' answers collected during the middle and the end of the study.

Interestingly, during the middle of the study, patients expressed a desire for modifications to the emotion section modified. After implementing the technical improvements, no further modifications were identified

Among the patients who continued to participate, some expressed an obligation to answer to the app and/or wished for more appointment controls.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

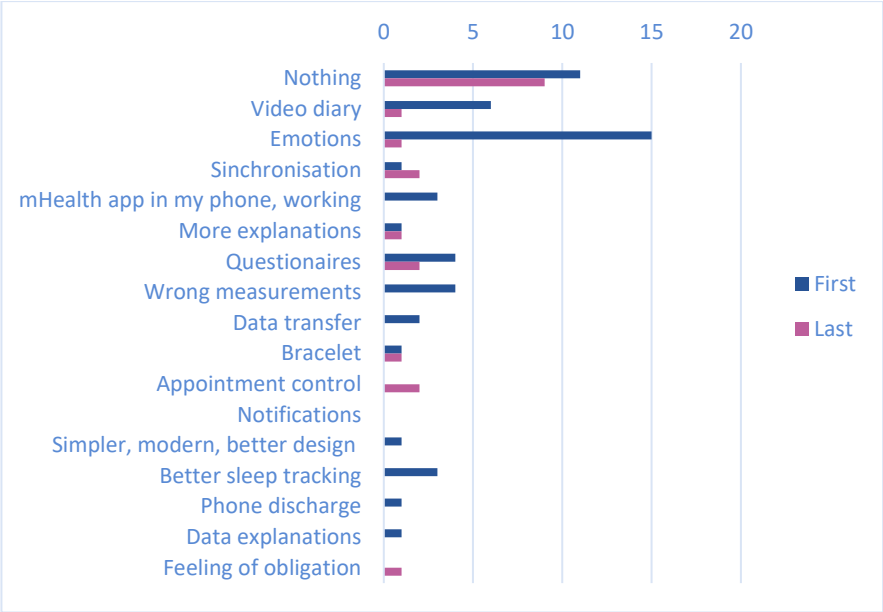

Figure 46 Comparison of patients answers from the middle to end of the study.

➔ Conclusions:

In conclusion, improvements to the emotion section of the app led to a decrease in patient complaints. However, towards the end of the study, new complaints emerged, such as the feeling of obligation, which could be attributed to the length of the clinical study. Additionally, some patients suggested improvements such as better sleep tracking and more detailed explanations for missing data. Despite these concerns, it appears that patients have adapted to using the app.

Part C: feedback about devices

Altogether 15 questionnaires were filled in three time points. 6 from CHU, 3 from SERGAS, 1 from UL and 5 from UKCM.

✓ **How do you rate your experience with smart bracelets?**

There is a statistically significant differences between any two time points (Friedman One-Way Repeated Measure Analysis of Variance by Ranks p=0,041; Conover's post-hoc pairwise comparisons: initial-mid p=0,035, initial-final p>0,999, mid-final p=0,035).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	First	Middle	Last
Mean	6,867	6,000	6,933
Median	7,000	6,000	7,000
Std. Deviation	2,232	2,104	1,534
Minimum	3,000	2,000	4,000
Maximum	10,000	10,000	10,000
25th percentile	5,500	4,500	6,000
50th percentile	7,000	6,000	7,000
75th percentile	8,500	7,500	8,000

Table 135: Descriptive statistics of" How do you rate your experience with smart bracelets?"

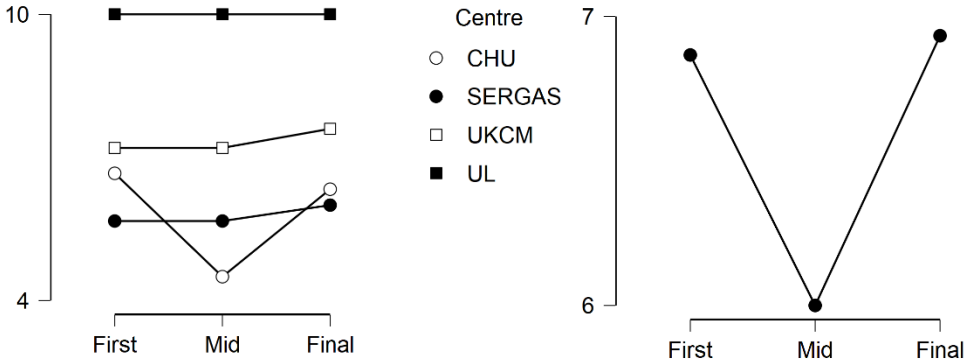

Figure 47: Descriptives plot of PERSIST block C 1st question

➔ Conclusions:

The results suggest that participants generally find their experience with smart bracelets to be satisfactory, and this perception tends to improve over time. This indicates that the use of smart bracelets in healthcare has the potential to enhance patient engagement and adherence to treatment plans. Interestingly, participants from the UL centre tended to rate their experience with smart bracelets higher, which may suggest that the use of smart bracelets is particularly effective for participants in this centre and might indicate cultural and social differences, as well as access to technology. However, it is important to note that no statistically significant differences were found, and further research is required to confirm these observations.

✓ **How do you rate your experience with mobile phone?**

As can be seen in Table 136 there are no statistically significant differences between any two time points (Friedman One-Way Repeated Measure Analysis of Variance by Ranks

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

p=0.227; Conover's post-hoc pairwise comparisons: initial-mid p=0.087, initial-final p=0.500, mid-final p=0.284).

	First	Middle	Last
Mean	6,800	7,333	6,867
Median	7,000	8,000	7,000
Std. Deviation	2,145	1,988	2,100
Minimum	2,000	2,000	2,000
Maximum	10,000	10,000	10,000
25th percentile	5,500	7,000	6,000
50th percentile	7,000	8,000	7,000
75th percentile	8,000	8,000	8,000

Table 136: Descriptive statistics of "How do you rate your experience with mobile phone?"

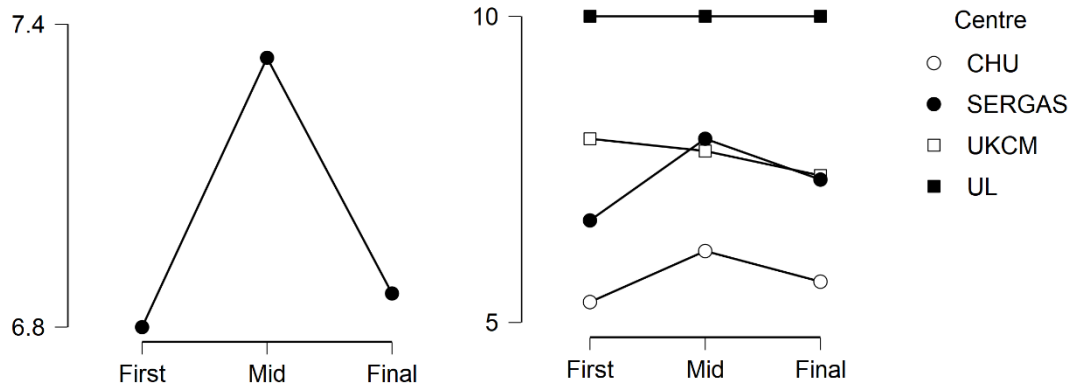

Figure 48: Descriptives plot of PERSIST block C 2nd question

➔ Conclusions:

Participants generally rated their experience with their mobile phones as satisfactory to good on average, indicating their contentment with the phones' functionality and usability.

Despite the absence of statistically significant differences between any two time points, the fact that participants' opinions did not decrease over time is a positive indication that the mobile phone experience did not deteriorate over time. The higher ratings given by participants from UL for their mobile phone experience suggest that they may have had a better experience or preference for the type of mobile phone used. The differences in ratings between centers may be influenced by various factors such as cultural and social differences, access to technology and healthcare, and personal preferences.

Overall, the study suggests that participants had a positive experience with their mobile phones, and their satisfaction levels remained stable over time.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

✓ **What would be your main complaints about using the devices?**

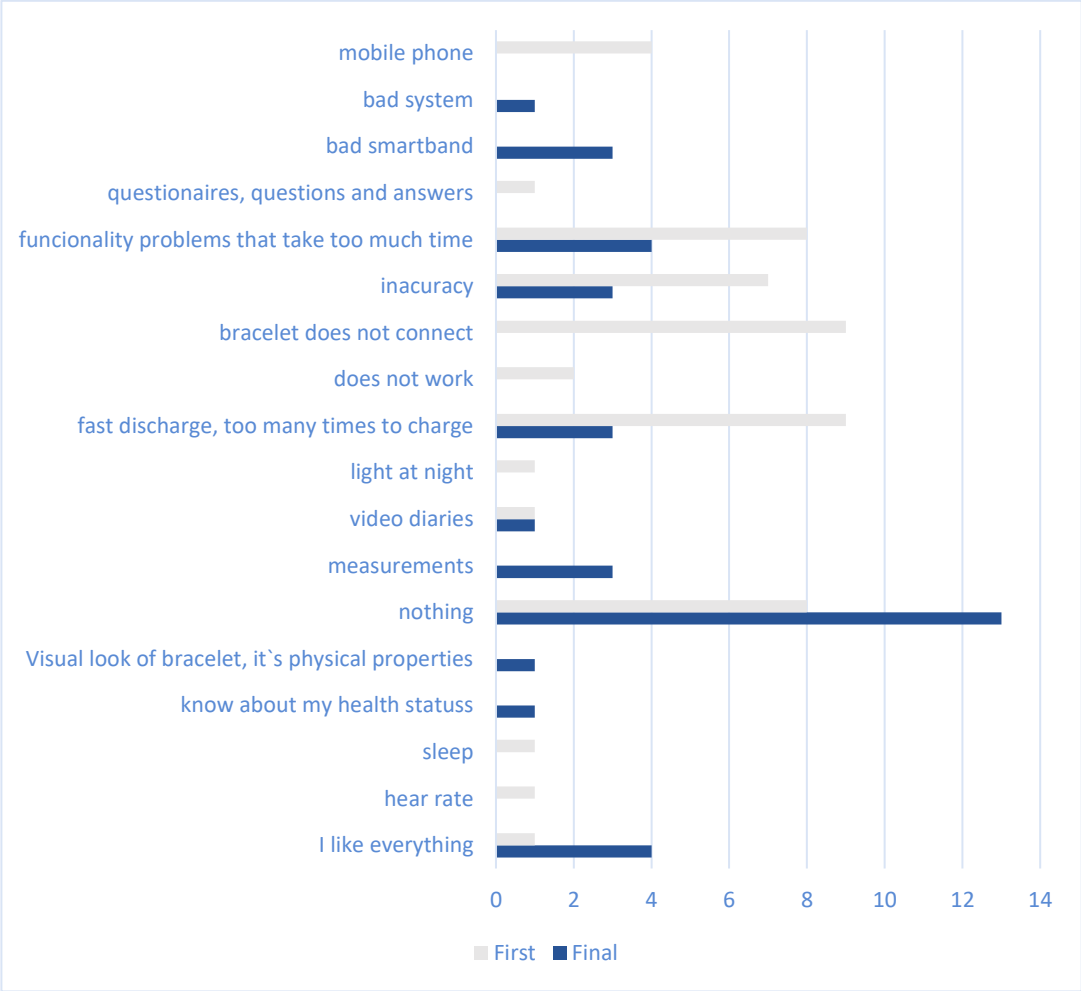

Figure 49: Patients' number of answers in the middle and at the end of study.

➔ Conclusions:

Towards the end of the study, there was an increase in patient complaints about the smart band and the overall system. This could be attributed to frequent problems with the smart band, such as blank screens and broken straps. However, it is worth noting that during the same period, the number of responses indicating a desire to maintain the system in its current state also increased. In fact, some patients even expressed enthusiasm towards the system (e.g., "I like everything") as the study progressed.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

9. Impact of conversational agent on quality of answers

Due to negative feedback received from patients during workshops it was decided not to use the conversational agent as an obligatory function of the mHealth app.

10. PHQ2 – depression sign questionnaire

In order to screen for depression among participants in PERSIST study, the Patient Health Questionnaire (PHQ-2) was used. The PHQ-2 is a shorter version of the PHQ-9 questionnaire that includes only the first two questions of the PHQ-9 [13]. The PHQ-2 questionnaire was validated by Kroenke et al. (2003) and asks about the frequency of symptoms of anhedonia and depressed mood, scoring those symptoms from 0 (not at all) to 3 (nearly every day). Thus, values lower than 2 indicate no signs of depression, while values ≥ 2 indicate caution since signs of depression are present [14; 13].

To analyse the results of the questionnaires and evaluate whether there was an evolution among the participants in PERSIST, descriptive statistics of the PHQ-2 questionnaire scores at three different time points: baseline (the first time the patient replied), month 8, and month 15 of the study can be seen in Table 137. In the case of patients who answered the questionnaire more than once at any time point, the mean of the scores was obtained.

PHQ2 Score		Baseline Score N= 66	8M ± 1 score N= 49	15M ± 1 score N= 39
Mean		1,06	1,04	0,93
Median		1,00	0,00	00,50
Standard deviation		1,108	1,568	1,050
Minimum		0	0	0
Maximum		4	6	3
Percentiles	25	0,00	0,00	0,00
	50	1,00	0,00	0,50
	75	2,00	2,00	2,00

Table 137: Descriptive statistics of the scores of the PHQ2 questionnaire at three time points (8 ± 1M and 15 ± 1M: after 7-9 and 14-16 months of baseline respectively.)

A Friedman test was performed to compare the median scores in the 3 time points. Compared to the baseline, a slight decrease in the scores, but not statistically significant, were observed at M8 ± 1 and M15 ± 1 (table 2). Neither a statistically significant difference was found with an independent Wilcoxon test comparison ($p = 0,888$ and $p = 0,200$, respectively).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Questionnaire	Basal Score N= 66	8 ± 1 M score N= 49	15 ± 1 M score N= 39	p
PHQ2 score median (IQR)	1 (0-2)	0 (0-2)	0,5 (0-2)	0,662

Table 138: Median scores comparison among three time points
8 ± 1M and 15 ± 1M: after 7-9 and 14-16 months of recruitment respectively

To verify these findings, the proportion of patients with and without signs of depression at the 3 time points were compared using the Cochrane Q test. Again, a decrease in the proportion of patients with signs of depression was observed after the baseline, at 8 ± 1 and 15 ± 1 months, although these differences were not statistically significant (Table 139). Likewise, no statistically significant differences were observed when comparing the data independently with a Mcnemar test (p= 0,302 for M 8 ± 1 and p= 0,092 for 15 ± 1M).

