EOSC-Life: Building a digital space for the life sciences

D14.3 – Report about use and user satisfaction of the COVID-19 repository, including a maintenance and sustainability plan

WP14 – Design, development, implementation and use of a repository for individual participant data from COVID-19 trials
Lead Beneficiary: ECRIN
WP co-leaders: Jacques Demotes (ECRIN), Gard Thomassen (UiO)

Authors of this deliverable: **Sergio Contrino (ECRIN), Gerd Felder (ECRIN), Christian Ohmann (ECRIN), Maria Panagiotopoulou (ECRIN), Swarnalathaa Kichenassamy (ECRIN), Sergei Gorianin (ECRIN), Jacques Demotes (ECRIN), Milen Kouylekov (UiO), Gard Thomassen (UiO)**
With contributions from: **Jonathan Ewbank (ERINHA)**

Contractual delivery date: **31 March 2022**
Actual delivery date: **17 April 2023**
H2020-INFRAEOSC-2018-2

Grant agreement no. 824087
Horizon 2020
Type of action: RIA

Table of Contents

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

Executive Summary

In EOSC-Life WP14, the European Clinical Research Infrastructure Network (ECRIN) has partnered with the University of Oslo (UiO) in order to commonly design, develop, implement and operate a Clinical Research Repository for individual participant data (IPD) from COVID-19 clinical studies that is compliant with European regulations and in particular with the GDPR. The functionality of the repository is split between two main systems – the service for sensitive data (TSD) infrastructure managed by the UiO on the one hand and a Repository Management System (RMS) developed by ECRIN, on the other. The TSD is used for the secure storage, access and reuse of controlled-access data objects (e.g. datasets). The RMS is designed to support and record the workflows associated with managing the repository and its interactions with data object providers, users, and requesters.

The first chapter of this report provides an introduction to the Clinical Research Repository, the actors involved, and the two main processes that it supports: the Data Transfer Process (when clinical study material is provided to the repository) and the Data Use Process (when clinical study material is accessed for secondary use). Terms around the operations of the repository such as *"data object"*, *"provider"* and *"secondary user"* are defined to facilitate the further reading of the report.

The second chapter describes the steps followed for evaluating the usability and user satisfaction of the Clinical Research Repository. Usability is defined as *"the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use"*. User satisfaction is part of the usability of a system and can be defined as *"a measure of how well a product or system meets the needs and expectations of its users, as well as how positively users feel about their overall experience with the product or system"*. To evaluate the usability and user satisfaction of the repository, two workshops were organised on the 14th and 15th of March 2023 with a small group of potential users (n=13). As part of the workshops, the participants tested the alpha version of the repository and completed a usability and user satisfaction survey. Overall, the respondents of the survey appreciated the Clinical Research Repository as a tool to facilitate clinical research data sharing and reuse. Usability scores according to the UMUX-LITE and SUS standards are provided. The respondents also raised points for improvement that are being addressed for the beta version of the tool in order to increase its usability and user satisfaction.

The third chapter details the requirements for the maintenance and sustainability of the Clinical Research Repository, taking into account the needs of both the RMS and the TSD. The cost estimations include needs in software, hardware, and personnel (e.g. for the repository manager, for developers, for legal support with contracts). In an effort to be precise, estimations are provided for a short-term, a mid-term and a longer-term period. Possible funding models for the repository are discussed. Beyond the EOSC-Life project, the repository's short-term sustainability has been secured via follow-up grants; however, for the longer-term sustainability plan to be translated into a concrete reality, extensive discussions will be needed with relevant stakeholders following the public launch of the repository.

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

Project Objectives

"D14.3 - Report about use and user satisfaction of the COVID-19 repository, including a maintenance and sustainability plan" contributed to the integration of user feedback into the Clinical Research Repository development process. It also provided the WP14 partners with valuable insights into how the system is perceived by "naive users" and how to improve future versions to better address users' needs. Maintenance and sustainability requirements have been collected for the short-term, mid-term and longer-term. Possible funding models have been identified and the short-term sustainability of the repository has been secured through EC project funding. The longer-term sustainability plan will need to be discussed with relevant stakeholders following the public launch of the repository. These discussions will be initiated within EOSC-Life but are expected to continue beyond the project's duration.

The above serve the WP14 objective of defining the specifications, developing, implementing, and routinely operating a repository for individual participant data (IPD) from COVID-19 trials, compliant with European regulations and in particular with GDPR, allowing clinical trial data sharing after completion of the trial.

Detailed Report on the Deliverable

1. Introduction

1.1 The Clinical Research Repository

The EOSC-Life COVID-19 Clinical Research Repository - in this document referred to simply as "the repository" - is designed to provide secure, file-based, storage for material generated by clinical research, including documents, individual participant datasets (IPD) and metadata, making that material available to others in accordance with the relevant EU ethical and legal frameworks.

In the first instance, the repository is intended to store material from clinical studies related to COVID-19, though there is nothing inherent in the repository's design or planned operation that limits it to this particular area of research. The stored material will remain as discrete files - there is no intention at this stage to aggregate data from different studies into a single platform, or curate data into a single structure. The repository is therefore similar in architecture to other file-based systems, e.g. Vivli[1], CSDR[2], and Dryad[3].

The repository is being developed jointly by the European Clinical Research Infrastructure

[1] https://vivli.org/
[2] https://www.clinicalstudydatarequest.com/
[3] https://datadryad.org/stash/

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

Network, ECRIN[4], and the University of Oslo, UiO[5]. UiO will provide a pre-existing secure infrastructure for storage and access control, called TSD, or the Service for Sensitive Data[6]. ECRIN will act as the main interface of the repository with the research community - i.e. negotiating with those providing the material and those requesting access to it, managing the processes of both data transfer and data use, ensuring the provision of adequate metadata, and monitoring compliance with legal and ethical obligations.

To improve the clarity and readability of this report some key terms used in the text are defined:

A ***Data Object*** is any file available in electronic form, of any type (document, data, media etc.). The repository will store and make available data objects, not simply "data" (or datasets, as defined below), because it is designed to contain protocols, analysis plans, consent forms, result summaries and other documents associated with a clinical research study, as well as IPD.

The term ***Dataset*** is used when referring to a data object that contains only data – e.g., a spreadsheet, CSV, JSON or XML file, database dump, etc. In the context of the repository, "dataset" will usually refer to the file or files of IPD derived from a clinical research study.

A Data Object Provider, or more often simply the ***Provider***, is an organisation that provides data objects to the repository, i.e. that enters into a Data Transfer Agreement with the repository. Unless those data objects are already explicitly in the public domain, the Data Object Provider is assumed to have the legal power to enter into that agreement (e.g. clinical trial sponsors). For datasets of sensitive personal data, the Provider would be the Data Controller as defined under the GDPR.

Individuals seeking to re-use data objects are called Data Object Secondary Users, or simply the ***Secondary Users***. The organisation arranging such re-use on their behalf, normally their employer, is the data object ***Requester.***

Figure 1 illustrates the roles of the various groups involved in the repository schematically.

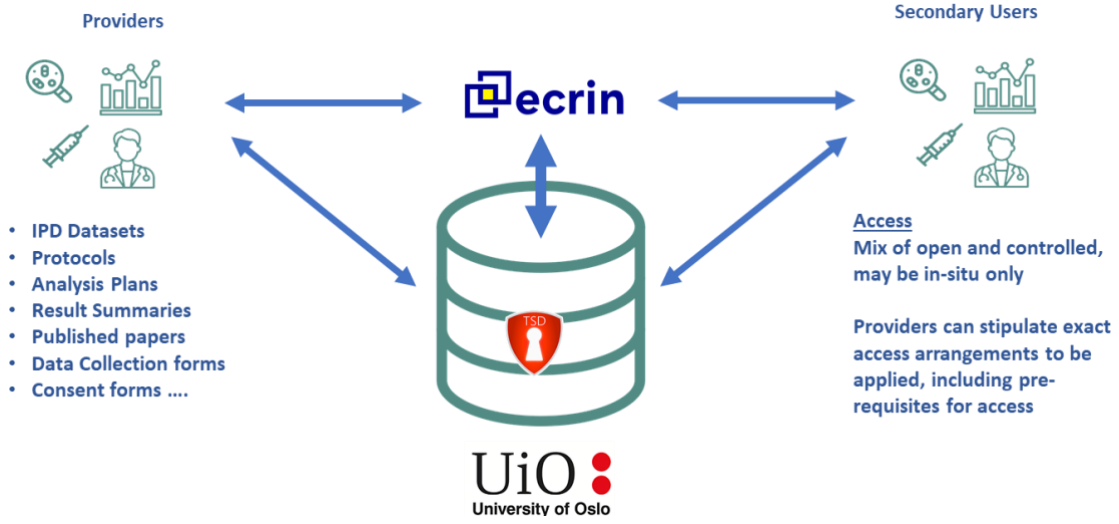


Figure 1: The main actors involved in the repository.

[4] https://ecrin.org/
[5] https://www.uio.no/english/
[6] https://www.uio.no/english/services/it/research/sensitive-data/

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

The transfer of objects to the repository will be governed by a ***Data Transfer Agreement*** (DTA), with that document covering the transfer of all data objects, including but not limited to datasets. Each DTA will include an appendix describing the data objects to be transferred to the repository, and the access arrangements desired for each of them. Further details (e.g. prerequisites for secondary use, categorisation as anonymised or pseudonymised) will be required for objects under controlled access. For data objects under managed access, requester organisations will be asked to sign a ***Data Use Agreement*** (DUA) on behalf of the secondary users. The DUAs will constrain the use of the managed access objects, typically explicitly prohibiting any attempt to re-identify individuals within a dataset, stipulating that datasets should be stored securely and destroyed after use, and requiring that the providers and repository are notified of the results of the secondary use. In some cases, providers may stipulate that access can only take place *in-situ*, within the TSD infrastructure.

