Serums

Project no. 826278

SERUMS

Research & Innovation Action (RIA)

SECURING MEDICAL DATA IN SMART-PATIENT HEALTHCARE SYSTEMS

Report on Refined Use Cases and Evaluation of Serums Technologies in POCs and pilots.

D7.5

Due date of deliverable: 31st January 2021

Start date of project: 1st January 2019

Type: Deliverable
WP number: WP7

Responsible Institution: Zuyderland Medisch Centrum (ZMC)
Editor and editor's address: Ivo Buil (i.buil@zuyderland.nl)
Partners Contributing: FCRB, ZMC, USTAN, ACC, IBM, SOPRA, SCCH, UCY, UCL, UD

Approved by:
Reviewers: Marios Belk (UCY)
Vladimir Janjic (UD)
Technical Manager: Juliana Bowles

Version 1.0

Project co-founded by the European Commission within the Horizon H2020 Programme		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Release History

Release No.	**Date**	**Author(s)**	**Release Description/Changes made**
V0.1	1-10-2020	Mark Mestrum (ZMC)	Created document template based on D7.3 with executive summary and questionnaire for PoC 2
V0.2	9-12-2020	Mark Mestrum (ZMC), Ivo Buil (ZMC), Leon van de Weem (ZMC)	Adjusting ZMC Use Story
V0.3	7-01-2021	Ivo Buil (ZMC)	Added AMPI method and weights for KPI's and SI calculations
V0.4	11-01-2021	Marios Belk (UCY), Andreas Pitisllides (UCY), Christophoros Christophorou (UCY), Ivo Buil (ZMC), Santiago Iriso (FCRB), Julio Burgos (FCRB), Juliana Bowles (USTAN), Thais Webber (USTAN), Annemarie Paton (USTAN)	Added User KPIs, added adjustments to FCRB and USTAN User Stories.
V0.5	20-01-2021	Sascha van der Vliet (ACC), Bram Elshof (ACC), Ivo Buil (ZMC)	Added information on the execution of PoC 2 and future planning of PoC 3
V0.6	21-01-2021	Ivo Buil (ZMC), Larissa Haen-Jansen (ZMC), Santiago Iriso (FCRB), Julio Burgos (FCRB), Thais Webber (USTAN), Emma Morley (USTAN), Sascha van der Vliet (ACC)	Added end user evaluations
V0.7	25-01-2021	Wanting Huang (ACC), Sascha van der Vliet (ACC) Michael Roßbory (SCCH), Ivo Buil (ZMC), Leon van de Weem (ZMC), Julio Burgos (FCRB), Santiago Iriso (FCRB), Marios Belk (UCY), Andreas Pitisllides (UCY), Christophoros Christophorou (UCY), Euan Blackledge (SOPRA), Thais Webber (USTAN)	Added extension for transnational user stories and results for technical KPIs

V0.8	27-01-2021	Marios Belk (UCY), Andreas Pitisllides (UCY), Christophoros Christophorou (UCY), Euan Blackledge (SOPRA), Wanting Huang (ACC), Sascha van der Vliet (ACC), Michael Roßbory (SCCH), Santiago Iriso (FCRB), Julio Burgos (FCRB), Ivo Buil (ZMC), Larissa Haen-Jansen (ZMC), Thais Webber (USTAN), Emma Morley (USTAN)	Final adjustments to chapters 2, 3 and 4
V0.9	28-01-2021	Marios Belk (UCY), Andreas Pitisllides (UCY), Christophoros Christophorou (UCY), Euan Blackledge (SOPRA), Wanting Huang (ACC), Sascha van der Vliet (ACC), Michael Roßbory (SCCH), Santiago Iriso (FCRB), Julio Burgos (FCRB), Ivo Buil (ZMC), Larissa Haen-Jansen (ZMC)	Added the conclusion. Send document for reviewers
V1.0	31-01-2021	Marios Belk (UCY), Christophoros Christophorou (UCY), Ivo Buil (ZMC), Larissa Haen-Jansen (ZMC), Vladimir Janjic (UD), Juliana Bowles (USTAN)	Reviewed and corrected end version

SERUMS Consortium

Partner 1	**University of St Andrews (USTAN)**
Contact Person	Name: Juliana Bowles Email: jkfb@st-andrews.ac.uk
Partner 2	**Zuyderland Medisch Centrum (ZMC)**
Contact Person	Name: Larissa Haen-Jansen Email: la.jansen@zuyderland.nl
Partner 3	**Accenture B.V. (ACC)**
Contact Person	Name: Bram Elshof, Wanting Huang Email: bram.elshof@accenture.com, wanting.huang@accenture.com
Partner 4	**IBM Israel Science & Technology Ltd. (IBM)**
Contact Person	Name: Michael Vinov Email: vinov@il.ibm.com
Partner 5	**Sopra-Steria (SOPRA)**
Contact Person	Name: Andre Vermeulen Email: andreas.vermeulen@soprasteria.com
Partner 6	**Université Catholique de Louvain (UCL)**
Contact Person	Name: Axel Legay Email: axel.legay@uclouvian.be
Partner 7	**Software Competence Centre Hagenberg (SCCH)**
Contact Person	Name: Michael Roßbory Email: michael.rossbory@scch.at
Partner 8	**University of Cyprus (UCY)**
Contact Person	Andreas Pitsillides Email: andreas.pitsillides@ucy.ac.cy
Partner 9	**Fundació Clínic per a la Recerca Biomèdica (FCRB)**
Contact Person	Name: Santiago Iriso Email: siriso@clinic.cat
Partner 10	**University of Dundee (UD)**
Contact Person	Name: Vladimir Janjic Email: VJanjic001@dundee.ac.uk

Table of Contents

Executive Summary

Serums aims to increase the efficiency of healthcare systems in Europe while ensuring patient safety and the privacy of sensitive health data using innovative techniques that will increase resilience to cyber-attacks and promote trust in the safe and secure operation of the system. In order to meet this challenge, Serums will develop and implement innovative methods, tools and technologies addressing the need for cybersecurity in hospitals including remote care and home-care settings. Through these developments, the Serums project expects to achieve a significant impact in each area that has been identified in the SU-TDS-02-2018 call, providing significantly more secure smart health care provision, with significantly reduced potential for data breaches, and significantly improved patient trust and safety.

Work, related to the demonstration of the Serums technologies' effectiveness on real-world use cases from the domain of using and analysing medical data, was planned to proceed in three phases. The work performed during the first phase is presented in D7.3. The current deliverable (D7.5) presents the work performed during the second phase.

More specifically, during the second phase, prototypes of the Serums technologies have been refined, integrated and evaluated against the overall project requirements and success criteria that were identified in D7.4. In addition, the Serums technologies are evaluated using use cases supporting mechanisms to share information between patients and hospitals/medical centres. During this phase, focus is also given on ensuring ownership and appropriate involvement of all stakeholders/end-users within the medical centres, educating end-users on the future Proof of Concepts (PoCs) and Pilots and measuring the change progress. It is worth mentioning that during the execution of the second phase of the evaluation (Month 24 of the project; December 2020), the consortium had to deal with the consequences of the Covid-19 pandemic, affecting mostly the smooth operation of the Serums Second Proof of Concept (PoC2). More specifically, since PoC2 was to be conducted with a group of people who are highly vulnerable to the Covid-19 virus, it became clear that for medical, practical, and ethical reasons, the pilot could not continue as planned. Thus, despite the early preparations of the second PoC, Covid-19 brought up some challenges. First of all, since it was not possible to physically perform PoC2, participants needed to be recruited upfront to which a digital appointment was scheduled. The digital appointments were very difficult to be scheduled among the participants (both for the patients and the caregivers), a fact that had a negative effect on the time needed for the preparation of the PoC2 and also on the sample size for ZMC and FCRB. Also, the digital interviews were short and difficult to extract proper results.

The lessons learned and the results extracted from the work performed during the second phase will provide feedback into different technical work packages, steering the development of final Serums technologies and that will be used during the third and final phase of the evaluation. In the third (final) phase, the final versions of the use cases will be produced, also extending them with mechanisms for information sharing between patients, hospitals/medical centres, local e-health providers and other caregiver organisations (general practitioners/paramedics). During this phase, the required educational/information/training materials and environments for the PoCs and pilots will be designed, developed and tested, before the actual deployment of the PoCs and the Pilots, of the Serums tool and technologies, with the end-users of the medical centres. The results will be reported in D7.6 on M36 of the project.

1 Introduction

1.1 Role of the Deliverable

The role of this deliverable is to present the results of the work performed during the second phase of the demonstration of the Serums technologies effectiveness. More specifically, this deliverable: **i)** defines a detailed specification of the use cases (supporting basic information sharing between patients and hospitals/medical centres) that have been used for the evaluation of the refined prototypes of the Serums technologies; and ii) evaluate the refined prototypes of the Serums technologies developed, against the overall project requirements and success criteria that were identified in D7.4. During this phase focus is also given on ensuring ownership and appropriate involvement of all stakeholders/end-users within all medical centres, educate end-users on the Proof of Concepts (PoCs) and Pilots, and measure the change progress.

The lessons learned and the results extracted from the work performed during the second phase will provide feedback into different technical work packages, steering the development of final Serums technologies.

1.2 Relationship to Other SERUMS Deliverables

The relationship of D7.5 (that builds on D7.3) with the other SERUMS deliverables is provided in the figure below:

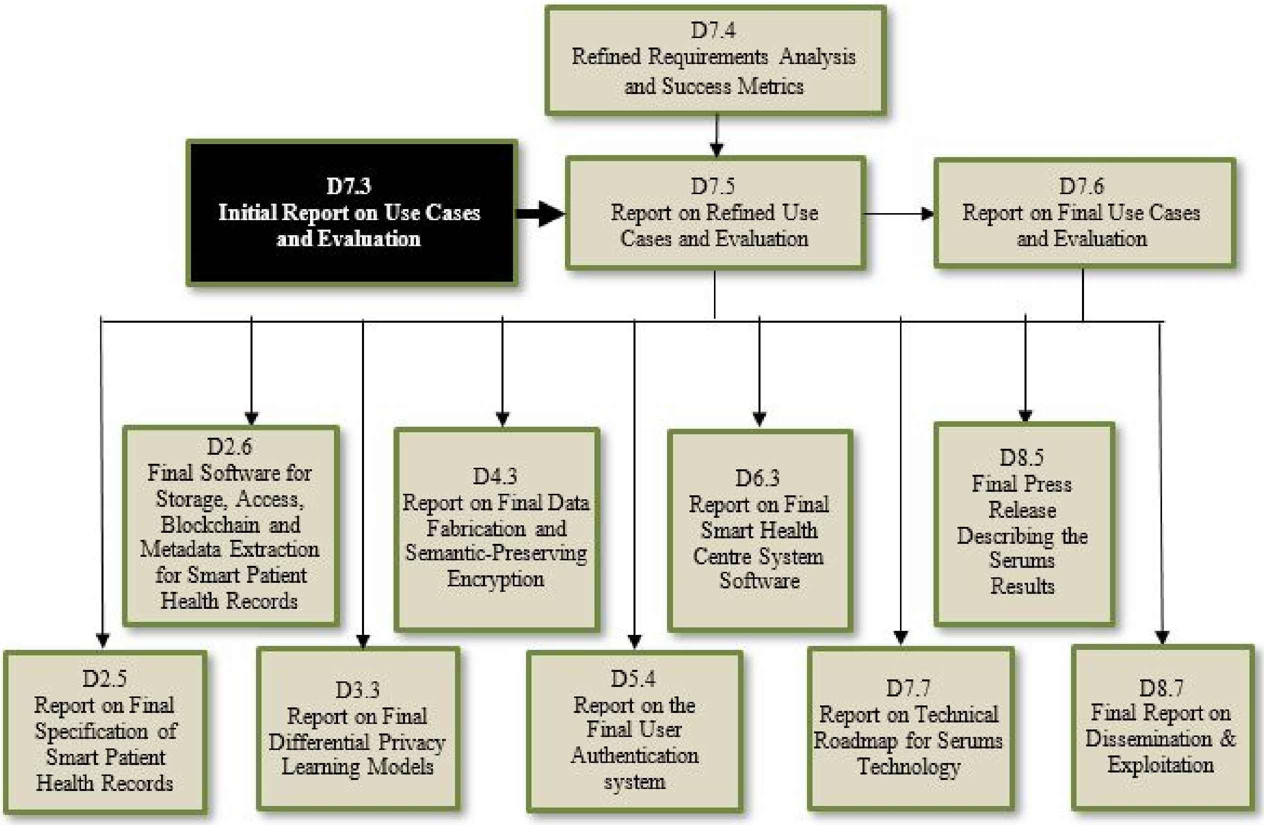

D7.4
Refined Requirements Analysis and Success Metrics
D7.3
Initial Report on Use Cases and Evaluation
D7.5
Report on Refined Use Cases and Evaluation
D7.6
Report on Final Use Cases and Evaluation
D2.6
Final Software for Storage, Access, Blockchain and Metadata Extraction for Smart Patient Health Records
D4.3
Report on Final Data Fabrication and Semantic-Preserving Encryption
D6.3
Report on Final Smart Health Centre System Software
D8.5
Final Press Release Describing the Serums Results
D2.5
Report on Final Specification of Smart Patient Health Records
D3.3
Report on Final Differential Privacy Learning Models
D5.4
Report on the Final User Authentication system
D7.7
Report on Technical Roadmap for Serums Technology
D8.7
Final Report on Dissemination & Exploitation

Figure 1. Table of relations between deliverables.

1.3 Structure of this Document

This document is organized having in mind the chronological order in which the different tasks described for WP7 and related to this Deliverable (T.7.2, T7.3 and T7.4) have been executed during the second phase of the demonstration of the Serums technologies effectiveness. Chapter 2 corresponds to T7.2 and elaborates on the use cases on which the different Serums technologies have been evaluated. Chapter 3 refers to T7.3, and provides information regarding the Proof of Concepts and pilots developed by Accenture and the respective Use case partners together with the remarks that have been provided by the users. Chapter 4 refers to T7.4, and reports on the evaluation of the Serums technologies effectiveness against the overall project requirements and success criteria that were identified in D7.4. Finally, Chapter 5 provides some conclusions.

2 Use Cases Specification

This chapter provides a detailed description of the use cases, supporting basic information sharing between patients and hospitals/medical centres, that have been used for the evaluation of the refined prototypes of the Serums technologies. Because the use cases described in D7.3 already contained the necessary refinements on the mechanisms to share information locally between patients, hospitals/medical centres and local e-health providers, during the second phase of the demonstration of the Serums technologies' effectiveness the same use cases described in D7.3 have been used. Note that the refined Serums technologies have been tested in realistic conditions with synthetic but realistic data produced using data-fabrication methods from WP4, which were obtained from private, confidential medical data. Additionally, in order to prepare for PoC 3 (September 2021) in the third and final phase of the evaluation, two of the three use cases (see the use cases of ZMC and FCRB in sections 2.1 and 2.2) have been extended with mechanisms to also share the same information on a transnational level.

The ZMC Smart Health Centre use case (see section 2.1) describes a system in which all the medical data of a patient (in this use case Peter, an 70-year old male) from hospitals, physiotherapists and wearables is stored, and where the patient can manage which caregiver has access to which part of his medical data. This use case exploits the smart patient records from WP2, privacy-preserving and secure communication mechanisms from WP4 for gathering data from different devices and authentication methods from WP5. The system developed was based on the Smart Health Centre System that will be developed in WP6.

The FCRB use case (see section 2.2) consists of the HCB - Smart Platform (HCB-SP). Together with the help of the technologies developed during the Serums project, we intend to give Joana, a 85 year-old patient with various chronic diseases, an easy way to gather her vital signs for the hospital using two wearable devices and the possibility to share them with all the professionals that care for her health, even in emergency situations abroad. As a use case, the HCB-SP exploits the smart patient records from WP2, privacy-preserving and secure communication mechanisms from WP4 for gathering data from different devices and authentication methods from WP5.

The USTAN use case (see section 2.3) mostly focuses on communication mechanisms for fetching the selected information to the central patient portal, displaying this information to the user and creating anonymous predictive outcome values in the future based on the stored information. Therefore, this use case mostly exploits mechanisms for smart patient records access from WP2, privacy preserving and secure communication mechanisms from WP4 and authentication methods from WP5.

2.1 ZMC - A New Hip

ZMC started the Serums project without any digital system with which patient data could be shared with the patient. The only way for the patient to acquire his medical records is to request a printable version of his records from the patient service center. Sharing medical data with other health care providers was thus also only done through printable PDF's.

The aim of ZMC with the Serums project is to research, test and develop a system that will be able to collect data from various sources and distribute them in a way that meets the criteria set up by the Dutch Government, GDPR and patient expectations. This means that ZMC expects the system to be secure, user-friendly, meet all the privacy-preserving regulations, able to obtain/share medical data from/to a wide scale of sources and distribute that data accordingly, including hospitals, patients, wearable devices, external physiotherapists and international clinics. We put a high emphasis on the correct use of blockchain and access/authentication rules. In addition, we encourage the use of only keeping the data at the original location, this way (i) the original data cannot be tampered with through this system and (ii) if a data breach does occur, not all the data is available in one place. Furthermore, the use of real fabricated data is essential in the development phase, as we can then see the capabilities of the system without violating privacy regulations. Despite the demanding technical requirements, it is also essential that the whole system will be extremely user-friendly, in such a way that most of our elderly patients do not find difficulties using the system.

In preparation for the second PoC, the first integrated solution for the Serums system was ready and tested by our patients. This first integrated version included the updated authentication system, a dual authentication option, the option for patients to view their real fabricated data in the data lakes and the ability to change the access rules on the blockchain.

In preparation for the third PoC in the final phase of the evaluation, the integrated solution will need to include mechanisms in order for the real fabricated data to be shared to an organisation in another country and to make sure that the patient can only give access to trustworthy sources. To meet this additional requirement, additional steps have been included in section 2.1.2.

2.1.1 ZMC User Story

Peter is a 70-year old male who has recently been provided with a new artificial hip at Zuyderland Medical Centre (ZMC). After a short stay at the hospital, Peter is dismissed and sent home to complete his recovery there. There he can already view his medical data related to his injury and operations in his account in the Personal Health Environment (PHE), because he arranged that before the operation.

Note: The medical files in the hospital regarding basic characteristics, injury, X-rays and operation details are easily accessible and compatible with the Personal Health Environment system.

Identification and authentication for sharing hospital data with the PHE needs to be done in a secure and user-friendly way according to the principles of the European GDPR guidelines and UAVG law in the Netherlands.

Note: Dutch law states that medical information can only be shared when the patient gives explicit permission for it. In addition, in consultation with his/her doctor, a patient is allowed to withhold access to certain parts of their medical information if they wish to.

To ensure Peter's recovery, the physician has ordered physiotherapy and the use of an Activity Monitor (AM) with an E-coach for 1 week. Prior to his surgery, Peter has already used the Activity Monitor, to measure his mobility before the hip replacement. The Activity Monitor is a very precise instrument that measures if and how well a patient is active at a validated clinical level.

Commentary: To comply with the GDPR, Peter must provide explicit permission to:

- Specifically allow the Activity Monitor to provide his personal activity data to: i) each medical practitioner that needs the results from it; and ii) the SHC.
- Specifically allow the E-Coach to share the guidelines that it provides with: i) each medical practitioner that needs the results from it; and ii) the SHC.
- Allow each medical practitioner to share their medical records with the SHC, and vice-versa.

Under the GDPR, Peter may revoke any of these permissions at any time, or choose to exclude some part of the information from being seen by any agent in the system, including historical information. However, doing this may be detrimental to his treatment, lead to false diagnoses, incur additional treatment costs, require him to take unnecessary drugs etc.

From the first session and the letter from the surgeon, the physiotherapist knows that Peter also wears an Activity Monitor. He knows the results will tell him how stable Peter's condition is with his new hip. Together with giving Peter exercises he can do at home, the physiotherapist asks Peter if he will allow him to see all relevant medical files from the hospital and the results from the Activity Monitor. They agree that Peter will share the files regarding the surgery and the daily results on the E-coach. Informed consent to share the data with the physiotherapist can be given via Peter's Personal Health Environment.

Note: Peter can choose in his Personal Health Environment which medical files and how long he wants to transfer the medical files from the Hospital and the E-coach to the physiotherapist. The rights, rules and communications of the data access will be ensured, logged, checked and tracked via Blockchain.

Peter knows that the Activity Monitor needs charging every day. Because of its accuracy it is a very energy consuming instrument and will last only 24 hours. Therefore, the nurse at the hospital explained to Peter that he needs to charge the battery every evening when he goes to bed. During charging, the measured data is transferred for analysis to show the results in the E-coach.

Note: The transfer needs to be secure. Data must be private and not tampered. The raw data of the Activity Monitor is transported to an external server and, analysed by a validated algorithm. Once the Activity Monitor receives a confirmation that the raw data is transferred successfully, it purges its data.

The results are available the next day to the physician via the E-coach and to Peter himself via his Personal Health Environment on his computer. Each morning Peter transfers the results from the Activity Monitor to the physiotherapist in his Personal Health Environment. Each day a trained nurse can then evaluate the stored results in the E-coach and can take actions accordingly.

Note:

1. Both the Activity Monitor and the validation algorithm do not know which patient is using which Activity Monitor. Yet in the E-coach and in the SHC the link between the sensor ID of the Activity Monitor and the patient must be made. This is a potential danger for the patient when not done correctly.
2. The E-coach needs to send the results to the PHE in a secure way so that a patient cannot be identified during transport.
3. Identification and authentication in the E-coach or portal needs to be done in a secure and user-friendly way.

Data transport must be private and secure. The physiotherapist can now view the results Peter has sent them. The data transport is logged in Peter's record in the Personal Health Environment.

Peter finds it hard to get up from a chair or bed. He is afraid that his new hip will hurt him, causing him to use his muscles wrong and his first steps to be unstable. After a while that feeling goes away, but the fear prevents him from exercising correctly. During his physiotherapy session on the fourth day Peter is told that he should try to exercise more and that he needs to put more pressure on the leg with the new hip. The results have shown that Peter did not do his exercises and that when he gets up, he is not standing straight which might cause Peter to fall. Peter promises to improve his exercises. The physiotherapist is able to explain the effect of this to Peter from the graphical image of the results.

On the fifth day the physician looks at the results of the last four days and concludes that Peter should have done better, but also sees an improvement on the fourth day. He tells the nurse to contact Peter and prolong the Activity Monitor until his 6 weeks follow-up session at the hospital.

Note: The extension of the Activity Monitor is administered in both the SHC and the E-coach and automatically visible in Peter's Personal Health Environment.