Questionnaire	Basal Score N= 66	8 ± 1 M score N= 49	15 ± 1 M score N= 39	p
PHQ2 n (%)				
Signs of depression (score ≥2)	28 (42,4)	14 (28,6)	11 (28,2)	0,301
No signs of depression (score <2)	38 (57,6)	35 (71,4)	28 (71,8)	

Figure 139: Comparison of the proportion of patients with and without signs of depression at 3 different time points
8 ± 1M and 15 ± 1M: after 7-9 and 14-16 months of recruitment respectively

The alerts sent to mHealth about patients' mental health (e.g. depression signs) in order for clinicians to make additional decisions about the possibility to send patients to a psychologist was not common, since the application was not fully developed. On the other hand, thanks to the statistical analysis of the results of the PHQ-2 scores obtained, it has been possible to observe a general decrease in the signs of depression among the patients. Previous studies showed that coping self-efficacy was related with depression. Thus, patients who showed higher levels of coping self-efficacy were less likely to report symptoms of depression when compared to those patients with lower levels of coping self-efficacy (Philp et al.2013). This fact invites us to think that the testing of a mHealth app still in development might have been able to increase the levels of self-efficacy of the patients participating in the project, favouring the reduction of the reported signs of depression.

11. GAD7 anxiety questionnaire

The Generalised Anxiety Disorder (GAD-7) questionnaire was used to monitor the level of anxiety of patients in the PERSIST project. The GAD-7 questionnaire measures anxiety through 7 questions that enquire about feeling nervous, anxious or on edge, the ability to control worry, the difficulty relaxing and the level of irritability and feelings of fear. The response options for each item are: 0 - Not at all, 1 - Several days, 2 - More than half the days, 3 - Nearly every day. The sum of all the answers generates a total score that ranges from 0 to 21. Indicating the degree of anxiety as follows [14; 3; 15]:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

- **Score 0-4**: Minimal Anxiety
- **Score 5-9:** Mild Anxiety
- **Score 10-14:** Moderate Anxiety
- **Score greater than 15:** Severe Anxiety

Table 135 shows the descriptive statistics of the GAD-7 questionnaire scores at the three study times.

		Baseline Score N= 59	8M ± 1 score N= 48	15M ± 1 score N=29
Mean		4,68	5,32	3,49
Median		3	2,5	2,75
Standard deviation		3,976	5,99	3,267
Minimum		0	0	0
Maximum		18	21	12
Percentiles	25	0,00	0,00	0,00
	50	3	2,5	2,75
	75	6,00	6,00	5,50

Figure 140: Descriptive statistics of the scores of the GAD-7 questionnaire at 3 different time points 8 ± 1M and 15 ± 1M: after 7-9 and 14-16 months of recruitment respectively

Patients were asked to answer GAD-7 questionnaires through the app along the project. Analysis were performed at three time points baseline (first time filled), 8 ± 1 months and 15 ± 1 months. For patients who answered the questionnaire more than once at any of these three time points, the mean score was used. A slight decrease in the scores were observed at M8 ± 1 and M15 ± 1 when compared to baseline, although this decrease was not statistically significant when using a Friedman test (Table 141). Neither were statistically significant differences found when analysed with the Wilcoxon test ((p = 0,290 and p = 0,056, respectively).

Additionally, comparison with the Wilcoxon test between the baseline score and the score at 8 $\pm$ 1 or 15 $\pm$ 1 M, did not find statistically significant differences (p = 0,290 and p = 0,056, respectively).

Questionnaire	Basal Score N= 59	8 $\pm$ 1 M score N= 48	15 $\pm$ 1 M score N= 29	p
GAD-7 score median (IQR)	3 (0-6)	2,5 (0-6)	2,75 (0-5,5)	0,375

Figure 141: Median scores comparison among 3 time points by Friedman test

Table 142 shows the frequencies and percentages of the different anxiety levels at each of the study time points. As it can be seen, frequencies and percentages of anxiety markedly decreased along the project.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Questionnaire	Basal Score N= 59	8 ± 1 M score N= 48	15 ± 1 M score N= 29
Minimal anxiety n (%)	30 (50,8)	23 (47,9)	17 (58,6)
Mild anxiety n (%)	22 (37,3)	19 (39,6)	11 (37,9)
Moderate anxiety n (%)	6 (10,2)	0 (0)	1 (3,4)
Severe anxiety n (%)	1 (1,7)	6 (12,5)	0 (0)

Figure 142: Frequencies and percentages of the different levels of anxiety at different times

In order to identify statistically significant differences in anxiety levels throughout the follow-up, the first two levels and the last two levels were assembled in two different groups, as follows:

- ➔ **Minimal or Mild Anxiety** (Total score: 0-9)
- ➔ **Moderate or Severe Anxiety** (Total score: >9)

The two groups were compared using a Chochrane Q test (Table 138). A decrease in the percentage of patients with moderate to severe anxiety from baseline (11.9%) to 15 ± 1M (3,4%) was observed. However, not statistically significant differences were found. Likewise, by comparing independently the same above groups with the McNemar test, no statistically significant differences were obtained (p= 0,625 baseline vs M8 ± 1 and p= baseline vs 0,375 M15 ±1).

Despite not having detected statistically significant differences in the anxiety levels of the patients throughout the project, it is necessary to consider the marked decrease in the levels of anxiety detected. As we have seen regarding the degree of depression of the patients hitherto, it has been proposed in previous studies that by increasing the levels of self-efficacy of the patients, it would be possible to reduce the levels of anxiety [16]. Thereby, the use of the mHealth app might be able to help patients reduce their anxiety levels by enhancing their self-efficacy. These results are encouraging, since the increase in self-efficacy and, with it, the decrease in the degree of depression and anxiety of cancer patients, help them to improve their adherence to treatment and follow-up, as well as communication with the health care professionals [3; 16; 17]. A further development of the PERSIST system would be necessary to show beyond doubt the ability of the mHealth app to improve the self-efficacy of patients.

Likewise, by comparing independently the same above groups with the McNemar test, no statistically significant differences were found (p= 0,625 baseline vs M8 ± 1 and p= baseline vs 0,375 M15 ±1).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Questionnaire	Basal Score N= 59	8 ± 1 M score N= 48	15 ± 1 M score N= 29	p
GAD-7 (%)				
Minimal or Mild Anxiety (score 0-9)	52 (88,1)	42 (87,5)	28 (96,6)	0,607
Moderate or Severe Anxiety (score >9)	7 (11,9)	6 (12,5)	1 (3,4)	

Table 143: Comparison of the proportion of patients with minimal-mild anxiety levels and moderate-severe anxiety at 3 different time points

➔ Conclusions:

Although statistically significant differences in anxiety levels were not detected throughout the PERSIST project, there was a marked decrease in the levels of anxiety observed. The frequency and percentage of anxiety markedly decreased along the project, and the percentage of patients with moderate to severe anxiety decreased from baseline to 15 + 1M. This suggests that the PERSIST study may have had a positive impact on the anxiety levels of the patients. However, further studies with larger sample sizes and longer follow-up periods may be necessary to confirm these findings.

12. Description of negative outcomes over time

Each hospital gathered a list of patient negative outcomes over time: hospitalisation, exacerbations, treatment adherence, depression, recurrence, and drug escalation etc., for patients who continued to participate till the end of the study.

Patients from UL

Among 47 UL patients in the course of the clinical study, 3 had cancer recurrence with metastasis, one patient died, and one presented a new tumour in a second localization (skin) (see table 144). Altogether 14 negative outcomes were registered for 13 patients while they were in study.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	Diagnosis	Event	Date	Patient left the trial
1	BRC	Second localisation cancer (skin basalioma). excision of skin tumour	05.2021.	
2	CRC	Tumour recurrence in the rectum. MTS in intramesorectal lymph nodes. MTS in the right iliac lymph nodes. MTS in the right iliac lymph nodes	02.2022.	10.2021.
3	CRC	Recurrence of cancer and metastasis in liver.	02.2022.	
4	BRC	Complication after treatment (hormone therapy/ Tamoxifen): endometrial polyposis Hysteroresectoscopy of endometrial polyp.	09.2021.	04.2022.
5	BRC	Progression of the disease. Multiple metastasis, intoxication, SARS COV2 infection, exitus letalis.	12.2021.-02.2022.	02.2022.
6	BRC	Complication after treatment (hormone therapy/ Tamoxifen): endometrial polyposis	09.2020.	
7	BRC	Complication after treatment (hormone therapy/ Tamoxifen): endometrial polyposis	01.2021.	
8	CRC	Treatment complications. Postoperative hernia.	09.2020.-10.2020.	
9	CRC	Treatment complications. Descendo-rectoastomosis stenosis.	10.2022.-11.2022.	
10	BRC	Toxic auditory nerve damage (after chemotherapy). Complication after treatment (platinum based chemotherapy)	year after operation	05.2021.
11	BRC	Complication after treatment (hormone therapy/ Tamoxifen): hot flashes ... (surgery): pain and sensory loss in the left arm.		
12	BRC	Complication after treatment / surgery-implant capsular contraction	09.-10.2022.	
13	BRC	Complication after treatment / surgery.... Late seroma around implant	07.2022.	05.2021.
		Implant evacuation because of infected seroma	08.-08.2022.	

Table 144. UL patients' negative outcome summary

Patients from CHU

Altogether 12 of CHU patients had some negative outcomes. Only one of them experienced recurrence and one suffered from prostatic cancer.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

	Diagnosis	Event	Date
1	CRC	Voluminous left inguino-scrotal hernia. Persistent perineal pain and fluctuating irritation	Hernia 05/2021
2	CRC	Dec 2021**:Recurrence** of pelvic lesions previously in complete response and the appearance of left hepatic lesion. February 2022 Surgery for two pelvic adenopathies as well as a hepatectomy and a hysterectomy with excision of the vaginal fundus. June 2022 new hepatic lesion .	Recurrence 05/2021
3	CRC	None	
4	CRC	April 2021 infectious phenomenon, whose origin is not clear.	Infection 04/2021
5	CRC	Internal thrombosis of the left lower limb	04.2022
6	CRC	June 2021 Intense hyperfixing bilateral prostatic **neoplasia.** Suspicion of invasion of the left seminal vesicle. Bilateral secondary iliac lymphadenopathy.	06.2021
7	CRC	Bile salt malabsorption . Recurrence of diarrhoea	04.2021
8	BRC	Joint pain	09.2021
9	BRC	Pain , Memory lost, Fatigue	03.2021
10	BRC	Fatigue with a lack of energy Increase cholesterol in the context of treatment with Femara.	07.2021
11	BRC	Joint stiffness. Fatty liver	08.2021
12	BRC	Pain upper limbs	06.2021
13	BRC	Hot flashes and perspiration Joint pain in the hands. Fatigue	09.2022
14	BRC	Persistent fatigue and difficulty concentrating. Recurrent headaches especially with fatigue. Articular pain. Jump right thumb. Cyclical hot flashes.	08.2022

Table 145. CHU patients' negative outcome summary

Patients from SERGAS

Altogether 8 of SERGAS patients had some negative outcomes. Two of them experienced recurrence and one a new cancer (see table 134).

	Diagnosis	Event	Date
1	BRC	Anxiety problems	05.21
2	BRC	Anxiety problems	11.21
3	BRC	Recurrence in lung, lymph nodes, and liver	05.22
4	BRC	Bone recurrence	01.22
5	CRC	Exploration laparotomy + flange section + adhesiolysis + incidental appendectomy	12.21
6	CRC	variations in CEA marker	Along the project
7	CRC	Aortic valve surgery. Bypass	10.22
8	CRC	Tubular adenoma with low histological grade dysplasia	12.21

Table 146. SERGAS patients' negative outcome summary

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Patients from UKCM

From 40 UKCM patients in the course of clinical trial 1 patient had recurrence of cancer and had metastasis and left the study. Altogether 4 negative outcomes have been registered for 4 patients while they were in study.

	Diagnosis	Event	Date	Patient left the trial
1	BRC	Recurrence of cancer and metastasis		6.2021.
2	CRC	Using different drugs: metformin, perindopril	08.2022	
3	BRC	Lumpectomy of breast, right breast, upper-outer quadrant of breast (aborted)	05.2022	
4	BRC	anxiety disorder	02.2022	

Table 147. CHU patients' negative outcome summary

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Clinicians' perspective of PERSIST system

1. User acceptance (SUS) for mClinician web and app versions

User acceptance questionnaires were also distributed to clinicians working with mClinician web and App in the 4 participating hospitals. Two rounds of answers (first only with mClinician web version in August 2022, second with mClinician app version at the end of the study in October 2022) were performed.

The sum score of the points of 10 questions in each round can be seen in Figure 50 and 51. According to the definition of the System Usability level (Table 148) most of the clinicians (81,55%) who replied thought that system was not easy to use and had some usability issues. The other (12,5%) considered that the system was acceptable to good and only one clinician thought that it had an excellent usability.

Level	Definition
<=50	Not easy to use
50-70	Experiencing usability issues
70-85	Acceptable to good
>85	Excellent usability

Table 148 The definition of system usability level

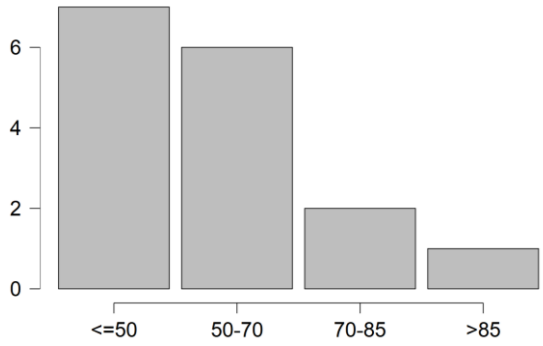

Figure 50. mClinician score group 1st round

	Frequency	Percent	Valid Percent
<=50	7	43,750	43,750
50-70	6	37,500	37,500
70-85	2	12,500	12,500
>85	1	6,250	6,250
Missing	0	0,000	
Total	16	100,000	

Table 149: Frequencies for score group 1.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

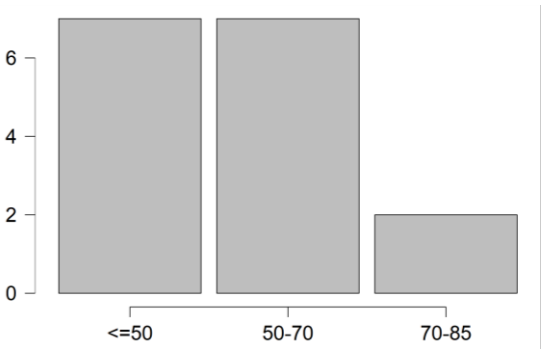

Figure 51. mClinician score group 2nd round

	Frequency	Percent	Valid Percent
<=50	7	43,750	43,750
50-70	7	43,750	43,750
70-85	2	12,500	12,500
>85	1	0,000	43,750
Missing	0	100,000	
Total	16	43,750	

Table150: Frequencies for score group 2.