The functionality of the repository is split between two main systems - the TSD infrastructure on the one hand and a ***Repository Management System*** (RMS), developed by ECRIN, on the other. The RMS is designed to support and record the workflows associated with managing the repository and its interactions with providers, users, and requesters. The TSD will be used for the secure storage, access and reuse of controlled-access data objects (e.g. datasets).

1.2 Functionality and workflows in the repository

The repository's functionality is based upon the interaction of the RMS and the TSD infrastructure. There are two main processes or workflows that need to be supported:

- ➔ the **Data Transfer Process (DTP)**, when clinical study material is provided to the repository.
- ➔ the **Data Use Process (DUP)** when clinical study material is accessed for secondary use.

During the DTP:

a) The initial information about the data objects to be uploaded, and the people who will upload them, will be collected within the RMS. This includes all necessary study and data object metadata and the access arrangements, as stipulated by the data object providers, required for non-public data objects.

b) Once this information has been collected a DTA can be constructed (based on a standard template) and agreed, and this too is recorded in the RMS. The RMS then communicates to TSD:
   - Who will be uploading the data objects, including their e-mail addresses, parent organisations, and Life Science AAI ID.
   - What data objects are to be uploaded.

c) The TSD staff can then arrange for access to be provided to the identified individuals, so that they can upload the material to a specific part of the TSD system. The TSD staff then inform the people uploading:
   - Of the credentials required for access.

   At the same time, they should inform the RMS, via ECRIN staff:
   - That these arrangements have been put in place.

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

- When the material has been uploaded.
- Where ECRIN staff can locate the uploaded material, for quality checks.

d) The ECRIN staff then carry out quality checks on the uploaded data objects and associated metadata, the results being stored in the RMS. These checks make sure, for example, that descriptive metadata has been applied, and de-identification of the data has taken place. They need to inform TSD of the results of the quality checks.
e) If checks are successful, ECRIN changes the status of the material so that it is advertised as available (even if under controlled access for some data objects).
f) If the check is unsuccessful, a dialogue needs to occur between ECRIN and the provider staff, with the possibility of uploading revised versions of some data objects. ECRIN will record this process within the RMS and keep TSD informed, in case re-application or extension of upload permissions are required.
g) Once the material has been established within the repository as "available", TSD informs:
    - ECRIN and the data uploaders that this is the case. The DTP is then complete.

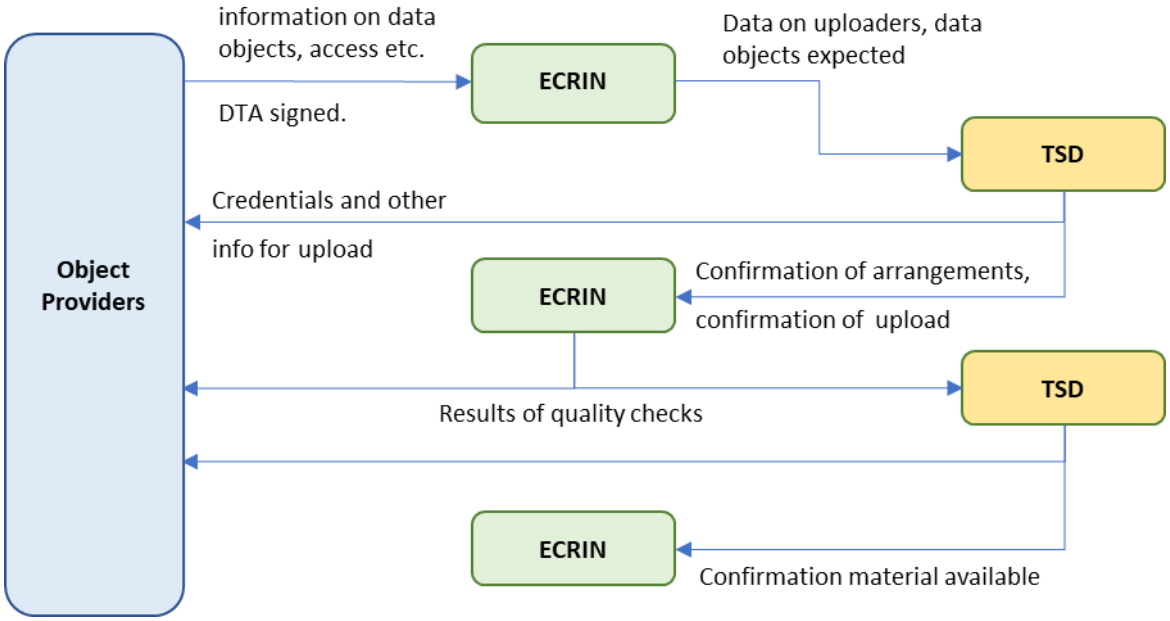


Figure 2: The Data Transfer Process.

During the DUP:

a) Initial information about the request – including the data objects sought, the people requesting and the reasons for their request – is collected by ECRIN within the RMS. Where prerequisites have been established by the data providers, these are also checked.
b) DUAs are established between the data requesters and either the data providers directly or between the requesters and the repository acting on behalf of the data providers (depending on the providers' previously stated preferences).
c) Once the DUA is agreed ECRIN communicates to TSD:
    - The data objects that are to be made available.

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

- The individuals (names, e-mail addresses, organisations, Life Science AAI ID) to whom permissions for access *in-situ* and / or download should be provided.

d) TSD staff then inform the people accessing / downloading the material:
   - Of any credentials (or processes for establishing credentials) required for access.
   - The specific area within the system where the material can be found (if this is not handled automatically by the system).
   - Time limits for access.

   ECRIN is informed once these arrangements are put in place and the RMS is updated.

e) The users confirm successful access / download to ECRIN, and ECRIN then:
   - Informs TSD that the DUP has been completed (so that permissions can be removed).

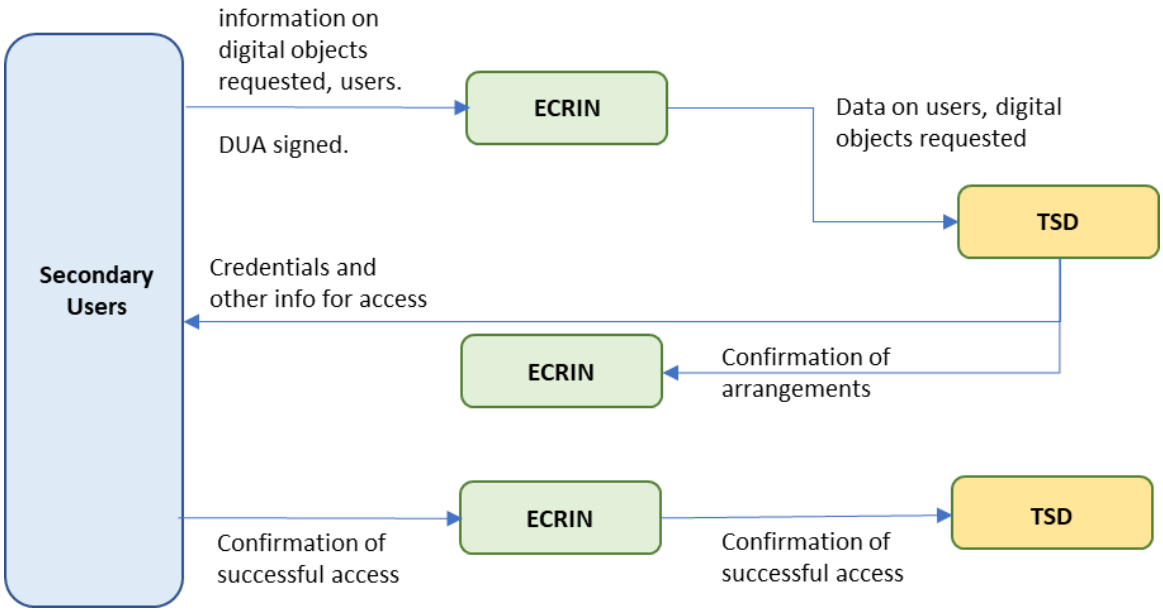


Figure 3: The Data Use Process.

2. Usability and user satisfaction

ISO (in standard ISO 9241-11[7]) defines usability as *"The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use"*. Usability is a quality attribute that assesses how easy user interfaces are to use. Usability is defined by 5 quality components[8]:

- **Learnability:** How easy is it for users to accomplish basic tasks the first time they encounter the design?
- **Efficiency:** Once users have learned the design, how quickly can they perform tasks?

[7] https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en
[8] https://www.nngroup.com/articles/usability-101-introduction-to-usability/

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

- **Memorability:** When users return to the design after a period of not using it, how easily can they re-establish proficiency?
- **Errors:** How many errors do users make, how severe are these errors, and how easily can they recover from the errors?
- **Satisfaction:** How pleasant is it to use the design?

Usability is important in User Experience (UX) design because it is a measure of how easy and intuitive a product or system is to use, and how well it meets the needs of its users. A product with high usability will be easy to learn, easy to use, and efficient in helping users achieve their goals. In contrast, a product with poor usability can frustrate users, increase the time and effort required to complete tasks, and ultimately lead to user dissatisfaction. User satisfaction can be defined as *a measure of how well a product or system meets the needs and expectations of its users, as well as how positively users feel about their overall experience with the product or system*. While usability is a key factor in user satisfaction, other factors such as aesthetics, functionality, and emotional engagement also play a role.