Peter improves his stability within the next four days. This is shown in the results of the Activity Monitor. The physician acknowledges this improvement and orders a digital consult with Peter for his standard 6-weeks follow-up. There is no need to see him physically. Peter will be asked to transfer the data from his physiotherapist to his physician when his physiotherapy has ended. Peter agrees and transfers the physiotherapy journal to his Personal Health Environment for the physician.

Note: This is administered in SAP. The request for sending the data from the physiotherapist is automatically visible in Peter's Personal Health Environment in a secure and private way.

Peter transfers in his Personal Health Environment the journal from his physiotherapist to his physician in a private and secure way. He has control over the timing and content of this transfer.

2.1.2 Additional Steps for PoC 3

Seven weeks after his operation, Peter goes on holiday to Barcelona. He is happy to be there, however the hotel has no elevator and his room is on the second floor. No worries, he has a new hip and an activity monitoring device. But after 5 days he notices that his hip is hurting him. Since his operation is not long ago, Peter worries. He therefore goes to the hospital clinic FCRB for a consult. He explains to the physician his complaints and tells him that he has currently received a new hip. Peter agrees with the physician from FCRB to share the necessary medical data with him from Peter's PHE.

Note: The data sharing now also needs to conform to the different national regulations and the Dutch data must become compatible with the Spanish system. This may include some translation of medical data.

The physician from FCRB evaluates the data and sees no abnormalities. He examines Peter walking and comes to the conclusion that Peter is not walking straight. This is due to the fact that he constantly wears flip flops during his holiday. The doctor checks his drug intolerance in his PHE, which only relates to iodine. He therefore prescribes a painkiller and gives the advice to wear normal shoes. His complaints are gone a few days later. The physician from FCRB asks Peter to share his findings with his Dutch physician and physiotherapist via Serums. Peter logs in and allows his physician and physiotherapist in the Netherlands to view the data from Barcelona.

Two weeks before the annual check-up, Peter is invited to the hospital for an X-ray of his hip and again receives the Activity Monitor and the E-coach from Zuyderland to monitor his recovery for 1 week. The results of the weekly Activity Monitor and X-Ray are positive. The physician orders the physician assistant to have a digital consult with Peter, as it isn't necessary to see him physically.

2.1.3 ZMC Serums system requirements

The table below shows which problems or needs arise in the ZMC use case, what solutions need to be implemented and which technical implications it gives.

No.	Problem/Need	Solution	Remarks/Notes	Technical Implication
1	**Under GDPR[1] Peter must provide explicit permission to share his health data across different caregivers, revoke these permissions or specify specific data to be shared in his Personal Health environment**			
a	Peters health data must be filling up in the personal health environment	Peter's Health environment is connected to Peter's health organisations to which data is sent or can be retrieved on the fly.	Aside from his health data, this includes the names and roles of the Care professional involved per organisation.	**Smart Patient Health Record**: This is a centralised data source that allows all of the patient's records to be accessed from a single source, regardless of the source system
b	Peter must be able to connect any external device and E-coach he wants	The results from the Activity Monitoring E-coach can be shared with Peter's Health environment		**Smart Patient Health Record:** The structure of the record allows for seamless integration of any additional data sources
c	Peter must be able to log in with the method and options he prefers	Peter logs in to his Personal Health environment using Picture Guessing.	Of course all types of his preferred authentication methods (e.g., graphical password or textual password, Two-Factor Authentication) need to be accessible	**Personalized User Authentication:** Based on the suggested flexible and personalized authentication approach, end-users have the option to choose their preferred authentication method (*i.e.*, graphical or textual) in order to login. After successfully entering the password secret, for adding an

[1] Information on GDPR can be found at https://gdpr-info.eu/

				additional layer of security, a push notification is sent to the end-user’s mobile device that (s)he needs to approve in order to complete the login process. After successful completion of the login process, the authentication system generates a security token (JWT) and sends it to the client that is used for subsequent requests to the Serums systems.
d	Patients need full control over which data is sent, to who (and who not), and for how long.	Peter sees in his Personal Health on a special page his health data grouped by device/organization (including external physiotherapy) and if it is shared, partly shared or not shared.	The view can be also the other way around. Peter selects an organisation or Care professional and then looks at which data is shared with that organisation. In the end it all comes down to Care professional --> permissions <-- data. It is a n to m relation The same way sharing data is allowed, so is revoking sharing the data.	**Blockchain:** The default permission for the caregiver to access the patient is defined by the hospital administrator. Patients have the possibility to view existing rules, create additional rules to permit or restrict access for a selected set of data.
e		Peter selects the Activity Monitor from the above mentioned list and allows sharing		
f		Peter then sees the organisations he can share the data with and selects his physiotherapist		
g		Peter now has the option to choose a certain period of time he wants to share his Activity Monitor data. Since Peter only uses the Activity Monitor for a week, Peter chooses this time frame for sharing. Furthermore he checks all data to be shared		
h		Peter needs to confirm this request for sharing and is then led back to the page where he can see in health data and if it is shared, partly shared or not shared		**Blockchain:** Patient’s confirmation triggers the creation of the rule to allow the caregiver specified in

				the rule to access his data.
k	Setting specific documents to be shared	Some of the medical data from the hospital contains subsets of data. Peter can choose whether he wants to share all data or specific data. Peter selects his hip operation details.	Since this a specific part of the data and a single document no time frame will be asked.	**Smart Patient Health Record:** The record stores data in a Data Vault structure, wherein only highly correlated data is stored in the same satellite. This works in conjunction with the Blockchain to ensure granular control over the access

Table 1. System requirements for ZMC User Story

2.2 FCRB - Chronic Disease Management (HCB-SM)

With the new strong need of securing data and making the patient the owner of it due to the new European regulations, the Hospital Clinic de Barcelona is in a strong need of innovations in their ICT infrastructure. This is especially true in terms of making the patients the owners of their data since this is a total change of how their data has been managed traditionally, where the patient has only been able to request it when needed in case that the data was needed to receive treatment in another health centre not included in the Catalan Public Health System. On the other hand, the increase in data leaks from health organizations has made the need for new secure systems scale to maximum priority.

These two urgent needs are expected to be satisfied during the Serums project, on one side FCRB expects a practical proposal on the access permit management by ACC, since if it was to be implanted would mean a dramatic improvement in our organization. On the other side, as the times of Big Data and Artificial Intelligence demand for an improved data organization of the patient record and all the Hospital data in general FCRB is interested in how SOPRA manages to create their Data Lake.

Finally on the side of UCY's Proposed Authentication Scheme, FCRB is indeed interested on it if it represents an improvement over the habitual passwords in security and usability, but it will have to prove that the system will help old people to interact with the Hospital systems since they are the bigger proportion of the population in the Catalan Health Assistance.

Although the technologies will be integrated in the Smart Health Centre System, FCRB and the Hospital Clínic de Barcelona intend to integrate the Serums technologies to create a patient-oriented distributed system of their own. This new system will allow the Hospital patients to upload data gathered by e-health devices that can be taken home, allowing them to monitor patients with multiple chronic diseases. Section 2.2.1 presents the user story that this system plans to satisfy.

On the other side, the Privacy Preserving Data Analytics is a completely new approach to security that at Hospital Clínic de Barcelona has never been considered, nevertheless it is not refused by any means, it will be considered and with this project we hope that SCCH does a great job of showing its advantages and novel features over the more traditional approaches.

As referred above, in preparation for the third PoC in the final phase of the evaluation, the integrated solution will need to include mechanisms in order for the real fabricated data to be shared to an organisation in another country and to make sure that the patient can only give access to trustworthy sources. To meet this additional requirement, additional steps have been included in section 2.2.2.

2.2.1 FCRB Use Story

Joana is 85 years old female with several chronic diseases: she has diabetes and chronic heart failure (for which she receives medication). Joana lives in a private apartment close to a Primary Care Centre. She is getting some care via the Primary Care Centre but wants to remain independent for as long as possible. For that reason, her Doctor, from the Hospital Clínic de Barcelona, specialist in Diabetes, has given her wearable medical devices: i) a wireless pulse oximeter, to monitor her oxygen blood percent and her cardiac frequency; and ii) a wireless glucometer to measure her own glycemia.

Note: The following sensors will be made available to the health professionals to give to the patients:

- Pulse oximeter
- Glucometer
- Thermometer
- Tensiometer

All these devices will connect wirelessly to a smartphone application through Bluetooth 5 in a secure way.

Serums Interaction: The user will gain access to the HCB-SP through the authentication system provided by UCY, which frontend will be embedded in the Patient application and Professional platform.

For the second device, Joana has been informed that she will have to periodically upload her glycemia and oxygen in blood results to the HCB-SP platform through a mobile phone application called Saludata which basic usage has been taught by the doctors.

Serums Interaction: All information concerning patient record data and the measurements taken by the eHealth devices will be securely stored on the Data Vault provided by SOPRA. None of the HCB-SP will ever store personal data; these will always be retrieved from the Data Vault when needed.

Joana is happy with this because she can control her progress in this matter. With this smartphone application Joana is totally in control of the data generated by the devices and her patient record. Joana has therefore given the doctor her permission to access her data on that platform.

Serums interaction: The access and modification permissions over the patient data will be stored in the Blockchain solution developed by ACC. This will include various levels of information access, from only accessing the Patient Record to the granular access to only the information related to an aspect of the Record History (e.g. Endocrinology Record, Quirurgic Operation, etc.)

The Doctor has also commented to Joana that her General Practitioner would also need to have access to the glycemia web portal to monitor her evolution and he will contact her to follow up on that, and also on the rest of her health issues.

Note: The Blockchain solution is not only for personal permission and professional, but also for departments, organizations and for the whole hospital.

In addition, more complex rules can be generated by Joana or the Hospital administrators.

Explanation: The Saludata application is to be in full compliance with the GDPR and by thus has to provide:

- Full control of who can access the patient data.
- Full control of which parts of the patient record each hospital, professional or service can access.

On the other side, a cardiology medical team is in charge of taking care of her chronic heart failure and is composed of two nurses and one cardiologist. One of their tasks is to monitor the evolution of the patients with chronic heart failure at home, they receive and monitor all the data generated by the wireless pulse oximeter through an application installed in local servers of the hospital, where they can review Joana's list of measurements and communicate with her through notes with her smartphone app if necessary.

The hospital nurse periodically generates a clinical note with the events that have occurred and sends this to the patients. With this information and the glycemic control from Joana's device, the General Practitioner can (with Joana's approval) collaborate to monitor, control and detect abnormalities not only in one of those two diseases but can merge all of Joana's health issues and provide her with a better quality of life, by taking an holistic approach of her health status.

In terms of the technical flow of the use case, first the patient will be told by the hospital or their caregivers to download a Smartphone application in order to communicate with the eHealth Devices and with the Central System. Patient's devices will be connected to the application in a secure standard way and all the health data generated for this application will be stored into the Central System's Data vault. The application can also retrieve the history of personal health measurements, grant or revoke permits to the professionals, groups or caregivers stored in the blockchain and send notes to them in a secure way.

The HCB-SP will have a second part that will be used by the caregiver to retrieve and review patient data to which it has given permits and send notes in the case it is considered necessary. As the smartphone application, this system communicates with the Authentication System, the Blockchain solution and the Datavault in order to perform adequately. Nevertheless, this platform won't be installed in the user system but will be integrated with our Information Communications and Technologic Systems (ICT) and will be presented to the patient as a web application only accessible through the Hospital Network.

Both systems will communicate with the Central System provided by the SERUMS project using the Authentication Schema (UCY) and retrieving and storing data from the Data Vault (SOPRA) depending on the permits each user has on the Blockchain Solution (ACC).

2.2.2 Additional Steps for PoC 3

During the summer, Joana decided to do an all-inclusive trip to Scotland. Her doctors have recommended that she'll take her wearable medical devices with her, and so she has. While enjoying an organized sightseeing tour in Edinburgh, she passes out, and is taken to the hospital with an ambulance where she remains unconscious. In the ambulance they find a card in her wallet that tells them she has a patient record in Serums (with her Patient-ID in Serums). Since Joanna remains unconscious the Emergency doctor activates the emergency option in Serums. The emergency doctor is thus granted access to all medical files from Hospital Clinic de Barcelona. They are also able to see the most recent results from her wireless pulse oximeter and wireless glucometer. Here they notice that her glucose levels have dropped too much and caused hypoglycemia. The hospital in FCRB is notified that Joanna's patient record has been lifted due to the use of the emergency button.

Note: Certain figures within care organisations should be allowed to issue an emergency response that will show the doctor all information about a certain patient in emergencies once the patient is unable to grant permission his-/herself. This event will indisputably be registered in the Blockchain and all direct caregivers will be shown that this event occurred.

After treatment she was able to recover that same evening. The doctors in Edinburgh were able to help Joana adjust the insulin dosages for the rest of her vacation and as long as she regularly checks her glucose values, she should be able to enjoy the rest of her vacation. The emergency doctor creates a record in his Hospital system/NHS and recommends Joana to share this with her local doctor via the Serums system.

Note: Starting an emergency response should automatically create a Personal Health Record for the hospital in question. The patient can then grant access to other caregivers to see this record.

2.2.3 FCRB Serums system requirements

The table below shows which problems or needs arise in the FCRB use case, what solutions need to be implemented and which technical implications it gives.

No.	**Problem/Need**	**Solution**	**Remarks/Notes**	**Technical Implication**
1	**Vital Sign Monitoring**			
a	The health professionals need to have all the vital signs stored in only one platform.	The Saludata smartphone application would gather all the measurements from different devices in only one platform facilitating a complete monitoring of the patients.	Professionals often find themselves having to access multiple platforms from different vendors and devices	**Smart Patient Health Record**: This is a centralised data source that allows all of the patient's records to be accessed from a single source, regardless of the source system
b	The ability to have a periodic stream of vital sign data from chronic disease patients would greatly help the health professionals to treat them	The Saludata smartphone application for patients is able to read vital signs measurements and store for review or for the professionals to see them.		**Personalized User Authentication**: patient and professional have to be authenticated and thus in possession of their security token (JWT), that will be used to utilize all the other technologies **Blockchain**: When health professionals need to access the new measurement data, it will be checked whether the requestor has the corresponding permission to access this patient's data. When positive, a request will be triggered to retrieve the data. **Smart Patient Health**

				Record: The data will be retrieved from this system and sent to the end-systems in a secure way using SFTP
2	**Improvement in Security**			
c	Data exchanged between health assistance actors and patients' needs to be secure.	In the whole platform securely communications, storage and access will be enforced	This includes each element in the communication chain or any component with which the system has relation	**Smart Patient Health Record:** When a rule is successfully triggered on the Blockchain, the corresponding set of data will be moved to a secure location in the Data Lake and encrypted by a unique public key provided by the request. Once it is encrypted, it can be passed to the Serums system, with only the correct private key allowing the decryption
3	**In compliance with the GDPR compliance and data protection**			
d	Joana needs to be able to grant and deny access to her data to the distinct actors in their health assistance (doctors, nurses, hospitals, services, etc)	Through the Saludata application Joana will be able to create and eliminate these permits, allowing her to manage granular access to her patient record.		**Personalized User Authentication**: patients and professionals have to be authenticated and thus in possession of their security token (JWT), that will be used to use all the other technologies. **Blockchain:** Permission rules to grant or restrict access can be defined by the patient for health organizations, individuals or groups.
e		Joana is able to remove access to certain professionals or assistance services that are part of an allowed organization.		**Blockchain:** Although default rules for the caregiver to access the patient is defined by the hospital administrator according to national regulations. Patients

				have the possibility to create specific rules to permit or restrict access.
4	**Improvement in Patient-Health Assistant communication**			
f	Communication between professionals and patients would be beneficial in the treatment of chronic diseases.	Both the patient application and the professional platform allow us to exchange messages through secure channels.		

Table 2. System requirements for FCRB User Story

2.3 USTAN - Chemotherapy Toxicity Predictor

There are always ways in which healthcare provision can be enhanced, particularly concerning the management and treatment of highly complex chronic conditions such as cancer. Such conditions are more prevalent in the elderly who often have additional comorbidities. For this use case, we consider breast cancer patients with comorbidities and undergoing chemotherapy. These patients only come to hospital for chemotherapy treatment approximately every three weeks, and stay at home between treatments. The Edinburgh Cancer Centre (ECC) and the Western General Hospital (WGH) within NHS Lothian[2], have an interest in exploring the use of technological solutions to improve the monitoring of the wellbeing of patients at home, and detect any changes in symptoms and side effects that need to be controlled. The vision of this use case is to add a way for patients to record symptoms daily whilst at home, and send their data to the hospital and their registered medical practice so that the team at the WGH and the patient's GP can monitor and intervene as necessary. This should be done securely and preserving the confidentiality of the data. The data can be further integrated with the treatment information to more accurately predict further evolution, and compare patient outcomes with cohorts of similar patients. Furthermore, giving digital solutions to patients also makes them more involved with their own treatment and health which would change the perception of cancer treatments in the future. In addition, the ability to fine tune treatments to individuals is a trend within personalised medicine. This is not how currently patients are treated in the WGH and NHS Lothian.

There are two main roles for this use case within Serums. One is for the creation of a chemocare toxicity predictor which uses fabricated data from the IBM Data Fabrication Platform (WP4), and the associated data that is passed to the data lake from SOPRA (WP2), and further visualised within the integration work in the Smart Health Centre System (SHCS) within WP6. In addition, the USTAN dataset will be applied as an evaluation dataset for the privacy-preserving machine learning approach developed by SCCH (WP3). A secondary role is the evaluation of the UCY's proposed authentication scheme. Even though it explores new ways of capturing passwords through images, which may be more memorable, this is a feature that we will not be able to evaluate at present with ECC/WGH participants for one main reason: ECC is not a direct partner of the project, and we are hence not able to hold meetings at the WGH and interact with patients and participants in that setting. The development of models for the toxicity predictor from large fabricated datasets will give us new main research insights which our colleague Dr Peter Hall, ECC/WGH, is very committed to.

2.3.1 USTAN User Story

Emma is a 38 years old patient in the Western General Hospital (WGH) who has recently been diagnosed with breast cancer. To prevent the spreading of the tumour, she underwent breast surgery. After her surgery, chemotherapy treatment is given as a follow-up to her surgery. She is now dismissed and only visits the hospital for her chemotherapy appointments every three weeks.

To ensure her wellbeing and best outcomes, and to be sure that the treatment plan is suitable and minimises further side-effects and further hospitalisations, a treatment plan and regimen have been established (this will be over several months with treatment in the hospital every three weeks). Emma also has a comorbidity. As any cancer patient on chemotherapy, she might have higher toxicity levels as a result, but it is crucial to guarantee that the scale does not go above level three. Toxicity levels range from 0 (no toxicity) to 5 (very high toxicity).

[2] https://www.nhslothian.scot

Emma agrees on using and sharing data between treatment visits via the cancer data gateway and patient portal. Emma determines who in the medical team sees this information: The oncologist/nurse and her GP. Emma is also informed about how to use the web application and pass on relevant information to the clinical team.

Via a user-friendly web application, Emma can provide information on symptoms daily throughout the treatment. These Patient Reported Outcome Measures (PROMS) are based on questionnaires. Severe reported symptoms can be picked up by the clinical team and acted upon as soon as possible.

Note: The conditions that are being monitored and provided by the patients are nausea, vomiting, diarrhoea, constipation, oral mucositis, oesophagitis, neurotoxicity, hypersensitivity and fatigue. With these symptoms, the oncologist can determine the level of patient toxicity.

The information Emma provides about patient characteristics, cancer information hospitalisation data, and information about comorbidities, are all combined.

This combined data will help clinicians adapt treatments better to Emma as an individual patient which results in controlled toxicity levels and improved health outcomes. It uses data from several patients treated over the years with comparable characteristics.

If during the treatment there are signs that toxicity levels are high or that Emma's condition is deteriorating, one of the members of the clinical team (e.g. oncologist, specialist consultant, nurse, GP) will identify the irregularities in Emma's data, and contact Emma to intervene.

During a phone call, a decision is made for the GP/nurse to visit Emma at home and provide some additional medication to alleviate symptoms. Admission to hospital is not necessary. As scheduled, Emma comes to the WGH for the next chemotherapy treatment. This procedure is iterative until the end of the chemotherapy treatment.

Overall, Emma can have more personalised treatment. If a complication arises, the clinical team can act more quickly. Furthermore, Emma's well-being increases as she gets more involved in her treatment plans.

We are developing a dashboard to help oncologists observe, monitor, and analyse the condition of their patients over time. It can also be used to analyse the effect of different chemotherapy treatments when given to patients with similar characteristics, and consequently influence future decisions to improve the well-being and survival rate of patients. Our ultimate aim is to have a system to predict the toxicity of chemotherapy treatments based on history and feedback from patients. The overall features of the system is shown below.