In comparison between both replies there was no statistically significant difference between scores at both time points (p=0,784, Wilcoxon test), see Table 151.

	1st score group	2nd score group
Mean	1,813	1,688
Median	2,000	2,000
Std. Deviation	0,911	0,704
Minimum	1,000	1,000
Maximum	4,000	3,000
25th percentile	1,000	1,000
50th percentile	2,000	2,000
75th percentile	2,000	2,000

Table 151: Descriptive statistics of both score groups.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

➔ Conclusions:

The results of the user acceptance questionnaire distributed to clinicians using the mClinician web and app versions show that there were some usability issues identified by most of the clinicians (81,55% in the first round and 87,5% in the second round). However, it is positive that the scores did not significantly differ between the two rounds, indicating that the app version did not introduce new usability issues. Additionally, one clinician rated the usability as excellent in the first round, which is a positive indication. Further investigation and improvement of the identified usability issues could potentially lead to increased acceptance and adoption of the mClinician system among clinicians.

The following subsections include the SUS questionnaire individual analysis.

✓ I think that I would like to use this system frequently (V1)

	V1 first				V1 second			
	CHU	SERGAS	UKCM	UL	CHU	SERGAS	UKCM	UL
Median	1,500	4,500	3,000	2,000	3,000	3,500	2,000	3,000
Mode	1,000	4,000	1,000	1,000	3,000	3,000	1,000	3,000
Minimum	1,000	4,000	1,000	1,000	2,000	3,000	1,000	1,000
Maximum	3,000	5,000	4,000	3,000	4,000	4,000	3,000	3,000
25th percentile	1,000	4,250	1,000	1,000	2,750	3,250	1,000	2,000
75th percentile	2,250	4,750	4,000	3,000	3,250	3,750	3,000	3,000

Table 152: Descriptive statistics of V1

Although there seem to be some slight differences between the two time points in some of the centres, overall, the difference between both responses are not statistically significant (p=0.725, Wilcoxon sign test).

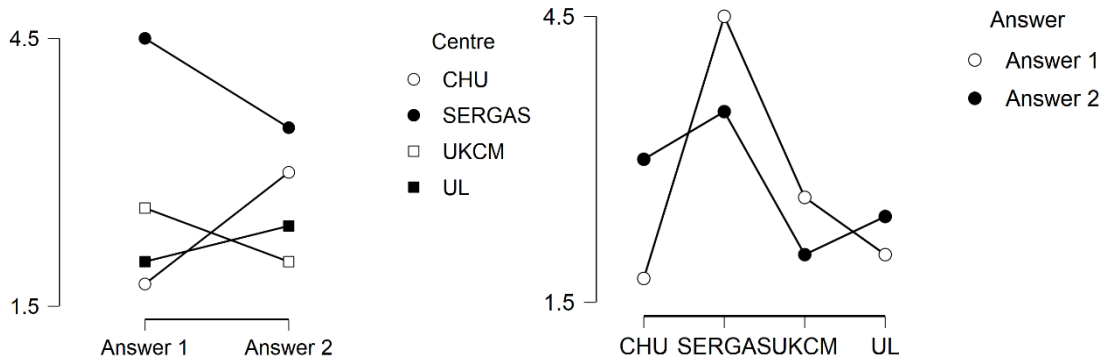

Figure 52: Comparison of answers to V1 by centres

➔ Conclusions:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

The data shows that, on average, clinicians are not enthusiastic about using the system frequently, and there were no statistically significant differences between the two time points. It is worth noting, however, that there were slight differences between responses from different centres, which could be explored further to identify possible reasons for these differences and potentially address them in future iterations of the system.

✓ I found the system unnecessarily complex (V2)

	V2 first				V2 second			
	CHU	**SERGAS**	**UKCM**	**UL**	**CHU**	**SERGAS**	**UKCM**	**UL**
Median	2,500	2,000	4,000	3,000	2,500	2,500	4,000	3,000
Mode	2,000	2,000	4,000	3,000	1,000	2,000	3,000	3,000
Minimum	2,000	2,000	1,000	2,000	1,000	2,000	3,000	3,000
Maximum	4,000	2,000	5,000	5,000	4,000	3,000	5,000	4,000
25th percentile	2,000	2,000	2,000	3,000	1,750	2,250	3,000	3,000
75th percentile	3,250	2,000	4,000	4,000	3,250	2,750	4,000	3,000

Table 153: Descriptive statistics of V2

Although some differences are shown between time points in some of the sites, no statistically significant differences were found overall (p=0,813, Wilcoxon sign test).

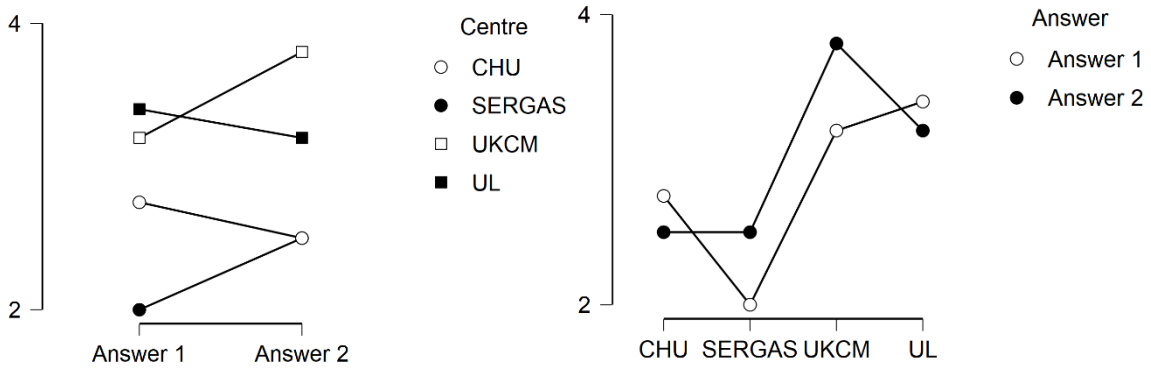

Figure 53: Comparison of answers to V2 by centres

➔ Conclusions:

The results show that clinicians, on average, did not rate the system positively in terms of complexity in both first reply and the second. However, it is worth noting that there were no statistically significant differences in the responses between the two time points and between the different centres. This suggests that the clinicians' perceptions of the system remained consistent over time and were similar across different hospitals. It may be

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

beneficial to further investigate the reasons behind the clinicians' perception of the system and make improvements to address their concerns.

- ✓ I thought the system was easy to use (V3)

	V3 first				V3 second			
	CHU	**SERGAS**	**UKCM**	**UL**	**CHU**	**SERGAS**	**UKCM**	**UL**
Median	4,000	4,000	4,000	2,000	4,000	4,000	3,000	3,000
Mode	4,000	4,000	4,000	2,000	4,000	4,000	3,000	3,000
Minimum	3,000	4,000	2,000	1,000	4,000	4,000	2,000	3,000
Maximum	4,000	4,000	5,000	4,000	4,000	4,000	3,000	4,000
25th percentile	3,750	4,000	3,000	2,000	4,000	4,000	2,000	3,000
75th percentile	4,000	4,000	4,000	3,000	4,000	4,000	3,000	4,000

Table 154: Descriptive statistics of V3

The difference between both responses is not statistically significant (p>0,999, Wilcoxon sign test).

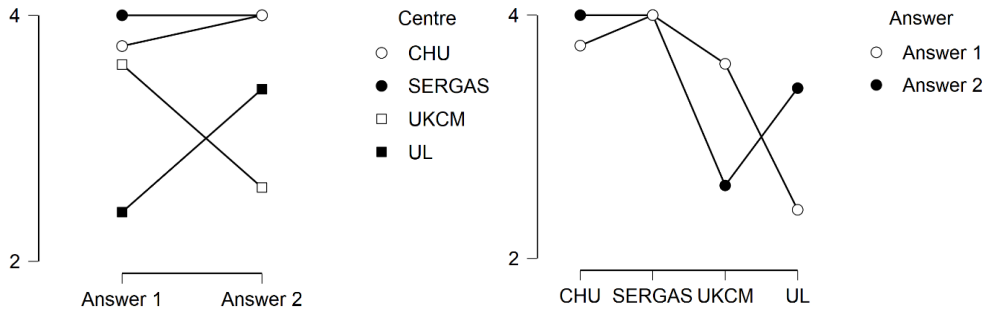

Figure 54: Comparison of answers to V3 by centres

- ➔ Conclusions:

The majority of clinicians found the system easy to use (with a median and mode score of 4 out of 5) in both third question rounds. Additionally, there were no statistically significant differences between the first and second rounds of responses, indicating that the system's ease of use remained consistent over time. However, it's worth noting that clinicians from UL initially had a lower evaluation for ease of use compared to clinicians from other hospitals, but this improved in the second round of responses.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

- ✓ I think that I would need the support of a technical person to be able to use this system (V4)

	V4 first				V4 second			
	CHU	**SERGAS**	**UKCM**	**UL**	**CHU**	**SERGAS**	**UKCM**	**UL**
Median	1,000	2,000	1,000	3,000	1,500	1,500	1,000	3,000
Mode	1,000	2,000	1,000	3,000	1,000	1,000	1,000	3,000
Minimum	1,000	2,000	1,000	3,000	1,000	1,000	1,000	1,000
Maximum	2,000	2,000	2,000	5,000	2,000	2,000	3,000	4,000
25th percentile	1,000	2,000	1,000	3,000	1,000	1,250	1,000	2,000
75th percentile	1,250	2,000	2,000	4,000	2,000	1,750	3,000	3,000

Table 155: Descriptive statistics of V4

The difference between both responses is not statistically significant (p=0,429, Wilcoxon sign test).

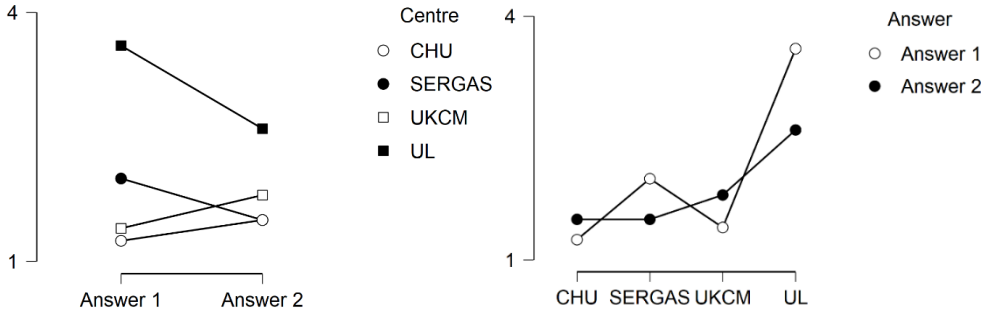

Figure 55: Comparison of answers to V3 by centres

➔ Conclusions:

Firstly, it seems that the system was generally perceived as easy to use by clinicians across all centres, with median and mode scores of 4.0 for ease of use. Additionally, there was no significant difference in perceptions of ease of use between the beginning and end of the study. Secondly, it's worth noting that while clinicians felt they needed technical support to use the system, there were no significant differences in perceptions of the system between the beginning and end of the study. This could suggest that the technical support provided was effective in helping clinicians become more comfortable with the system over time. Overall, these findings could be seen as positive indicators that the mClinician app and web systems are generally user-friendly and can be effectively implemented with appropriate technical support.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

✓ I found the various functions in this system were well integrated (V5)

	V5 first				V5 second			
	CHU	**SERGAS**	**UKCM**	**UL**	**CHU**	**SERGAS**	**UKCM**	**UL**
Median	1,000	3,000	3,000	4,000	3,000	3,500	2,000	4,000
Mode	1,000	3,000	1,000	2,000	3,000	3,000	2,000	4,000
Minimum	1,000	3,000	1,000	2,000	3,000	3,000	1,000	2,000
Maximum	2,000	3,000	5,000	5,000	4,000	4,000	3,000	5,000
25th percentile	1,000	3,000	2,000	2,000	3,000	3,250	2,000	3,000
75th percentile	1,250	3,000	4,000	4,000	3,250	3,750	3,000	4,000

Table 156: Descriptive statistics of V5

The difference between both responses is not statistically significant (p=0,235, Wilcoxon sign test).

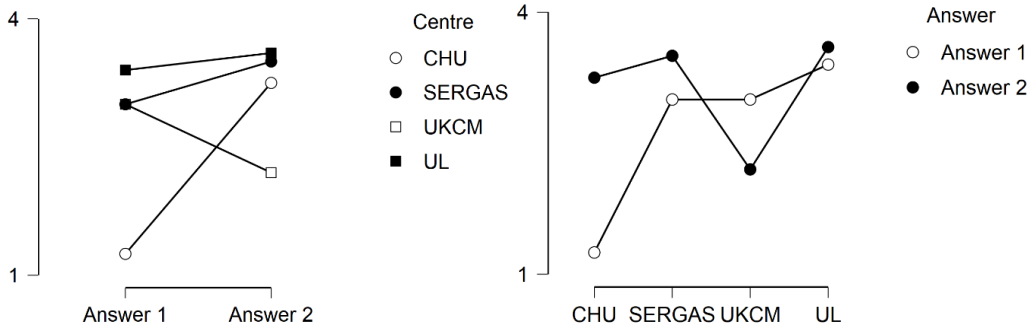

Figure 56: Comparison of answers to V5 by centres

➔ Conclusions:

Despite some initial variation in opinions on the integration of various functions in the system, all clinicians improved their opinion over time except for those from UKCM. Additionally, there were no statistically significant differences detected between the initial and final responses or between centres, which suggests that the system was well received overall and that its various functions were well integrated.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

✓ I thought there was too much inconsistency in this system (V6)

	V6 first				V6 second			
	CHU	**SERGAS**	**UKCM**	**UL**	**CHU**	**SERGAS**	**UKCM**	**UL**
Median	2,500	2,500	3,000	4,000	2,000	3,000	4,000	3,000
Mode	2,000	2,000	3,000	4,000	2,000	2,000	3,000	2,000
Minimum	2,000	2,000	1,000	2,000	2,000	2,000	3,000	2,000
Maximum	3,000	3,000	5,000	5,000	3,000	4,000	5,000	5,000
25th percentile	2,000	2,250	2,000	4,000	2,000	2,500	3,000	2,000
75th percentile	3,000	2,750	3,000	4,000	2,250	3,500	4,000	4,000

Table 157: Descriptive statistics of V6

The difference between both responses is not statistically significant (p=0,832, Wilcoxon sign test).