Nowadays, several models for measuring usability and user satisfaction are available. Doll and Torkzadeh[9] developed a 12-item End-User Computing Satisfaction (EUCS) instrument by contrasting the traditional data processing environment and end-user computing environment, which comprised of 5 components: content, accuracy, format, ease of use, and timeliness (Figure 4). The construct was developed with a five-point Likert-type scale (1 = almost never; 2 = some of the time; 3 = about half of the time; 4 = most of the time; and 5 = almost always).

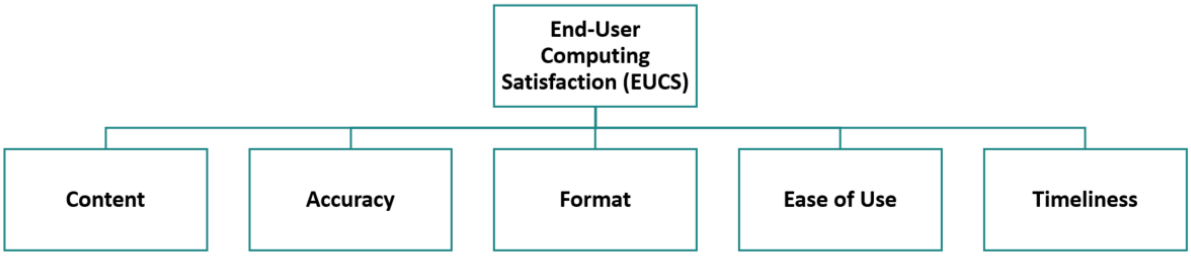


Figure 4: End-User Computing Satisfaction (EUCS) instrument by Doll and Torkzadeh (1988).

Another example is the System Usability Scale (SUS)[10] questionnaire, which uses 10 five-point scales (see Appendix A). SUS yields a single number representing a composite measure of the overall usability of the system being studied. To calculate the SUS score, first the score contributions from each item need to be summed. Each item's score contribution will range from 0 to 4. For items 1,3,5,7, and 9 the score contribution is the scale position minus 1. For items 2,4,6,8 and 10, the contribution is 5 minus the scale position. Then, the sum of the scores needs to be multiplied by 2.5 to obtain the overall value of usability. SUS scores have a range of 0 to 100. Based on research, a SUS score above a 68 would be considered above average and anything

[9] Doll W.J. and Torkzadeh, G. (1988) The Measurement of End-User Computing Satisfaction. MIS Quarterly, 12, 259-272. https://doi.org/10.2307/248851

[10] Brooke J. (1996). SUS: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (Eds.), Usability Evaluation in Industry. London: Taylor and Francis.

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

below 68 is below average, however the best way to interpret results involves "normalising" the scores to produce a percentile ranking.

The Usability Metric for User Experience (UMUX)[11] was designed to get a measurement of perceived usability consistent with the SUS, but using only four (rather than 10) items. The primary purpose for its development was to provide an alternate metric for perceived usability for situations in which it was critical to reduce the number of items while still getting a reliable and valid measurement of perceived usability (e.g., when there is a need to measure more attributes than just perceived usability leading to limited "real estate" for any given attribute). Like the standard SUS, UMUX items vary in tone but unlike the SUS, have seven rather than five scale steps from 1 (strongly disagree) to 7 (strongly agree). The four UMUX items are: 1. This system's capabilities meet my requirements. 2. Using this system is a frustrating experience. 3. This system is easy to use. 4. I have to spend too much time correcting things with this system.

UMUX-LITE is a short version of the UMUX, consisting of its positive-tone (odd-numbered) items (maintaining the use of 7-point scales). Thus, for the UMUX-LITE, the items are: 1. This system's capabilities meet my requirements. 2. This system is easy to use. A reason for including the specific two items of the UMUX-LITE was their connection to the content of the items in the Technology Acceptance Model (TAM)[12], a questionnaire from the market research literature that assesses the usefulness (e.g., capabilities meeting requirements) and ease-of-use of systems, and has an established relationship to likelihood of future use. According to TAM, good ratings of usefulness and ease of use (perceived usability) influence the intention to use, which influences the actual likelihood of use.

UMUX-LITE was chosen to evaluate the perceived usability of the COVID-19 Clinical Research Repository because of its parsimony (two items), reliability, validity and structural basis (usefulness and usability). UMUX-LITE offers a promising alternative to SUS when it is not desirable to use a long 10-item instrument for measuring usability[13].

2.1 Design of the usability and user satisfaction study

Subjects

The participants of the usability and user satisfaction study were 13 volunteers: 9 females and 4 males. Out of the 13 volunteers, 7 are part of the ECRIN core team, 5 are working as European Correspondents (EuCos) and 1 as senior data consultant. The ECRIN core team is based in Paris (France), and the EuCos are present in all the ECRIN member countries to coordinate the work with the national scientific networks and partners. ECRIN has 10 member countries (Czech Republic, France, Germany, Hungary, Ireland, Italy, Norway, Portugal, Spain, and Poland) and 2 observer countries (Switzerland and Slovakia). The age of the respondents ranged from 27 to 74 years. The respondents differed in their level of computer experience with 5 having "high"

[11] Finstad K.A. (2010). The Usability Metric for User Experience.Interacting with Computers, 22 (5), 323–327, https://doi.org/10.1016/j.intcom.2010.04.004

[12] Davis D. (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly. 13, 319–339

[13] Lewis J.R., Utesch B.S., Maher D.E. (2015). Investigating the Correspondence Between UMUX-LITE and SUS Scores. In: Marcus, A. (eds) Design, User Experience, and Usability: Design Discourse. Lecture Notes in Computer Science, vol 9186. Springer, Cham. https://doi.org/10.1007/978-3-319-20886-2_20

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

computer experience and 8 having "moderate" computer experience (self-evaluation). They also have different knowledge levels of the clinical research field, with their involvement in clinical trials ranging from 0 to 22 years.

Table 1 below provides key descriptive statistics of the participants of the usability and user satisfaction study.

Gender	*Female*	*Male*	*Total*
	9 (69,2%)	4 (30,8%)	13
Computer experience	*Low*	*Moderate*	*High*
	0 (0%)	8 (61,5%)	5 (38,5%)
Years of experience in clinical trials	*0-9 years*	*10-15 years*	*>15 years*
	6 (46,1%)	5 (38,5%)	2 (15,4%)
Location of testing	*Face-to-face workshop*	*Online workshop*	*Total*
	6 (46,2%)	7 (53,8%)	13

Table 1: Descriptive statistics of the study participants.

Materials and Procedure

The usability and user satisfaction study took place on the 14th and 15th of March 2023 (Figure 5). The 13 respondents were asked, based on their availability, preferred testing mode, and testing location to join one of the two workshops that were organised in the frame of EOSC-Life WP14. On the 14th of March, a face-to-face workshop was organised at the ECRIN office in Paris and it was joined by 6 respondents. On the 15th of March, an online workshop was organised using Microsoft Teams and it was joined by 7 respondents. Both workshops took place from 10:00-14:00 CET (4 hours). Apart from the 13 respondents, representatives from the ECRIN and UiO EOSC-Life WP14 teams joined the workshops to oversee the procedure and provide information and clarifications to the study participants.

The participants were given prior to the workshops background information material about the repository and its development. This information material consisted of the EOSC-Life WP14 deliverables "*D14.1 - Strategic plan for the development of a COVID-19 repository including specification of technical requirements, policies and procedures"*[14] and "*D14.2 - Report about technical implementation and validation of the COVID-19 portal"*[15] and of the repository's Data Sharing Policy[16]. Reading the background information material prior to joining the usability and user satisfaction study was not mandatory, but it was recommended.

[14] https://zenodo.org/record/4341385#.ZBm_uHbMKUk
[15] https://zenodo.org/record/6555207#.ZBm_93bMKUk
[16] https://zenodo.org/record/5519122#.ZBnAQ3bMKUk

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

Figure 5: Group photo from the face-to-face workshop in Paris (A). Screenshot from the online workshop in Microsoft Teams (B).

On the day of the usability and user satisfaction study, the respondents were given a brief (~15 minutes) introduction to EOSC-Life and the WP14 "Design, development, implementation and use of a repository for individual participant data from COVID-19 trials" and then they were provided with i) the url of the Repository Management System[17], ii) the url of the user guide[18], iii) a survey implemented with Google forms to capture the feedback of the respondents. They were asked to test in the alpha version of the repository the 2 main processes that the Clinical Research Repository supports: **1) The Data Transfer Process** (= putting clinical study material in the

[17] https://ecrin-rms.org/
[18] https://crr.gitbook.io/crr/

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

repository) and **2) The Data Use Process** (= accessing clinical study material already stored in the repository).

The alpha version of a software product is a pre-release early version that is part of a dedicated testing process. Most software products move through a multi-step process before being released to the public. An alpha version is part of that system for developing efficient, accurate and bug-free software programs. The feedback collected during the user satisfaction study will serve to improve the beta version of the system and correct any remaining bugs before launching the repository to the public.

2.2 Evaluating the user guide

To support the "seamless" navigation and use of the Clinical Research Repository, a user guide was developed using GitBook and it is publicly available[19]. The first item in the usability and user satisfaction survey consisted of evaluating the usefulness of the user guide for navigating the Clinical Research Repository portal (Figure 6).