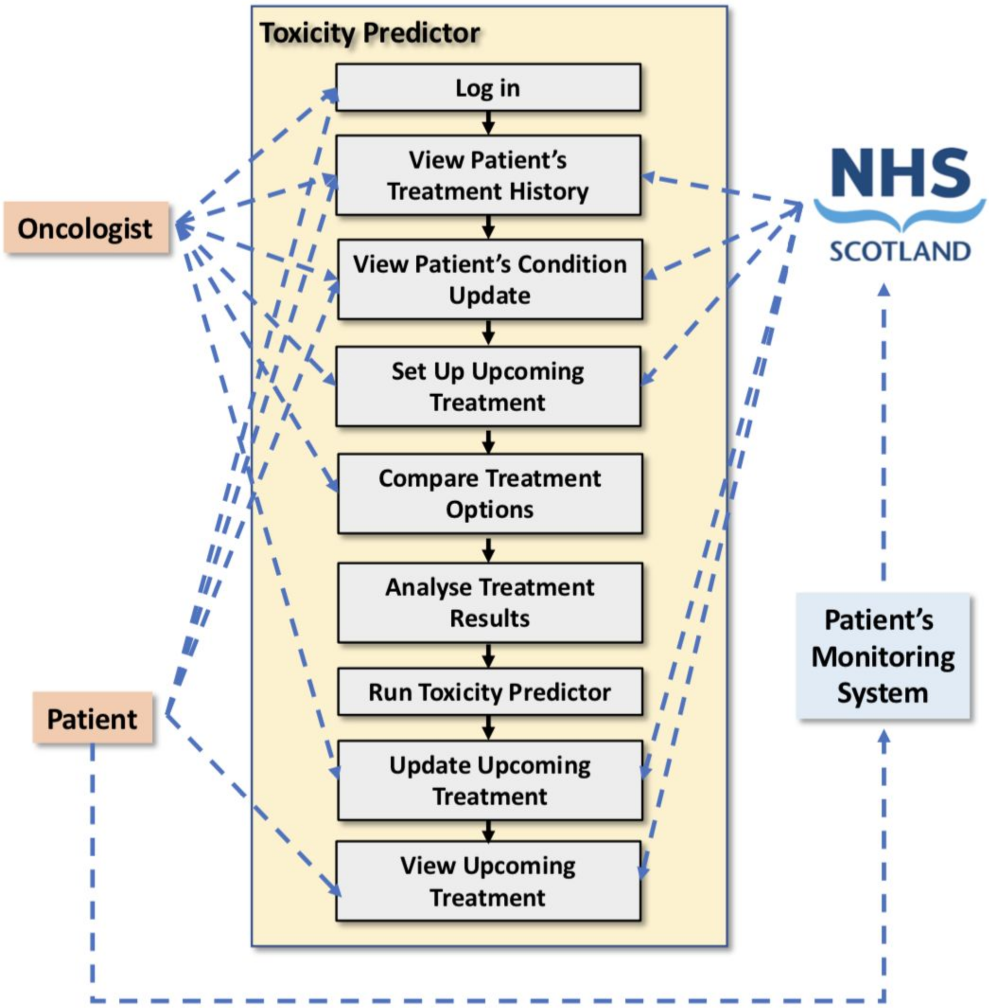

Toxicity Predictor
Log in
View Patient's Treatment History
View Patient's Condition Update
Set Up Upcoming Treatment
Compare Treatment Options
Analyse Treatment Results
Run Toxicity Predictor
Update Upcoming Treatment
View Upcoming Treatment
Oncologist
Patient
NHS SCOTLAND
Patient's Monitoring System

Figure 2. Use case analysis for USTAN

2.3.2 USTAN Serums system requirements

The table below shows which problems or needs arise in the USTAN use case, what solutions need to be implemented and which technical implications it gives.

No.	Problem/Need	Solution	Remarks/Notes	Technical Implication
1	**More personalised treatment with improved and more regular monitoring of side effects. This will enable the clinical team to act more quickly when complications arise.**			
a	The oncologist needs to be able to observe the patient's condition before giving them the next	The developed system (SESO Gateway) provides the patient timeline visualisation, which	The oncologist may need to access multiple platforms for monitoring the patient's condition.	**Smart Patient Health Record**: This is a centralised data source that allows all of the

	chemotherapy treatment. They can see the medications prescribed by the GP for other diagnosed conditions as well. Some of the information is not accessible by the oncologist as it is held by the system of the medical practice.	shows the overall patient's cancer care journey. It allows the oncologist to see the latest patient's toxicity/condition measurement result as well as how it may have changed over time as a consequence of the chemotherapy. There is information on medications currently being taken and for what conditions.		patient's records to be accessed from a single source, regardless of the source system. This will work as an extension for the SESO Gateway system.
b	Streamline the process of providing information on the patient's condition between hospital visits.	The monitoring application allows the patient to give an update on her symptoms anytime, anywhere. NHS Lothian started developing a smartphone application which allows the patients to input their condition. The data is directly stored and collected in the NHS database.	The SESO Gateway can directly access the information from the database.	**Smart Patient Health Record:** The SERUMS data lake would have to access the same information as the SESO Gateway.
c	The oncologist can access tools that provide second opinions regarding the upcoming treatment of the patient.	The SESO Gateway has a feature for predicting the upcoming treatment result by inputting the treatment into machine learning models.	The accuracy of the predictor needs to be improved. At the moment, we develop it as proof of concept due to data scarcity.	**Smart Patient Health Record:** The SERUMS platform through the data lake could integrate recommendations from different GPs/oncologists reviewing the case. Additionally, the Privacy Preserving Machine Learning module of SERUMS can help to analyze existing treatments and recommend alternatives based on

				existing and/or synthetic data.
2	**In compliance with the GDPR and data protection**[3]			
d	The patient gives or withdraws her consent to specific professionals (oncologists, GPs, professionals) for her data access at any time, in a secure and transparent fashion. Data from hospitals databases and the data collected from devices, or the patient's home environment, are integrated into the smart patient record.	Assuring the patients that the application will maintain *data privacy* by showing medical data only to authorised users based on custom access rules. Also, the application guarantees *ownership verification* as only reliable and verified sources will be included. *Data integrity* is achieved by means of encryption through all the transmission channels, and only the final user will have the private key needed to decode the data. *Transparent and traceable transactions* are registered in a ledger supported by a blockchain.	Different healthcare providers have varied sets of static information and dynamic information stored in their *own proprietary formats* in several databases. Personal monitoring devices from patients also acquire medical information in different formats. Patients are allowed to give or withdraw access under their *own personal reasons* and following *government laws and institutions policies*. In principle, the *patient will be the owner of her data*, for that she will decide who has access to it, as long as it complies with the national and institutional regulations.	**Blockchain:** The SERUMS platform through an integrated interface to define *access rules*, links the features of the *blockchain and data lake components* to retrieve requested data for authenticated and authorised users. This interface enables the user to interact with formal definition of access rules through a more intuitive natural language version of such rules. **Personalized User Authentication**: Patients are able to authorise access over well-defined data categories (called tags), thus providing high granularity for rules. **Smart Patient Health Record:** In both cases, data from hospitals and from out-of-hospitals environments are integrated in the data lake with specific tags. Devices are able to connect with Serums API to securely transfer medical data.

[3] Information on GDPR can be found at https://gdpr-info.eu/

				Blockchain: The blockchain module also maintains a registry of each transaction that will ensure accountability through audit trails.

Table 3. System requirements for USTAN User Story

3 Proof of Concept

The second Proof of Concept (PoC2) took place in December 2020 at the three end user locations, with all consortium partners involved, based on the previously drawn up hospital-specific use cases. The measurements were carried out both qualitatively, through semi-structured interviews and a focus group, and quantitatively, through usability metrics and questionnaires. In this section, we will further dedicate these measurements and by whom the measurements were carried out.

In this consortium, three hospitals have been designated as end users to carry out the PoC2 measurements and to test the future Serums policy. These three end user locations are:

- Fundació Clínic per a la Recerca Biomèdica (FCRB)
- Zuyderland Medical Centre (ZMC)
- Edinburgh Cancer Centre (ECC) [4]

All end user locations developed a use case, as described in Chapter 2. Those use cases were used as the foundation for the PoC and the recruitment of participants.

3.1 Measurement design

The measurements were carried out both qualitatively, through semi-structured interviews and a focus group, and quantitatively, through usability metrics and questionnaires. For a comprehensive data collection, three stakeholder groups were included, which were patients, healthcare professionals, and IT staff.

3.1.1 Patients

The initial focus in the collection of this stakeholder group was on finding patients that belong to the proposed hospital-specific use cases. Due to Covid-19, all interviews with patients needed to be performed digitally. The sessions with the patients took approximately 60 minutes. The first part of the session was a guide through the user system. The second part of the session focused on answering the questionnaire. An example of the planning is provided in Figure 3.

Patient data has been measured using questionnaires. These questionnaires were performed in each of the hospitals at the same time to minimize bias. Also, an interview with one patient was conducted per end-user location to obtain more in-depth information in addition to the results of the questionnaires. The interview guide and questionnaire can be found in respectively Appendix 1 and 2. The patient chosen for the interview has been randomly selected from the patient participant population voluntarily.

The questions for both the questionnaires and the semi-structured interviews were prepared by UCY, SOPRA and the end-user locations. The questions have been used to gain information about Success indicator 3 and included information on perceived usability, perceived memorability, perceived security, trust in the proposed Personalized User Authentication (PUA) system, and patient trust towards the overall SHCS system (KPI 3.1 till KPI 3.5). During the interview, all interactions with the PUA was measured indirectly and used to substantiate Success indicator 1. In addition, other metrics were retrieved for the other Serums technologies.

3.1.2 Care professionals

Care professionals were carefully selected by the hospitals themselves and included in the PoC measurement after voluntary consent. The healthcare professionals who were included for the measurements were

[4] The Edinburgh Cancer Centre (ECC) is not a direct partner in the Serums project

specifically selected per hospital based on the proposed use case and included medical specialists as well as other healthcare providers (e.g., physiotherapists and nurses).

Data of the care professionals was collected using semi-structured interviews per end-user location, in which the questions corresponded to the questions from the questionnaire in Appendix 2, which was used for interviewing patients. These interviews took approximately 60 minutes. KPI 3.5 was not measured in the interviews with the care professional since this KPI is focused on patient trust.

3.1.3 IT staff

The IT staff participating in this PoC measurement were measured through a small focus group facilitated by the end user locations and the integration partners. The IT staff was the final stakeholder group measured and their input mainly focused on the security of data processing, for the whole process and further developments of the Serums policy. Feedback from the security and IT experts is used for further refining the Serums technologies in the next development life cycle, e.g., feedback on security aspects of the user authentication technology will be used as input and reported in D5.4 Report on Implementation and Verification of Final User Authentication System, Security Metrics and Authentication Policies (due on M34).

The focus group with hospital IT staff took place at the end user locations and could be seen as a homogeneous internal focus group. Before all measurements, permission was requested from all participants by means of an Informed Consent drawn up in their language. In addition, participants received an informative briefing upfront the focus group to be able to have a more in-depth discussion about the Serums Policy. Integration partners were available to deepen possible discussions. The focus group took approximately one hour. The research was approved for each hospital by the ethics committee before the baseline measurement was carried out.

3.2 PoC execution at the End users locations

This section further explains the PoC at the specific end user locations. Table 4 shows the amount of participants that were included at all end user locations.

At ZMC, the second PoC took place between November 30th and December 18th 2020. ZMC started earlier with the PoC to work as an example for the other locations. In those two weeks, interviews with medical personnel and patients and a focus group with IT personnel were executed. Recruitment of the patients was done via the relevant use case physicians. When patients had a consultation with the orthopedic physician, they were introduced to the research. If patients were open to participate in the research, an informed consent was asked by the Serums project leader and an appointment for the PoC participation was scheduled. During the PoC, additional patients that did not fit the use case were recruited, by use of the internal patient panel, in order to reach a more sufficient number of participants. This is a group of patients connected to ZMC who have indicated that they are willing to participate in research. Participants received a €10 voucher from a major Dutch online store (Bol.com) after input was received on the 6th day. In total, ZMC was able to include 14 patients. In addition, four people, part of the medical staff, were interviewed. The last component of the PoC at ZMC was the focus group conducted with two IT professionals.

At FCRB, the PoC took place between the 3rd of December and December 18th 2020. In order to be able to include the lessons learned in the execution, both FCRB and ECC started a few days later than ZMC. In addition, FCRB’s patients were recruited via an existing list of patients that were willing to participate in research. The focus groups and interviews with medical personnel took place in the last week of the PoC. 15 patients and 4 caregivers were interviewed, and 2 IT professionals attended the focus group.

At USTAN, the PoC took place between the 7th and 18th of December 2020 and tested the system on a group of 26 volunteers from the area local to ECC. Participants were recruited by the USTAN team via a public information flyer and via social media. The recruitment announcement was shared in local Facebook groups across Fife, Tayside, and Edinburgh to ensure only local patients and medical personnel were interviewed. Potential participants expressed their interest by emailing the Serums local email address. Interviews were executed via Microsoft Teams with at least two members of the USTAN team. Following two successful (or attempted) logins, each participant was emailed a link to activate their £10 Amazon Voucher in retribution for their time testing the SHCS and the use of their broadband. In total, USTAN was able to include 26 participants in the research, of which 3 were working in healthcare settings. The focus group with the IT staff was not executed at USTAN.

	FCRB	ZMC	ECC	Proposed
Patients	15	14	23	25
Medical personnel	4	4	3	5
IT staff	2	2	0	5

Table 4. Participants included in PoC 2

3.3 Planning

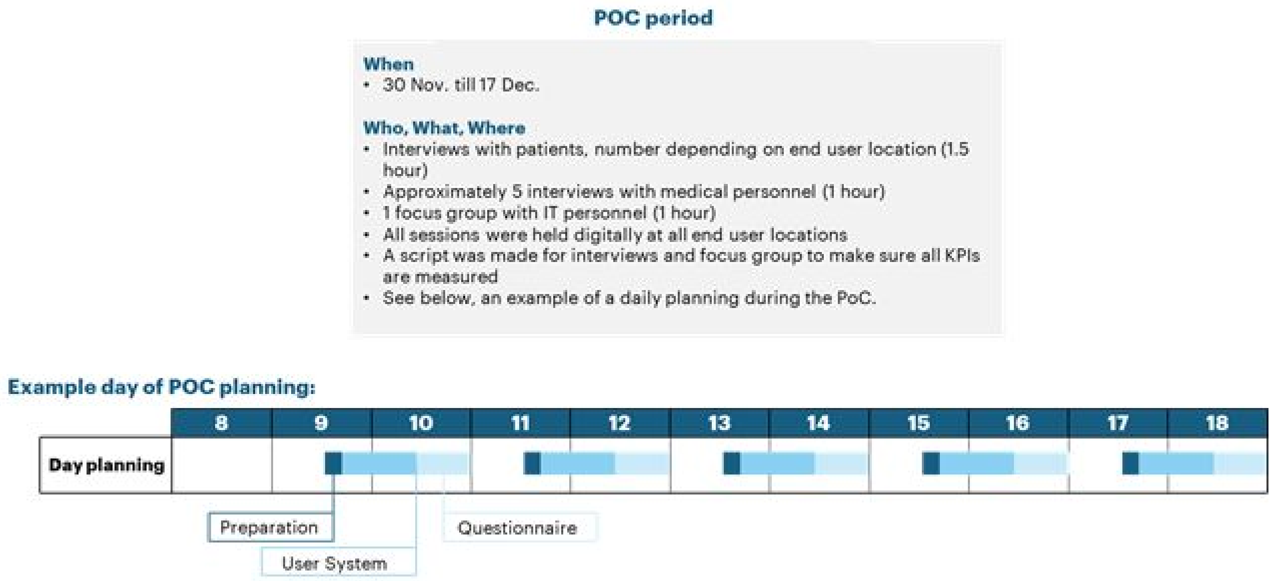
In total three PoC measurements were planned to be carried out throughout the Serums project lifetime. The first PoC (which is described in D7.3) was carried out in Month 13 of the project (January and early February 2020). The second PoC (that is described above) was carried out in Month 24 of the project (December 2020).The final PoC measurement will be performed at all three end-user locations in Month 33 (September 2021).

The first PoC was measured in two of the three previously agreed locations. In order to make sure that the second PoC could take place at all three locations, preparations started early. Despite the early preparations, Covid-19 brought up some challenges. For example, unlike previous year, it was not possible to physically perform the PoC. Participants needed to be recruited upfront, to which a digital appointment was scheduled. This required more time and therefore, instead of having the PoC take 2 days per location, it was decided to extend it to three weeks. To minimize bias, all PoC's were planned between the same dates.

At the end of the year, we expect that the impact of the Covid-19 pandemic is minimised. However, then the care systems need to catch up on the non-critical care that was downscaled or even paused due to the pandemic. Therefore, it might become problematic to measure the final PoC (M33) as planned. To minimize this impact, we will be in continuous contact with the responsible doctors at the end-user locations. Also, we will start recruitment early and might use some creative ways to recruit participants, just as in PoC2.

After the first PoC (M13), the results per end-user were evaluated and the lessons learned were identified to improve the design of the second PoC (M25), including the questionnaires and interview guides developed by UCY. Also, the results of the different end-users have been compared and evaluated for further improvements of the next and final PoC (M33). Similarly, after this PoC, lessons learned will be gathered to further improve the execution of PoC3 (M33).

To include the different stakeholders, two months after each PoC the initial results are shared via a conference call (M26). In addition, a newsletter (M26) will be available for all those interested.



POC period

When

- 30 Nov. till 17 Dec.

Who, What, Where

- Interviews with patients, number depending on end user location (1.5 hour)
- Approximately 5 interviews with medical personnel (1 hour)
- 1 focus group with IT personnel (1 hour)
- All sessions were held digitally at all end user locations
- A script was made for interviews and focus group to make sure all KPIs are measured
- See below, an example of a daily planning during the PoC.

Example day of POC planning:

	8	9	10	11	12	13	14	15	16	17	18
Day planning											

Figure 3. Example of PoC Schedule. In this image an example of the PoC (M24) planning is provided

4 Evaluation of the Second Proof of Concept

Through the development of the second PoC, we gathered results using interviews and questionnaires that jointly were used to measure the metrics defined for each KPI mentioned in the Deliverable 7.4. These KPIs allowed the end-users to report the values of the evaluation of the Second Version of the Serums technologies.

In the case of the Baseline measurements, due to the inexistence of equivalent technologies in the Organizations of the use case partners, values provided by the literature in the corresponding field has been reported. Moreover, the results of this second PoC let us track progress and compare them with the first PoC.

During the evaluation of the second PoC, we thought that it would be insightful to evaluate more KPIs compared to the first PoC. Specifically we evaluated one more KPI, namely KPI 2.6 measuring the Efficiency of cross-country patient data sharing. Furthermore, we have been able to evaluate different metrics that we couldn’t evaluate in the first PoC, such as the Granular access to patient records, among others.

As defined in the Deliverable 7.4, three types of Success Indicators, that pertain to the next impacts, will be reported in the following pages

- Quantifiable improvement in the secure provision of health and care services
- Significantly reduced risk of data privacy breaches
- Increased patient trust and safety

4.1 Remarks on the evaluation method

During the elaboration of the tasks that are reported in this Deliverable, some issues arose on the matter of how the results of the Metrics and KPIs had to be generated and merged to obtain the Success Indicators.

These issues were primarily two, the first being that in any previous Deliverable it was described the way in which the metrics and the KPIs with different units had to be merged and the second being that it was considered that all the KPIs had the same importance. The solution to these two problems is explained in the following sections (see sections 4.1.1 and 4.1.2).

4.1.1 The AMPI method

As can be seen in the following pages and in the Deliverable 7.1 the KPI consist of metrics, each of these having different units and ranges. This resulted in the problem of having to merge these numbers, sometimes being as diverse as 20 bits and 1.38E-23, into one single number (the KPI). For obvious reasons this was impossible to do by a simple arithmetic addition. The chosen method to achieve the calculations of the KPI has been the AMPI Index [1] (De Muro et al., 2011).

$$r_{ij} = \left[\frac{\left(y_{ij} - min\, y_i \right)}{\left(max\, y_i - min\, y_i \right)} \right]$$

Figure 4. AMPI index formula

As illustrated in Figure 4, to use this formula two limits have to be chosen, and in a very thoughtful way, since these will set the maximum and minimum range of the improvement. These limits will not have to change and will be used also for the Final Evaluation of the Serums project.

Each of the intervals chosen for the measurements and KPIs (since some measurements are KPI by themselves) can be found on the following chart:

KPI/Measurement	**Minimum value**	**Maximum value**
KPI 1.1: Guessability / Theoretical Entropy	0 bits	105.4 bits
KPI 1.1: Guessability / Practical Entropy	0 bits	105.4 bits
KPI 1.1: Guessability / Guess Number	1	5.35256E+31
KPI 1.1: Guessability / Textual password complexity	0%	100%
KPI 1.1: Guessability / Graphical password complexity	0%	100%
KPI 1.1: Guessability / Push notification accuracy	0%	100%
KPI 1.2: Password leaks (through social engineering) / Memory time	0	168
KPI 1.2: Password leaks (through social engineering) / Human guessing attack	0%	100%
KPI 1.3: System vulnerability	24	240
KPI 2.1: Password cracking resistance	0%	100%
KPI 2.2: Data Breaches	10	110
KPI 2.3: Enhanced model privacy	0	100
KPI 2.4: Granular access to patient record	1	4
KPI 2.5: Authorization data integrity	1	4
KPI 2.6: Efficiency of cross-country patient data sharing	1	4
KPI 3.1: Perceived usability	1	5
KPI 3.2: Perceived memorability	1	5
KPI 3.3: Perceived security	1	5
KPI 3.4: Trust in the proposed PUA scheme	1	5
KPI 3.5: Data Analytics Model Utility	1	5
KPI 3.6: Patient trust	0%	100%
KPI 3.7: Perceived Usability of SERUMS System	1	5
KPI 3.8: Perceived Data Ownership in the SERUMS System	1	5
KPI 3.9: Perceived Security in the SERUMS System	1	5

KPI 4.1: Data Analytics Model Utility	1	100

Table 5. Intervals used to calculate the AMPI indicator

4.1.2 KPI and metric weights

During the gathering and calculation of the metrics and KPIs two issues arose. The first one was the finding that there were differences in the importance of the different measurements (KPIs and metrics) and that reporting them with the same importance into the Success Indicators would be a great mistake. Secondly, after reviewing some of the metrics it was found out that there could be some of them that weren't necessary or were not the object of the study. Only one of the metrics is considered to share these circumstances, and it is the Theoretical Entropy on KPI 1.1.

To solve both of these problems, the technical partners with technologies evaluated by the KPIs directly agreed on a list of coefficients that would be used to weigh all their metrics/KPIs. These will grade the importance of the measurements in the evaluation of the Serums technologies.

These weights can be in the rage from 0 to 3:

0: This metric should not affect the Success Indicator.

1: This measurement does not affect the Success Indicator in a very noticeable manner.

2: This measurement does affect the Success Indicator in an important way.

3: This measurement does have great importance in the Success Indicator.

Success Indicator	**Coefficient**	**KPI**	**Coefficient**	**Metric**
SI 1	2	KPI 1.1: Guessability	0	Theoretical Entropy
			2	Practical Entropy
			2	Guess Number
			3	Textual password complexity
			3	Graphical password complexity
			2	Push notification accuracy
	1	KPI 1.2: Password leaks (through social engineering)	2	Memory time \ Memory time
			1	Human Guessing attack
	2	KPI 1.3: System vulnerability		
SI 2	2	KPI 2.1: Password cracking resistance		
	3	KPI 2.2: Data Breaches		
	1	KPI 2.3: Enhanced model privacy		

	2	KPI 2.4: Granular access to patient record	
	1	KPI 2.5: Authorization data integrity	
	1	KPI 2.6: Efficiency of cross-countr data sharing	
SI 3	2	KPI 3.1: Perceived usability	
	1	KPI 3.2: Perceived memorability	
	2	KPI 3.3: Perceived security	
	3	KPI 3.4: Trust in the proposed PUA scheme	
	1	KPI 3.5: Data Analytics Model Utility	
	2	KPI 3.6: Patient trust	
	2	KPI 3.7: Perceived Usability of SERUMS System	
	3	KPI 3.8: Perceived Data Ownership in the SERUMS System	
	3	KPI 3.9: Perceived Security in the SERUMS System	
SI 4	1	KPI 4.1: Data Analytics Model Utility	

Table 6. Success indicator and KPI definition

After reporting these weights some justification has to be made on the numbers chosen. These have been chosen after a discussion about the importance of each of the KPIs and metrics for the goals of the project. In some cases the decision was made depending on the ability of the partners to obtain a trusted value for the metric, like in the case of the metrics Human Guessing Attack and Memory Time for the KPI 1.2 Password Leaks (through social engineering) in which the experiments required where impossible to be conducted, in the first case because of the great difficulty of doing physical experiments on patients.