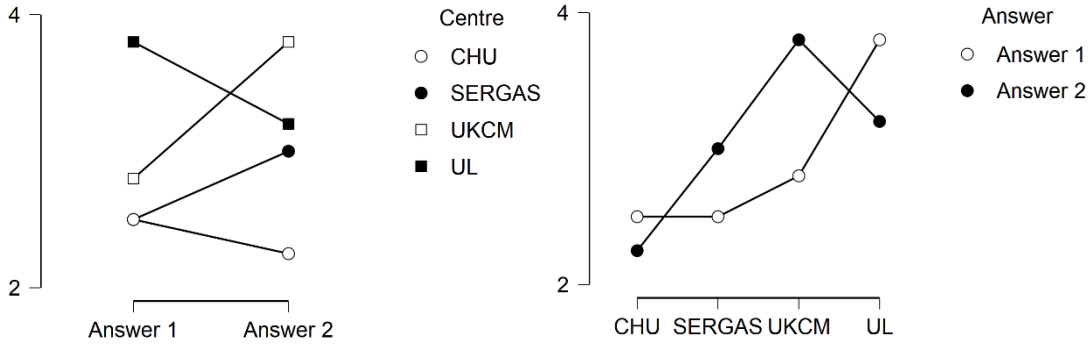

Figure 57: Comparison of answers to V6 by centres

➔ Conclusions:

It's important to note that there were no statistically significant differences between the responses from the different hospitals, which suggests that the level of inconsistency was similar across all centres. Additionally, it's possible that the feedback provided by clinicians during the study may help improve the consistency of the system in the future. And it should be noted that the system was a co-creation process meaning that it was possible to experience some level of inconsistency.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

- ✓ I would imagine that most people would learn to use this system very quickly (V7)

	V7 first				V7 second			
	CHU	**SERGAS**	**UKCM**	**UL**	**CHU**	**SERGAS**	**UKCM**	**UL**
Median	4,000	4,000	4,000	3,000	2,500	3,500	3,000	4,000
Mode	4,000	4,000	4,000	3,000	2,000	3,000	4,000	4,000
Minimum	3,000	4,000	2,000	1,000	2,000	3,000	1,000	3,000
Maximum	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000
25th percentile	3,750	4,000	2,000	3,000	2,000	3,250	2,000	3,000
75th percentile	4,000	4,000	4,000	3,000	3,250	3,750	4,000	4,000

Table 158: Descriptive statistics of V7

The difference between both responses is not statistically significant (p=0,587, Wilcoxon sign test).

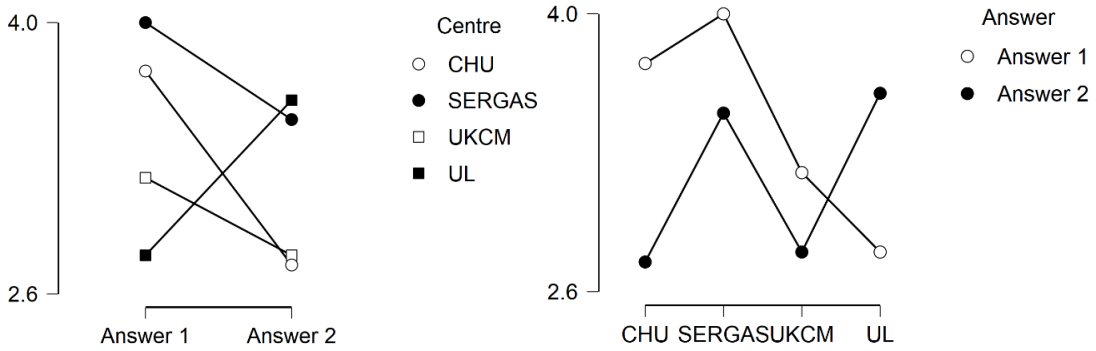

Figure 58: Comparison of answers to V7 by centres

➔ Conclusions:

The median response across all four hospitals was 4,0, indicating that most clinicians believe that people would learn to use the system quickly. The mode response was also 4.0, which reinforces the idea that the majority of clinicians hold this belief. The difference in responses between the beginning and end of the study was not statistically significant, suggesting that the clinicians' views on this question did not change over time. Overall, these findings suggest that the clinicians who participated in the study generally believe that people would be able to learn to use the system being evaluated in this question quickly and easily.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

✓ I found the system very cumbersome/awkward to use (V8)

	V8 first				V8 second			
	CHU	**SERGAS**	**UKCM**	**UL**	**CHU**	**SERGAS**	**UKCM**	**UL**
Median	4,500	2,500	4,000	4,000	3,500	2,000	4,000	3,000
Mode	4,000	2,000	5,000	5,000	4,000	1,000	4,000	1,000
Minimum	4,000	2,000	1,000	2,000	2,000	1,000	3,000	1,000
Maximum	5,000	3,000	5,000	5,000	4,000	3,000	4,000	4,000
25th percentile	4,000	2,250	2,000	3,000	2,750	1,500	4,000	1,000
75th percentile	5,000	2,750	5,000	5,000	4,000	2,500	4,000	4,000

Table 159: Descriptive statistics of V8

The difference between both responses is not statistically significant (p=0,164, Wilcoxon sign test).

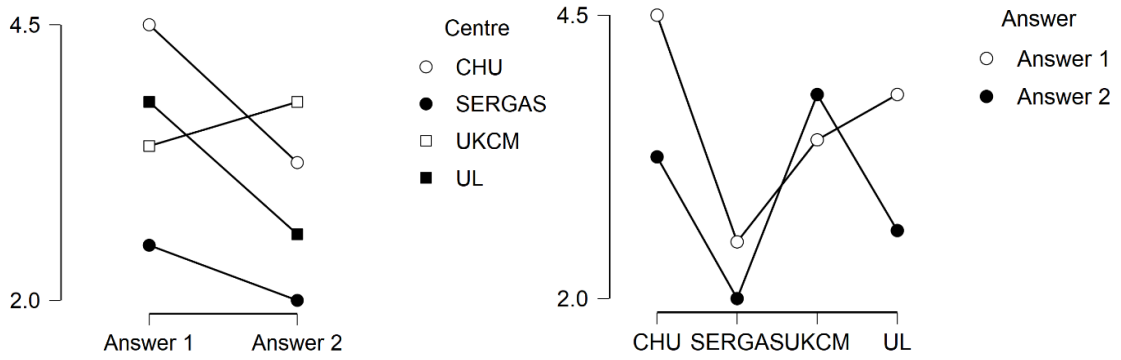

Figure 59: Comparison of answers to V8 by centres

➔ Conclusions:

The median score is relatively low, and all hospitals showed a decrease in evaluation over time. However, it is worth noting that there were no statistically significant differences between hospitals in terms of their evaluation of the system's usability, and it is possible that further improvements or adjustments could be made to the system to address some of the usability issues identified.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

✓ I felt very confident using the system (V9)

	V9 first				V9 second			
	CHU	**SERGAS**	**UKCM**	**UL**	**CHU**	**SERGAS**	**UKCM**	**UL**
Median	5,000	3,500	4,000	3,000	3,000	3,500	3,000	3,000
Mode	5,000	3,000	4,000	4,000	3,000	3,000	2,000	3,000
Minimum	4,000	3,000	3,000	1,000	2,000	3,000	2,000	2,000
Maximum	5,000	4,000	4,000	4,000	3,000	4,000	4,000	4,000
25th percentile	4,750	3,250	3,000	2,000	2,750	3,250	2,000	3,000
75th percentile	5,000	3,750	4,000	4,000	3,000	3,750	4,000	4,000

Table 160: Descriptive statistics of V9

The difference between both responses is not statistically significant (p=0,132, Wilcoxon sign test).

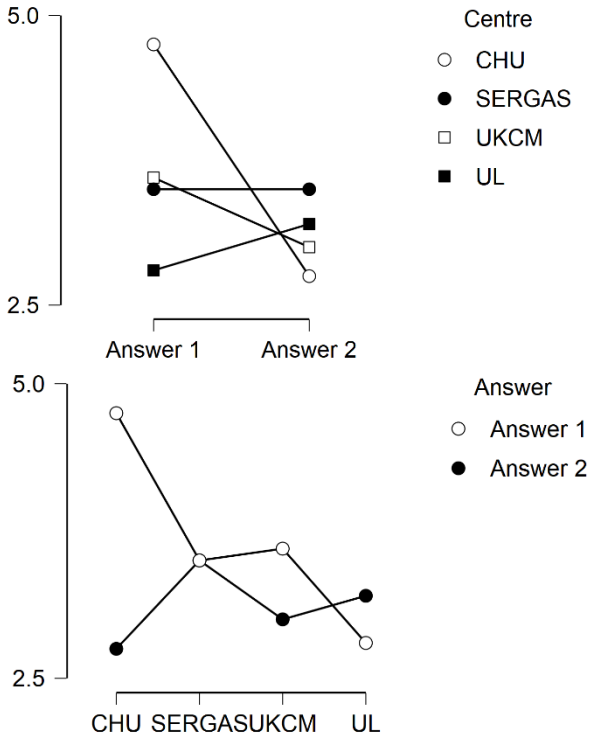

Figure 60: Comparison of answers to V9 by centres

➔ Conclusions:

The median and mode scores for all hospitals were not particularly high, and there was no statistically significant difference between the initial and final responses. However, it is worth noting that the CHU clinicians initially reported feeling very confident using the

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

system, which could suggest that PERSIST has the potential to inspire confidence in users. Further research may be needed to explore this possibility.

- ✓ I needed to learn a lot of things before I could get going with this system (V10)

	V10 first				V10 second			
	CHU	**SERGAS**	**UKCM**	**UL**	**CHU**	**SERGAS**	**UKCM**	**UL**
Median	2.000	3.000	1.000	3.000	2.000	2.500	1.000	2.000
Mode	2.000	2.000	1.000	3.000	2.000	2.000	1.000	2.000
Minimum	1.000	2.000	1.000	3.000	2.000	2.000	1.000	2.000
Maximum	2.000	4.000	2.000	5.000	2.000	3.000	2.000	4.000
25th percentile	1.750	2.500	1.000	3.000	2.000	2.250	1.000	2.000
75th percentile	2.000	3.500	1.000	4.000	2.000	2.750	2.000	3.000

Table 161: Descriptive statistics of V10

The difference between both responses is not statistically significant (p=0,430, Wilcoxon sign test).

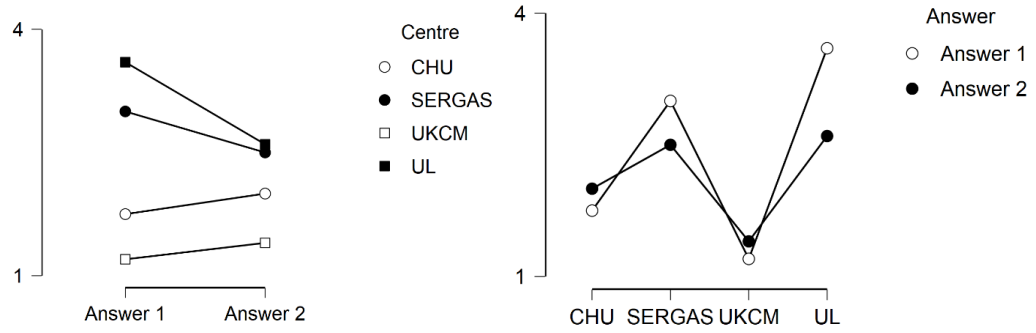

Figure 61: Comparison of answers to V10 by centres

➔ Conclusions:

The clinicians from UL and SERGAS, who initially reported needing to learn a lot of things before they could get going with the system, improved their opinion slightly by the end. This suggests that with some training and experience, the clinicians were able to gain confidence in using the system. It also indicates that the system may have potential to be effective with further refinement and development.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

2. Clinicians workshops and/or individual feedbacks

UL clinicians workshops

Workshop 1 (clinicians involved in study and with University of Latvia - 6)

Videos from EMODA shown (additional video from Technical advisory board sent later by email; mClinician web and app versions shown).

Workshop 2 (clinicians outside clinical trial - 21) – information about the project and its results given in a ppt to a broad range of clinicians (not only oncologists, but including, for example, gastroenterologists and ophthalmologists).

Summary of comments received:

Possible use of the PERSIST system into clinical practice could be following the alerts and checking whether a patient needs additional consultation with a clinician. In case of that additional examinations and tests could be done to patients allowing them to detect whether patients` health has gotten worse or the cancer has recurred. This allows timely detection of any problems and the treatment in early phases not only means better recovery chances for patients but reduces the costs. Suggestion: to avoid adding the workload for clinicians, the specifically trained nurses help for evaluation of alerts could be used.

In case of higher cardiovascular risk detection additional consultations could be done with cardiologists to evaluate patients` health and minimise the further risk of actual negative cardiovascular events.

Analysis of emotions and video-diaries lets to timely detect mental health problems patient can experience. For example, in case symptoms of depression occur, the patient can be referred to a psychologist or psychotherapist.

Summary of patients` biological parameters gained from smart-band (e.g., blood pressure, heart rate, physical activity, sleep) evaluated in visit with clinician allows to gain insight of patients` lifestyle before last check-up. In case of worsening measurements suggestions and reminders to improve them could be made. Getting back on the track of a healthier lifestyle allows us to avoid possible health problems in future. Suggestion: built in the app reminders about following a healthy lifestyle, for example, to walk at least 7000 steps a day.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

CHU clinicians workshops

1-Workshop 1 Colorectal Cancer Healthcare professionals of CHUL

Concerning cancer recurrence, colorectal cancer survivors in Belgium are followed for 5 years. Therefore, the mHealth of PERSIST is especially interesting in the risk of recurrence beyond 5 years. An increased number of consultations and exams should be offered if the risk shows a statistically significant difference compared to the normal population. Once again, the recurrence risk of cancer survivors should be as personalised as possible by adjusting it to normal populations of the same age, gender, ethnicity, etc.