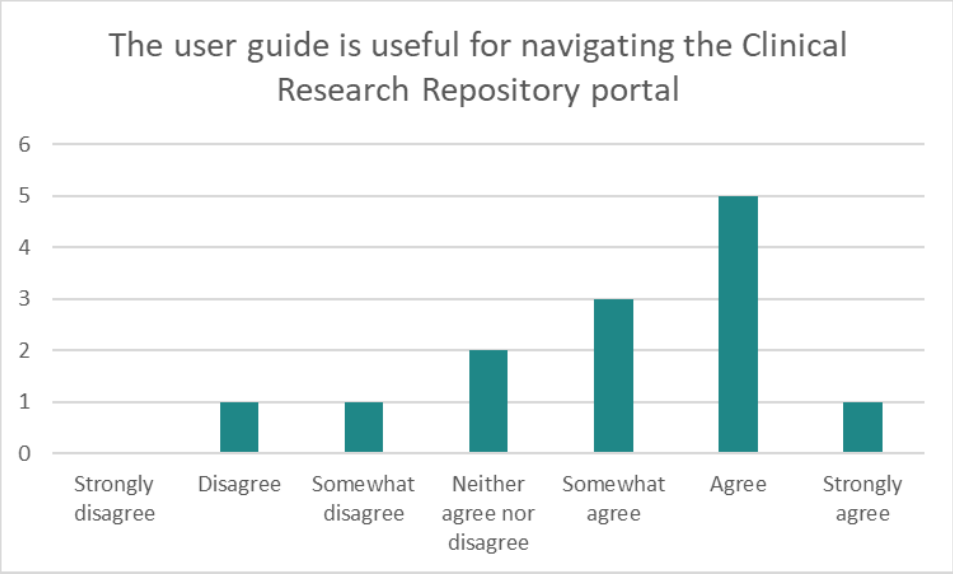


Figure 6: Evaluating the usefulness of the user guide[20] for navigating the Clinical Research Repository portal.

Out of the 13 respondents, 9 agree to some extent (1 strongly agrees, 5 agree, 3 somewhat agree) that the user guide is useful for the navigation of the Clinical Research Repository portal, especially for "first-time" users. 2 respondents are neutral (neither agree nor disagree) with regards to the usefulness of the current version of the user guide and 2 are critical of its usefulness (1 disagrees and 1 somewhat disagrees).

[19] https://crr.gitbook.io/crr/
[20] https://crr.gitbook.io/crr/

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

The respondents were also asked to suggest points for improvement. Some of their comments were:

- *"The user guide is still very broad. One would expect to find information that allows the user to quickly understand what the user must fill in at each field, through each stage of the different processes. In addition, not all the sections found in the Clinical Research Repository portal are searchable in the user guide."*
- *"Having more detailed text in the user guide for all the sections of the Clinical Research Repository would help the user understand better the steps to be followed. Accompanying screenshots showing how to fill out the forms could also be useful."*
- *"The user guide is very helpful. When completing the information in the section "Studies" some fields were not easy to understand. Specific fields that caused confusion were Study enrolment, topic value, CT code, and import and export for uploading files. Maybe consider including definitions either in the user guide or directly at the appropriate sections where the request to the user is unclear."*
- *"There should be an option to download the user guide as a pdf file so that it can be viewed as users fill in the data."*
- *"People tend to not read instructions and user guides, the sections and the fields that need to be completed by the user should be self-explanatory or the information for the fields that may cause confusion should be provided directly in the portal."*

Apart from the above, a couple of typos were spotted in the user guide that have been now corrected. The respondents suggested also including more information about the exact security measures that the TSD applies in the section of the user guide called "The service for sensitive data". The respondents highlighted that the repository is disease agnostic and this should be reflected in the user guide and in all the material around the repository. Despite the repository being developed with the COVID-19 use case in mind, there is nothing inherent that prevents it from serving any type of disease area and this is something that should be clear in the documentation.

2.3 Evaluating the landing page

Next, the 13 respondents of the survey were asked to evaluate the landing page of the Clinical Research Repository. As the current landing page is being updated with the help of ECRIN's communications team, we asked the respondents to evaluate the design proposed by the communications team (Appendix B) so that we can integrate their feedback. People were asked to evaluate whether the landing page design is aesthetically appealing (Figure 7A) and whether it seems to provide the information that they would expect to see "at a glance" on the landing page of the repository (Figure 7B).

Out of the 13 respondents, all of them agree to some extent (5 strongly agree, 7 agree, 1 somewhat agrees) that the design provided by the ECRIN communications team was visually appealing. When asked to evaluate whether the landing page seems to provide useful information to the user "at a glance" 10 respondents agree to some extent (8 respondents agree, 2 respondents somewhat agree), 2 are neutral (neither agree nor disagree) and 1 respondent somewhat disagrees.

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

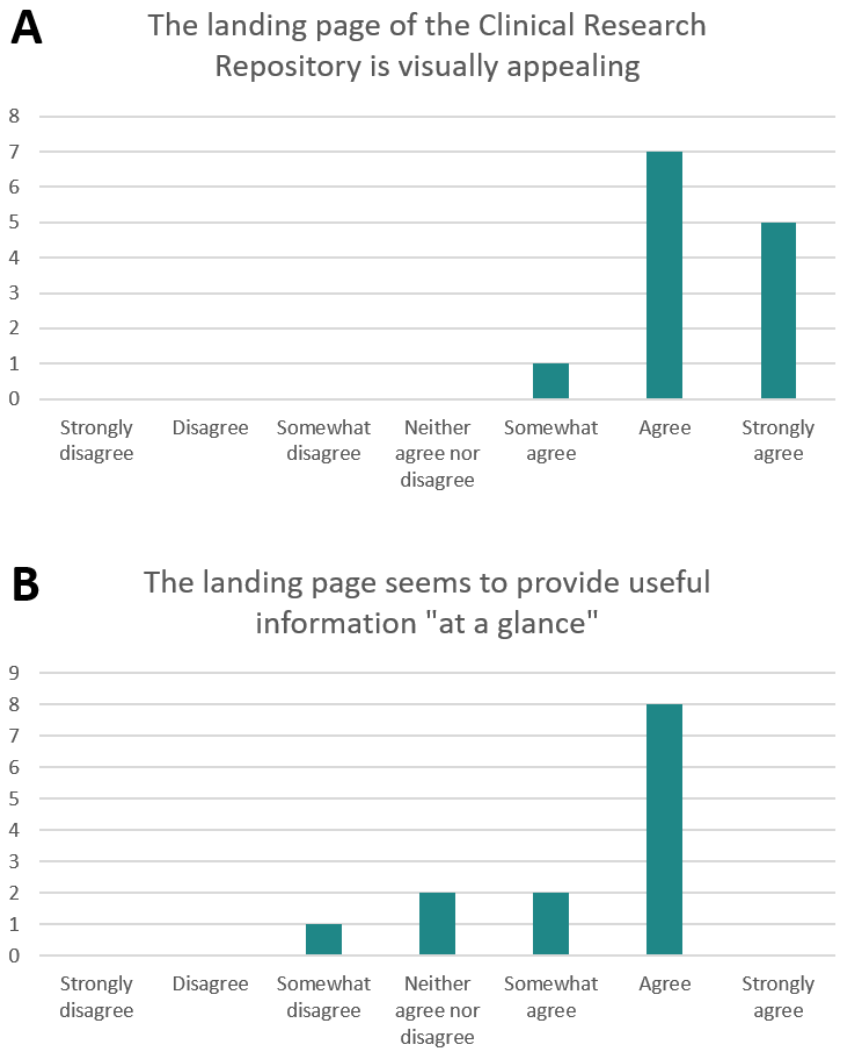


Figure 7: Evaluating the landing page of the Clinical Research Repository.

The respondents were also asked to suggest points for improvement. Some of their comments were:

- *"The landing page would look nicer with more colours. Also, the "University of Oslo" text is currently in black, which makes it hard to read in contrast with the dark blue in the background."*
- *"A link leading to an explanatory page about the repository should be easily visible and accessible. Something like a "Who we are" section."*
- *"There are two types of potential users: the ones who want to share data for secondary use (data providers) and the ones who want to reuse data provided by others. Maybe this distinction could already become visible on the landing page and each type of user would only see the information regarding his/her request and needs."*

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

2.4 Evaluating the Data Transfer Process

The third item that we wanted to evaluate through the survey was the user experience of the Data Transfer Process. This includes the RMS portal fields relevant to data transfer, the TSD fields relevant to data transfer, and the implementation of the Data Transfer Process as a whole (Figure 8).

Out of the 13 respondents, 8 agree to some extent (5 agree, 3 somewhat agree) that the Data Transfer Process is implemented in a user-friendly manner, 1 respondent is neutral (neither agrees nor disagrees) and 4 respondents somewhat disagree.