On the other hand, as can be seen on KPI 1.1 there is a metric that won't affect the SI since its weight is 0. This metric is Theoretical Entropy. The idea behind giving a weight of zero to a metric is that it has been considered that in this case the results obtained should not be included in the computation of the KPI or SI. This is mostly due to the fact that is not relevant to the project goals or that it may be misleading. In the case of the Theoretical Entropy the reader may find that the doubling of the metric if we consider the passphrase a great increase and a great achievement for the project, but although this can be noted as an improvement, the real indicator to consider would be the increase in the Practical Entropy, that measures the real possibilities that the users consider as passwords. As said, the Theoretical Entropy may not be of great importance, but may be interesting to keep the metric.

It is important to mention that in the case of the metrics Practical Entropy and Guess Number the metrics can not be measured during the project due to security and privacy issues with patient data and the values will be taken from literature.

4.2 Evaluation: Use Case perspective

4.2.1 Evaluation: ZMC

Due to the pandemic of the Coronavirus, the inclusion and interview sessions could not take place physically but needed to take place virtually. This made recruitment planning and execution hard. We therefore needed to extend the period of PoC2 from 2 days to three weeks. Even with the extension and the help of the patient service center from Zuyderland and recruitment by the caregiver, we were only able to recruit half of the planned numbers of participants.

In the first week of studies, some minor cosmetic issues were spotted and fixed. Also, some deployment and cross-browser compatibility issues arose with regards to the user authentication system, in which a few of the first testing patients could not successfully complete the authentication task. All issues were quickly spotted and fixed. Despite these difficulties, patients had a positive outlook on all the Serums technologies and their potential use.

The focus groups with caregivers and IT staff were a success. Due to the comprehensive explanation beforehand and the presence of several technical partners, a lot of additional insight that otherwise could have been missed were gained. These insights, mostly regarding accessibility, ethics, and the law will be considered for improvements to the third and final PoC.

At ZMC, improving the structure in which the data is shown and the scaling issues that occurred with the picture password creation are seen as the biggest improvement towards the final PoC

4.2.2 Evaluation: FCRB

The main obstacle to the correct execution of the second PoC has been the Covid-19 local situation. The results were the following:

- The recruitment of patients was done remotely by the Diabetes Medical Specialist via e-mail or telephone call.
- The participation was lower than the recruitment because of the complexity of the Tele-Interview: finding time-slots for patients was complex, and patients had difficulty utilizing Tele-Interview
- Further remote contact, via email, to measure memorability was complex and did not give the expected results.
- Because the caregivers were busy due to Covid-19, we were unable to find a time-slot that worked for all caregivers. Because of this, only one smaller interview could be done.
- The same case for the IT focus group
- The process of obtaining ethical consent was delayed because several internal meetings of the Ethical Committee were canceled or delayed due to the Covid-19 situation.

In general terms, the solution has been very well accepted. The users also still trusted the system, despite PoC2 being executed completely remotely. It was expected that results would become lower because the meetings weren't physical anymore. As an improvement for the next PoC, many patients recommend that the interview questions be better worded, especially because the PoC was performed remotely and giving clarification was difficult.

4.2.3 Evaluation: USTAN

Due to the ongoing pandemic, recruitment and interview sessions with participants took place virtually with volunteers from the area local to the Edinburgh Cancer Centre (ECC). Since ECC is not a direct partner of the project, we were not able to hold meetings within the hospital, with medical personnel or IT staff, or with their patients locally. Therefore, the USTAN team recruited 26 volunteers via social media in local Facebook groups across Fife, Tayside and Edinburgh to ensure only local patients (23) and potential healthcare personnel (3) were interviewed in real time by video link on Teams. This allowed a high degree of confidence in both the participant's legitimacy and locality.

Participants were scheduled across days, starting on Monday 7th December 2020, with at least two members of the USTAN team (including the system developer available online for technical support if necessary) available at each interview. Participants were asked to willingly share their screen with the interviewers during the FlexPass password creation and SHCS testing part of the process, which allowed support to be offered if required and allowed early identification of any technical issues.

Overall, participants had a positive outlook on the Serums system and their potential uses, despite some remarks on the system layout and few unsuccessful attempts in the authentication module at the beginning of PoC2 execution. They envisaged the system being potentially useful in many aspects of their lives, especially concerning health evaluation by professionals in different healthcare facilities and professions. They also found the questionnaire quite long, occasionally confusing, and few times repetitive, in specific sections; some participants also needed assistance with some questions wording. In general terms, participants overwhelmingly trusted the system despite PoC2 being executed completely remotely and only a few participants were not familiar with technologies such as 'Teams meetings', screen sharing feature and video/audio options.

PoC2 also provided information and lessons for the future version of SHCS and to reach adequate interoperability among technologies. Integration tasks were updated to add one important feature to the client application: the error handling and coding refinement that has been overlooked before PoC2. During PoC2, the system only provided errors' occurrence documented in the system log. In future, the client application will also inform the user if a given error occurs. A positive outcome of performing the PoC2 before reporting D6.2 is that all SHCS modules and APIs have important feedback from a users' perspective, which is the major goal of SERUMS project.

4.3 Impact I, Success Indicator 1

The Success Indicator that will be used for measuring SERUMS progress and specific impact in terms of "Secure provision of health and care services", is:

- **S1)** Quantifiable improvement in secure provision of health and care services (by at least a factor of 2), evidenced by reduced vulnerability of the Smart Health Centre to common cyber-attacks, as measured by standard indexes determining system resilience, robustness and availability during and after the attacks.

Below, the various SERUMS tools/technologies and techniques contributing to S1, clear definitions of the Key Performance Indicators (KPIs) along with their corresponding metrics, as well as the Baseline and the Trial measurements that will be used for measuring S1, are provided.

S1) Quantifiable improvement in secure provision of health and care services (by at least a factor of 2), evidenced by reduced vulnerability of the Smart Health Centre to common cyber-attacks, as measured by standard indexes determining system resilience, robustness and availability during and after the attacks.

SERUMS' Technologies Contributing in Achieving the Success Indicator:

- **Personalized User Authentication (PUA)**
- **Smart Patient Record (SPR)**
- **Verification Technologies (VOT)**

KPI 1.1: Guessability	**PUA**

__Theoretical entropy:__

Entropy is a measure on how difficult it is to guess a password. Entropy is measured as the expected value (in bits) of the information contained in a string, and can be related to authentication key strength by providing a lower bound on the expected number of guesses to find a text. The primary difference between key space and entropy is that key space is an absolute measure of maximum combinations, whereas entropy is related to how users select from the key space. The password key space (k_p) can be related directly to the maximum entropy as follows:

$$H_{max} = log_2 k_p \text{ [bits]}$$

The minimum and maximum value that could be achieved are 0 and 105.4 bits respectively.

Baseline Measurement:

The baseline is determined by the current authentication system for all end user systems. For measuring theoretical entropy, we followed state-of-the-art predictions reported in [2, 3]. Accordingly, considering that at all the end users, the password consisted of a minimum of 8 characters, this resulted in an entropy value of 52.7 at all end users. The difference between FCRB and both ZMC and USTAN is that the former does not have restrictions/rules bound to them, while the latter two do.

	USTAN	***ZMC***	***FCRB***	

Baseline	***52.7***	***52.7***	***52.7***	

Trial Measurement:

Trial measurement (PoC) 1: Contrary to the baseline, the trial results are based on the authentication rules created for the Serums user authentication system developed by UCY. Because the system consisted of two types of authentication, a picture password (3 gestures) and a passphrase (minimum 16 characters), each login credentials has its own theoretical entropy. According to [2-5], these are 53.7 bits and 105.4 bits for the picture password and passphrase respectively.

Trial measurement (PoC) 2: Given that this is a theoretical analysis of the same user authentication policies, the values remain the same for the second evaluation study.

	Picture	***Passphrase***
1st Trial	***53.7***	***105.4***
2nd Trial	***53.7***	***105.4***

Commentary on results:

Results reveal that the theoretical entropy of the PoC authentication system for both picture password and textual password are significantly higher than the entropy of the textual password systems of the baseline for all three end-user organizations.

Practical entropy:

A true measure of theoretical entropy cannot be computed in cases of user-chosen authentication keys since users tend to choose more memorable than random keys. For measuring practical entropy, we have considered the work and results described in [2-5] which provide estimates of practical entropy of different password policies.

The minimum and maximum value that could be achieved are 0 and 105.4 bits respectively.

Baseline Measurement:

The baseline is determined by the current authentication system for all end user systems. At all the end users, the password consisted of a minimum of 8 characters. The difference between FCRB and both ZMC and USTAN is that the former does not have restrictions/rules bound to them, while the latter two do. Because of this, the practical entropy for FCRB is 29.43 bits, while for ZMC and USTAN it is 34.3 bits.

	USTAN	***ZMC***	***FCRB***
Baseline	***34.3***	***34.3***	***29.43***

Trial Measurement:

Trial measurement (PoC) 1: Contrary to the baseline, the trial results are based on the authentication rules created for the Serums user authentication system developed by UCY. Because the system consisted of two types of authentication, a picture password (3 gestures) and a passphrase (minimum 16 characters), each login credentials has its own practical entropy, which are 35 bits and 44.67 bits respectively.

Trial measurement (PoC) 2: Given that we applied the same user authentication policy as in the first evaluation study, the values remain the same for the second evaluation study.

	Picture	*Passphrase*
1st Trial	***35***	***44.67***
2nd Trial	***35***	***44.67***

Commentary on results:

Results indicate that practical entropy is lower than theoretical entropy for all user authentication systems (both baseline and PoC). Furthermore, the practical entropy of the PoC authentication system for the textual password type is significantly higher than the entropy of the textual password systems of the baseline for all three end-user organizations. The picture password of the PoC authentication system has similar levels of practical entropy as the textual password systems of the USTAN and ZMC case study, while significantly larger than the textual password systems of the FCRB case study.

Guess number:

Guess number refers to how many guesses a particular password-cracking algorithm with particular training data would take to guess a password.

The actual number of guesses is typically calculated by applying a certain brute-force attack on the actual user passwords in a database. However, due to security restrictions at each end-user organization, we could not get access to an actual database which includes the hashed user passwords, and hence we could not run an actual brute-force attack on the user passwords. Due to this, we are reporting the predicted guess numbers by following existing state-of-the-art studies and reports in [2-5] for similar policies.

The minimum and maximum number of guesses are 1 and 5.235256E+31 respectively.

Baseline Measurement:

The baseline is determined by the current authentication system for all end user systems. At all the end users, the password consisted of a minimum of 8 characters. For measuring the guess number, we have considered the work and results described in [2-5] which provide estimates for guess number of different password policies. The difference between FCRB and both ZMC and USTAN is that the former does not have restrictions/rules bound to them, while the latter two do. Because of this, the guess number for FCRB is 723,290,519, while for ZMC and USTAN it is 21,150,899,968.

	USTAN	*ZMC*	*FCRB*
Baseline	***21,150,899,968***	***21,150,899,968***	***723,290,519***

Trial Measurement:

Trial measurement (PoC) 1: Contrary to the baseline, the trial results are based on the authentication rules created for the Serums user authentication system developed by UCY. Because the system consisted of two types of authentication, a picture password (3 gestures) and a passphrase (minimum 16 characters), each login credentials has its own guess number. According to [2-5], these are 34,359,738,368 and 2.79905E+13 guesses for the picture password and passphrase respectively.

Trial measurement (PoC) 2: Given that we applied the same user authentication policy as in the first evaluation study, the values remain the same for the second evaluation study.

	Graphical	*Passphrase*
1st Trial	***34,359,738,368***	***2.79905E+13***
2nd Trial	***34,359,738,368***	***2.79905E+13***

Commentary on results:

Results indicate that the guess number of the PoC authentication system for the textual and picture password types is significantly higher than the guess number of the textual password systems of the baseline for all three end-user organizations.

Additional/Updated measures used in PoC2 Evaluation:

Textual password complexity:

For measuring textual password complexity, we have implemented Dropbox's *zxcvbn*, which is a widely applied and realistic password strength estimator. The minimum and maximum value of textual password complexity is 0% and 100% respectively. The higher the score, the more complex the password is. Textual password complexity is not applicable for the baseline study since this information was not available by the organizations.

	USTAN	***ZMC***	***FCRB***	***Overall***
2nd Trial	***69.47%***	***80%***	***80%***	***76.49%***

Graphical password complexity:

An additional measure for graphical passwords is graphical password complexity, which describes how complex a graphical password is based on the users' image selections and gestures. For measuring graphical password complexity, we used a heuristic approach by considering state-of-the-art knowledge on picture gesture authentication [4, 5, 6]. Taking into consideration that the tap (click) is the least complex gesture, while the line is the more complex gesture [4], we set our initial complexity heuristic as follows:

Combination of gestures -> Complexity

3 taps -> 40%

3 circles -> 80%

3 lines -> 100%

# Disregarding order

1 tap & 2 circles -> 50%

2 taps & 1 circle -> 50%

1 tap & 1 line % 1 circle -> 70%

2 taps & 1 line -> 70%

1 tap & 2 lines -> 70%

2 circles & 1 line -> 70%

2 lines & 1 circle -> 80%

Taking into consideration that the proximity of different gestures impacts the overall complexity of the password (*e.g.*, different gestures on the same *x*, *y* segment on the grid are less secure), we take an extra step to either penalise (-20%) password combinations that include gestures that are in close proximity (as defined by the threshold of a circle of 3 segments radius [4]), or reward (+20%) password combinations that do not include gestures in close proximity. The minimum and maximum value of textual password complexity is 0% and 100% respectively. The higher the score, the more complex the password is.

Graphical password complexity is not applicable for the baseline study since all three end-user organizations do not implement a graphical password system.

	USTAN	***ZMC***	***FCRB***	***Overall***
2nd Trial	***76.19***	***73.07***	***77***	***75.45***

Push notification accuracy:

In the second evaluation study, we deployed and evaluated the two-factor authentication system (2FA) that was implemented as a mobile application (Serums Authenticator) and optionally installed by the participants on their smartphones. The effectiveness of the 2FA system is measured through push notification accuracy, which measures the accuracy of the users' approval of push notifications.

The table below summarizes the number of users that enabled the 2FA option and installed the Serums Authenticator app and whether they successfully authenticated or not using the mobile application. Although a limited number of users downloaded the mobile application, push notification accuracy scored a 100% success rate.

	USTAN		***ZMC***		***FCRB***	
	# of users	***Success rate***	***# of users***	***Success rate***	***# of users***	***Success rate***
2nd Trial	***2***	***100%***	***6***	***100%***	***6***	***100%***

KPI 1.2: Password Leaks (through Social Engineering)	**PUA**

Memory time:

Memory time will be measured over time by considering actual login attempts of the end-users. In particular, memory time refers to the greatest length of time between a password creation and a successful password login using the same password. Large memory times indicate higher memorability. Memorable passwords lead to potentially less social engineering-based password leaks because users will not need to follow coping strategies (*e.g.*, write down their passwords).

Memory time data could not be measured for the baseline study since the relevant data was not supported by the existing authentication systems at the end-user organizations (or not available due to privacy regulations and policies of the corresponding organization). In addition, given that memory time requires participants using the system over time, we did not measure this in PoC1 since the aim of the first evaluation of the user authentication system was to elicit the users' perceptions and likeability towards the first PoC authentication system.

Another metric for memorability relates to the number of password resets as well as time needed to login. For the baseline authentication system, we received summarized password reset data from ZMC. The table below summarizes the amount of resets and the average amount of days between the resets at ZMC starting from January 01, 2019 until October 31, 2019.

Baseline Measurement:

Number of resets at ZMC	**Average amount of days between resets**	**Total number of occurrences**
1	0	1893
2	91	738
3	69	222
4	64	92
5	42	20
6	47	6

Trial Measurement:

Trial measurement (PoC) 1: During the PoC study, participants interacted with the user authentication prototype by creating a textual and picture password and then using their password to login. Within this session, we measured the number of resets required by the end-users. The number of resets for each user authentication type are summarized in the table below.

	ZMC		**FCRB**	
	Passphrase	**Picture**	**Passphrase**	**Picture**
# resets	0/15	1/16	0/4*	1/18
Login time (sec)	15.41	6.19	n/a	6.14

**note: the number of occurrences for each password type varies since users chose and logged in with their preferred authentication type (textual vs. graphical). In the case of the FCRB study, 4 users chose a passphrase to login, while 18 users chose a picture password.*

Trial measurement (PoC) 2: Memory time (seconds) is the greatest length of time between a password creation and a successful password login using the same password. In order to measure authentication memory time (over time) after the user password creation phase, following an accredited method reported in Stobert and Biddle (2013) [7], we have sent three notification emails on Day 1, Day 3, and Day 6 to the participants. Each email directed the participants to the Serums study Website and it instructed them to access the Serums system.

Memory time (in hours) - maximum 168 hours (7 days * 24 hours)

	USTAN	***ZMC***	***FCRB***	***Overall***
2nd Trial	***149.27***	***119.73***	***124.04***	***135.39***

Login time for passphrase (sec)

	USTAN	***ZMC***	***FCRB***	***Overall***
2nd Trial	***14.55***	***13.46***	***N/A***	***13.98***

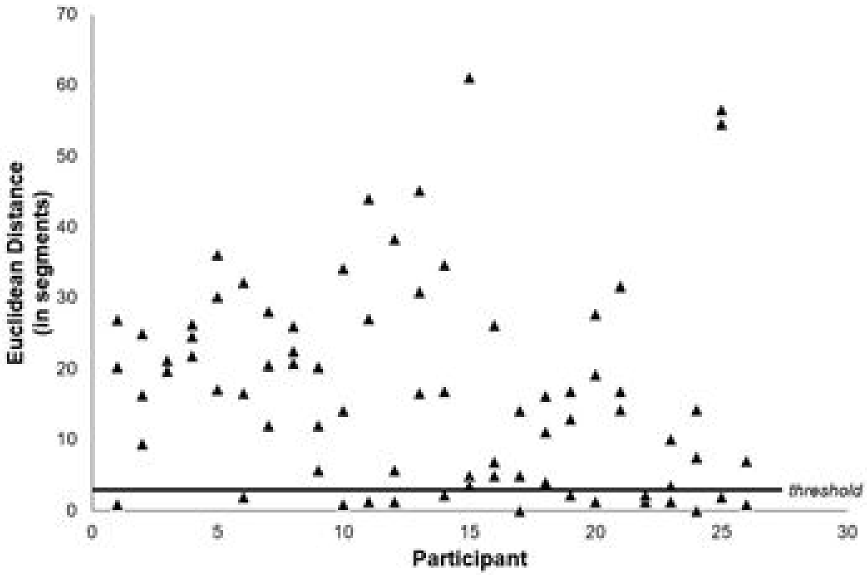
Login time for picture password (sec)

	USTAN	***ZMC***	***FCRB***	***Overall***
2nd Trial	***7.04***	***12.16***	***5.24***	***7.62***

Human Guessing Attack

Bearing in mind that when using the suggested personalized and retrospective approach, graphical password selections are based on the users' existing sociocultural experiences, it is probable that the individuals who share common experiences with the end-users might be able to guess their selections. In order to shed light on this aspect, we have conducted a human attack study focusing on guessing vulnerabilities of the approach among people sharing common sociocultural experiences. Each session of the study embraced pairs of participants that were closely related (e.g., friends, couples, relatives, etc.) and who shared common experiences. In each session, we asked both participants to first create a graphical password, and then each participant was asked to guess the password selections of the other participant. A total of 26 individuals (12 females) participated in the study, ranging in age between 25-60 years old (m=40.03, sd=10.23). *Note: A thorough analysis of the method and results reported below can be found in Constantinides et al. (2021) [8].*

To investigate how far the attackers' guessing selections were from the end-users' actual secret selections, we calculated the Euclidean distance between the 3 x, y segments provided by the attacker and the 3 x, y segments of the end-user. The figure below depicts the Euclidean distance of each gesture of each participant by disregarding the type and the exact order of the attackers' gestures and the end-user's gestures. Accordingly, among 78 gestures (3 gestures x 26 participants), 16 gestures (20%) were in close proximity with the attacker's guessed selections.



Euclidean distance between attackers' gestures and end-users' gestures, by disregarding the type and the exact order [8].

Furthermore, we compared the attempts of each attacker with the end-user's stored password from the same pair of participants, simulating in principle an online guessing attack. From a total of 78 attacking guesses (3 attempts of each attacker x 26 participants), there was only 1 successful attempt, yielding an online success guessing rate of 0.01%. It is worth noting that the successful online attack contained 3 gestures on 3 hot-spots areas.

KPI 1.3: System Vulnerability	**SPR**

<u>**Metrics**</u>**:**

The measure of how susceptible the system is via penetration testing as well as the security of the authentication methods. The types of penetration that we will use will be both external network and internal network penetration testing. This will allow us to see how vulnerable the system is from the outside as well as once they have gained some form of access. Additionally, we will score the security of the programming languages used, as well as the lifespan of security support that is left for these.

<u>**Baseline Measurement:**</u>

This took form as a questionnaire that was given to the use case partners. Due to the sensitive nature of the questions, it was only possible to receive a complete and usable set of responses from FCRB.

Each result of the questionnaire was scored on a scale of 1 - 10, with 1 being critical and 10 being no known issues. These scores were calculated through known knowledge of vulnerabilities as well as length of time left of support for the various programming languages and frameworks. This gave a minimum score of 24 and a maximum score of 240, with FCRB scoring 109.

	USTAN	***ZMC***	***FCRB***	***TOTAL***
Baseline	***N/A***	***N/A***	***109***	***109***

Trial Measurement:

Trial measurement (PoC) 1: At the time that the first PoC took place, the development of the system was very early in development so we were unable to record measurements for the system.

Trial measurement (PoC) 2: With the refined versions of the software for work package 2 integrated into a single system, we were able to take measurements. The same criterias used for the baseline were measured against the integrated Serums system. As above, each result was scored on a scale of 1 -10 with 1 being critical and 10 being no known issues. The results for the current Serums systems were a score of 131.