Despite some limitations and recommendations, the workshop's participants generally agree that an mHealth app can be a valuable tool for healthcare professionals treating colorectal cancer. The app's ability to personalise alerts and offer closer physical and psychological follow-up during chemotherapy breaks would be particularly beneficial for patients undergoing treatment.

They also pointed out that the tool would be of great help when by replacing a colleague, treatment and follow-up need to be decided for cancer survivors they have not seen before.

Physicians who participated in these workshops are eager to see an enhanced version with alerts even more personalised by adjusting them with age, sex, jobs, etc. They emphasised that mHealth alerts could help them to be even closer to the real situation of the cancer survivors they follow.

Moreover, the specialists suggest that the app could be more useful to general practitioners who see a high number of patients with various pathologies every day. They believe that such a solution could save time while increasing the accuracy of treatment and follow-up decisions. Regarding cancer survivors, general practitioners need to periodically monitor the cancer survivors who underwent chemotherapy or radiotherapy with cardiotoxicity.

The app's ability to provide personalised cancer recurrence risk information is seen as a particularly valuable feature, especially beyond the five-year follow-up period. The mClinician of PERSIST would offer an increased number of consultations and examinations if the risk presents a statistically significant difference compared to the normal population.

Finally, Oncologists have expressed a desire to have a mHealth app similar to PERSIST one for patients undergoing colorectal cancer treatment. This would allow them closer physical and psychological follow-up during chemotherapy breaks. They are excited to participate in a future project establishing this type of mHealth.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Workshop 2 Breast Cancer Physician of CHUL

Focus on emotional aspects of quality of life: The physicians recognize the importance of monitoring not just the physical aspects of treatment, but also the quality of life of patients. PERSIST helps identify patients who struggle with emotional issues. They practise a holistic approach to medicine and care as much about the physical follow-up as the quality of life of their patients. Aware that some patients have difficulty reporting dissatisfaction to them, PERSIST allows identifying these patients. The physicians are convinced that some survivors participating in the study were more comfortable with the digital PERSIST system for reporting emotional issues.

Risk of recurrence: Physicians recognize the great added value of the mClinician app in its ability to predict the risk of recurrence. This capacity will help clinicians develop more tailored follow-up plans based on the risk levels of each survivor. The outcomes of the following months will allow us to assess the accuracy of the risk of recurrence predicted by the mClinician. The next version may need adjustments for enhancing even more targeted and effective care for survivorship. In line with this, the AI systems should propose personalised thresholds of vital parameters for adjusting the alerts to each patient physiology. This will allow for a more tailored approach to patient care, which will improve outcomes.

Thus patients with a higher risk of recurrence will be offered an increased number of follow-up consultations. This will ensure that patients receive the necessary support and care they need to manage their condition effectively.

Overall, the workshop highlights the potential benefits of incorporating AI into breast cancer treatment to provide personalised care and support for patients.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

SERGAS clinicians workshops

Colorectal cancer oncologists

Clinicians from CRC value the PERSIST as pretty good, generally speaking. The follow-up of the symptoms that the system does for detecting relapse have been highlighted as one of the most useful things of the referenced PERSIST system. Moreover, oncologists perceive that the PERSIST system has the potential to make patients aware about a healthy lifestyle and objectively detect its benefit. Considering this, clinicians specialised in CRC consider that some of the lifestyle factors that can be improved thanks to this system are physical activity, blood pressure, heart rate and even reduce tiredness. Additionally, they think that using the PERSIST mHealth app makes patients more aware of important symptoms which can make them consult about them and, thus, proper actions to change habits earlier can be taken.

Taking a look at the cons, to oncologist opinion, the PERSIST system might be useful for primary medicine since it can alert about decompensations of the general pathology that are normally followed in this field. All the information about diagnosis and therapy is already known by oncologists, so it is not that useful for them, but it could be for primary care clinicians.

As a conclusion, those are general conclusions about clinicians opinions by using the mClinician web and app, but they feel that they did not have time enough to evaluate the system properly.

Breast cancer oncologists

BRC oncologists have rated the PERSIST system as notably good. In general, they have transmitted their interest in using this tool in their regular practice, especially when it comes to monitoring the functional status of the patients. To their view, the most useful aspect of the PERSIST system would be its capacity of recording and monitoring the biological parameters of the patients, such as blood pressure and heart rate. Together with that, they consider an important pro the ability of the mHealth app to encourage patients to exercise regularly. It is precisely physical activity that is most valued as something that can be personalised for each breast cancer patient. Generally speaking, breast cancer oncologists find the PERSIST system quite useful and would be interested in using it in their general practice. However, some of them have opined that it might be more comfortable to work if the mClinician app were integrated in the web version, all together.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

UKCM clinicians workshops

Clinical workshop: 13th September 2022 - CTC and data collection.

In the workshop, the clinicians' made a plan for the last CTCs collections. They overview the data that they need to insert in the mHealth clinicians. Still, some commented on the usability of the mClinician platform (time-consuming manual calculation of the measurements, data saving problem etc.) but they expressed overall satisfaction. The clinicians pointed out that the system can be a good supporting tool for the patients and clinicians. From the patient's point of view, they can follow their overall physical health (more exercising) and mental health, to help them process their emotional state with help of the questionnaires. Clinicians would use tools in general practice to follow up with the patients, especially useful alerts in the long term period.

Clinical workshop: 16th February 2023

In the last of the 4 clinical workshops, we made the final overview of the project. Clinicians pointed out that many activities have been reduced due to the COVID-19 pandemic in the crucial moments of the project. They have foreseen a better peer-to-peer experience than we as the hospital could provide them since we were obliged to follow strict COVID-19 rules and restrictions.

The goal of this project was to correlate the new PERSIST data with the clinical results, to understand if there was any prediction value that could be applied in the future. Although the study was not interventional but observational, the patients and the medical doctors expected a review of results and a clear protocol for the follow-up based on PERSIST results.

Medical doctors have also seen some mental improvement, better overall spirit and focus on their overall grasp on personal health in patients. Medical doctors and the majority of patients expressed the willingness to re-engage in the following similar projects (even asking if there will be a continuation of the PERSIST project). The medical doctors see the future of medical practice in the field of cancer prevention, treatment, and follow up the patients after the active period of treatment in tools that have been tested in the project. They are overall satisfied they were part of this study and are willing to continue to support similar, future projects.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

3. Clinicians questionnaire

A generic questionnaire was developed in order to standardise the feedback from clinicians of the 4 hospitals. Altogether 11 clinicians involved in PERSIT research replied (4 from UL, 2 from SERGAS, 2 from CHU and 3 from UKCM).

- ***How would you rate the PERSIST system in general*** *(from 1 bad to 10 excellent)?*

  Average points given were **6,27**. Clinicians from all hospitals had similar average points.

- ***Would you like to use the PERSIST system as it is in general in your clinical practice?*** *(Yes/No / Part of it / Hard to say)*

  "Part of it" was the favourite answer (9 clinicians selected it) and 2 did not agree to use the system.

- ***Why?*** *(Free text answer)*

  The 2 clinicians who disagree to use the system reported that the system is running too slowly and is not aligned with use in oncology practice as it would be more useful for general practitioners. For part of it – biological marker follow-up, additional monitoring and innovation were marked as positive things, but device quality, data duplication and effort provided by their patients were considered as a burden.

- ***What is the most useful thing the PERSIST system would help you with in your practice?*** *(Free text answer)*

  Feedback from patients, data on vital parameters (summary), patients' subjective feelings, patient's statistics, risk factors etc. were marked.

- ***What kind of other medicine field uses a PERSIST system?*** *(Free text answer)*

  "*General practice*" was mentioned most of the times (5), followed by "*psychology*", "*infections*" and "*inflammatory diseases*".

- ***What, in your opinion, are the most important potentially modifiable lifestyle factors for cancer survivors that PERSIST detects?***

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Clinicians could choose from: Physical activity / Blood Pressure / Heart Rate / Sleep / Mood / Tiredness / Depression signs /other (what?).

"*Physical activity*" was the most chosen response (9); followed by "Blood Pressure" (6) and "Heart Rate" (6) and "depression signals" (3).

➔ ***What do you see as PERSIST overall added value?*** *(Free text answer)*

In general the **"option of monitoring the patients"** was considered as the best value.

➔ ***How would you rate PERSIST usability?*** *(from 1 bad to 10 excellent)*

Average points – **7.**

➔ ***How would you rate the precision of PERSIST system to identify risks in advance for cancer survivors?*** *(from 1 bad to 10 excellent)*

Average values – **6,9**

➔ ***Is PERSIST helping to personalise care plans/treatments for cancer survivors (yes, no, hard to say,)***

5 clinicians chose part of it, 4 marked yes, but 1 – hard to say.

➔ ***Why?*** *(Free text answer)*

- For part of it – limited clinical data, questionable performance, makes patients to be more aware of themselves, allows to change habits , risk factor detection questionable.
- For yes - monitoring patients' parameters over time, accordingly adapting recommendations, identifying patients who need interventions.

➔ ***What would be the best way to implement preventive strategies taking into account the individual patients trajectories? For example check the app once a week, inform patients by email or automatize this all by the App?*** *(Free text answer)*

Automatization of the App, checking once a week or every six months, the involvement of trained assistants was highlighted.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

➔ ***How would you rate mClinician web in general*** *(from 1 bad to 10 excellent)*

On average **6,1**

➔ ***Would you like to use the mClinician web version as it is in your clinical practice?*** *(yes, no, hard to say)*

5 - hard to say, 4 – no, 1-yes.

➔ ***Why?*** *(Free text answer)*

- Hard to say – too much information, too much work, no automation.
- For no - too slow, data entry is extremely time consuming, not practical, not user friendly.
- For yes – usability for monitoring biological parameters of patients

➔ ***Which parts of the mClinician web version seems most useful for clinical practice to you? (You can choose more than one)***

mHealth data (8); Tests (6); General and medical history (5), Diagnosis and symptoms (5), Cancer treatment (4)

➔ ***What parts of mClinician web should be changed or removed? Why?*** *(free text answers)*

Tests, diagnostic and therapeutic parts were mentioned the most.

➔ ***How would you rate the mClinician app in general*** *(from 1 bad to 10 excellent)?*

Average points - **6**.

➔ ***Would you like to use the mClinician app version as it is in your clinical practice?***

Hard to say – 9, yes -1, no - n1.

➔ ***Why?*** *(Free text answer)*

- For hard to say: slow, not user friendly, trained assistant needed;
- For yes – only some technical improvements needed,

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

- For no – EHR already completed and updated at the hospital .

➔ ***Which parts of the mClinician app version seems most useful for clinical practice to you?*** *(You can choose more than one)*

Parts	Points
Alerts	7
Patient overview	6
Appointments	5
Recurrence prediction	3
Cardiovascular Disease Risk	2
Usage stats	1
Trajectories	1

Table 162 Evaluation of mClinician app parts

➔ ***What parts of the mClinician app should be changed or removed? Why?*** *(free text answers)*

Trajectories and duplication of EHR.

➔ Conclusions:

he PERSIST system received an average rating of 6.27 out of 10, indicating that clinicians generally found it to be useful.

The majority of clinicians (9 out of 11) would like to use some part of the PERSIST system in their clinical practice.

The most useful aspects of the PERSIST system for clinicians include feedback from patients, data on vital parameters, and risk factors.

Clinicians believe that the PERSIST system could be used in various medical fields, such as general practice, psychology, infections, and inflammatory diseases.

Physical activity was identified as the most important potentially modifiable lifestyle factor for cancer survivors that the PERSIST system detects.

The usability of the PERSIST system was rated as average to good by clinicians.

The mClinician app received an average rating of 6 out of 10, and some clinicians (1 out of 11) would like to use it in their clinical practice.

The most useful aspects of the mClinician app for clinicians include alerts, patient overview, and appointments.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Overall, the feedback from clinicians suggests that the PERSIST system and mClinician app have potential to be useful tools in clinical practice, particularly for monitoring patient parameters and providing personalised care plans.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Data support for patient care

1. Trajectories

The possible impact of patient trajectories (extracted from retrospective data) and patient gathered health data on prediction of possible secondary disease or negative outcome and on improved intervention against the appearance of secondary diseases, worsening late toxicities or development of fatal events (sudden death, suicide).

Patient trajectories analysis aims at enhancing diagnosis, treatment, and prognosis decisions. Patient's trajectories are integrated into the mClinician app, which provides (i) patient's trajectories, (ii) the list of feature importance per patient's trajectory, and (iii) patient's cohorts. Patient's trajectories are based on statistical and AI-based models, which leverage retrospective and prospective data providing accurate estimations of the patients' trajectories. Clinicians can group several patients by categories (e.g., cancer stage or recurrence) enriching the set of tools for clinical decision support.

The trajectory is based on several variables such as cancer stage. However, variables that contain the same information may be under different code systems (e.g., ICD 9 and ICD 10). Therefore, a pre-processing process is necessary to harmonise the data before processing them. Moreover, in the case of staging, the PERSIST datasets present "clinical staging" and "pathological staging". Clinicians on the PERSIST app can decide which one to select for comparison. Finally, a "global staging" variable has been introduced. This variable takes the "pathological staging" when available, otherwise the "clinical staging". The variable "global staging" is used when the population is small due to missing values.