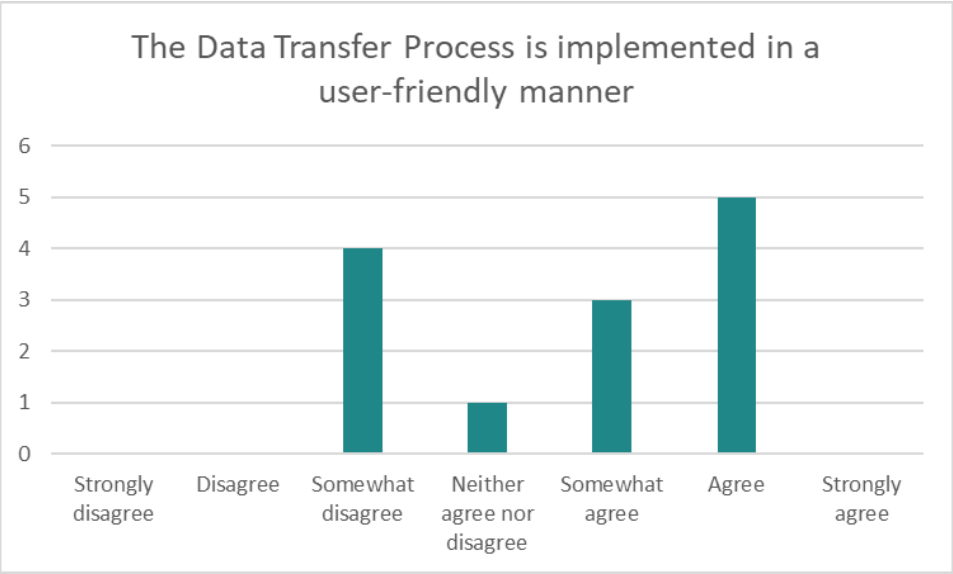


Figure 8: Evaluating the implementation of the Data Transfer Process.

The respondents were also asked to suggest points for improvement. Some of their comments were:

- *"The forms to describe a study and a data object have many fields and in some of them the user does not understand what is expected. For example, in the study section: What is a "data sharing statement"?; What is meant by "study enrolment"?; And as regards enrolment, what is meant by "units"?; What is meant by "CT code"?; What is meant by "Study identifier(s)/Identifier Organisation"?; Or in the data object section: What/who is the "Managing Organisation"? One would expect that the user guide gives precise definitions of what is expected in each field. Otherwise, information bubbles could be added and by hovering over them, one is provided a definition/explanation of what must go into the field."*
- *"It is possible to mark a study as completed, and then register a starting date that is in the future - this should be blocked. Also, when registering a data object currently it is not possible to include any URLs."*
- *"In the future, it should be checked how much time a "naive user" needs to complete each section and ensure that this is reduced as much as possible; clinical researchers are quite busy,*

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

and if we do not provide them with a user-friendly and not cumbersome tool it might further discourage data sharing."

- *"For the user, it is not clear when the process of including the study and the data objects is finished and what is the next step. Maybe the user could receive an e-mail alert or a message from the portal informing him/her for example when the repository manager needs to approve the quality checks step. Also, after adding a data object, it would be great to have at the bottom an "add another object" button more visible."*
- *"The mandatory fields for the description of the study and the data objects need to be clearly marked."*
- *"The layout is appealing and easy to understand, but the process of completing the information for studies and data objects is not intuitive for some fields. It would help for example to know if what you are asking for is text or a number. When you ask "study enrolment" is it to describe the process of registering or entering the participants or is it the status of the study, and months of recruitment? The format of months given in numbers is not intuitive and the format JAN, FEB would be preferable. The start date to me is the full date DD-MM-YYYY, currently, the portal asks separately for day, month, and year. At the end of the pages for studies and data objects, there is no button to return to the main menu and we are left without knowing if we have completed all the fields. The access to TSD upload needs to be explained step by step in the user guide instructions."*
- *"It should be specified that DOI number should not be a URL and just a number. In the data objects section also, there seems to be a bug and when a real URL is entered in the field called "URL" the tab closes and an error is caused."*
- *"Being able to link your data object to any study in your organisation could be problematic; you could end up linking to a study you don't know, especially in big organisations like INSERM[21]."*

2.5 Evaluating the Data Use Process

Following the evaluation of the Data Transfer Process, we asked the respondents to evaluate whether the Data Use Process has been implemented in a user-friendly manner. This includes the relevant fields in the RMS, the relevant fields in the TSD, and the implementation of the Data Use Process as a whole (Figure 9).

Out of the 13 respondents, 7 agree to some extent (1 strongly agrees, 3 agree and 3 somewhat agree) that the Data Use Process is implemented in a user-friendly manner, 1 respondent was neutral (neither agrees nor disagrees) and 5 respondents disagree to some extent (4 somewhat disagree and 1 disagrees).

[21] referring to the French National Institute of Health and Medical Research

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

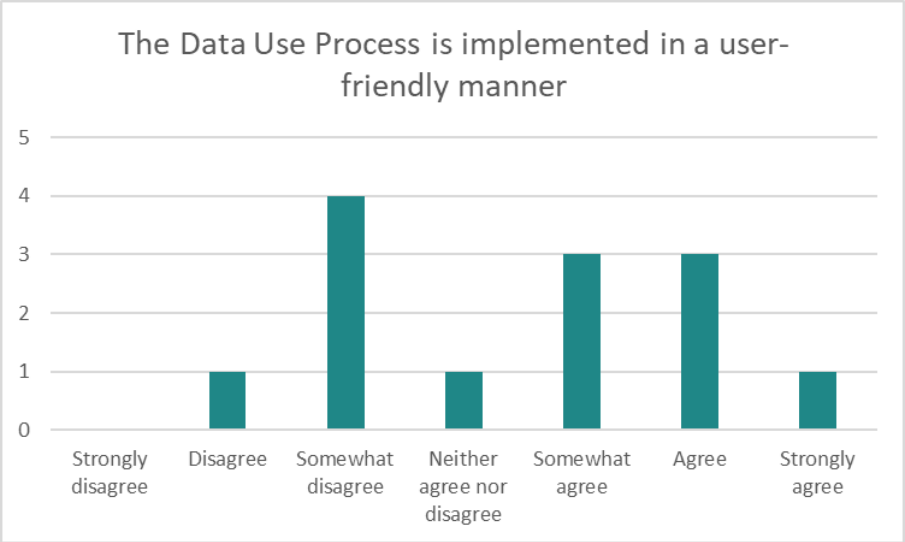


Figure 9: Evaluating the implementation of the Data Use Process.

The respondents were also asked to suggest points for improvement. Some of their comments were:

- *"After testing the Data Transfer Process, the Data Use Process becomes intuitive. Nevertheless, for first-time users, the same issues as described in the Data Transfer Process will apply."*
- *"Some definitions and detailed/clear guidelines in the user guide for the Data Use Process are still missing."*
- *"Log-in process to the secure space of TSD definitely needs to be simplified. This will require a step-by-step outline in the user guide. It would also be useful to have a section of "common mistakes to avoid". In addition, having to use a personal phone for authentication for entering TSD can be a problem for some users."*[22]
- *"I find the change from a website to a virtual machine (VM) a little strange. It's somewhat difficult to use a desktop machine inside a browser window. Maybe this is necessary for privacy and security, but I feel it is somewhat overkill to give access to a VM, only for people to access a file folder (e.g. in case of download and not in situ use)."*
- *"The default selection options for data objects and names need to be deleted."*

2.6 Evaluating the added value of the repository

Next, we wanted to evaluate how the respondents perceived the usefulness of the repository for improving clinical research data sharing and reuse.

Currently, and despite the efforts of different stakeholders (journal editors, funders, patient organisations) to promote clinical research data sharing and reuse, in practice, only a small percentage of studies declare in their data sharing statements (DSS) that IPD will become

[22] referring to the need to use the Google Authenticator App for entering TSD

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

available for sharing (literature reports that the studies with a positive DSS do not exceed the 5-15% of the registered ones[23] a percentage that drops even further when actual data sharing requests are received[24])·

First, we asked the 13 respondents of the survey if they were aware of other repositories that serve for clinical research data sharing and reuse (Figure 10). 8 respondents answered that they do not know other repositories, while 5 respondents answered that they already knew existing repositories that could serve the purpose of clinical research data sharing and reuse.

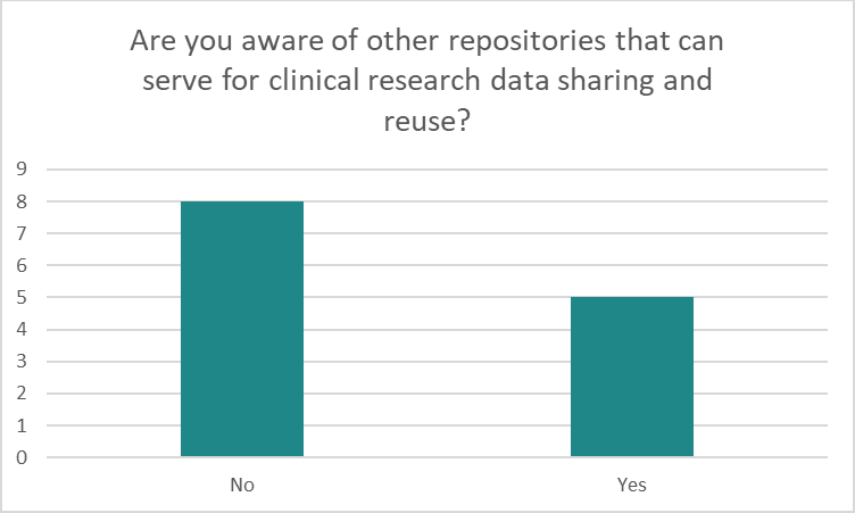


Figure 10: Evaluating whether the respondents are aware of other repositories for clinical research data sharing and reuse.