	TOTAL
1st Trial	*N/A*
2nd Trial	*131*

Commentary on results:

As expected, the security of the existing systems is high. However, by utilizing more recent versions of programming languages and by choosing an operating system with many years of long term support still ahead of it, we minimise the risk of vulnerabilities. In addition, by using best practice standards for passwords and encryption, we have already seen a noticeable increase in the achieved score for the Serums system. For the third PoC we are aiming to improve our logging and activity monitoring which we believe will see another measured increase in the system security.

4.4 Impact II, Success Indicator 2

The Success Indicator that will be used for measuring SERUMS progress and specific impact in terms of "Less risk of data privacy breaches caused by cyber-attacks", is:

- **S2)** Significantly reduced risk of data privacy breaches (at least 75%), evidenced by quantitative metrics showing the quantity of private data that is revealed through a number of common cyber-attacks.

Below, the various SERUMS tools/technologies and techniques contributing to S2, clear definitions of the Key Performance Indicators (KPIs) along with their corresponding metrics, as well as the Baseline and the Trial measurements that will be used for measuring S2, are provided.

S2) Significantly reduced risk of data privacy breaches (at least 75%), evidenced by quantitative metrics showing the quantity of private data that is revealed through a number of common cyber-attacks.

SERUMS' Technologies Contributing in Achieving the Success Indicator:

- **Credential Hardening (CH)**
- **Smart Patient Record (SPR)**
- **Privacy-preserving Data Analytics (PDA)**
- **Verification Technologies (VOT)**
- **Distributed Ledger Technology (DLT)**

KPI 2.1: Password Cracking Resistance	**CH**

Metrics:

Password cracking rate will be measured in a leaked database storing hardened credentials through an offline brute-force attack. Normally, the credential hardening offered by the system does not allow an attacker to crack a password database. Assuming that the private key of the system is stolen, then the cracking rate is depicted below. Numbers are presented with comparison to other popular hashing schemes. Cracking rate is the power you have to crack a password, which is proportional to how many hash computations you can do per second. Hence, we report cracking rate, which is measured to hash computations/sec.

Trial Measurement:

Trial measurement (PoC) 2:

Hashing Scheme	***Hash (ms)***	***Rate (h/sec)***
WordPress (8,192 iterations of MD5)	***2.22***	***450***
bcrypt (cost 12)	***249.60***	***4***
bcrypt (cost 11)	***124.68***	***8***

bcrypt (cost 10 - default)	***62.42***	***16***
bcrypt (cost 9)	***31.29***	***32***
bcrypt (cost 8)	***15.72***	***64***
Drupal (65,537 of SHA1)	***65.16***	***15***
SERUMS Credential Hardening	***50.23***	***20***

Hash (ms) is the time needed to calculate the hash based on each hashing scheme.

Rate (h/sec) is the number of hashed passwords that are computed per second.

KPI 2.2: Data Breaches	**SPR**

Metrics:

The measure of data that will be able to be accessed by unauthorised or inappropriate sources. Through the use of the log files for the database we will take measurements on how much data can be accessed by both an unknown user and a known user for unauthorised reasons. Additionally, we can apply a score against the ease at which physical copies of the data can be generated.

Baseline Measurement:

This took form as a questionnaire that was given to the use case partners. Due to the sensitive nature of the questions, it was only possible to receive a complete and usable set of responses from FCRB.

Each result of the questionnaire was scored on a scale of 1 - 10, with 1 being critical and 10 being no known issues. The questions covered the access that staff have to patients' records as well as the options available to create physical copies of the data. This gave a minimum score of 11 and a maximum score of 110, with FCRB scoring 49.

	USTAN	***ZMC***	***FCRB***	***TOTAL***
Baseline	***N/A***	***N/A***	***49***	***49***

Trial Measurement:

Trial measurement (PoC) 1: At the time that the first PoC took place, the development of the system was very early in development so we were unable to record measurements for the system.

Trial measurement (PoC) 2: With the refined versions of the software for work package 2 integrated into a single system, we were able to take measurements. As above, each result was scored on a scale of 1 -10 with 1 being critical and 10 being no known issues. The results for the current Serums systems were a score of 47.

It is worth noting, however, that one of the criteria was not applicable to the Serums system (ability to copy to removable media). As such, the potential maximum score that the Serums system could achieve is 100 as opposed to the existing systems' 110. This difference in potential score is mitigated by the AMPI results in the section below.

	USTAN	*ZMC*	*FCRB*	*TOTAL*
1st Trial	*N/A*	*N/A*	*N/A*	*N/A*
2nd Trial	*N/A*	*N/A*	*47*	*47*

Commentary on results:

The baseline results were to be expected. The hospital must balance the ability to see relevant patient data in an emergency with the potential for data breaches. There are sensible policies in place, however, the Serums system would remove the need for physical copies of data to be made which will result in a higher score.

During the second trial, we were able to measure the Serums system and saw a very small improvement over the existing system when the results were applied via the AMPI method. While the improvement was negligible, we are happy that we have at least matched the existing system's score. By improving the logging and monitoring of the Serums system, we are confident that we will see this score increase and the difference between the existing technology and the Serums technology should become more apparent.

	USTAN	*ZMC*	*FCRB*	*TOTAL*
Baseline AMPI	*N/A*	*N/A*	*0.38*	*0.38*
1st Trial AMPI	*N/A*	*N/A*	*N/A*	*N/A*
2nd Trial AMPI	*N/A*	*N/A*	*0.41*	*0.41*

KPI 2.3: Enhanced Model Privacy	PDA

Metrics:

Model Privacy measures the ability of a model to preserve the privacy of the data used to train the model when releasing the model's output. Since there is always a tradeoff between a models privacy and a models utility, we need to define the level of accuracy, in this case the percentage of correct predictions, at which we want to measure the privacy level in order to be able to compare different approaches of privacy preservations.

The measure the level of privacy we use the well-established mathematical framework of (e; δ)-differential privacy. Enhanced model privacy is the factor of increase in differential privacy when comparing two models at the same level of utility.

Since the actual level of privacy that a model can achieve is on the one hand dependent on the level of accuracy that needs to be achieved and on the other hand on the training dataset itself, no general statement can be made about the factor of model privacy enhancement of an privacy preserving model.

Baseline Measurement:

The baseline that we compare our developed model to is a state-of-the-art model that uses the classical Gaussian mechanism to achieve differential privacy at a prediction accuracy of 95%.

Trial Measurement:

In the trial measurement we calculate the level of privacy of our privacy preserving mechanism that uses an optimal-noise adding mechanism at a prediction accuracy of 95%. Enhanced model privacy is the factor of increased differential privacy of the baseline compared to the trial.

At the time of writing we have no access to appropriate use case data from the USTAN use case to be able to measure this KPI. But in order to initially evaluate the enhancement of our newly developed approach above the state-of-the-art, we calculated the factor of increase for we use the MNIST data as a benchmark dataset.

For this benchmark dataset the classical Gaussian mechanism achieved an (e; δ)-differential privacy of (2; 1-e5). Our proposed mechanism on contrary resulted in a differential privacy of (1.14; 1-e5), which is an increase of privacy by a factor of 1.7544.

KPI 2.4: Granular access to patient record	**DLT**

Metrics:

This KPI will measure how granular the solution will offer the ability to manage the access to the patient record.

We have defined 4 levels of permission granularity of patient record access with a scale of 1-4 where level 4 is the most satisfactory level. The DLT solution aims to reach level 4.

1. No digital access management of the patient record
2. Access can be managed by the organization (e.g. hospital) at patient record level. which means the record can be accessed or not as a whole for the caregiver.
3. Access can be managed by the organization (e.g. hospital) at a granular level (e.g. a subset of the patient record)
4. Access can be managed by the organization (e.g. hospital) and the patients themselves at a granular level (e.g. a subset of the patient record)

Baseline Measurement:

The baseline measurement was collected based on the assessment of the current systems in place in the hospitals.

	USTAN	***ZMC***	***FCRB***	***TOTAL***
Baseline	***1***	***1***	***1***	***1***

Trial Measurement:

Trial measurement (PoC) 1: At the time that the first PoC took place, there were no integrated Serums system available, thus we were unable to record measurements for this KPI.

Trial measurement (PoC) 2: At the current state of the blockchain solution, it is able to manage access to the patient record; this access can be both at patient record level – the complete record – or a subset of data in the patient record, meeting the level 4 of this KPI. When an integrated system including end-user frontend is available, a patient is able to create rules to manage access to his/her patient record by himself. Without a patient end-user frontend, it is also possible for the patient to request the hospital to create the rules on their behalf. Meaning reaching the level 3 of this KPI.

	USTAN	*ZMC*	*FCRB*	*TOTAL*
1st Trial	*N/A*	*N/A*	*N/A*	*N/A*
2nd Trial	***4***	***4***	***4***	***4***

KPI 2.5: Authorisation Data Integrity | **DLT**

Metrics:

This KPI will be measuring how resilience the current system is handling the authorisation data.

In case a party on the DLT network is compromised, and it has been identified that data has been tampered with. The solution is able to identify the exact data that has been tampered with and retrieve the original value. We have defined 4 levels with a scale of 1-4 where level 4 is the most satisfactory level. The DLT solution aims to reach level 4.

1. No means to traceback when (authorization) data has been compromised.
2. Is able to retroactively track when data is compromised but cannot track which specific data was compromised.
3. Retroactive tracking when data is compromised, and is able to identify which data has been compromised but cannot restore the original value.
4. Retroactive tracking when the data is compromised, and is able to identify and restore the data that has been compromised.

Baseline Measurement:

The baseline measurement was collected based on the assessment of the current systems in place in the hospitals.

	USTAN	*ZMC*	*FCRB*	*TOTAL*
Baseline	***1***	***1***	***1***	***1***

Trial Measurement:

Trial measurement (PoC) 1: At the time that the first PoC took place, there were no integrated Serums system available, thus we were unable to record measurements for this KPI.

Trial measurement (PoC) 2: For this KPI, the inherent characteristics of a successful blockchain implementation itself to meet part of this requirement. Blockchain allows one to see the complete history of the ledger. It means that in the event of data being compromised (e.g. via stolen credentials), it is always possible to identify the historical values of the record. With the patient-Id/doctor-Id, combined with a timestamp of when the party is compromised, you would be able to see all transactions created related to the user. A manual process will need to be initiated to restore the data: all newly created rules could easily be deleted and invalidated, while rules that have been changed can be easily restored by looking at their previous value in the blockchain.

	USTAN	*ZMC*	*FCRB*	*TOTAL*
1st Trial	*N/A*	*N/A*	*N/A*	*N/A*
2nd Trial	*4*	*4*	*4*	*4*

KPI 2.6: Efficiency of cross-country patient data sharing **DLT**

Metrics:

This KPI will be measuring the efficiency of cross-country patient data sharing. The 2 types of data sharing are out-bound and in-bound. out-bound: one has data to be shared to another hospital; in-bound: patient data to be received from another hospital.

A scale of 1-4 for new KPI to measure how cross-country permissions for patient data sharing is managed**.**

1. Cross-country permission is managed manually
2. Only out-bound is managed in an automated way
3. Both in-bound and out-bound are managed in an automated way
4. Patient is in control and can initiate the process in an automated way

Baseline Measurement:

The baseline measurement was collected based on the assessment of the current systems in place in the hospitals.

	USTAN	*ZMC*	*FCRB*	*TOTAL*
Baseline	*1*	*1*	*1*	*1*

Trial Measurement:

Trial measurement (PoC) 1: At the time that the first PoC took place, the development of the system was very early in development so we were unable to record measurements for the system.

Trial measurement (PoC) 2: With the successful implementation of the integrated Serums system including the end-user front end for the patient, the access permission of the patient data can be across partner countries within the Serums network. This can be either initiated by a patient or by the hospital. Although the cross-country patient data use case is not yet demonstrated as part of the patient journey in PoC2, the functionality is already available as part of the integrated solution for such scenarios.

	USTAN	***ZMC***	***FCRB***	***TOTAL***
1st Trial	***N/A***	***N/A***	***N/A***	***N/A***
2nd Trial	***4***	***4***	***4***	***4***

4.5 Impact III, Success Indicators 3 and 4

The Success Indicators that will be used for measuring SERUMS progress and specific impact in terms of "Increased patient trust and safety" are:

- **S3)** Quantifiable improvement in levels of patient trust in the provision of smart health care (at least a factor of 2), evidenced by patient surveys and questionnaires.
- **S4)** Quantifiable improvement in patient safety (at least a factor of 2), evidenced by reduced risk of harm through incorrect treatments or medicines mediated by reduced risk of tampering with medical records, and measured vulnerabilities of connected medical systems.

Below, the various SERUMS tools/technologies and techniques contributing to S3 and S4, clear definitions of the Key Performance Indicators (KPIs) along with their corresponding metrics, as well as the Baseline and the Trial measurements that will be used for measuring S3 and S4, are provided.

S3) Quantifiable improvement in levels of patient trust in the provision of smart health care (at least a factor of 2), evidenced by patient surveys and questionnaires

SERUMS' Technologies Contributing in Achieving the Success Indicator:

- **Personalized User Authentication (PUA)**
- **Smart Patient Record (SPR)**
- **Privacy-preserving Data Analytics (PDA)**

KPI 3.1: Perceived Usability	**PUA**

Metrics:

One of the primary aims of the PoC evaluations of the user authentication system was to get feedback from end-user patients on aspects such as likeability towards the suggested flexible and personalized approach in authentication, and the end-users' perceptions towards usability, memorability, security and trust. For this purpose, we have designed a questionnaire by following state-of-the-art works and guidelines on usability, user experience, security and trust (*e.g.*, SUS, AttrakDiff, Technology Acceptance models, etc.).

With regards to perceived usability, we have asked questions that relate to the password creation process and login, *e.g.*, *"Overall, how difficult or easy do you find the password creation task?"*, *"Overall, how difficult or easy do you find the login task?"*, *"I could easily log on to the FlexPass password system"*, etc. Users rated the statements through a 5-point Likert scale (*e.g.*, 1: Not at all - 5: Absolutely).

Baseline measurement:

	USTAN	*ZMC*	*FCRB*	*TOTAL*
Baseline	*N/A*	*3.8*	*3.71*	*3.75*

ZMC baseline (patients)

With regards to questions that relate to the password creation difficulty, the majority of users find the creation process medium-to-easy (18/19 - Difficult: 1; Medium: 7; Easy: 6; Very easy: 5). With regards to login difficulty, 16 out of 19 find the login task as easy to use. When users were asked to report on the password reset difficulty, responses varied, with the majority stating that the reset process has moderate difficulty (Difficult: 2; Medium: 10; Easy: 3; Very easy: 4). Also, users have reported mixed methods for resetting their password (Email: 7; Mobile app: 5; Reset tool: 7).

ZMC (feedback from professionals)

Similar to the patient responses, the majority of responses received from the ZMC professionals reveal that the baseline user authentication system is usable. In particular, 13/19 professionals perceive the password creation as easy to use, 17/19 professionals login in the system without any difficulties, and 16/19 professionals find the reset process as easy to use.

FCRB baseline (patients)

With regards to questions that relate to the password creation difficulty, the majority of users find the creation process medium-to-easy (17/21 - Difficult: 2; Difficult-Medium: 2; Medium: 3; Easy: 4; Very easy: 10). With regards to login difficulty, 16 out of 21 find the login task as easy to use. When users were asked to report on the password reset difficulty, responses similarly varied as in the ZMC case, with the majority stating that the reset process has moderate difficulty (10/21). Also, users have reported mixed methods for resetting their password (Email: 9; Mobile app: 7).

FCRB (feedback from professionals)

Mixed responses were received from professionals with regards to perceived usability. A total of 8/19 professionals find the password creation task as easy to use, 5 as moderate difficulty, and 6 professionals find the task as difficult to use. Similar findings are observed in the case of login task with 11/19 professionals rating the login task as easy to use, while 4 professionals find the login task as moderate and difficult to use respectively. With regards to the password reset process, 11/19 professionals find the task as easy to use, 7 rated moderate difficulty and 1 user rated the task as difficult.

Trial measurement:

Trial measurement (PoC) 1:

	USTAN	***ZMC***	***FCRB***	***TOTAL***
1st Trial	***N/A***	***4.06***	***3.65***	***3.85***
2nd Trial*	***3.96***	***4.02***	***4.52***	***4.11***

**Beyond the SUS results reported below, we have also included the results based on the users' answers to the same usability questions that were used in the baseline and PoC1 to compare them with the ones of PoC2.*

ZMC PoC1 (patients)

Patients at ZMC found the password creation task as easy to use (password creation - easy: 21/31; moderate: 8/31). Similarly, with regards to password login usability, the majority of users found the login task as easy to use (25/31). When users were asked whether they would like to use the Serums user authentication system as their alternative password system, *the majority of users (25/31) were positive and would be willing to use it as an alternative authentication system*.

When ZMC end-user were asked on whether they like to personalized and flexible approach for user authentication, the majority of users extremely (18/31) or very much (9/31) liked the idea, with 5 users either moderately (2/31), slightly (2/31) or not liking (1/31) the idea.

ZMC PoC1 (feedback from professionals)

All ZMC professionals (4) like the flexible and personalized authentication paradigm, 3 professionals believe that the Serums authentication technology would be a good alternative method for patients, 1 showed moderate interest. Overall, 3 professionals find the authentication system as easy to use while 1 professional rated the system as difficult to use.

FCRB PoC1 (patients)

Patients at FCRB found the password creation task as easy to use (easy: 14/24; moderate: 6/24) and fast to use (fast: 14/24; moderate: 6/24). Nonetheless, 4 patients found the password creation task as both difficult and slow to use. Similarly, in the case of login usability, the majority of users found the login task as easy to login (easy: 21/24; moderate: 5) while 3 users found the login task as difficult to use.

With regards to likeability towards the flexible and personalized approach, the majority of users very much (11/24) liked the idea, with 2 users extremely like the idea, 7 moderately, 3 slightly and 1 user did not like the idea.

FCRB PoC1 (feedback from professionals)

The majority of FCRB professionals (4/5) like the flexible and personalized authentication paradigm, as well as believe that the Serums authentication technology would be a good alternative method for patients, 1 showed moderate interest. Overall, 4 professionals find the authentication system as easy to use while 1 professional rated the password creation task ease of use as moderate. All professionals believe that the creation task is fast to use. A total of 2 professionals believe that patients will easily log in while 3 professionals rated ease of use as moderate.

Likeability	**Extremely**	**Very much**	**Moderately**	**Slightly**	**Not at al**
ZMC	18	9	2	2	1
FCRB	2	11	7	3	1
Total	***20***	***20***	**9**	**5**	**2**

Commentary on results:

Results are encouraging for further research on the idea of flexible and personalized user authentication since the majority of users liked the proposed approach, as well as perceived both the password creation process and login task as easy to use. In comparison to the baseline measurements, the PoC authentication system has improved usability in the ZMC case study (3.8 *vs.* 4.06) whereas in the case of FCRB, the usability value has been slightly decreased from 3.71 to 3.65. Based on qualitative feedback received from the end-users this can be accredited to some users (*n*=4) that had difficulties in creating gestures on the image during the graphical password creation task.

Trial measurement (PoC) 2:

Given that in PoC2 we evaluated the complete functional FlexPass system, we have utilized the System Usability Scale (SUS) [9, 10], which is an accredited and widely applied system usability instrument. The table below summarizes the SUS scores across the end-user organizations.

	USTAN	***ZMC***	***FCRB***	***Overall***
SUS Score	***72.14%***	***74.58%***	***80%***	***74.77%***

Based on the literature, the average SUS score is 68%. In case the score is under 68%, the system entails various usability issues that need improvement, while a score above 68%, indicates that the system entails good usability practices. Accordingly, the scores across the three end-user organizations ranged between 72.14% for USTAN, 74.58% for ZMC and 80% for FCRB, with an overall score of 74.77%.

Such results are encouraging for further investigating and improving the system since the score suggests that the FlexPass system scores very well in usability, end-users like the system
and they can easily complete the authentication-related tasks. Nevertheless, given that the score is below 80%,there are still aspects that require improvements, e.g., during PoC2, some patients had difficulties in entering their graphical password through the developed gesture input mechanism. We will further improve the gesture input functionality during the final development life-cycle of the FlexPass

Furthermore, when end-users were asked on whether they like the personalized and flexible approach for user authentication, the majority of users extremely (26/44) or very much (14/44) liked the idea, with 4 users either moderately (1/44) and slightly (3/44) liking the idea.

Likeability	**Extremely**	**Very much**	**Moderately**	**Slightly**	**Not at all**
USTAN	12	6	1	2	0
ZMC	9	3	0	0	0
FCRB	5	5	0	1	0
Total	**26**	**14**	**1**	**3**	**0**

KPI 3.2: Perceived Memorability	**PUA**

Metrics:

Similar to perceived usability, we have asked participants questions on whether they recalled effectively their passwords and whether the login process was mentally demanding. Users rated the statements through a 5-point Likert scale (*e.g.*, 1: Not at all - 5: Absolutely).

Baseline measurement:

	USTAN	*ZMC*	*FCRB*	*TOTAL*
Baseline	*N/A*	***4.36***	***3.63***	***3.49***

ZMC baseline (patients)

The majority of patients from the ZMC case study responded positively to perceived password memorability; 8 users find the mental demand to recall their password as very low: 8; 10 require a low mental demand, and 1 users requires a moderate mental demand.

ZMC baseline (feedback from professionals)

All professionals responded that they require low mental demand to recall their password (Very low: 12; Low: 6).

FCRB baseline (patients)

Mixed responses on password memorability were received in the FCRB case study with 11 users having very low or low mental demand to recall their password, 5 users a moderate demand and 5 users high or very high demand.

FCRB baseline (feedback from professionals)

A total of 9/19 professionals require low mental demand during password recall, 6 professionals rated mental demand as moderate while 4 users require a high mental demand during password recall.

Trial measurement:

Trial measurement (PoC) 1:

	USTAN	*ZMC*	*FCRB*	*TOTAL*
1st Trial	*N/A*	***4.09***	***4.16***	***4.12***
2nd Trial	***4.09***	***4.20***	***4.45***	***4.21***

ZMC PoC1 (patients)

The majority of patients from the ZMC case study responded positively to perceived password memorability after their interaction with the Serums user authentication system. The majority of users reported a very low or low mental demand (21/31; moderate: 5/31) in recalling their password. In addition, 25/31 users could effectively recall their password.