What trajectories are used for? Trajectories can be used to inform patients about the statistics behind their conditions. For instance, for a cancer patient at stage 1, the clinician can ensure the patient by showing the trajectories of that specific population (e.g., 99 percent of surviving more than ten years). Hopefully, these precise statistics can enhance patients' physical comfort and emotional well-being. Trajectories can also raise alerts based on the combinations of particular features. Those trajectories can be exploited in order to understand the cause of the alert. Indeed, by checking the following plot, clinicians can retrieve the features (the present ones) that affect the survival probability the most. On the other hand, clinicians can retrieve the features that enhance the overall survivor probability (e.g., treatment: local excision of a breast lesion). Detailed trajectory explanations can be seen in Figure 62.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

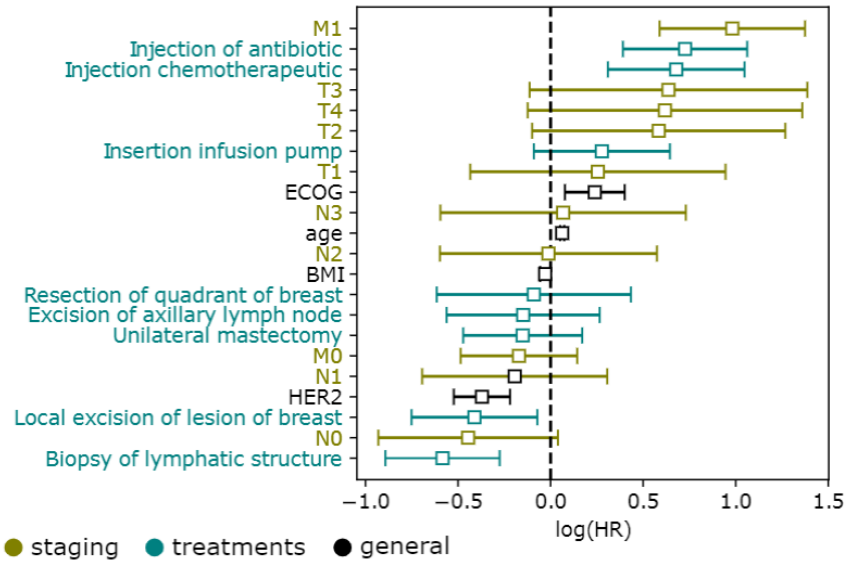

Figure 62: Detailed explanation of trajectories impacting factors

➔ Conclusions:

Improved prediction of possible secondary diseases or negative outcomes: By analysing patient trajectories (looking into detailed versions), clinicians can identify patterns and risk factors that may lead to secondary diseases or negative outcomes. This information can be used to develop personalised interventions and treatment plans that are tailored to the patient's specific needs, potentially leading to better health outcomes.

Enhanced diagnosis, treatment, and prognosis decisions: Trajectories can provide clinicians with a more comprehensive understanding of a patient's condition, including the progression of their disease and the effectiveness of previous treatments. This information can inform diagnosis, treatment, and prognosis decisions, leading to more accurate and personalised care.

Improved patient comfort and emotional well-being: Trajectories interpreted by clinicians can provide patients with statistics about their condition, including survival rates and the likelihood of recurrence. This information can help patients feel more informed and empowered, potentially improving their emotional well-being and overall quality of life.

Clinical decision support: Trajectories can be used to group patients into categories based on factors such as cancer stage or recurrence, providing clinicians with a set of tools for clinical decision support. This can help ensure that patients receive appropriate and timely care.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

2. Recurrence prediction

The effectiveness of Big Data prediction models to support decision-making in patient follow-up. Although the cancer may be in remission thanks to early detection and improvements in treatment, some patients may experience a relapse of the cancer. The recurrence is one of the main causes of death related to cancer, so its early detection can be highly valuable to support decision making and enable timely and more appropriate interventions that prevent the onset of recurrence.

For this reason, the development of a non-invasive computational system to predict the risk of relapse of breast and colon cancer, based on the clinical and treatment information of the patients available in the electronic medical record (EHR) was proposed. This is the colon and breast cancer recurrence prediction.

The colon and breast cancer recurrence prediction service offers an application that predicts the probability of relapse for a given patient using Artificial Intelligence (AI) and clinical data from different hospitals. For this purpose, a platform that, given a patient ID and a type of cancer (colon or breast), can recover the relevant data from the FHIR servers, restructure it and serve it to the recurrence prediction models, was developed. This service consists of three distinct components: the AI recurrence models, the FHIR servers that host the clinical patient data, and the API that recovers and serves the relevant information.

There are two AI recurrence models: one for colon cancer and one for breast cancer. These models were trained with data extracted by CHU de Liège. Several machine learning algorithms were tested to find the most optimal one for each cancer. In the end, Gradient Boosting (GB) and eXtreme Gradient Boost (XGB) were chosen for colon and breast, respectively.

The service is connected to the FHIR server and obtains the needed data for a certain patient to calculate their recurrence prediction. The collection of data obtained is different for each type of cancer, and can be seen in the following tables:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Shared data	Gender
	BMI (weight/height)
	Drinking behaviour
	Number of surgeries
	Age at diagnosis
	ECOG performance status
	TNM codes
	Histologic grade
	Body site codes
	Number of radiotherapies
	Number of chemotherapies
Breast cancer data	Estrogen receptor
	Progesterone receptor
	HER2 by ICH
	Ki67
Colon cancer data	CEA

Table 163 Variables considered in the recurrence prediction models

Shared comorbidities	Hypertension
	Chronic pulmonary disease
	Diabetes
	Hypothyroidism
	Metastatic cancer
	Solid tumour
	Obesity
Colon Cancer comorbidities	Cardiac arrhythmia
	Valvular disease
	Renal failure
	Lymphoma
	Weight loss
	Fluid and electrolyte disorders
	Deficiency anaemia

Table 164 Comorbidities considered in the recurrence prediction models

Recurrence prediction for UL patients

Altogether UL colorectal cancer patients average values of recurrence (Table 165) was 36,27, but for breast cancer patients – 23,79.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

No.	Recurrence prediction value
1	x
2	38
3	38,69
4	37,9
5	34,76
6	38,21
7	x
8	33,21
9	37,9
10	32,44
11	35,57
12	43,2
13	x
14	33,21
15	33,84
16	34,59
AVG	**36,27**

Table 165 UL colorectal cancer patient recurrence prediction values

No.	Recurrence prediction value
1	19,5
2	x
3	27,89
4	14,41
5	28,51
6	12,25
7	56,86
8	23,37
9	39,61
10	20,46
11	23,83
12	12,39
13	19,13
14	11,11
AVG	**23,79**

Table 166 UL breast cancer patient recurrence prediction values

Recurrence prediction for CHU patients

Altogether CHU cancer patients average values of recurrence was 26,69.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

No.	Prediction value
1	44,54
2	45,52
3	28,97
4	x
5	47,15
6	31,31
7	53,56
8	37,9
9	47,7
10	60,89
11	x
12	x
13	x
AVG	**44,12**

Table 167 CHU colorectal cancer patient recurrence prediction values

No.	Prediction value
1	17,46
2	9,48
3	11,51
4	18,65
5	25,93
6	38,04
7	11,57
8	12,64
9	9,46
10	7,29
11	38,53
12	25,44
AVG	**18,83**

Table 168 CHU breast cancer patient recurrence prediction values

Recurrence prediction for SERGAS patients

Altogether SERGAS breast cancer patients average values of recurrence was 23,25, but for colorectal cancer patients – 35,47.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

No.	Prediction value
1	26,4
2	30,48
3	14,88
4	14,35
5	28,63
6	27,69
7	32,87
8	11,39
9	24,48
10	11,76
11	21,68
12	35,76
13	26,66
14	39,43
15	8,57
16	16,94
AVG	**23,25**

Table 169 SERGAS breast cancer patient recurrence prediction values

No.	Prediction value
1	43,95
2	35,26
3	37,32
4	42,97
5	38,21
6	29,28
7	32,87
8	40,93
9	38,73
10	33,21
11	31,11
12	42,01
13	27,32
14	28,97
15	38,73
16	38,44
17	29,26
18	29,92
AVG	**35,47**

Table 170 SERGAS colorectal cancer patient recurrence prediction values

Recurrence prediction for UKCM patients

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Altogether UKCM breast cancer patients average values of recurrence was 31,54, but for colorectal cancer patients – 28,46

No.	Prediction value
1	42,03
2	37,56
3	29,26
4	38,44
5	42,01
6	29,56
7	32,87
8	40,93
9	38,65
10	33,21
11	29,11
12	42,01
13	27,32
14	28,98
15	34,56
16	11,98
17	41,78
AVG	**31,54**

Table 171 UKCM breast cancer patient recurrence prediction values

No.	Prediction value
1	27,35
2	26,87
3	11,39
4	23,48
5	12,79
6	21,23
7	35,85
8	28,77
9	40,42
10	7,57
11	18,85
12	29,27
13	28,25
14	9,45
15	38,53
16	23,45
17	50,14
AVG	**28,46**

Table 172 UKCM colorectal cancer patient recurrence prediction values

The average prediction data reveals that breast cancer patients according to GRAD prediction models have ~12% lesser risk of recurrence than colorectal cancer patients. When comparing the average prediction values for patients in each hospital, the cancer patients have similar values in UL and SERGAS. In CHU breast cancer patients have around 5% lower values than other centres, but 8% higher values for colorectal cancer patients. UKCM colorectal cancer patients had the lowest possible risk, but breast cancer patients the highest possible risk, compared to patients from other hospitals.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Site	Breast cancer patients	Colorectal cancer patients
UL	23,79	36,27
SERGAS	23,25	35,47
CHU	18,83	44,13
UKCM	31,54	28,46
AVG	24,35	36,08

Table 173 Average prediction values for each hospitals` patients

The actual incidence of recurrence has been compared to the risk prediction for the 5 patients who had recurrence events recorded and who remained in the study when the recurrence predictive models were released (see Tables 165-173). As shown in Table 174, the risk profile was not very high for those patients who had recurrence. However, the number of patients who had recurrence is so low and the results do not have statistical significance.

No.	Risk identified
1	34,76%
2	45,52%
3	35,76%
4	26,66%
5	23,45%

Table 174. Recurrence risk for patients who indeed had recurrence

Longitudinal studies should be carried out to understand the actual accuracy and value of the recurrence predictive models for the recurrence screening and risk assessment.

3. Cardiovascular Disease Risk CDS

Cardiovascular disease is one of the most common secondary diseases among cancer survivors. The treatments and the medications can lead to heart failures, so in addition to regular patients' cardiovascular risk, cancer survivors' additional factors affect cardiovascular disease risk. Knowledge base for this service was built upon these facts from clinical guidelines and the results of the work done in 5.1. The knowledge and the data collected in PERSIST are paired and rules and conditions are prepared accordingly for the inference engine.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Input	Source
Can you maintain your normal physical activity?	Questionnaires from mHealth App
Are you having worsening shortness of breath with activity?	Questionnaires from mHealth App
Have you experienced increased swelling of legs, feet, and ankles?	Questionnaires from mHealth App
Do you feel persistent fatigue despite a good night's sleep?	Questionnaires from mHealth App
Does fatigue interfere with your usual activities?	Questionnaires from mHealth App
How would you rate your fatigue on a scale of 0 (none) to 10 (extreme) over the past week?	Questionnaires from mHealth App
Presence of chronic such diseases: Diabetes, High Cholesterol	mClinician
Age	mClinician
Cardiovascular History	mClinician
Exercise	mClinician
Smoking	mClinician
Alcohol	mClinician
Sleep	mClinician
Diabetes	mClinician
BMI	mClinician
HDL	mClinician
LDL	mClinician
Blood Pressure	mClinician
C-reactive protein	mClinician

Table 175 Inputs of Cardiovascular disease risk and their sources

The cardiovascular risk scores for all CHU and all SERGAS patients were marked as low.

The cardiovascular risk score for almost all UL patients was marked as low (except one UL patient, which was marked as high due to negative effects shown in the App such as current smoking, BMI over 30 kg/m2, LDL between 96.7 mg/dL and 127.6 mg/dL, and age). However, we have not recorded any negative outcomes connected with this patient and cardiovascular risks so far

The cardiovascular risk score for almost all UKCM patients was marked as low (except one UKCM patient, which was marked as high due to negative effects in the App such as BMI over 30kg/m2, existing diabetes, and age 52). Nevertheless, there have been no recorded negative outcomes connected with this patient and cardiovascular risks so far.

Cardiovascular risk scores could be useful for general practitioners and cardiologists to suggest optional examinations and/or therapy for patients with a high cardiovascular risk. If the App shows high risk for a patient, he or she may be referred to a cardiologist for a risk evaluation and additional examinations.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

4. Colon/breast cancer recurrence CDS

➔ *Colon Cancer Risk assessment*

The clinical practice guideline NCCN Clinical Practice Guidelines was used as a basis to build the knowledge base for determining a breast cancer survivor's likelihood of a recurrence. Additionally, extensive study was done on the factors that influence breast cancer recurrence. As a result, a set of patient information and conditional justifications for breast cancer recurrence are identified. The rules and information required for building the knowledge base for breast cancer recurrence have become complete as this set of data matched with the input list established in accordance with patient data acquired in PERSIST.

Input	Source
Age	mClinician
Exercise	mClinician
Smoking	mClinician
Alcohol	mClinician
Diabetes	mClinician
BMI	mClinician
Blood Pressure	mClinician
C-reactive protein	mClinician
faecal immunochemical test (FIT)	mClinician
Faecal occult blood test (FOBT)	mClinician
CA19-9	mClinician
CEA	mClinician
Have you seen red blood in your faeces?	Questionnaires from mHealth App
Blood in your faeces, are few drops of bright red (fresh) blood only occasionally?	Questionnaires from mHealth App
Blood in your faeces, is it mixed with it?	Questionnaires from mHealth App
Has this been happening to you for more than 3 weeks?	Questionnaires from mHealth App
Do you have recent onset rectal or perianal discomfort or pain?	Questionnaires from mHealth App
Have your faeces become darker, blacker, tarry or maroon in colour?	Questionnaires from mHealth App
Have you noticed changes in your usual bowel rhythm in recent weeks, such as diarrhoea, constipation, or a combination of both?	Questionnaires from mHealth App
Do you have an adequate daily intake of healthy nutrients (at least five servings of fruits or vegetables daily, avoidance of processed foods, limitations of too much red meat and alcohol)?	Questionnaires from mHealth App

Table 176 Colon cancer recurrence service inputs and their source

➔ *Breast Cancer Risk assessment*

To create the knowledge base for assessing the risk of a recurrence in a breast cancer survivor the clinical guideline “NCCN Clinical Practice Guidelines in Oncology - Breast Cancer Version 5.2020 — July 15, 2020" was taken as a foundation. On top of that the

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 875406

research conditions affecting breast cancer recurrence was studied thoroughly. As a result, a set of patient data and conditional arguments were revealed for breast cancer recurrence. This set of data matched with the input list created according to patient data collected in the PERSIST environment. As a result, rules and information necessary for creating the knowledge base for breast cancer recurrence was completed.