In the case that the respondents declared knowing other clinical research repositories they were asked to name them. The 5 people that answered "yes" named the following repositories:

- Vivli
- Health Data Hub
- MIP, OHDSI
- current project
- Vivli, BioLINCC, YODA, DRYAD, FigSHARE

Although Vivli, BioLINCC, YODA, DRYAD and FigSHARE can be used for the sharing and reuse of clinical research data, the rest of the answers listed (Health Data Hub, MIP, OHDSI) have a primary focus on routinely collected health data (as opposed to clinical research data that the question was evaluating). This means that in reality, only 2 out of the 13 respondents were aware of existing solutions for clinical research data sharing other than the EOSC-Life repository.

[23] Larson K., Sim I., von Isenburg M. et al. (2022) COVID-19 interventional trials: Analysis of data sharing intentions during a time of pandemic. Contemporary Clinical Trials. 115:106709. https://doi.org/10.1016/j.cct.2022.106709
[24] Merson L., Ndwandwe D., Malinga T., Paparella G., Oneil K., Karam G., Terry R.F. (2022) Promotion of data sharing needs more than an emergency: An analysis of trends across clinical trials registered on the International Clinical Trials Registry Platform. Wellcome Open Research. 7:101. https://doi.org/10.12688/wellcomeopenres.17700.1

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

Then, we evaluated how the respondents perceived the potential use of the Clinical Research Repository to identify data objects belonging to clinical studies, support secondary analysis of clinical trial data, support systematic reviews/meta-analysis and support re-analysis of clinical trial data (Figure 11).

Figure 11: Evaluating the perceived potential of the Clinical Research Repository for: A) identifying data objects belonging to clinical studies; B) supporting secondary analysis of clinical trial data; C) supporting systematic reviews/meta-analyses; D) supporting re-analysis of clinical trial data.

All 13 respondents agree to some extent (5 strongly agree, 7 agree, 1 somewhat agrees) that the repository is useful for identifying data objects belonging to clinical studies. Similarly, all the respondents agree to some extent (4 strongly agree, 9 agree) that the repository is useful for supporting secondary analysis of clinical trial data. All 13 respondents agree also to some extent (3 strongly agree, 6 agree, 4 somewhat agree) that the repository is useful for supporting systematic reviews/meta-analyses. Finally, all 13 respondents agree to some extent (4 strongly agree, 6 agree, 3 somewhat agree) that the repository is useful for supporting the re-analysis of clinical trial data.

The respondents were also asked to answer whether they can identify other applications that the repository could support in the future. Some of the feedback received:

- *"In the middle to long term the repository could also help to inform overall practice and policy on data sharing and re-use; it could also contribute to the overall knowledge and technology solutions for data sharing, storing, re-use, etc."*

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

- *"I see the potential for integration of the Clinical Research Repository into the data access procedure for the EHDS2. It could also play a role in approaches dealing with generalised-evidence synthesis using different study types (e.g. clinical trials, observational studies, cohorts, epidemiologic studies)."*

2.7 Evaluating overall usability according to UMUX-LITE

Following the UMUX-LITE questionnaire the respondents were asked to evaluate the following statements:

1) The Clinical Research Repository's capabilities meet my expectations.
2) The Clinical Research Repository is easy to use.

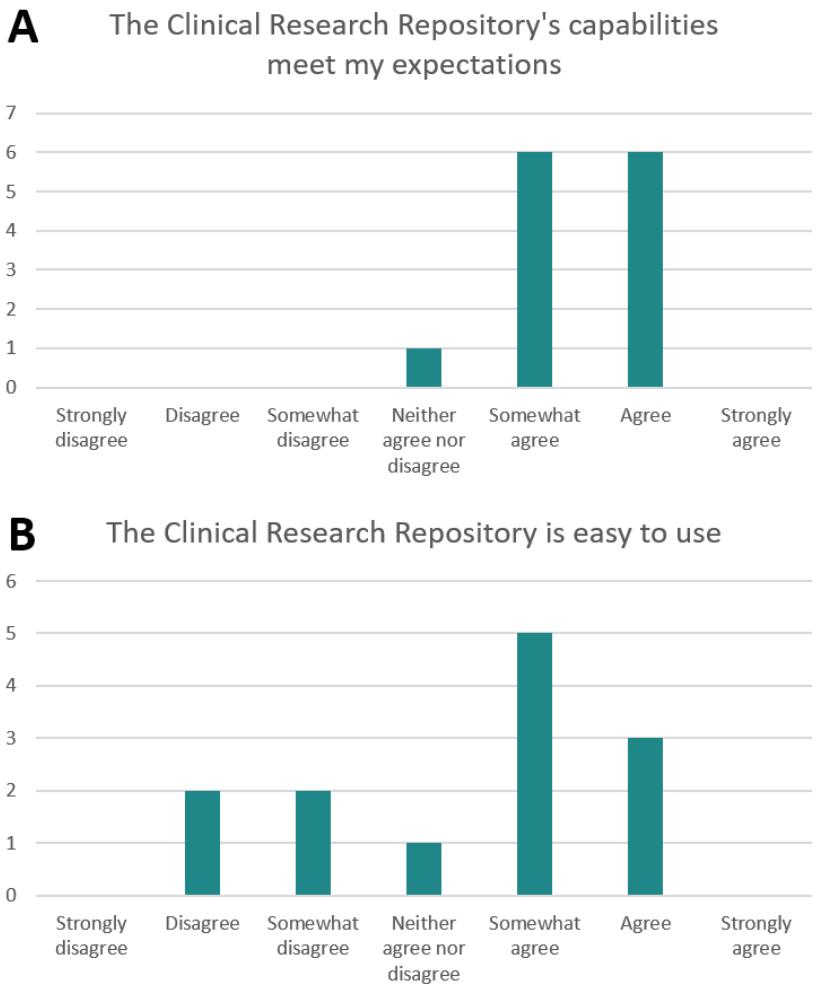


Figure 12: A) Evaluating whether the repository meets the users' needs. B) Evaluating whether the repository is easy to use.

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

Out of the 13 respondents, 12 agree to some extent (6 agree and 6 somewhat agree) that the Clinical Research Repository's capabilities meet their expectations; 1 respondent is neutral (neither agrees nor disagrees). Although this is a quite promising result, what proved to be more troublesome was the difficulty of some respondents in using the repository. More precisely, out of the 13 respondents only 3 agree that the repository is easy to use, 5 somewhat agree, 1 is neutral and neither agrees nor disagrees, 2 somewhat disagree and 2 disagree. This is not necessarily surprising given that this was the first time that "naive users" evaluated the Clinical Research Repository (*alpha version*) and their feedback for improvements was actually sought as part of the whole development process so that we can deliver an improved system to the public (*beta version*).

For the calculation of the UMUX-LITE score of the Clinical Research Repository, the following process was followed:

i. The 2 items of the questionnaire were scored by subtracting one from the respondent response: [respondent score - 1]
ii. The two adjusted scores were added and the sum was divided by 12 (the highest possible score).
iii. The quotient was multiplied by 100.
iv. The results were averaged across the respondents.

Following the above steps, the UMUX-LITE score for the Clinical Research Repository is **64,74**.

Lewis et al.[25] provided a regression equation to predict SUS scores from the two UMUX-LITE items and found that the UMUX-LITE could predict SUS scores with about 99% accuracy. The regression equation is below:

$$SUS\ Score = 0.65 * ((UMUX\text{-}Lite\ Item1 + UMUX\text{-}Lite\ Item2 - 2) * (100/12)) + 22.9$$

According to the Lewis equation, the SUS Score for the Clinical Research Repository is **64,98**.

For the SUS Score to be meaningful, we compared how the obtained value related to the grades proposed by Lewis and Sauro[26] presented in Table 2. The grade of the Clinical Research Repository according to the grading scale is C. For reference, in the 2020 benchmark report of Measuring U, Microsoft Word received a SUS score of about 75 (grade B)[27].

[25] Lewis J.R., Utesch B.S., Maher D.E. (2013). UMUX-LITE - when there's no time for the SUS. In: Proceedings of CHI 2013. 2099-2102. ACM Paris. https://doi.org/10.1145/2470654.2481287
[26] Lewis J.R. and Sauro J. (2018)Item benchmarks for the system usability scale. Journal of Usability Studies, 13, 158-167 (available at: https://uxpajournal.org/item-benchmarks-system-usability-scale-sus/)
[27] https://measuringu.com/sample-sizes-for-sus-ci/

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

Grade	SUS	Percentile range
A+	84,1-100	96-100
A	80,8-84,00	90-95
A-	78,9-80,7	85-89
B+	77,2-78,8	80-84
B	74,1-77,1	70-79
B-	72,6-74,00	65-69
C+	71,1-72,5	60-64
C	65,00-71,00	41-59
C-	62,7-64,9	35-40
D	51,7-62,6	15-34
F	0-51,6	0-14

Table 2: Curved Grading Scale for the SUS.