ZMC PoC1 (feedback from professionals)

All ZMC professionals (4) believe that the login task will require low mental demand from patients, 3 believe that patients will easily remember their password while 1 professional believes that patients will have difficulties to login.

FCRB PoC1 (patients)

Similar to the FCRB case study, the majority of patients responded positively to perceived password memorability after their interaction with the Serums user authentication system. The majority of users reported a very low or low mental demand (19/24) in recalling their password. In addition, 19/24 users could effectively recall their password. Nonetheless, 4 patients reported that they found the authentication system as mentally demanding.

FCRB PoC1 (feedback from professionals)

A total of 3 professionals believe that the login task will require low mental demand from patients (1 rated moderate mental demand), 2 believe that patients will easily remember their password while 3 professionals believe that patients will moderately remember their passwords.

Trial measurement (PoC) 2:

USTAN PoC2: The majority of patients from the ZMC case study responded positively to perceived password memorability; 17 users could effectively remember their password, 2 users could moderately recall their password, while 2 users could not easily recall their password.

ZMC PoC2: The majority of patients from the ZMC case study responded positively to perceived password memorability; 10 users could effectively remember their password, while 2 users could moderately recall their password.

FCRB PoC2: The majority of patients from the ZMC case study responded positively to perceived password memorability; 10 users could effectively remember their password, while 1 user could not easily recall their password.

Commentary on results:

Overall, in the ZMC case, users were positive towards perceived memorability in both the baseline and PoC user authentication system with the ZMC baseline system scoring higher levels of memorability. In the case of FCRB, results reveal a significant increase of perceived memorability for the PoC system compared to the baseline system (4.16 *vs.* 3.63). Results are encouraging for further investigating the proposed flexible and personalized user authentication system since in both PoC case studies, patients reported high levels of perceived memorability. Furthermore, results of the second PoC indicate an increase of perceived memorability (PoC1: 4.12 vs. PoC2: 4.21), which is encouraging for further investigating and improving the FlexPass system.

KPI 3.3: Perceived Security | **PUA**

Metrics:

Following state-of-the-art user studies in usable security research [11, 12], for perceived security, we have asked participants questions on whether they believe the user authentication system is secure, whether they believe their password is strong, etc. Example questions include *"Overall, how secure do you find the FlexPass password system?"*, *"How strong do you believe a FlexPass password is?"*, etc. Users rated the statements through a 5-point Likert scale (*e.g.*, 1: Very insecure - 5: Very secure).

Baseline measurement:

	USTAN	*ZMC*	*FCRB*	*TOTAL*
Baseline	***N/A***	***3.82***	***3.59***	***3.71***

ZMC baseline (patients)

The majority of the ZMC participants perceived the baseline user authentication system as secure (Very secure: 6; Secure: 10; Medium: 3) and commented that their password is strong (Very strong: 4; Strong: 6; Medium: 8; Weak: 1). Finally, with regards to the password reset process, 6 users find the process very secure, while the majority of users (13) believe that the password reset process has a moderate security.

ZMC baseline (feedback from professionals)

Professionals perceive the baseline user authentication system as secure (14/19) while 4 professionals perceive it as moderately secure. With regards to password strength, 10/19 professionals believe their password is strong, while 8 professionals believe it has medium strength. Similarly, 14/19 professionals find the reset process as secure, while 4 professionals believe it has medium security.

FCRB baseline (patients)

The majority of the FCRB participants perceived the baseline user authentication system as secure (Not secure at all: 1; Not secure: 2; Medium: 5; Secure: 11; Very secure: 2) and commented that their password is strong (Very weak: 1; Weak: 0; Medium: 8; Strong: 10; Very strong: 2). Similarly, with regards to the password reset process, the majority of users believe that the password reset process is secure (Not secure at all: 1; Not secure: 1; Medium: 7; Secure: 9; Very secure: 3).

FCRB baseline (feedback from professionals)

Mixed responses were received from professionals with regards to perceived security. A total of 4/19 users perceive the system as not secure, 8 as moderately secure and 7 as secure. The majority of professionals believe their password has medium strength (12/19), 2 believe it is weak while 5 believe it is strong. Similarly, the majority of professionals believe that the password reset process has medium security (11/19), 2 believe it is weak while 6 believe it is strong.

Trial measurement:

Trial measurement (PoC) 1:

	USTAN	***ZMC***	***FCRB***	***TOTAL***
1st Trial	***N/A***	***4.02***	***3.75***	***3.88***
2nd Trial	***3.85***	***4.41***	***4.59***	***4.20***

ZMC PoC1 (patients)

The majority of ZMC participants perceived the Serums user authentication system as secure (25/31) while 5 users and 1 user rated it to have moderate and limited security. With regards to password strength, 25 users believe that their password in the Serums user authentication is strong.

ZMC PoC1 (feedback from professionals)

Mixed responses were received on perceived security; 2 professionals find the system as secure, while 1 user believes it has medium security and another 1 believes it is not secure. With regards to password strength, 2 professionals believe the passwords are strong, while another 2 believe the passwords have medium strength.

FCRB PoC1 (patients)

The majority of FCRB participants perceived the Serums user authentication system as secure (17/24), 2 participants rated it to have moderate security, while 5 participants believe the system is not secure. With regards to password strength, 16 users believe that their password in the Serums user authentication is strong, 4 participants believe it has moderate strength and 4 participants believe the generated passwords are weak.

FCRB PoC1 (feedback from professionals)

All FCRB professionals (5) perceive the authentication system as secure as well as the passwords as strong.

Trial measurement (PoC) 2:

USTAN PoC2: The majority of the USTAN participants perceived the user authentication system as secure with a considerable number participants perceiving the security as medium (Very secure: 4; Secure: 8; Medium: 9) and commented that their password is strong (Very strong: 8; Strong: 7; Medium: 3; Weak: 3).

ZMC PoC2: The majority of the ZMC participants perceived the user authentication system as secure (Very secure: 7; Secure: 4; Medium: 1) and commented that their password is strong (Very strong: 7; Strong: 3; Medium: 1; Weak: 1).

FCRB PoC2: The majority of the FCRB participants perceived the user authentication system as secure (Very secure: 7; Secure: 2; Medium: 1) and commented that their password is strong (Very strong: 7; Strong: 4).

Commentary on results:

Overall, users' responses with regards to perceived security were positive towards both user authentication systems (baseline and PoC). A comparison between the two systems reveals that in both case studies, users perceived the PoC system as more secure than the baseline system.

KPI 3.4: Trust in the proposed PUA scheme	**PUA**

Metrics:

For perceived trust, we have asked participants questions that relate to their trust towards the user authentication system technology, its ability to protect their data privacy, their trust on security and trust to keep their data safe from cybercriminals. Users rated the statements through a 5-point Likert scale (*e.g.*, 1: Not at all - 5: Absolutely).

Baseline measurement:

	USTAN	***ZMC***	***FCRB***	***TOTAL***
Baseline	***N/A***	***4.01***	***3.56***	***3.78***

ZMC baseline (patients)

The majority of ZMC participants indicated that they trust the baseline user authentication system. Specifically, 15/19 trust the authentication technology while 4 participants have moderate trust towards the technology. With regards to trust on privacy, 16/19 participants trust the authentication system to protect their data privacy, 2 users have moderate trust while 1 user has no trust towards the system with regards to privacy. With regards to trust towards security, 13/19 participants are not worried about the security of the authentication system while 3 participants are mildly worried, and another 3 participants are worried about the authentication system security. Finally, when participants were asked whether they trust the authentication system to protect their account and data from cybercriminals, 15/19 participants trusted the system, while 2 participants indicated moderate trust and another 2 indicated not trust.

ZMC baseline (feedback from professionals)

Overall, professionals trust the baseline authentication technology (17/19), as well as trust the system to protect their privacy (13/19). However, in the case of trust towards its security and safety against cybercriminals, a considerable number of professionals (6) commented that they are worried about the security of the authentication technology (12/19 trust the security), and 5 professionals do not trust that their data is safe. Nonetheless, 11/19 professionals showed trust towards the system to keep their data safe against cybercriminals.

FCRB baseline (patients)

The majority of FCRB participants indicated that they trust the baseline user authentication system. Specifically, 14/21 trust the authentication technology while 5 participants have moderate trust towards the technology, and 2 participants do not trust the technology. With regards to trust on privacy, 13/21 participants trust the authentication system to protect their data privacy, 4 users have moderate trust while 4 users have no trust towards the system with regards to privacy. With regards to trust towards security, 13/21 participants are not worried about the security of the authentication system while 4 participants are mildly worried, and another 4 participants are worried about the authentication system security. Finally, when participants were asked whether they trust the authentication system to protect their account and data from cybercriminals, 13/21 participants trusted the system, while 5 participants indicated moderate trust and another 3 indicated not trust.

FCRB baseline (feedback from professionals)

Mixed responses were received from professionals with regards to perceived trust. A total of 12/19 users trust the technology, 5 show moderate trust and 4 do not trust the technology. With regards to trust towards protecting their privacy, 8/19 professionals trust the system, 6 show moderate trust and 5 do not trust the system. A total of 9/19 professionals (9/19) are somewhat worried about the security of the authentication system, while 9 trust the system, 1 user does not trust the system. Finally, 10/19 professionals trust the system to keep their data safe against cybercriminals, 7 have moderate trust while 2 professionals do not trust the system.

Trial measurement:

Trial measurement (PoC) 1:

	USTAN	***ZMC***	***FCRB***	***TOTAL***
1st Trial	***N/A***	***3.96***	***4***	***3.98***
2nd Trial	***3.80***	***3.90***	***4.75***	***4.10***

ZMC PoC1 (patients)

The participants of the ZMC PoC study responded positively with regards to trust towards the Serums user authentication technology. With regards to trust towards the technology, 22/31 trust the Serums authentication technology, 7 have moderate trust while 2 users do not trust the technology. Furthermore, 22/31 participants have trust towards the Serums authentication system to protect their privacy, while 8 users have moderate trust and 1 user has no trust. A total of 18/31 users are not worried about the security of the authentication system, 8 participants are somewhat worried while 5 users are worried about the security. Finally, 21/31 users trust the authentication system to keep their data safe against cybercriminals, 8 have moderate trust while 2 users do not trust the system.

ZMC PoC1 (feedback from professionals)

Mixed responses were received on perceived trust; 2 professionals trust the technology, while 2 users do not trust the technology. With regards to trust to protect privacy, 3 professionals trust the system and 1 does not trust the system. A total of 2 professionals are not worried about the system's security, while another 2 professionals are worried. Finally, 3 professionals trust the authentication system to keep the patients' data safe against cybercriminals, and 1 professional does not trust the system.

FCRB PoC1 (patients)

The participants of the FCRB PoC study similarly responded positively with regards to trust towards the Serums user authentication technology. With regards to trust towards the technology, 18/23 trust the Serums authentication technology, 2 have moderate trust while 3 users do not trust the technology. Furthermore, 18/23 participants have trust towards the Serums authentication system to protect their privacy, while 2 users have moderate trust and 3 users have no trust. A total of 16/23 users are not worried about the security of the authentication system, 2 participants are somewhat worried while 5 users are worried about the security. Finally, 16/23 users trust the authentication system to keep their data safe against cybercriminals, 4 have moderate trust while 3 users do not trust the system.

FCRB PoC1 (feedback from professionals)

All FCRB professionals trust the Serums authentication technology across all trust dimensions (technology, privacy, security, safety). In the case of trust with regards to safety against cyber criminals, 1 professionals rated moderate trust.

Trial measurement (PoC) 2:

USTAN PoC2: The majority of USTAN participants indicated that they trust the user authentication system. Specifically, 14/21 trust the authentication technology while 5 participants mildly trust and 2 participants do not trust the technology. With regards to trust on privacy, 17/21 participants trust the authentication system to protect their data privacy, 2 users have moderate trust while 2 users have no trust towards the system with regards to privacy. With regards to trust towards security, 14/21 participants are not worried about the security of the authentication system while 7 participants are worried about the authentication system security. Finally, when participants were asked on whether they trust the authentication system to protect their account and data from cybercriminals, 15/21 participants trusted the system, while 3 participants indicated moderate trust and another 3 indicated no trust.

ZMC PoC2: The majority of ZMC participants indicated that they trust the user authentication system. Specifically, 10/12 trust the authentication technology while 2 participants do not trust the technology. With regards to trust on privacy, 10/12 participants trust the authentication system to protect their data privacy, 1 user has moderate trust while 1 user has no trust towards the system with regards to privacy. With regards to trust towards security, 8/12 participants are not worried about the security of the authentication system while 1 participant is mildly worried, and another 3 participants are worried about the authentication system security. Finally, when participants were asked on whether they trust the authentication system to protect their account and data from cybercriminals, 8/12 participants trusted the system, while 2 participants indicated moderate trust and another 2 indicated no trust.

FCRB PoC2: The majority of FCRB participants indicated that they trust the user authentication system. Specifically, all 11 users trust the authentication technology. With regards to trust on privacy, similarly all 11 users trust the authentication system to protect their data privacy. With regards to trust towards security, 10/11 participants are not worried about the security of the authentication system while 1 participant is mildly worried. Finally, when participants were asked on whether they trust the authentication system to protect their account and data from cybercriminals, 10/11 participants trusted the system, while 1 participant indicated moderate trust.

Commentary on results:

Overall, participants in the baseline and PoC1evaluation studies at both end-user organizations (ZMC and FCRB) trust the user authentication technologies. A comparison between system versions suggest that in the case of the FCRB study, patients scored higher trust levels for the PoC authentication system compared to the baseline, while in the ZMC case, trust levels were similar for both systems. Furthermore, in PoC2, results indicate an increase in overall perceived trust (PoC1: 3.98 vs. PoC2: 4.1) towards the Serums authentication technology compared to PoC1. A cross organization comparison indicates that USTAN scored comparably lower trust level towards the technology (3.8), ZMC scoring 3.9, and FCRB scoring the highest trust level (4.75).

KPI 3.5: Data Analytics Model Utility and Model Privacy	**PDA**

Metrics:

In this metric we measure the model utility and the model privacy the same way as we do in KPI 2.3 and KPI 4.1, respectively. Even if we look at these properties of a model from a different point of view, how they impact the patient trust, the results are the same as in KPI 2.3 and KPI 4.1.

Baseline Measurement:

For the baseline of this metric we refer to the baseline of KPI 2.3 and KPI 4.1.

Trial Measurement:

For the trial measurement of this metric we refer to KPI 2.3 and KPI 4.1

At the time of writing we have no access to appropriate use case data from the USTAN use case to be able to measure this KPI. But in order to initially evaluate the enhancement of our newly developed approach above the state-of-the-art, we calculated the factor of increase for we use the MNIST data as a benchmark dataset.

The results for baseline and trial measurements for model utility and privacy are the same as for KPI 2.3 and KPI 4.1. For this benchmark dataset the classical Gaussian mechanism achieved a prediction accuracy of 95% for a (e; δ)-differential privacy level of (2; 1-e5). Our proposed mechanism on contrary resulted in a prediction accuracy of 96.84%, which is an increase of 1.84%. The (e; δ)-differential privacy level for a given accuracy of 95% is (2; 1-e5) for the classical Gaussian mechanism. Our proposed mechanism resulted in a differential privacy level of (1.14; 1-e5), which is an increase of privacy by a factor of 1.7544.

KPI 3.6: Patient Trust	**SPR**

Metrics:

The goal of this KPI is to measure the trust that patients have in how their data is stored and shared. We are interested in this as a metric as there is a chance that the solution is less trusted than the existing systems and thus be less likely to be opted into by patients.

Baseline Measurement:

We provided two questions to the end users to be asked that were graded on a scale of 1 - 5. These were designed to measure the patients' trust in the current system. These questions were asked in conjunction with a series of other questions related to other KPIs throughout this deliverable. Both participating hospitals were able to ask 19 patients which gave a maximum potential score of 190 and a minimum of 38. The questions themselves were: "*How comfortable or uncomfortable would you be with this system managing your medical data?*" and "*How capable or incapable do you consider this system in handling medical data securely?*".

	USTAN	***ZMC***	***FCRB***	***TOTAL***
Baseline	***N/A***	***136/190***	***128/190***	***264/380***

Trial Measurement:

Trial measurement (PoC) 1:

At the time that the first PoC took place, the development of the system was very early in development so we were unable to record measurements for the system.

Trial measurement (PoC) 2:

With the refined versions of the software for work package 2 integrated into a single system, we were able to take measurements. The same criterias used for the baseline were measured against the integrated Serums system. As above, we provided the same two questions to the end users which were graded on a scale of 1 - 5. These were designed to measure the patients' trust in the current system. These questions were asked in conjunction with a series of other questions related to other KPIs throughout this deliverable.

	USTAN	***ZMC***	***FCRB***	***TOTAL***
1st Trial	***N/A***	***N/A***	***N/A***	***N/A***
2nd Trial	***210/520***	***111/280***	***99/220***	***420/1020***

Commentary on results:

As expected, the patients exhibited a high trust in the current system. From speaking to the patients, it became clear that this is based on a lifetime of using the hospitals and the in-built assumption that the data within the hospitals' systems is safe. This is an area we may not be able to improve upon with the Serums system.

As can be seen from the tables above, the existing systems are more highly trusted than the Serums system. This is something that we will continue to monitor and attempt to address.

KPI 3.7: Perceived Usability of SERUMS System	**SHCS**

Metrics:

In PoC2 we have included a new KPI for measuring users' perceived usability of the SHCS system with the aim to receive feedback from end-user patients on important usability dimensions such as perceived usefulness and perceived ease of use towards the complete SHCS system. For this purpose, we have designed a questionnaire by following state-of-the-art works and guidelines on usability, user experience and technology acceptance (*i.e.*, [13]).

With regards to perceived usefulness, following the work reported in [13], we have asked questions such as, *"Using the Serums technology would make it possible to share and get insight in the patient's medical data"*, *"Using the Serums technology would make finding and sharing the patient's medical information more efficient"*, *"Using the Serums technology would enhance my ability to retrieve and share all patient's medical files"*, etc. Users rated the statements through a 5-point Likert scale (e.g., 1: Strongly disagree - 5: Strongly agree).

With regards to perceived ease of use, following the work reported in [13], we have asked questions such as, *"Learning to operate the Serums technology would be easy for me"*, *"I would find it easy to get the Serums technology to do what I want it to do"*, *"I would find the Serums technology easy to use"*, etc. Users rated the statements through a 5-point Likert scale (e.g., 1: Strongly disagree - 5: Strongly agree).

Trial Measurement:

Trial measurement (PoC) 2:

We provided a series of questions to the end-users to be asked that were graded on a scale of 1-5. These were designed to measure the patients' perceived usefulness and ease of use towards the overall SHCS system.

Perceived Usefulness:

The patients found the SHCS system highly useful with USTAN scoring an overall mean of 4.57/5, ZMC 4.04/5 and FCRB 4.68/5. Specifically, the majority at all organizations responded that they find the Serums technology useful, i.e., at USTAN, 24/26 patients find the system very useful, 1/26 does not find it useful while 1/26 finds it moderately useful; at ZMC 12/14 patients find the SHCS system useful, while 1/14 respectively find it moderately and not useful; and at FCRB all 11 patients find the system as useful.

	USTAN	***ZMC***	***FCRB***	***TOTAL***

2nd Trial	***4.57***	***4.04***	***4.68***	***4.43***

Perceived Ease of Use.

The patients found the SHCS system as easy to use with USTAN scoring an overall mean of 4.35/5, ZMC 4.2/5 and FCRB 4.32/5. Specifically, the majority at all organizations responded that they find that the Serums technology as easy to use, i.e., at USTAN, 24/26 patients find the system as easy to use, 1/26 does not find it as easy to use while 1/26 finds it moderately easy to use; at ZMC 11/14 patients find the SHCS system as easy to use, 2/14 patient do not find it as easy to use, while 1 patient finds it moderately easy to use; and at FCRB 8/11 patients find the system as easy to use, and 3/11 patients find it moderately easy to use.

	USTAN	***ZMC***	***FCRB***	***TOTAL***
2nd Trial	***4.35***	***4.20***	***4.32***	***4.29***

KPI 3.8: Perceived Data Ownership in the SERUMS System	**SHCS**

Metrics:

In PoC2 we have included a new KPI for measuring users' perceived data ownership when using the SHCS system. For this purpose, we have designed a questionnaire by following state-of-the-art works and guidelines in the field of usable privacy and security (*i.e.*, [14]).

Following the work reported in [14], we have asked questions such as, *"I believe the patient's personal medical information is accessible only to those authorized to have access"*, *"I think the patient has control over what personal information he or she can share via Serums"*, etc. Users rated the statements through a 5-point Likert scale (e.g., 1: Strongly disagree - 5: Strongly agree).

Trial Measurement:

Trial measurement (PoC) 2:

We provided a series of questions to the end-users to be asked that were graded on a scale of 1-5. These were designed to measure patients' data ownership when using the SHCS system. Accordingly, the patients exhibited a high data ownership when using the current system with USTAN scoring an overall mean of 4.37/5, ZMC 3.92/5 and FCRB 4.67/5. Furthermore, when patients were asked *"I think the patient has control over what personal information he or she can share via Serums"*, the majority at all organizations responded that they have control over their data, i.e., at USTAN, 25/26 patients believe they have control over their data, while 1 user has a moderate feeling towards data control; at ZMC 10/14 patients believe they have control over their data, 3/14 believe they do not have control, and 1 user has a moderate feeling towards data control; and at FCRB all 11 patients believe they have control over their data.

	USTAN	***ZMC***	***FCRB***	***TOTAL***
2nd Trial	***4.37***	***3.92***	***4.67***	***4.32***

KPI 3.9: Perceived Security in the SERUMS System	**SHCS**

Metrics:

Following state-of-the-art user studies in usable security research [11, 12], for perceived security, we have asked participants questions on whether they believe the overall SHCS system is secure. Example questions include *"Overall, how secure do you find the Serums system?"*, *"I am not worried about the security of the Serums system"*, etc. Users rated the statements through a 5-point Likert scale (*e.g.*, 1: Very insecure - 5: Very secure).

Trial Measurement:

Trial measurement (PoC) 2:

We provided a series of questions to the end-users to be asked that were graded on a scale of 1-5. These were designed to measure the patients' perceived security towards the overall SHCS system. Accordingly, the majority of patients perceive the SHCS system as secure with USTAN scoring an overall mean of 4.04/5, ZMC 3.73/5 and FCRB 4.48/5. Furthermore, when patients were asked *"Overall, how secure do you find the Serums system?"*, the majority at all organizations responded that they found the system secure, i.e., at USTAN, 21/26 patients found the system secure, 3/26 found the system insecure, while 2/26 scored moderate system security; at ZMC 8/14 patients found the system secure, while 3/14 respectively found the system as insecure or with moderate security; and at FCRB 10/11 patients found the system secure, while 1 patient responded that the system is moderately secure.