Input	Source
Alcohol Consumption	mClinician
Smoking	mClinician
BMI	mClinician
CEA	mClinician
CA 15-3	mClinician
TNM Stage	mClinician
Tumour Size	mClinician
Physical Activity	mClinician
Estrogen and Progesterone Receptors	mClinician
BRCA1	mClinician
Age and Menopausal Status	mClinician
Cancer Subtype	mClinician
Anxiety	Questionnaires from mHealth App

Table 177 Breast cancer recurrence service inputs and their source

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Colon/breast cancer recurrence for each centre was calculated by mClinician app and can be seen below:

No.	Recurrence	Cancer type
1	medium	CRC
2	medium	CRC
3	medium	CRC
4	high	CRC
5	high	CRC
6	high	CRC
7	high	CRC
8	high	CRC
9	high	CRC
10	high	CRC
11	high	CRC
12	high	CRC
13	high	CRC
14	medium	BRC
15	high	BRC
16	medium	BRC
17	medium	BRC
18	high	BRC
19	medium	BRC
20	high	BRC
21	high	BRC
22	high	BRC
23	medium	BRC
24	high	BRC
25	high	BRC

Table 178 Breast cancer recurrence risk for CHU patients

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

No.	Recurrence	Cancer type
1	Medium	BRC
2	Medium	BRC
3	Medium	BRC
4	Medium	BRC
5	Low	BRC
6	Medium	BRC
7	Medium	BRC
8	High	BRC
9	Medium	BRC
10	Medium	BRC
11	Medium	BRC
12	Medium	BRC
13	High	BRC
14	High	BRC
15	Medium	BRC
16	High	CRC
17	High	CRC
18	High	CRC
19	High	CRC
20	Medium	CRC
21	Medium	CRC
22	High	CRC
23	High	CRC
24	High	CRC
25	High	CRC
26	High	CRC
27	High	CRC
28	Medium	CRC
29	High	CRC
30	Medium	CRC
31	Medium	CRC

Table 179 Breast cancer recurrence risk for UL patients

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

No.	Recurrence	Cancer type
1	high	BRC
2	medium	BRC
3	high	BRC
4	medium	BRC
5	high	BRC
6	medium	BRC
7	high	BRC
8	medium	BRC
9	medium	BRC
10	medium	BRC
11	high	BRC
12	high	BRC
13	medium	BRC
14	medium	BRC
15	medium	BRC
16	high	BRC
17	medium	CRC
18	medium	CRC
19	medium	CRC
20	high	CRC
21	medium	CRC
22	low	CRC
23	medium	CRC
24	medium	CRC
25	medium	CRC
26	medium	CRC
27	medium	CRC
28	medium	CRC
29	medium	CRC
30	high	CRC
31	medium	CRC
32	high	CRC
33	medium	CRC
34	medium	CRC

Table 180 Recurrence risk for UKCM patients

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

No	Recurrence	Cancer type
1	High	BRC
2	High	BRC
3	Medium	BRC
4	Medium	BRC
5	Medium	BRC
6	Medium	BRC
7	Medium	BRC
8	High	BRC
9	Medium	BRC
10	Medium	BRC
11	Low	BRC
12	Medium	BRC
13	Medium	BRC
14	Medium	BRC
15	Medium	BRC
16	Medium	BRC
17	High	CRC
18	Medium	CRC
19	High	CRC
20	Medium	CRC
21	Meidum	CRC
22	Medium	CRC
23	High	CRC
24	High	CRC
25	High	CRC
26	Medium	CRC
27	High	CRC
28	High	CRC
29	Medium	CRC
30	High	CRC
31	Medium	CRC
32	Medium	CRC
33	High	CRC
34	High	CRC

Table 181 Recurrence risk for SERGAS patients

In order to see whether these prediction values match with the future patient situation, further follow-up will be necessary. Within the time scale of this study 4 of the patients who had recurrence (Table 182) mClinician had calculated the risk from medium to high.

No.	Risk identified
1	high
2	medium
3	medium
4	medium
5	medium

Table 182. Recurrence risk levels from mClinicians for patients who indeed had recurrence

➔ *CDSS alerts received in mClinician*

Combining patient replies to questionnaires and their medical history (laboratory results etc.) the PERSIST system generated alerts or recommendations. Altogether in the course

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

of T6.4. clinical validation phase 6 recommendations and 19 alerts have been created and sent to mClinician app (see Table183).

last_updated	site	category_text	code_text
17.09.2022	sergas	Recommendation	A healthy diet and exercise regimen helps maintain a good bowel habit
17.10.2022	sergas	Alert	Not having adequate daily intake of nutrients regarding quantity and quality (vegetables and fruits half the volume on the plate, whole grains 30%, protein 20%) , Not having adequate daily intake of healthy nutrients (at least five servings of fruits or vegetables daily, avoidance of processed foods, limitations of too much red meat and alcohol)
21.10.2022	sergas	Alert	Not having adequate daily intake of nutrients regarding quantity and quality (vegetables and fruits half the volume on the plate, whole grains 30%, protein 20%) , Not having adequate daily intake of healthy nutrients (at least five servings of fruits or vegetables daily, avoidance of processed foods, limitations of too much red meat and alcohol), Not having adequate daily intake of fluids (more than 1,5 litres)
30.09.2022	sergas	Alert	Pain of scale rating > 4 - (5 > 4)
15.10.2022	sergas	Alert	Pain of scale rating > 4 - (6 > 4)
15.10.2022	sergas	Recommendation	A healthy diet and exercise regimen helps maintain a good bowel habit
16.09.2022	sergas	Recommendation	A healthy diet and exercise regimen helps maintain a good bowel habit
09.09.2022	sergas	Recommendation	A healthy diet and exercise regimen helps maintain a good bowel habit
26.09.2022	sergas	Recommendation	A healthy diet and exercise regimen helps maintain a good bowel habit
30.10.2022	sergas	Recommendation	A healthy diet and exercise regimen helps maintain a good bowel habit
16.10.2022	ul	Alert	Pain of scale rating > 4 - (5 > 4)
28.09.2022	ul	Alert	Rating your fatigue (7) is more than 6 over the past week 0–10
12.10.2022	ul	Alert	Rating your fatigue (8) is more than 6 over the past week 0–10
24.10.2022	ul	Alert	Rating your fatigue (8) is more than 6 over the past week 0–10
30.10.2022	ul	Alert	Pain of scale rating > 4 - (5 > 4)
07.09.2022	liege	Alert	Pain of scale rating > 4 - (7 > 4)
29.10.2022	liege	Alert	Rating your fatigue (7) is more than 6 over the past week 0–10
06.10.2022	liege	Alert	Rating your fatigue (7) is more than 6 over the past week 0–10
09.10.2022	liege	Alert	Pain of scale rating > 4 - (5 > 4)
22.11.2022	liege	Alert	Rating your fatigue (9) is more than 6 over the past week 0–10
10.10.2022	ukcm	Alert	Rating your fatigue (7) is more than 6 over the past week 0–10
23.10.2022	ukcm	Alert	Pain of scale rating > 4 - (7 > 4)
16.12.2022	ukcm	Alert	Pain of scale rating is 8
22.09.2022	ukcm	Alert	Rating your fatigue (7) is more than 6 over the past week 0–10
18.09.2022	ukcm	Alert	WHO questionnaire S1-12: if the answer is “at least severe”, WHO questionnaire H1-3 if the answer is “more than 10 days”

Table 183 Notifications about patients in mClinician app

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 875406

Alert analysis. In case of two UL`s patients the alert was only once and as the system was still updated patients were not contacted. In the case of the third patient there were 3 alerts, the patient was contacted. Patient explained tiredness with greater work-load and refused to meet the oncologist. Recommendations to visit a general practitioner were made. Negative outcomes reported above (table 144 for UL patients) showed that only one had complications after treatment, but those were not connected with recurrence. None of the patients who received alerts in UL had cancer recurrence. None of the patients from CHU, SERGAS and UKCM who received alerts had registered recurrence.

Further evaluations of existing data can be made comparing patients' negative outcomes with alerts. In case of alerts, clinician should contact the patient for further discussions and/or recommendations. Due to great workload clinicians have suggested that this should be done by a specialised nurse (that is previously trained how to use the PERSIST system). She can evaluate the individual case and decide whether the clinician needs to be specifically informed.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

CTC results effects on healthcare of cancer survivors

With the aim to understand the relevance of CTCs in early stage BRC and CRC, blood samples were collected from the patients enrolled in the PERSIST study at different time points for CTC enumeration and phenotypic analysis.

For this purpose, RUBY had to scale up the fabrication of their RUBYchip™. Under the scope of PERSIST, RUBY was able to convert the initial RUBYchip™ prototype manufactured in their laboratory by soft-lithography in PDMS, into an industrial version of the device in COP using injection moulding.

Within the PERSIST clinical study, RUBY received peripheral blood samples from breast and colorectal cancer patients from UKCM and SERGAS. Blood samples were collected at three time points in the project (Fig 63).

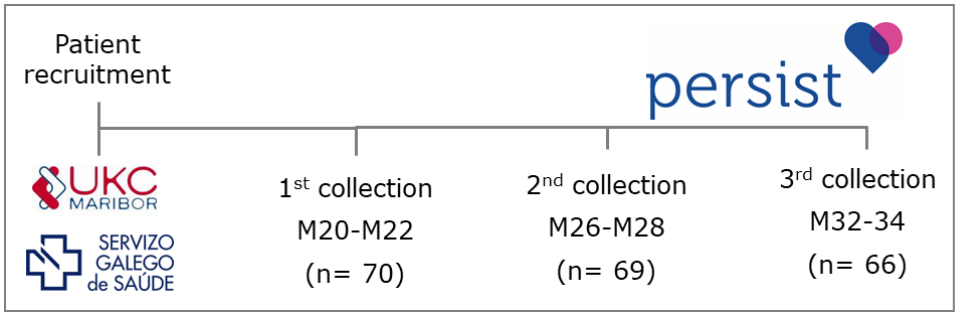

Figure 63: Summary of the clinical samples collection calendar.

In total, RUBY analysed 205 samples, 105 CRC and 100 BRC, from 36 and 34 patients, respectively. Between M20-22, 70 samples were collected at baseline and analysed. We performed a second collection 6 months after the baseline but, unfortunately, one of the BRC patients from SERGAS was unavailable on the sample collection day. Twelve months after the baseline, between M32-34, we performed the third collection. Unfortunately, we only received 66 samples because 2 CRC patients from UKCM decided to step out of the study and 1 from SERGAS one couldn't not attend the last CTC collection appointment because he underwent a heart surgery. Samples were expected to be processed within 24 h of collection, and a logistics network and protocol was put in place for this purpose. However, some samples suffered delays. Table 184 summarises the samples collected, and their time of processing. Interestingly, in the last sample collection period all samples were processed in 24hrs, as a consequence of the improved management with transportation company, packaging and organisation of the process during the project.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

		Baseline (M20-M22)		6 month (M26-M28)		12 month (M32-M34)	
	Sample type	# samples	% analysed in 24hrs	# samples	% analysed in 24hrs	# samples	% analysed in 24hrs
UKCM	CRC	18	50%	18	92%	16	100%
	BRC	18		18		18	
SERGAS	CRC	18	100%	18	100%	17	100%
	BRC	16		15		15	
Total number of samples		70		69		66	

Table 184. Summary of the samples collected longitudinally from the two hospitals, and percentage of samples that were processed the day after collection.

Processing over 200 clinical samples allowed RUBY to evaluate and test the newly industrialised version of the RUBYchip™, providing relevant evidence for the usability of the devices produced.

As the beginning of the clinical study was delayed due to the COVID-19 pandemic and prolonged recruitment phase, sample collections were delayed and the third was only finished very close to the end of the project. Full evaluation still requires a more detailed analysis.

As a general result, using the RUBYchip™, it was possible to detect CTCs in early-stage patients after treatment with curative purposes. Numbers were however low, as initially expected (Figure 64).

Thresholds to discriminate between good and bad prognosis in metastatic colorectal cancer (mCRC) and metastatic breast cancer (mBRC) using RUBY`s technology were previously defined at ≥7 and ≥5, respectively [18]. However, thresholds in early-stage disease using the aforementioned technology have not yet been defined.

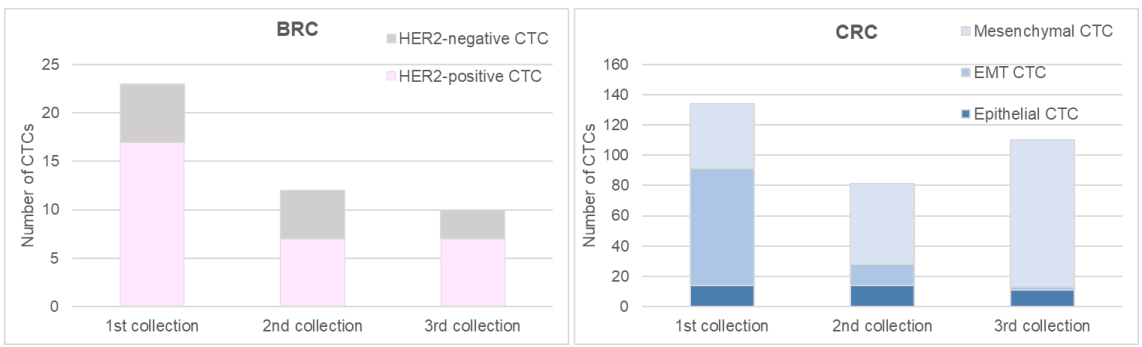

Figure 64: Total number of CTCs found in the BRC (left) or CRC (right) cohort, per sample collection.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Out of the 70 patients analysed in the first collection, 35 (50%) had some CTCs (≥1). 14/34 BRC (41%) and 21/36 CRC (58%). The average number of CTCs in the positive population was 1.6 in BRC and 6.4 in CRC, while the median was 1 and 5, respectively.

Looking at the longitudinal study, only 13 patients (19%) were positive in all 3 collections, while 17 cases (23%) were negative in all 3. Important to mention that all cases positive in the 3 collections were CRC, while 82% of the cases negative in the 3 collections were BRC. Of the remaining patients that changed status, 22 (31%) were positive in 2 collections and negative in one, and 19 (27%) negative in one, and positive in 2. Out of the patients that were CTC positive at some point, 20 (29%) had increasing numbers, while 10 (14%) were stable, and 22 (31%) had decreasing numbers of CTCs (comparing the baseline and the third collection).

Using the previously defined threshold of 5 CTCs in BRC and 7 in CRC, no BRC (0%) patients and 6 (9%) CRC were positive at baseline. The average number of CTCs in the positive population was 14.7 and the median was 15.5. Only 1 BRC (3%) and 2 CRC (6%) patients were positive in the second collection. In the third collection 4 (11%) CRC patients were positive.