2.8 Next steps for improving usability and user satisfaction

Taking into consideration the feedback from the respondents to the survey, the following actions have been identified for improving the usability and user satisfaction of the Clinical Research Repository:

- Some terms in the sections “Studies” and “Data objects” confused the respondents (e.g. “CT code”, “Data Sharing Statement”, “Study enrolment”). To address this issue, different actions were proposed like updating the user guide to include a glossary of terms or adding a FAQ for the terms that are not self-explanatory. In addition to updating the user guide, it was proposed to include “information bubbles” where relevant directly in the RMS sections to help users gain time by directly finding the information they need. The user guide should somehow be accessible when the user logs in to the RMS (e.g. added to the left menu, downloaded as a pdf); alternatively, a “Glossary” section can be added to the left menu of the RMS.
- The steps for the Data Transfer Process and the Data Use Process that the user must perform should be clearly stated in the user guide, presented in order, and numbered. It should also be made clear early in the process what kind of applications/tools are needed for going through the Data Transfer Process or the Data Use Process (e.g. Google Authenticator, Google Chrome as a browser) and propose alternatives in the user guide in case the proposed tools are incompatible with the user’s settings.
- At some points, it was difficult for the users to understand what they had to do next in the process. It would be good to include some features in the system informing the user on what has been done successfully and what remains to be done. For example, if the further entry of

This project has received funding from the *European Union’s Horizon 2020 research and innovation programme* under grant agreement No 824087.

information is dependent on a validation step of the repository manager, the user should be clearly informed. For example, users found it weird that without being informed, the added study-data objects would only be displayed in the RMS after validation of the repository manager.
- The branding of the Clinical Research Repository as currently proposed by the ECRIN communications team needs to be improved according to the feedback of the respondents and integrated into the current portals and landing pages of the RMS[28] and TSD[29]. This includes alignment on the naming of the repository across documents and portals. Some respondents raised the fact that we should not call the repository "COVID-19 repository" because it is a repository that can host studies and data objects from any disease. The agreed naming should be applied in all documentation and portals.
- The fields that are mandatory in the "Study" and "Data objects" sections should be clearly marked as such and one should try to keep this number to a reasonable minimum in order to make the processes "lighter" for the users.
- A short online video tutorial could be created for the users demonstrating the "Data Transfer Process" and the "Data Use Process".
- Some respondents reported that it is not clear how to move back from the TSD to the RMS system when some action in TSD has been performed (e.g. data upload). All actions performed in TSD should be visible in the RMS to allow the user an overview. The interface between RMS and TSD needs further improvement in this direction.

3. Maintenance and sustainability

In software engineering "maintenance" refers to the ongoing activities that are necessary to keep a software system functioning properly over time. This might include fixing bugs, addressing security vulnerabilities, and adding new features or functionality based on user feedback. The goal of maintenance is to ensure that the software remains usable and effective.

In a software lifetime, the type of maintenance may vary based on its nature. It may be just a routine maintenance task as some bugs are discovered by users or it may be a large event in itself based on maintenance size or nature. Some types of maintenance are[30]:

- Corrective Maintenance: This includes modifications and updates done in order to correct or fix problems, which are either discovered by users or reported in user error reports.
- Adaptive Maintenance: This includes modifications and updates applied to keep the software product up-to-date and tuned to the ever-changing world of technology and business environment.
- Perfective Maintenance: This includes modifications and updates done in order to keep the software usable over a long period of time. It includes new features and new user requirements for refining the software and improving its reliability and performance.

[28] https://ecrin-rms.org/login
[29] https://covid-19-repo.usit.uio.no/
[30] https://www.tutorialspoint.com/software_engineering/software_maintenance_overview.htm

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

- Preventive Maintenance: This includes modifications and updates to prevent future problems with the software. It aims to attend to problems, which are not significant at this moment but may cause serious issues in the future.

"Sustainability", on the other hand, comes from the latin sustineō, sustinēre meaning "to uphold". In the context of a European project like EOSC-Life "sustainable" means that the project outcome (here the Clinical Research Repository) continues to deliver benefits to the project beneficiaries and / or other constituencies for an extended period after the European Commission's financial assistance has been terminated.

3.1 Requirements for maintenance and sustainability

When it comes to the maintenance of the Clinical Research Repository, we would need to take into account the requirements of the RMS developed by ECRIN and the requirements for the use of the secure server in TSD for the storing of sensitive data. The maintenance and sustainability requirements of the two parts are briefly described below.

RMS requirements

The ECRIN RMS cloud implementation relies on a duplicated environment, useful for continuous development and continuity of service. It utilises a Linux server for the database operations and a Windows server for the user interface. The cost with the current provider is ~400€/month. All the servers have a similar configuration: 32 GB RAM, 880 GB storage. The domain name for the RMS costs 15€/year. The cost of software licences (JetBrains both for front end (WebStorm) and back end (Rider)) used for the development of the RMS is estimated at 1500€/year.

TSD requirements

The secure storage is provided by the TSD at the University of Oslo, assigning a standard "project" to the Clinical Research Repository. The composition of such a "project" is summarised in Table 3. The University of Oslo will be providing in-kind a standard "project" in TSD for the operations of the Clinical Research Repository.

Characteristics of a "project" in TSD	Maintenance considerations
1 TiB storage and backup	additional storage: 450 NOK/year per TiB
1 Windows server VM with 2CPUs and 4GiB RAM	can be changed to 4CPUs and 16GiB RAM, no extra cost
1 Linux VM with 2CPUs and 4GiB RAM	can be changed to 4CPUs and 16GiB RAM, no extra cost

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

Standard Software[31]	Windows: PDF-reader, Kleopatra, PuTTy, 7-zip Linux: Libre-office, R, Emacs, Perl, Python, xdg-open, VLC media player
Access to use the Consent system	
TSD Dataloader	
Access to use "Nettskjema"[32]	

Table 3: Characteristics of a "project" within TSD.

In case more storage is required than the 1 TiB provided within the standard TSD "project" option, this will be assigned with an extra cost of 450 NOK (about 40€) per year per TiB.

Personnel costs

The estimation of personnel costs is very much dependent on the number of studies and data objects stored in the Clinical Research Repository. Table 4 contains personnel costs estimations in the short-term (1-15 studies), mid-term (15-50 studies), and longer-term (>50 studies). The starting time-point for the calculations is marked by the end of the EOSC-Life project (31/08/2023).

Organisation	Short-term sustainability (1-15 studies)	Mid-term sustainability (15-50 studies)	Long-term sustainability (>50 studies)
ECRIN	- *3-5 PM*: For the Repository Manager to support the DTP and DUP. - *8 PM*: For developers to fix bugs after the Clinical Research Repository's public release. Improving functionality following users' requests. - *1-2 PM*: For the legal	- *5-7 PM*: For the Repository Manager to support the DTP and DUP. - *6 PM*: For developers to support the RMS maintenance. - *2-3 PM*: For the	- *7-9 PM*: For the Repository Manager to support the DTP and DUP. - *5 PM*: For developers to support the RMS maintenance. - *3-4 PM*: For the

[31] https://www.uio.no/english/services/it/research/sensitive-data/help/software.html
[32] https://www.uio.no/english/services/it/adm-services/nettskjema/

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

	officer to assist with the contract preparation (DTA, DUA).	legal officer to assist with the contract preparation (DTA, DUA).	legal officer to assist with the contract preparation (DTA, DUA).
UiO (TSD)	- *5 PM*: For developers to fix bugs in the RMS-TSD interface after the Clinical Research Repository's public release. Improving functionality following users' requests. Support for the DTP and DUP.	- *2-3 PM*: For developers to support the RMS-TSD interface maintenance. Support for the DTP and DUP.	- *2-3 PM*: For developers to support the RMS-TSD interface maintenance. Support for the DTP and DUP.

Table 4: Personnel costs estimations for ECRIN and UiO (TSD) in the short-term, mid-term, and long-term.

The above personnel costs estimations refer to the short-term costs of improving the first versions of the Clinical Research Repository according to users' feedback and the longer-term costs of supporting technical maintenance and operations for the Data Transfer Process and the Data Use Process, and the related contracting (DTA, DUA). These estimations do not include the provision of major additional services or functionality other than the one that the system currently supports. It might be desirable in the future to include additional services (e.g. support for data anonymisation, conversion to specific data standards, trainings); the design, development and operations of these are not budgeted above as these are not seen as maintenance costs of the current system but rather costs for extensions.

3.2 Sustainability plan

Nowadays, there are several ways for making online tools sustainable. The authors of this report reviewed different sustainability models used in practice that could be applied for the maintenance and sustainability of the repository:

- **Fee-based models**: Users can be charged a fee for using the tool and its features. In the case of the repository, fees could be implemented per Data Transfer Process (*to data object providers*) and/or per Data Use Process (*to data object secondary users*).
- **Freemium model**: A basic version of the tool could be offered to users for free and charges could apply for additional features/premium services. In the case of the repository, the free version could include the Data Transfer Process and the Data Use Process, while premium services could include anonymisation support, conversion to CDISC standards, training etc.
- **Sponsorship**: Sponsorship could be sought from organisations that would benefit from the operations of the tool. In the case of the repository, this could include the WHO or pharmaceutical/medical device companies. CDISC could also be interested in sponsoring the

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

repository, especially if a premium service is created to support conversion to CDISC standards.

- **Donations**: Allow users to donate to support the ongoing development and maintenance of the tool. Wikipedia is a great example of how users, by contributing an amount of money of their choice can support the maintenance of tools.
- **Grants**: Continue to seek grants to support the development and maintenance of the tool. In the case of the repository, short-term funding has been secured within the BY-COVID[33] and the canSERV[34] projects. Another option here could be to budget the fees for using the Clinical Research Repository in the grant that supports the clinical study/clinical trial.
- **Crowdfunding**: In some cases, crowdfunding platforms are used to raise funds for projects for the "common good". This includes continued development and support of tools.

While these models are being extensively used in practice, the authors of this report believe that in the long-term the best-suited sustainability model for the repository would be **sustainable public infrastructure funding**. This will need extensive discussion with stakeholders such as representatives of the European Commission or the Member States and monitoring of the EOSC and EHDS2 progress to identify opportunities for sustaining (even partly) the repository in the mid-term to long-term.