	USTAN	***ZMC***	***FCRB***	***TOTAL***
2nd Trial	***4.04***	***3.73***	***4.48***	***4.08***

S4) Quantifiable improvement in patient safety (at least a factor of 2), evidenced by reduced risk of harm through incorrect treatments or medicines mediated by reduced risk of tampering with medical records, and measured vulnerabilities of connected medical systems.

SERUMS’ Technologies Contributing in Achieving the Success Indicator

- **Privacy-preserving Data Analytics (PDA)**

KPI 4.1: Data Analytics Model Utility	**PDA**

Metrics:

Model utility measures the ability of a model to make correct predictions. It is measured as the percentage of corrected predictions by the total number of predictions.

As already explained in KPI 2.3 Enhanced model privacy there is always a tradeoff between a model's privacy and a model's utility. Therefore we need to define the level of privacy, in this case the level of (e; δ)-differential privacy at which we want to measure the utility level in order to be able to compare different approaches of privacy preservations.

Again, no general statement can be made about the increase of model utility, since it depends on the selected level of model privacy and the dataset itself. As a reasonable value of privacy we selected a (e; δ)-differential privacy level of (2; 1-e5) to compare different models.

Baseline Measurement:

The baseline that we compare our developed model to is a state-of-the-art model that uses the classical Gaussian mechanism to achieve differential privacy at a (e; δ)-differential privacy level of (2; 1-e5).

Trial Measurement:

In the trial measurement we calculate the level of utility of our privacy preserving mechanism that uses an optimal-noise adding mechanism at a (e; δ)-differential privacy level of (2; 1-e5). Enhanced model utility is the difference between prediction accuracy of the baseline compared to the trial.

At the time of writing we have no access to appropriate use case data from the USTAN use case to be able to measure this KPI. But in order to initially evaluate the enhancement of our newly developed approach above the state-of-the-art, we calculated the factor of increase for we use the MNIST data as a benchmark dataset.

For this benchmark dataset the classical Gaussian mechanism achieved a prediction accuracy of 95%. Our proposed mechanism on contrary resulted in a prediction accuracy of 96.84%, which is an increase of 1.84%.

4.6 Summary

Success Indicator	KPI	Technology	Baseline	1st Trial	2nd Trial
S1	**1.1: Guessability**	PUA	N/A	N/A	58
	1.2: Password Leaks (through social engineering)	PUA	N/A	N/A	79
	1.3: System Vulnerability	SPR	45	N/A	55
S2	**2.1: Password Cracking Resistance**	CH	N/A	N/A	0
	2.2: Data Breaches	SPR	38	N/A	41
	2.3: Enhanced Model Privacy	PDA	N/A	N/A	N/A
	2.4: Granular Access to Patient Record	DLT	0	N/A	100
	2.5: Authorisation Data Integrity	DLT	0	N/A	100
	2.6: Efficiency of Cross-Country Patient Data Sharing	DLT	33	N/A	100
S3	**3.1: Perceived Usability**	PUA	69	71	78
	3.2: Perceived Memorability	PUA	75	78	80
	3.3: Perceived Security	PUA	68	72	80
	3.4: Trust in the Proposed PUA Scheme	PUA	70	74	78
	3.5: Data Analytics Model Utility	PDA	N/A	N/A	41
	3.6: Patient Trust	SPR	69	N/A	41
	3.7: Perceived Usability of SERUMS System	SHCS	N/A	N/A	86
	3.8: Perceived Data Ownership in the SERUMS System	SHCS	N/A	N/A	77
	3.9: Perceived Security in the SERUMS System	SHCS	N/A	N/A	83
S4	**4.1: Data Analytics Model Utility**	PDA	N/A	N/A	N/A

Table 7. KPI metrics for Baseline, PoC1 and PoC2

Success Indicator	Baseline	1st Trial	2nd Trial
S1	45	N/A	61
S2	38	N/A	58
S3	70	74	74
S4	N/A	N/A	N/A

Table 8. Success Indicator values for Baseline, PoC1 and PoC2

Our findings show that, from a technical standpoint, we are indeed on the right track. Our technical KPIs have seen us achieve similar or better scores than our baseline measurements. These have helped us identify areas which still have room for improvement and will help guide the final set of improvements and developments to be in place by PoC3.

One measurement which has not shown improvement was on Patient Trust (KPI 3.6). This was always expected to be a tough metric to improve upon the baseline. This is due to the fact that patients have trusted their existing healthcare providers for their entire lives. We are happy with the measurement, however, and will continue to monitor this. This has highlighted our need to increase our communication about the Serums solution.

The technical component of blockchain was developed prior to PoC1, but it was not ready to be demonstrated to the end-user because the integration did not take place yet. This means the score of the KPIs (2.4, 2.5, 2.6) can only be measured as of PoC2. The outcome values for the Blockchain KPIs demonstrate an improvement from the baseline to PoC2 thanks to the additional changes which were made in between PoC1 and PoC2. Looking ahead, we do not expect significant changes in these three KPIs, as the focus will be on optimising the existing functionalities and the integration with the overall solution.

With regards to the Serums user authentication technology (FlexPass), results of the second PoC evaluation are promising for further investigating the suggested flexible and personalized user authentication approach since the relevant KPIs of the technology achieved similar or better results in PoC2 compared to PoC1. Results indicate an increase of perceived usability (PoC1: 3.85 vs. PoC2: 4.11), perceived memorability (PoC1: 4.12 vs. PoC2: 4.21), perceived security (PoC1: 3.88 vs. PoC2: 4.2) and perceived trust (PoC1: 3.98 vs. PoC2: 4.1) towards the Serums authentication technology. Memory time scored also very well (135.4/168), which suggests that end-users were able to effectively recall their passwords over a period of one week. Furthermore, the overall system usability score of FlexPass (74.77%) scores well based on the guidelines of the System Usability Scale (SUS), which suggests that a score above 68% indicates that the system entails very good usability practices. Areas for improvement relate to the graphical password creation tasks given that some participants had issues with the gesture input functionality. Comments received from the participants and issues that were spotted in the password creation task will be considered for improving the final FlexPass technology.

Furthermore, with regards to the Two-Factor Authentication (2FA) mobile application that was introduced in PoC2, although a rather limited number of users downloaded and used the mobile application, push notification accuracy scored a 100% success rate. Security-related scores of PoC2 remain unchanged compared to PoC1 given that the same password policies were applied in the second evaluation.

Finally, when end-users were asked on whether they like the proposed personalized and flexible approach for user authentication, the majority of users extremely (26/44) or very much (14/44) liked the idea, with 4 users either moderately (1/44) or slightly (3/44) liking the idea.

Next steps include improving usability issues with regards to the graphical password creation task and the accuracy of gesture inputs. We also plan to investigate the effects of different graphical password policies and graphical password estimation and feedback (as a password strength meter) towards the investigated dependent variables of the FlexPass system.

Concerning privacy preserving data analytics we have not yet been able to calculate the related KPIs due to delays in fabrication of use case data. Nevertheless we have already evaluated our approach on several benchmark data sets in D3.1 and D3.2 and reported the results of one representative data set in this deliverable. These results clearly show that we have made great progress in research on privacy preserving data analytics. The enhancements in the privacy-utility-tradeoff enable us to raise the utility level of machine learning models to improve the provision of health care while still keeping the privacy level sufficiently high.

5 Conclusions

This deliverable presents the work performed during the second phase of the demonstration of the Serums technologies effectiveness. More specifically: i) it evaluates the refined prototypes of the Serums technologies developed against the overall project requirements and success criteria that were refined/updated in D7.4; and ii) reports the progress achieved in the different Success Indicators compared to the work performed during the first phase and presented in D7.3.

Compared to the first phase of the evaluation (during which the integration of the Serums Technologies has not been achieved), for the second phase of the evaluation considerable strides for the Serums Technology have been achieved by the consortium, with an initial integrated and coherent version ready and tested. Because of this, many KPIs and baseline values that could not be measured during the first phase and remained unanswered in D7.3, have now been given outcome values enabling both the end users and the related technical partners to measure the majority of the metrics defined and obtain the impact of the first three Success Indicators, related to the secure provision health and care services, risk of data privacy breaches and patient trust in the provision of smart health care. Furthermore, due to the additions of four new KPIs, namely KPI 2.6 (described in section 4.4) related to the efficiency of cross-country patient data sharing and KPIs 3.7, 3.8 and 3.9 (described in D7.4) related to the perceived usability, data ownership and security of the coherent Serums system, broader aspects of the Serums system could be evaluated, demonstrating the overall impact the system can have. Unfortunately, the outcome of Success Indicator 4 could not be measured during this phase of the evaluation even though the (e; δ)-differential privacy is completed, because the End User Organisations have not completed the data fabrication necessary for implementing and testing analytical models. This Success Indicator will be measured during the third phase.

The main findings extracted by the second phase of the evaluation demonstrates an improvement in all technologies that are used in the integrated Serums Technology. These improvements are especially visible on the technical front. The Blockchain system has shown major improvements in all three domains (granular access to patient record, authorisation data integrity and efficiency of cross-country data sharing) compared to baseline values. In addition, we were able to decrease system vulnerability and mechanisms to both preserve privacy and correct data analysis are ready for implementation with higher accuracy than classical mechanisms. The next major step for these mechanisms is to incorporate large real fabricated data sets into the Data Lakes which then can be used to create the analytical models. The impressions we've received from patients, caregivers and IT staff are also very encouraging, with increased or at least steady perceptions in almost all aspects but some usability aspects (i.e., graphical password creation) of the PUA and the general patient trust in the system. The lower usability of the PUA is to be expected, as the participants were able to identify several imperfections with regards to the gesture input during the graphical password creation and login task. On the other hand, we are satisfied with the feedback we've received on the aspect of patient trust. This will aid us in further improvement of the system.

It is worth mentioning that during the execution of the second phase of the evaluation, the consortium had to deal with the consequences of the Covid-19 pandemic, affecting mostly the smooth operation of the Serums Second Proof of Concept (PoC2). More specifically, since PoC2 was to be conducted with a group of people who are highly vulnerable to the Covid-19 virus, it became clear that for medical, practical, and ethical reasons, the pilot could not continue as planned. In addition, since it was not possible to physically perform PoC2, participants needed to be recruited upfront to which a digital appointment was scheduled. The digital appointments were very difficult to be scheduled among the participants (both for the patients and the caregivers), a fact that had a negative effect on the sample size for ZMC and FCRB. Also, the digital interviews were short and difficult to extract proper results. Finally, we measured higher computer literacy skills from the patients who did participate. This may have skewed a.o. the perceived security and usability results of the Serums Technology in our favour.

6 References

[1] De Muro, P., Mazziotta, M. & Pareto, A. Composite Indices of Development and Poverty: An Application to MDGs. *Soc Indic Res* 104, 1–18 (2011). https://doi.org/10.1007/s11205-010-9727-z

[2] Burr, W., Dodson, D., Polk, W. (2006). Electronic authentication guideline. Technical report, NIST

[3] Komanduri, S., Shay, R., Kelley, P., Mazurek, M., Bauer, L., Christin, N., Cranor, L., Egelman, S. (2011). Of passwords and people: measuring the effect of password-composition policies. In ACM CHI '11, ACM Press, 2595-2604

[4] Microsoft Developers' Blog. Signing in with a picture password. https://docs.microsoft.com/en-us/archive/blogs/b8/signing-in-with-a-picture-password

[5] Zhao, Z., Ahn, G., Seo, J., Hu, H. (2013). On the security of picture gesture authentication. In USENIX Security (SEC'13), USENIX Association, 383–398

[6] Zhao, Z., Ahn, G.J. and Hu, H., 2015. Picture gesture authentication: Empirical analysis, automated attacks, and scheme evaluation. ACM Transactions on Information and System Security (TISSEC), 17(4), pp.1-37.

[7] Stobert, E., & Biddle, R. (2013). Memory retrieval and graphical passwords. In Proceedings of the symposium on usable privacy and security (p. 15). ACM

[8] Constantinides, A., Fidas, C., Belk, M., Pietron, A., Han, T., Pitsillides, A. (2021). From hot-spots towards experience-spots: Leveraging on users' sociocultural experiences to enhance security in cued-recall graphical authentication. International Journal of Human-Computer Studies, Elsevier (to appear)

[9] John Brooke. 1996. SUS-A quick and dirty usability scale. Usability evaluation in industry 189, 194: 4--7.

[10] John Brooke. 2013. SUS: A Retrospective. J. Usability Studies 8, 2: 29--40

[11] Sonia Chiasson, P. C. van Oorschot, and Robert Biddle. 2006. A usability study and critique of two password managers. In Proceedings of the 15th conference on USENIX Security Symposium - Volume 15 (USENIX-SS'06). USENIX Association, USA, Article 1, 1.

[12] Ken Reese, Trevor Smith, Jonathan Dutson, Jonathan Armknecht, Jacob Cameron, and Kent Seamons. 2019. A usability study of five two-factor authentication methods. In Proceedings of the Fifteenth USENIX Conference on Usable Privacy and Security (SOUPS'19). USENIX Association, USA, 357–370.

[13] Fred D. Davis. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. 13, 3 (September 1989), 319–340. DOI:https://doi.org/10.2307/249008

[14] Oshrat Ayalon and Eran Toch. 2019. Evaluating users' perceptions about a system's privacy: differentiating social and institutional aspects. In Proceedings of the Fifteenth USENIX Conference on Usable Privacy and Security (SOUPS'19). USENIX Association, USA, 41–59.

7 Abbreviations

2FA	Two-Factor Authentication
AM	Activity Monitor
CH	Credential Hardening
DLT	Distributed Ledger Technology
ECC	Edinburgh Cancer Centre
FlexPass	The Serums user authentication technology
KPI	Key Performance Indicators
MNIST	Modified National Institute of Standards and Technology
PDA	Privacy-preserving Data Analytics
PHE	Personal Health Environment
PoC	Proof of Concept
PoC2	second Proof of Concept
PoC3	third Proof of Concept
PROMS	Patient Reported Outcome Measures
PUA	Personalized User Authentication
SHC	Smart Health Centre
SI	Success Indicator
SPR	Smart Patient Record
SUS	System Usability Scale
VOT	Verification Technologies
WGH	Western General Hospital

Appendixes

This appendix contains:

1. User Authentication Questionnaire for Proof of Concept Study - Participant Information
2. INSTRUCTIONS FOR PARTICIPANTS - Serums Second Proof of Concept Study
3. QUESTIONNAIRE FOR PATIENTS - Serums Second Proof of Concept Study
4. QUESTIONNAIRE FOR PROFESSIONALS - Serums Second Proof of Concept Study

APPENDIX 1 User Authentication Questionnaire for Proof of Concept Study - Participant Information

What is the study about?

We invite you to participate in a research project about Securing Medical Data in Smart Patient-Centric Healthcare Systems (Serums) which deals with security and privacy of future-generation healthcare systems, putting patients at the centre of future healthcare provision, enhancing their personal care and maximizing the quality of treatment they receive.

Why have I been invited to take part?

The main purpose of this study is to elicit the end-users opinions, preference and likeability with regards to FlexPass, a novel user authentication system that aims to improve usability and memorability of passwords and at the same time preserve security.

Do I have to take part?

This information sheet has been written to help you decide if you would like to take part. It is up to you and you alone whether you wish to take part. If you do decide to take part you will be free to withdraw at any time without providing a reason, and with no negative consequences.

What would I be required to do?

Today you will test that system by performing several tasks in it. Afterwards you will be provided with a questionnaire that will ask your opinion. Please be honest, as your opinion can still cause us to improve the system in the last phase of the project.

The user study will take about 45-75 minutes. Your answers will be treated confidentially and anonymously

Part 1: You will interact with the Proof of Concept (PoC) Web-based authentication system by creating a password and then logging into the system.

Part 2: You will then be given a PoC questionnaire to get feedback on aspects like perceived usability, security, acceptance, and trust towards the PoC authentication system.

After the study, we will send you three notification emails on Day 1, Day 3, and Day 6. Each email will direct you to the Serums system and it will instruct you to access the system

Are there any risks associated with taking part?

There are no risks to individuals participating in this study beyond those that exist in daily life.

Informed consent

It is important that you are able to give your informed consent before taking part in this study and you will have the opportunity to ask any questions in relation to the research before you provide your consent.

For further questions about this study, the project or about the way your contribution will be used, please feel free to contact us.

Who is funding the research?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 826278.

For more information about the project, please visit the project's official Website: www.serums.h2020.org

What information about me or recordings of me ('my data') will you be collecting?

We will explicitly collect your opinions with regards to the user authentication scheme questionnaires to measure the perceived usability, memorability, security, and trust with regards to the user authentication scheme.

During user interaction, we will track the following data for the purpose of the project. *All the data will be anonymously stored without any binding information to the identity of the participants*:

User Interaction and Usage Data

- *Authentication usage data: i)* timed events of user interaction, *i.e.*, time to create each gesture (seconds), time to create password (seconds); *ii)* number of attempts to create and confirm password (ordinal); *iii)* time to login (seconds); *iv)* number of attempts to login (ordinal); and *v)* second factor response (true/false) along with the timestamp of occurrence.
- *Authentication memory data: i)* memory time (seconds) which is the greatest length of time between a password creation and a successful password login using the same password; *ii)* number of password resets (ordinal).
- *False Acceptance Rate of 2FA:* the percentage of identification instances in which unauthorized persons are incorrectly accepted.
- *False Rejection Rate of 2FA:* the percentage of identification instances in which authorized persons are incorrectly rejected.
- *Effectiveness of 2FA:* Percentage of sent push notifications that arrived at the correct smartphone.
- *Failure to Enroll to 2FA:* The percentage of the population which fails to complete enrollment of the mobile application.

User-created Password Data

- Security-enhanced textual and graphical password data based on credential hardening.
- Selected images of the recognition-based picture password without any binding information to the identity of the end-user.
- Gesture type (*i.e.*, tap, line circle) and selections on a background image (*i.e.*, x, y coordinates, image semantics of the selection, whether the selection is a hotspot *vs.* non-hotspot region) of the recall-based picture password without any binding information to the identity of the end-user.

We will also securely store email address for the follow up for the memorability evaluation your email will be stored at our organization for a period of 7 days after your participation, and it will be then deleted permanently. Your email will be used solely for the purpose of sending the above-mentioned 3 notification emails for instructing you only to access the Serums system.

How will my data be securely stored, who will have access to it?

Your data will be stored in an anonymised form, which means that parts of your data will be edited or deleted such that no-one, including the researchers, could use any reasonably available means to identify you from the data. Your un-anonymised data will then be permanently deleted. Your data will be stored in secure location, and only relevant members on the project will be able to access it.

How will my data be used, and in what form will it be shared further?

Your research data will be analysed as part of the research study. It will then be used in various research publications and in the project Reports. It will also be shared i.e. by placing it in a database accessible by other members of the consortium. All data will be anonymised for processing, which means that no-one could use any reasonably available means to identify you from the data and will be stored on a secure server which will be encrypted.

When will my data be destroyed?

Data will only be used during the duration of the project which will end on Dec 2021 and will be destroyed.

International data transfers – Personal data

No identifiable data will be shared. Only anonymised data based on user opinion will be shared and stored by other members of the group in Barcelona, Cyprus and The Netherlands where it will be stored on a secure encrypted server.

Will my participation be confidential?

Yes, your participation will only be known to the relevant members on the project. This data will be kept only for the use of the project and will not be shared out with the members of the consortium.

Use of your personal data and data protection rights

The University of St Andrews (the ‘Data Controller’) is bound by the UK 2018 Data Protection Act and the General Data Protection Regulation (GDPR), which require a lawful basis for all processing of personal data (in this case it is the ‘performance of a task carried out in the public interest’ – namely, for research purposes) and an additional lawful basis for processing personal data containing special characteristics (in this case it is ‘public interest research’). You have a range of rights under data protection legislation. For more information on data protection legislation and your rights visit https://www.st-andrews.ac.uk/terms/data-protection/rights/. For any queries, email dataprot@st-andrews.ac.uk.

Ethical Approvals

This research proposal has been scrutinised and subsequently granted ethical approval by the University of St Andrews Teaching and Research Ethics Committee.

What should I do if I have concerns about this study?

In the first instance, you are encouraged to raise your concerns with the researcher. However, if you do not feel comfortable doing so, then you should contact The School of Computer Science Ethics Administrator. Ethics-cs@st-andrews.ac.uk A full outline of the procedures governed by the University Teaching and Research Ethics Committee is available at https://www.st-andrews.ac.uk/research/integrity-ethics/humans/ethical-guidance/complaints/.

For more information about the project, please visit the project's official Website: www.serums-h2020.org

Thank you for taking your time to support this project!

APPENDIX 2 INSTRUCTIONS FOR PARTICIPANTS - Serums Second Proof of Concept Study

Thank you for participating in this user study for the EU Horizon 2020 research project Serums.

The main purpose of this study is to evaluate the usability and security, as well as to elicit the end-users' opinions, preference and likeability with regards to Serums, a system that allows patients and caregivers to have a central hub place to view and share medical data.

It includes a novel authentication system, FlexPass, which aims to improve usability and memorability of passwords and at the same time preserve security. Moreover, it includes a central place to view your medical data from your hospital and it enables you to set-up rules on who is allowed to see your data and who is not.

Today you will test that system by performing several tasks in it. Afterwards you will be provided with a questionnaire that will ask your opinion. Please be honest, as your opinion can still cause us to improve the system in the last phase of the project.

About FlexPass

FlexPass is a user authentication system that allows users to create secret picture passwords. Instead of remembering complex text passwords, the only thing you need to remember is 3 secret spots on an image by drawing them on the image.

In order to make your picture password more memorable, secure and easier to use, FlexPass provides images tailored to each user's prior daily life activities and experiences.

In addition, in case you like to use textual passwords, you can also create a secret passphrase which you can use to flexibly switch between your picture password in order to login.

Finally, to add an additional layer for security, end-users have the option to install and enrol to a mobile application, which is used as a second factor for authentication through easy-to-use push notifications aiming to increase the security of the login task.

Study Procedure Instructions

> Imagine yourself being Mary Smith, a patient at Edinburgh Cancer Centre. You just have been in the hospital for a consultation with your oncologist.
>
> For your own interest and your future treatment plan, you would like to get insight into your medical data and share it with your general practitioner (GP) and other healthcare professionals.
>
> In order to do that you will use the Serums system.

Please find below the main steps you need to follow for completing this action.