Looking at the longitudinal study, and considering the threshold no patients (0%) were positive in all 3 collections, while 60 cases (86%) were negative in all 3. Of the remaining patients that changed status, 2 (3%) were positive in 2 collections and negative in one, and 8 (11%) negative in one, and positive in 2. Out of the patients that were CTC positive at some point, 4 (6%) had increasing numbers, while 6 (9%) had decreasing numbers of CTCs (comparing the baseline and the third collection).

In the overall results, we found CTCs that meet the classification criteria for the all the CTC classification: Mesenchymal, epithelial and epithelial-to-mesenchymal transition. A relevant increase in the number of detected mesenchymal CTCs is observed in the 3rd collection (above 80% of the total number of CTCs) (Figure 64). In case of breast cancer patients, the occurrence of HER2-positive cells was found to be more frequent (69%) than CTCs that do not express HER2 (31%) (Figure 64). This is a very relevant aspect to be analysed in the context of the patient clinical data.

As a summary, and in good agreement with previous studies and with the prognosis of BRC versus CRC patients, in the analysed patient cohort the number of CTCs was higher in CRC patients than in BRC patients, but the correlation of these results with clinical data is currently under analysis.

➔ Conclusions:

Processing over 200 clinical samples allowed RUBY to evaluate and test the newly industrialised version of the RUBYchip™, providing relevant evidence for the usability of the devices produced.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

The results showed that it was possible to detect CTCs in patients after treatment, although the numbers were low. Out of the 70 patients analysed in the first collection, 35 had some CTCs (≥1). Of the longitudinal study, only 13 patients were positive in all 3 collections, while 17 cases were negative in all 3. Out of the patients that were CTC positive at some point, 20 had increasing numbers, while 10 were stable, and 22 had decreasing numbers of CTCs.

The successful conversion of the RUBYchip™ prototype into an industrial version using injection moulding is a positive outcome for the future of CTC detection in cancer patients. The study also provides a basis for further exploration of the relevance of CTCs in patients.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

Main conclusions

The study conducted in the PERSIST project provided valuable insights about the physical activity levels, heart rate, and emotional well-being of cancer survivors using the mHealth app and smart bracelets. These findings suggest that the patients were moderately active and had a positive outlook on life after cancer treatment. The study also identified potential reasons for differences in physical activity levels among hospitals, such as environmental factors (e.g., walkability) and weather conditions. This information can be used to develop targeted interventions to promote physical activity and improve the health outcomes of patients. Furthermore, the use of the mHealth app and smart bracelets led to a decrease in signs of depression and anxiety among patients, indicating a positive impact on the psychological well-being of cancer patients. This highlights the potential of mHealth apps and smart bracelets to engage patients, monitor their health, and provide valuable data that can inform strategies to promote healthy behaviour and prevent chronic disease.

The PERSIST tool provides options for personalised cancer survivor care plans, including an alert system and a different parameter overview about a patient for clinicians. The system also promotes self-management via the possibility for patients to follow their activity and facilitates communication with healthcare providers. The mClinician app and web systems were generally perceived as easy to use from the clinicians' side, and technical support provided was effective in helping clinicians become more comfortable with the system over time. PERSIST has the potential to improve the clinical outcomes of cancer survivors following suggestions for healthy lifestyle adjustments and due to timely recurrence risk detection, improve the coordination and continuity of care among healthcare providers, promote patient empowerment and engagement, and contribute to broader social goals around cancer survivorship. However, further investigation and improvement of the identified usability issues could potentially lead to increased acceptance and adoption of the mClinician system among clinicians.

The PERSIST project highlights the potential of big data in cancer survivorship care. By integrating different types of data within the big data platform, PERSIST can provide personalised care plans for cancer survivors, promoting self-management, and facilitating communication with healthcare providers. The system also has the potential to improve the coordination and continuity of care among healthcare providers, ultimately leading to better outcomes for cancer survivors.

One critical aspect to consider is the need for data harmonisation to achieve this goal. The success for implanting big data technologies depends on the ability to collect, store, and analyse data from different sources, including wearables, electronic health records, and patient-reported outcomes. The overall harmonisation of data across different sources is essential for ensuring that the data collected is meaningful and can be used to inform clinical decision-making. This, in turn, can lead to the development of new standards of care for cancer survivorship and the design of future interventions and tools for survivorship care.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

However, while the PERSIST project provided valuable insights into the potential of digital therapies in cancer's survivor care, more high-quality, larger scale studies are still needed to demonstrate their effectiveness. Continuing the PERSIST system research within a wider patient population and involving more clinicians from different countries would provide more data for evaluation and additional evidence for PERSIST benefits in cancer survivorship care.

1. Practical suggestions:

The effective use of the PERSIST big data platform requires prompt and accurate transfer of lifestyle data, laboratory test results and other electronic health record (EHR) data. This information should be entered into the system on a daily basis to ensure timely updates to the warning system and to optimise the performance of the prediction models and clinical decision system. Automating this process is recommended to prevent delays and ensure accuracy.

Patients are encouraged to use qualitative devices such as smart bracelets and their mobile phones to provide valuable data to the PERSIST system. The data transfer process should be automated to encourage patients to regularly update their data without delay. Reminders should be sent to patients to complete questionnaires within a specific timeframe to enable AI models to detect any significant changes. Additionally, mHealth apps can include reminders for healthy habits and gamification elements to create a positive patient experience.

Clinicians and their assistants can promptly check alerts and communicate with patients for further check-ups and necessary adjustments in treatment by accessing the patient's profile in mClinician during their visit. This can facilitate discussions regarding deviations from the usual and referrals to other specialists, such as cardiologists or psychologists.

The PERSIST system has the potential to be used in other fields of medicine, especially beneficial for general practitioners who attend to a large number of patients with different health conditions daily. By implementing these measures, the benefits of the PERSIST system can be maximised, improving patient outcomes and clinical decision-making.

2. Lessons learnt:

To prevent technical difficulties from burdening participants, it is crucial to carefully select and test appropriate devices before distribution. Simplifying technologies to make them user-friendly can also mitigate complaints and difficulties with device usage among larger groups of patients. Additionally, using standardised questionnaires to assess digital literacy during recruitment can aid in selecting suitable participants for co-creation activities with technical partners. It is important to avoid overwhelming patients with too many daily tasks or prolonged participation to prevent dropouts from the study. Indeed, the phasing of PERSIST clinical study with two different groups of patients might have had a positive

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

impact in the gathering of results. Some patients dropped-out from the study because of their frustration and struggle with technologies and user interfaces that were not still mature during the co-creation phase, in which bugs and failures still occurred. Hence, they perceived PERSIST as useless and cumbersome. Therefore, a phased design, with one group of patients participating in the co-creation and another group participating in the clinical validation might have been desirable.

In addition, clinical partners should align with the goals of the European Health Data Space (EHDS), which seeks to promote better exchange and access to different types of health data, including electronic health records (EHRs) and genomics data, for healthcare delivery, research, and policy-making purposes. This will simplify the overall collection and analysis of health data in this kind of studies.

3. PERSIST advantages identified:

The PERSIST tool represents a significant step forward in survivorship care for cancer patients. Its ability to deliver personalised and dynamic care plans based on individual survivor needs has the potential to disrupt traditional models of survivorship care, which rely on printed or generic care plans that may not adequately address the unique needs of each survivor. By personalising care plans, the PERSIST tool can improve patient outcomes and overall quality of life while also reducing healthcare costs.

Overall, the PERSIST system is user-friendly and not unnecessarily complex, which makes it easy for participants to learn and use without requiring significant training. Participants had a neutral to slightly positive attitude towards using the system frequently, indicating its potential usefulness. The high rate of adherence in almost all hospitals suggests that patients found the app easy to use and manage on a daily basis. This suggests that the system was designed in a user-friendly way and was not a source of frustration or confusion for the participants, so it can be widely used in clinical practice.

This transition towards more personalised medicine strategies for cancer survivorship care plans is aligned with the goals of the Precision Medicine Initiative (PMI), which aims to tailor medical treatments and preventive strategies to an individual's unique genetic and environmental profile. The PMI recognizes that traditional "one-size-fits-all" approaches to healthcare may not be effective for all patients and that a more personalised approach can lead to better outcomes and cost savings.

The PERSIST tool has the potential to be a key driver in this transition towards personalised survivorship care. Its ability to adapt to the changing needs of each individual survivor and deliver personalised care plans can lead to better health outcomes and increased patient satisfaction. Additionally, by optimising follow up decisions while reducing hospital readmissions and emergency room visits, the tool can provide significant cost savings for healthcare providers and insurers.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

The PERSIST tool has the potential to have a positive socio-economic impact on survivorship care by providing personalised care plans for cancer survivors and promoting self-management. It can potentially reduce healthcare costs associated with cancer survivorship care by improving patient outcomes and reducing the need for costly medical interventions. Additionally, PERSIST can improve the coordination and continuity of care among healthcare providers, which can lead to more efficient use of healthcare resources and reduced healthcare costs. Furthermore, the PERIST`S potential to engage patients and monitor their health can lead to the prevention of chronic disease, reducing the burden of chronic disease on the healthcare system and society as a whole. The use of mHealth apps and smart bracelets can also provide valuable data that can inform strategies to promote healthy behaviour and prevent chronic disease, potentially leading to long-term cost savings for healthcare systems and society.

Overall, the PERSIST tool represents an exciting opportunity to improve survivorship care for cancer patients and transition towards more personalised medicine strategies. With its potential to disrupt traditional models of care and improve patient outcomes while also reducing costs, the PERSIST tool is a promising development in the field of cancer survivorship care.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

References

1. Roberts AL, Potts HW, Koutoukidis DA, Smith L, Fisher A. Breast, Prostate, and Colorectal Cancer Survivors' Experiences of Using Publicly Available Physical Activity Mobile Apps: Qualitative Study. JMIR Mhealth Uhealth. 2019 Jan 4;7(1):e10918. doi: 10.2196/10918. PMID: 30609982; PMCID: PMC6329432.

2. L. Fallowfield et al. Effective communication skills are the key to good cancer care. Eur J Cancer (1999).

3. Wolf, M. S., Chang, C. H., Davis, T., & Makoul, G. (2005). Development and validation of the Communication and Attitudinal Self-Efficacy scale for cancer (CASE-cancer). Patient education and counselling, 57(3), 333-341.

4. M. Petticrew et al. Influence of psychological coping on survival and recurrence in people with cancer: systematic review. BMJ. (2002).

5. Hibbard, J. H., Mahoney, E. R., Stockard, J., & Tusler, M. (2005). Development and testing of a short form of the patient activation measure. Health services research, 40(6p1), 1918-1930.

6. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study. JMIR Mhealth Uhealth. 2019 Apr 11;7(4):e11500. doi: 10.2196/11500. PMID: 30973342; PMCID: PMC6482399.

7. Althoff T, Sosič R, Hicks JL, King AC, Delp SL, Leskovec J. Large-scale physical activity data reveal worldwide activity inequality. Nature. 2017 July 20;547(7663):336-339. doi: 10.1038/nature23018. Epub 2017 Jul 10. PMID: 28693034; PMCID: PMC5774986.

8. Eurostat https://sport.ec.europa.eu/sites/default/files/physical-activity-factsheet_who-eu-201811_en.pdf

9. Taylor D. Physical activity is medicine for older adults. Postgrad Med J. 2014 Jan;90(1059):26–32. doi: 10.1136/postgradmedj-2012-131366. http://europepmc.org/abstract/MED/24255119. [PMC free article] [PubMed] [CrossRef] [Google Scholar].

10. Samani N. British Heart Foundation. Impact. 2017 Aug;2017(6):40–1. doi: 10.21820/23987073.2017.6.40.

11. Drillon P, Desvergée A, Prevost V, Blaizot X. Impact de l'activité physique adaptée sur les douleurs articulaires induites sous hormonothérapie adjuvante du cancer du sein : une revue de la littérature [Impact of adapted physical activity on joint pain induced

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406

under adjuvant hormone therapy for breast cancer: A review of the literature]. Ann Pharm Fr. 2023;81(1):1-12. doi:10.1016/j.pharma.2022.06.003.

12. Bluethmann SM, Truica C, Klepin HD, et al. Study design and methods for the using exercise to relieve joint pain and improve AI adherence in older breast cancer survivors (REJOIN) trial. J Geriatr Oncol. 2021;12(7):1146-1153. doi:10.1016/j.jgo.2021.05.011.

13. Arroll, B., Goodyear-Smith, F., Crengle, S., Gunn, J., Kerse, N., Fishman, T., ... & Hatcher, S. (2010). Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. The annals of family medicine, 8(4), 348-353.

14. . Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalised anxiety disorder: the GAD-7. Archives of internal medicine, 166(10), 1092-1097.

15. Plummer, F., Manea, L., Trepel, D., & McMillan, D. (2016). Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. General hospital psychiatry, 39, 24-31.

16. Mystakidou, K., Tsilika, E., Parpa, E., Gogou, P., Theodorakis, P., & Vlahos, L. (2010). Self-efficacy beliefs and levels of anxiety in advanced cancer patients. European journal of cancer care, 19(2), 205-211.

17. Philip, E. J., Merluzzi, T. V., Zhang, Z., & Heitzmann, C. A. (2013). Depression and cancer survivorship: importance of coping self-efficacy in post-treatment survivors. Psycho-Oncology, 22(5), 987-994.

18. S Ribeiro-Samy, et al. Sci Rep 2019, 9; C Lopes et al. Cancers 2021, 13(17): 4446

19. Pan LC, Wu XR, Lu Y, Zhang HQ, Zhou YL, Liu X, Liu SL, Yan QY. Artificial intelligence empowered digital health technologies in cancer survivorship care: A scoping review. Asia Pac J Oncol Nurs. 2022 Aug 23;9(12):100127. doi: 10.1016/j.apjon.2022.100127. PMID: 36176267; PMCID: PMC9513729.

20. Kelley MM, Kue J, Brophy L, Peabody AL, Foraker RE, Yen PY, Tucker S. Mobile Health Applications, Cancer Survivors, and Lifestyle Modification: An Integrative Review. Comput Inform Nurs. 2021 Jun 2;39(11):755-763. doi: 10.1097/CIN.0000000000000781. PMID: 34074873; PMCID: PMC8578050.

21. S. Pearce et al. 'More than just money'—widening the understanding of the costs involved in cancer care. J Adv Nur. (2001).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875406