Afterwards a questionnaire starts to ask you several questions about the system.

Starting the test

Requirements:

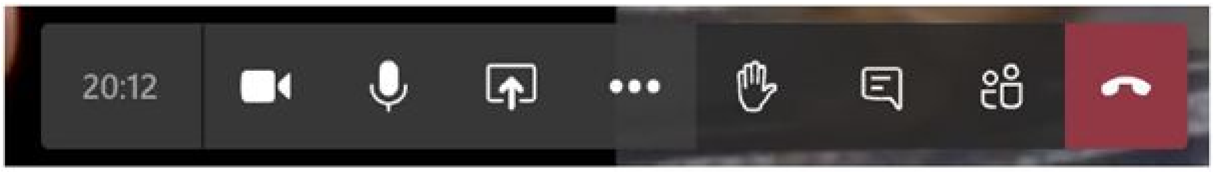
- PC or laptop with conexion to the Internet.
- Chrome browser installed.

Preparation steps:

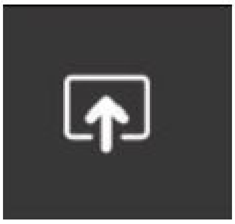
1) Please, access the Microsoft Teams meeting sent to you by email.
2) Open your Chrome browser and type the following address for the Serums system; or click on the link: https://shcs.serums.cs.st-andrews.ac.uk

Sharing your screen

Once you have accessed the Microsoft Teams meeting and in contact with a member of the Serums, you will be asked if you would like to share your screen. In order to do that, first click on the screen, then the menu below will appear:



Then, please select the option indicated below, after that, select the screen you would like to share (the one with the Serums system link, your browser):



Step 1 – Create Picture Password

1) Click on the Sign up button; a user will be provided to you during the Microsoft Teams' meeting.
2) A set of background images will be displayed on the screen. The images will depict content that you are familiar with. You are required to select one image from the set of images, on which you will then create your picture password.
3) Next, you will create a picture password by drawing 3 gestures on an image. You could use any combination of circles, straight lines and taps (clicks).
4) Memorize the size, the position, the directionality, and the ordering of your gestures. These gestures will be your secret picture password.

Step 2 – Create Textual Password (optional)

In case you like to use textual passwords, you can also create a secret passphrase (minimum 16 characters long), which you can use to flexibly switch between your picture password in order to login.

In order to make your password more memorable, we suggest reflecting the secret you created in the picture password as your passphrase. For example, *"the day I had lunch with my friends at the cafeteria"* .

Step 3 – Setup Second Factor for Authentication – requires installation of mobile application (optional)

In case you would like to make the access to your account more secure, you could set up an additional authentication factor by receiving a push notification during login on your mobile device.

In order to make access to your account more secure, we strongly encourage you to set up the second factor for authentication.

Step 4 – Login and Approval

In order to login in the system, you need to choose your preferred authentication method (picture or text) and then proceed to login by entering your secret password.

If you selected a second factor for authentication for increased security during the login process, you also need to approve your login through a push notification that will show up on your mobile device.

Step 5 – Find your medical information

In order to be able to share the necessary medical data with your general practitioner, you need to know where it is. Access your Smart Patient Record (SPHR) to find the information about diagnosis and personal information.

Step 6 – Find your treatment data

You also want to consult the information available on your upcoming treatments. Please find the information and view the details.

Step 7 – Allow a healthcare professional to see your data

You want to share your general medical information and diagnosis, for instance, with the general practitioner (GP) or other professional of your choice. Please allow the professional to see your data.

Step 8 – Deny your acquaintance to see your medical data

Within the oncologists group you have an acquaintance which you rather would not share your data with. Please deny a specific professional access to all of your data.

Last Step 9 – Questionnaire

In the last step, please answer a questionnaire to indicate your opinions, preference and likeability with regards to Serums' system.

Post-study User Interaction

After the study, we will send you three notification emails on Day 1, Day 3, and Day 6. Each email will direct you to the Serums system and it will instruct you to access the system. For doing so, your email will be stored at our organization for a period of 7 days after your participation, and it will be then deleted permanently. *Your email will be used solely for the purpose of sending the above-mentioned 3 notification emails for instructing you only to access the Serums system.*

Thank you for participating in this user study and help us improve Serums!

About Processed Data

During user interaction, we will track the following data for the purpose of the project. *All the data will be anonymously stored without any binding information to the identity of the participants*:

User Interaction and Usage Data

- *Authentication usage data: i)* timed events of user interaction, *i.e.*, time to create each gesture (seconds), time to create password (seconds); *ii)* number of attempts to create and confirm password (ordinal); *iii)* time to login (seconds); *iv)* number of attempts to login (ordinal); and *v)* second factor response (true/false) along with the timestamp of occurrence.
- *Authentication memory data: i)* memory time (seconds) which is the greatest length of time between a password creation and a successful password login using the same password; *ii)* number of password resets (ordinal).
- *False Acceptance Rate of 2FA:* the percentage of identification instances in which unauthorized persons are incorrectly accepted.
- *False Rejection Rate of 2FA:* the percentage of identification instances in which authorized persons are incorrectly rejected.
- *Effectiveness of 2FA:* Percentage of sent push notifications that arrived at the correct smartphone.
- *Failure to Enroll to 2FA:* The percentage of the population which fails to complete enrollment of the mobile application.

User-created Password Data

- Security-enhanced textual and graphical password data based on credential hardening.
- Selected images of the recognition-based picture password without any binding information to the identity of the end-user.
- Gesture type (*i.e.*, tap, line circle) and selections on a background image (*i.e.*, x, y coordinates, image semantics of the selection, whether the selection is a hotspot *vs.* non-hotspot region) of the recall-based picture password without any binding information to the identity of the end-user.
- Textual password complexity, which describes how complex (*e.g.*, low, or highly complex) a textual password is based on the users' selection of characters. Textual password complexity will be calculated at run-time based on state-of-the-art password strength meters.
- Graphical password complexity, which describes how complex (*e.g.*, low, or highly complex) a graphical password is based on the users' image selections and gestures. Graphical password complexity will be calculated at run-time based on state-of-the-art graphical password strength meters.

About the User Study

The user study will take about 15 minutes. Your answers will be treated confidentially and anonymously. Participation in the study is voluntary and can be cancelled at any time. You can terminate your participation at any time. In doing so, you also object to the use of your data collected up to that point.

The data collected as part of this study and described above will be treated confidentially.

Furthermore, the results of the study will be published in anonymous form, i.e., without your data being personally identifiable. There are no risks to individuals participating in this study beyond those that exist in daily life.

For further questions about this study, the project or about the way your contribution will be used, please feel free to contact us.

About Serums and Contact Information

The Serums project (Securing Medical Data in Smart Patient-Centric Healthcare Systems) deals with security and privacy of future-generation healthcare systems, putting patients at the centre of future healthcare provision, enhancing their personal care and maximizing the quality of treatment they receive.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 826278.

For more information about the project, please visit the project's official Website:

https://www.serums-h2020.org/

How to contact us:

The Serums Team, School of Computer Science, University of St Andrews

Serums-local@st-andrews.ac.uk

APPENDIX 3 QUESTIONNAIRE FOR PATIENTS - Serums Second Proof of Concept Study

Thank you for taking your time to support this project!

Consent

By clicking the "Next" button you declare that you

1) understand the purpose of the study,

2) are over 18 years old,

3) voluntarily participate in this study, and

4) have taken note and understand the study information presented above.

User ID

1. Please enter your User ID that was provided by the researcher

   Text field

General Background

2. What is your Age range (in years)?

   18-25; 26-35; 36-45; 46-55; 56-65; 66 and above

3. What is your highest degree of education?

   Ph.D. Studies; Master Studies; Bachelor Studies; High School; Primary School

4. How would you rate your computer literacy?

   Beginner 1 2 3 4 5 Advanced

5. Do you currently have regular access to a computer?

   Yes; No

FlexPass Password System Usability

Please rate the usability of the FlexPass Password System

6. I think that I would like to use the FlexPass system frequently.

   Strongly disagree 1 2 3 4 5 Strongly agree

7. I found the FlexPass system unnecessarily complex.

   Strongly disagree 1 2 3 4 5 Strongly agree

8. I thought the FlexPass system was easy to use.

   Strongly disagree 1 2 3 4 5 Strongly agree

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

   Strongly disagree 1 2 3 4 5 Strongly agree

10. I found the various functions in the FlexPass system were well integrated.

    Strongly disagree 1 2 3 4 5 Strongly agree

11. I thought there was too much inconsistency in the FlexPass system.

    Strongly disagree 1 2 3 4 5 Strongly agree

12. I would imagine that most people would learn to use the FlexPass system very quickly.

    Strongly disagree 1 2 3 4 5 Strongly agree

13. I found the FlexPass system very cumbersome to use.

    Strongly disagree 1 2 3 4 5 Strongly agree

14. I felt very confident using the FlexPass system.

    Strongly disagree 1 2 3 4 5 Strongly agree

15. I needed to learn a lot of things before I could get going with the FlexPass system.

    Strongly disagree 1 2 3 4 5 Strongly agree

Password Creation

Please rate your experience and perceptions with regards to the FlexPass password creation system and process

16. Overall, how difficult or easy did you find the password creation task in FlexPass?

    Very difficult 1 2 3 4 5 Very easy

17. Overall, how slow or fast did you find the password creation task in FlexPass?

    Slow 1 2 3 4 5 Fast

18. How long (in seconds) did you need to create your password in FlexPass?

    Text field

19. Overall, how secure do you find the FlexPass password system?

    Very insecure 1 2 3 4 5 Very secure

20. How strong do you believe your FlexPass password is?

    Very weak 1 2 3 4 5 Very strong

21. Did the image scenery impact your password selections (i.e., did you create a certain story when selecting points on the image, did you consider any past experiences as part of your selections)? If yes, please explain how the image scenery impacted your password selections (optional)

    Text field

22. How did you decide where to draw the gestures on the image? (optional)

    Text field

23. How did you decide which gesture (tap, line, or circle) to draw? (optional)

Text field

24. What strategy did you follow to create your password? (optional)

    Text field

25. What type of background image would you prefer? (optional)

    Text field

Password Login

Please rate your experience and perceptions with regards to the FlexPass login system

26. Overall, how difficult or easy did you find the login task in FlexPass?

    Very difficult 1 2 3 4 5 Very easy

27. How mentally demanding was the login task?

    Very low 1 2 3 4 5 Very high

28. I could easily log on to the FlexPass password system

    Strongly disagree 1 2 3 4 5 Strongly agree

29. I effectively remembered my password

    Strongly disagree 1 2 3 4 5 Strongly agree

Two-factor Authentication Mobile Application

In case you have used the two-factor authentication mobile application, please rate your experience and perceptions with regards to the two-factor authentication system

30. Did you successfully install and enroll to the two-factor authentication mobile application?

    Yes No

31. If your answer was "Yes", which two-factor authentication method did you use to login?

    Push notification message; Secret code (Time-based One-Time Password - TOTP)

32. Did you successfully access the system after using the two-factor authentication method?

    Yes No

33. Overall, how difficult or easy did you find the installation and enrollment to the two-factor authentication mobile application?

    Very difficult 1 2 3 4 5 Very easy

34. Overall, how difficult or easy did you find the two-factor authentication approval task (push notification or secret code)?

    Very difficult 1 2 3 4 5 Very easy

35. Overall, how secure do you find the two-factor authentication mobile application?

    Very insecure 1 2 3 4 5 Very secure

36. I would be willing to use the two-factor authentication mobile application in my everyday tasks

Strongly disagree 1 2 3 4 5 Strongly agree

Password Reset

In case you have reset your password, please rate your experience and perceptions with regards to the FlexPass password reset system and process

37. Overall, how difficult or easy did you find the password reset process of the FlexPass system?

    Very difficult 1 2 3 4 5 Very easy

38. Overall, how secure did you find the password reset process of the FlexPass system?

    Very insecure 1 2 3 4 5 Very secure

Trust

Please rate your trust towards the FlexPass password system

39. I trust in the technology the FlexPass password system is using

    Strongly disagree 1 2 3 4 5 Strongly agree

40. I trust in the ability of the FlexPass password system to protect my privacy

    Strongly disagree 1 2 3 4 5 Strongly agree

41. I am not worried about the security of the FlexPass password system

    Strongly disagree 1 2 3 4 5 Strongly agree

42. I trust the FlexPass password system to protect my account and data from cybercriminals

    Strongly disagree 1 2 3 4 5 Strongly agree

Password Experience and Preference

Please explain your overall experience, preference and opinions with regards to the FlexPass password system

43. Do you like the idea of creating picture passwords with personalized images tailored to the users' prior daily life activities and experiences?

    Not at all 1 2 3 4 5 Extremely

44. Do you like the idea of allowing users to flexibly choose their preferred authentication method (picture or text password)?

    Not at all 1 2 3 4 5 Extremely

45. What are the positive aspects you like in the FlexPass password system? (optional)

    Text field

46. What are the negative aspects you do not like in the FlexPass password system? (optional)

    Text field

47. I would be willing to use the FlexPass password system as an alternative user authentication system to login to my user account

Strongly disagree 1 2 3 4 5 Strongly agree

48. Explain the reasoning behind your answer in the previous question

Text field

Patient trust medical data

49. How comfortable (1) or uncomfortable (5) would you be with this system managing your medical data?

Very comfortable 1 2 3 4 5 Very uncomfortable

50. How capable (1) or incapable (5) do you consider this system in handling medical data securely?

Very capable 1 2 3 4 5 Very incapable

51. Please rate your agreement with the following statement: “I trust this system to handle my medical data in a safe and secure manner”

Strongly disagree 1 2 3 4 5 Strongly agree

Perceived Usefulness Questions (PU)

52. Using the Serums technology would make it possible to share and get insight in my medical data

Strongly disagree 1 2 3 4 5 Strongly agree

53. Using the Serums technology would make finding and sharing my medical information more efficient

Strongly disagree 1 2 3 4 5 Strongly agree

54. Using the Serums technology would enhance my ability to retrieve and share my medical files

Strongly disagree 1 2 3 4 5 Strongly agree

55. I would find the Serums technology useful

Strongly disagree 1 2 3 4 5 Strongly agree

Perceived Ease of Use Questions (PEU)

56. Learning to operate the Serums technology would be easy for me

Strongly disagree 1 2 3 4 5 Strongly agree

57. I would find it easy to get the Serums technology to do what I want it to do

Strongly disagree 1 2 3 4 5 Strongly agree

58. It would be easy for me to become skillful in the use of the Serums technology

Strongly disagree 1 2 3 4 5 Strongly agree

59. I would find the Serums technology easy to use

Strongly disagree 1 2 3 4 5 Strongly agree

Behavioural Intention to use (BI)

60. I would intend to use the Serums technology when I need access to my medical files

Strongly disagree 1 2 3 4 5 Strongly agree

Data ownership

61. I believe my personal information is accessible only to those authorized to have access.

Strongly disagree 1 2 3 4 5 Strongly agree

62. It is clear what information about me Serums keeps in the system.

Strongly disagree 1 2 3 4 5 Strongly agree

63. It is clear who is the audience of my shared information.

Strongly disagree 1 2 3 4 5 Strongly agree

64. I think Serums allows me to restrict the access to some of my personal information to some people.

Strongly disagree 1 2 3 4 5 Strongly agree

65. I think I have control over what personal information I can share via Serums.

Strongly disagree 1 2 3 4 5 Strongly agree

66. It is clear what information about me caregivers can see on Serums.

Strongly disagree 1 2 3 4 5 Strongly agree

Perceived security of Serums system

67. Overall, how secure do you find the Serums system?

Very insecure 1 2 3 4 5 Very secure

68. I am not worried about the security of the Serums system

Strongly disagree 1 2 3 4 5 Strongly agree

69. I trust in the ability of the Serums system to protect my privacy

Strongly disagree 1 2 3 4 5 Strongly agree

70. I trust in the technology the Serums system is using

Strongly disagree 1 2 3 4 5 Strongly agree

APPENDIX 4 QUESTIONNAIRE FOR PROFESSIONALS - Serums Second Proof of Concept Study

General Background

1. What is your Age range (in years)?

   18-25; 26-35; 36-45; 46-55; 56-65; 66 and above

2. What is your highest degree of education?

   Ph.D. Studies; Master Studies; Bachelor Studies; High School; Primary School

3. What is your occupation?

   Doctor; Nurse; Caregiver; IT Expert; Security Expert; Other

4. How would you rate your computer literacy?

   Beginner 1 2 3 4 5 Advanced

5. Do you currently have regular access to a computer?

   Yes; No

General Preference and Opinion about FlexPass

Please explain your overall preference and opinions with regards to the FlexPass password system

6. Do you like the idea of creating picture passwords with personalized images tailored to the users' prior daily life activities and experiences?

   Not at all 1 2 3 4 5 Extremely

7. Do you like the idea of allowing users to flexibly choose their preferred authentication method (picture or text password)?

   Not at all 1 2 3 4 5 Extremely

8. Do you believe that FlexPass would be a good alternative authentication method for patients?

   Not at all 1 2 3 4 5 Extremely

9. What are the positive aspects you like in the FlexPass password system?

   Text field

10. What are the negative aspects you do not like in the FlexPass password system?

    Text field

11. Would you be willing to use the FlexPass password system as an alternative user authentication system to login to your user account?

    Yes; No

12. Explain the reasoning behind your answer in the previous question

    Text field

Password Creation

Please rate your perceptions with regards to the FlexPass password creation system and process

13. Overall, how difficult or easy do you find the password creation task in FlexPass?

    Very difficult 1 2 3 4 5 Very easy

14. Overall, how slow or fast do you find the password creation task in FlexPass?

    Slow 1 2 3 4 5 Fast

15. Overall, how secure do you find the FlexPass password system?

    Very insecure 1 2 3 4 5 Very secure

16. How strong do you believe a FlexPass password is?

    Very weak 1 2 3 4 5 Very strong

Password Login

Please rate your perceptions with regards to the FlexPass login system

17. Overall, how difficult or easy do you find the login task in FlexPass?

    Very difficult 1 2 3 4 5 Very easy

18. How mentally demanding do you believe the login task is?

    Very low 1 2 3 4 5 Very high

19. Patients will easily log on to the FlexPass password system

    Strongly disagree 1 2 3 4 5 Strongly agree

20. Patients will effectively remember their password

    Strongly disagree 1 2 3 4 5 Strongly agree

Two-factor Authentication Mobile Application

Please rate your perceptions with regards to the two-factor authentication system

21. Overall, how difficult or easy do you find the installation and enrollment to the two-factor authentication mobile application?

    Very difficult 1 2 3 4 5 Very easy

22. Overall, how difficult or easy do you find the two-factor authentication approval task (push notification)?

    Very difficult 1 2 3 4 5 Very easy

23. Overall, how secure do you find the two-factor authentication mobile application?

    Very insecure 1 2 3 4 5 Very secure

Password Reset

Please rate your perceptions with regards to the FlexPass password reset system and process

24. Overall, how difficult or easy do you find the password reset process of the FlexPass system?

    Very difficult 1 2 3 4 5 Very easy

25. Overall, how secure do you find the password reset process of the FlexPass system?

    Very insecure 1 2 3 4 5 Very secure

Trust

Please rate your trust towards the FlexPass password system

26. I trust in the technology the FlexPass password system is using

    Strongly disagree 1 2 3 4 5 Strongly agree

27. I trust in the ability of the FlexPass password system to protect the patients' privacy

    Strongly disagree 1 2 3 4 5 Strongly agree

28. I am not worried about the security of the FlexPass password system

    Strongly disagree 1 2 3 4 5 Strongly agree

29. I trust the FlexPass password system to protect my account and data from cybercriminals

    Strongly disagree 1 2 3 4 5 Strongly agree

Patient trust medical data

30. How comfortable (1) or uncomfortable (5) would you be with this system managing the patient’s medical data?

    Very comfortable 1 2 3 4 5 Very uncomfortable

31. How capable (1) or incapable (5) do you consider this system in handling medical data securely?

    Very capable 1 2 3 4 5 Very incapable

32. Please rate your agreement with the following statement: “I trust this system to handle medical data in a safe and secure manner”

    Strongly disagree 1 2 3 4 5 Strongly agree

Perceived Usefulness Questions (PU)

33. Using the Serums technology would make it possible to share and get insight in the patient’s medical data

    Strongly disagree 1 2 3 4 5 Strongly agree

34. Using the Serums technology would make finding and sharing the patient’s medical information more efficient

    Strongly disagree 1 2 3 4 5 Strongly agree

35. Using the Serums technology would enhance my ability to retrieve and share all patient's medical files

Strongly disagree 1 2 3 4 5 Strongly agree

36. I would find the Serums technology useful

Strongly disagree 1 2 3 4 5 Strongly agree

Perceived Ease of Use Questions (PEU)

37. Learning to operate the Serums technology would be easy for me

Strongly disagree 1 2 3 4 5 Strongly agree

38. I would find it easy to get the Serums technology to do what I want it to do

Strongly disagree 1 2 3 4 5 Strongly agree

39. It would be easy for me to become skillful in the use of the Serums technology

Strongly disagree 1 2 3 4 5 Strongly agree

40. I would find the Serums technology easy to use

Strongly disagree 1 2 3 4 5 Strongly agree

Behavioural Intention to use (BI)

41. I would intend to use the Serums technology when I need access to all patients medical files

Strongly disagree 1 2 3 4 5 Strongly agree

Data ownership

42. I believe the patient's personal medical information is accessible only to those authorized to have access.

Strongly disagree 1 2 3 4 5 Strongly agree

43. It is clear what information about the patient Serums keeps in the system.

Strongly disagree 1 2 3 4 5 Strongly agree

44. It is clear who is the audience of the patient's shared information.

Strongly disagree 1 2 3 4 5 Strongly agree

45. I think Serums allows the patient to restrict the access to some of his personal information to some people.

Strongly disagree 1 2 3 4 5 Strongly agree

46. I think the patient has control over what personal information he or she can share via Serums.

Strongly disagree 1 2 3 4 5 Strongly agree

42. It is clear what patient information caregivers can see on Serums.

Strongly disagree 1 2 3 4 5 Strongly agree

Perceived security of Serums system

48. Overall, how secure do you find the Serums system?

Very insecure 1 2 3 4 5 Very secure

49. I am not worried about the security of the Serums system

Strongly disagree 1 2 3 4 5 Strongly agree

50. I trust in the ability of the Serums system to protect my privacy

Strongly disagree 1 2 3 4 5 Strongly agree

51. I trust in the technology the Serums system is using

Strongly disagree 1 2 3 4 5 Strongly agree