Currently, ECRIN is the organisation leading the Clinical Research Repository development, with the University of Oslo being the technical partner providing secure storage in TSD. In the short-term, the sustainability of the Clinical Research Repository is guaranteed thanks to funding received in BY-COVID (expected project end September 2024) and canSERV (expected project end August 2025). ECRIN and the University of Oslo, have participated in INFRA-EOSC-01-06 "Trusted environments for sensitive data management in EOSC" to expand the current functionalities of the repository by federating with other real-world data sources.

While these solutions support the short-term sustainability of the Clinical Research Repository, a long-term sustainability plan is still to be elaborated. This is because the public launch of the Clinical Research Repository is still pending. Upon its public launch to the community, ECRIN has engaged to a big communication, dissemination and training campaign to relevant stakeholders through its network and the University of Oslo is willing to contribute in-kind a "TSD standard project" (see Table 3) for the Clinical Research Repository even outside of grant funding. Nevertheless, there remain fixed maintenance expenses (e.g. in terms of hardware) and, more importantly, personnel costs that would need to be covered through different funding sources.

The specific question of the sustainability of this repository enters also the wider debate of *"who should pay the cost for data sharing and reuse"* and especially taking into account the costs that the acknowledged impact of ever-changing rules and regulations in Europe have to sensitive data sharing[35].

[33] https://by-covid.org/
[34] https://www.canserv.eu/
[35] Peloquin D., DiMaio M., Bierer, B. et al. (2020) Disruptive and avoidable: GDPR challenges to secondary research uses of data. Eur J Hum Genet 28, 697–705. https://doi.org/10.1038/s41431-020-0596-x

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

4. Conclusions and next steps

"D14.3 - Report about use and user satisfaction of the COVID-19 repository, including a maintenance and sustainability plan" has successfully contributed to a first evaluation of the Clinical Research Repository by naive users and to the first listing of requirements for maintenance and sustainability.

For the usability and user satisfaction study, 13 volunteers were recruited and invited to join one of the two workshops organised on the 14th (face-to-face) and 15th (online) of March to test the Clinical Research Repository and fill in a survey evaluating different aspects of the system (e.g., the user guide, the landing page, the implementation of the Data Transfer Process). Based on the survey results, a usability score for the repository was calculated following the UMUX-LITE and SUS standards. This score should be looked at with precaution given rather low sample size (n=13). More important than the score itself is the feedback received by the testers which helped us identify and fix some remaining bugs but also revealed some issues regarding the user-friendliness of the system, which will need to be addressed over the following months and for the beta release of the Clinical Research Repository.

For the maintenance and sustainability of the repository, the EOSC-Life WP14 teams calculated expected costs in the short-term, mid-term, and longer-term. The personnel costs around the repository operations are very much dependent on the number of studies and data objects that the repository will manage to attract. Taking into account that the short-term sustainability of the repository has been secured thanks to funding from BY-COVID and canSERV (and potential funding through INFRA-EOSC-01-06), we plan to initiate a series of bilateral discussions with potential sponsors/funders of the repository upon the public launch of its beta version. Some of these discussions will happen within EOSC-Life but are expected to continue beyond the project's duration.

For example, as part of *MS46 2nd Workshop of the Repository Stakeholder Forum* ECRIN and UiO are organising this summer a workshop inviting all relevant stakeholders for the repository; clinical trial sponsors, biostatisticians, medical Research Infrastructures (BBMRI, EATRIS), funders, journal editors, industry representatives, ELSI and technical experts with the aim of introducing the Clinical Research Repository and its functionality to a wider audience of potential users or sponsors/funders of the system.

Abbreviations

AAI	Authentication and Authorization Infrastructure
CDISC	Clinical Data Interchange Standards Consortium
CPU	Central Processing Unit
CSDR	Clinical Study Data Request

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

COVID-19	Coronavirus disease 2019
CT	Controlled Terminology
DOI	Digital Object Identifier
DSS	Data Sharing Statement
DTA	Data Transfer Agreement
DTP	Data Transfer Process
DUA	Data Use Agreement
DUP	Data Use Process
ECRIN	European Clinical Research Infrastructure Network
EHDS2	European Health Data Space for secondary use *(also HealthData@EU)*
EOSC	European Open Science Cloud
EuCos	European Correspondents
EUCS	End-User Computing Satisfaction
FAQ	Frequently Asked Questions
GB	Gigabyte *(=10^9 bytes)*
GiB	Gibibyte *(=2^{30} bytes)*
GDPR	General Data Protection Regulation
IPD	Individual Participant Data
ISO	International Organization for Standardization
MIP	Medical Informatics Platform
MS	Milestone
NOK	Norwegian Krone *(~0,088€)*
OHDSI	Observational Health Data Sciences and Informatics
PM	Person Month

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

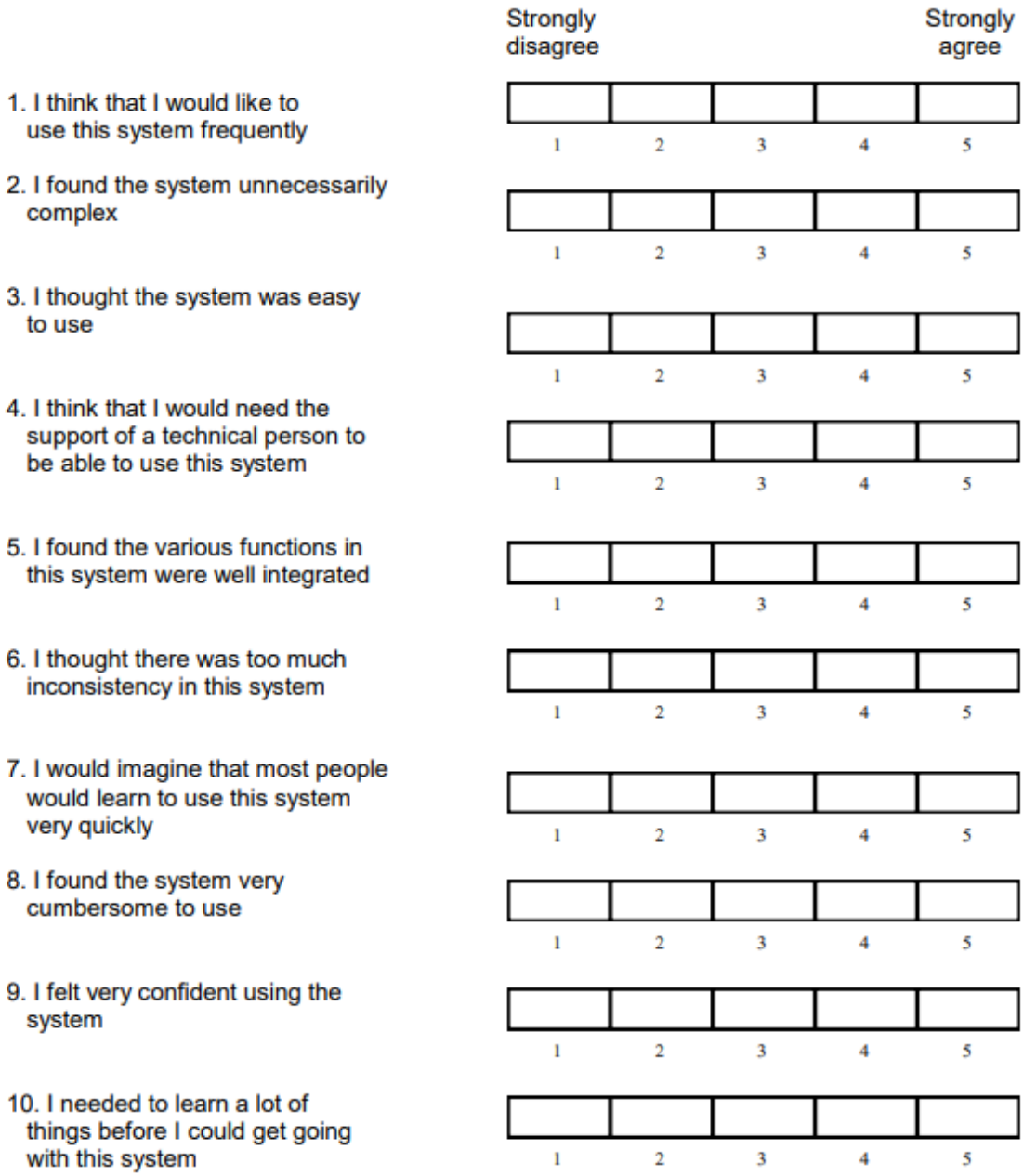
RAM	Random-Access Memory
RMS	Repository Management System
SUS	System Usability Scale
TAM	Technology Acceptance Model
TiB	Tetibyte *(=2^{40} bytes)*
TSD	Services for sensitive data *(a UiO service)*
UiO	University of Oslo
UMUX	Usability Metric for User Experience
URL	Uniform Resource Locator
UX	User Experience
VM	Virtual Machine
WHO	World Health Organisation

Delivery and Schedule

The delivery of *"D14.3-Report about use and user satisfaction of the COVID-19 repository, including a maintenance and sustainability plan"* (initially foreseen in March 2022) has been delayed. This is mainly due to personnel turn-over and some difficulties that ECRIN encountered to recruit software engineers to work on the EOSC-Life project. In addition, the development of the Repository Management System has taken longer than initially estimated, which has delayed all WP14 deliverables and milestones. Finally, some difficulties were also encountered for recruiting naïve users to participate in the usability and user satisfaction study; thus a special thanks is due to the 13 volunteers.

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

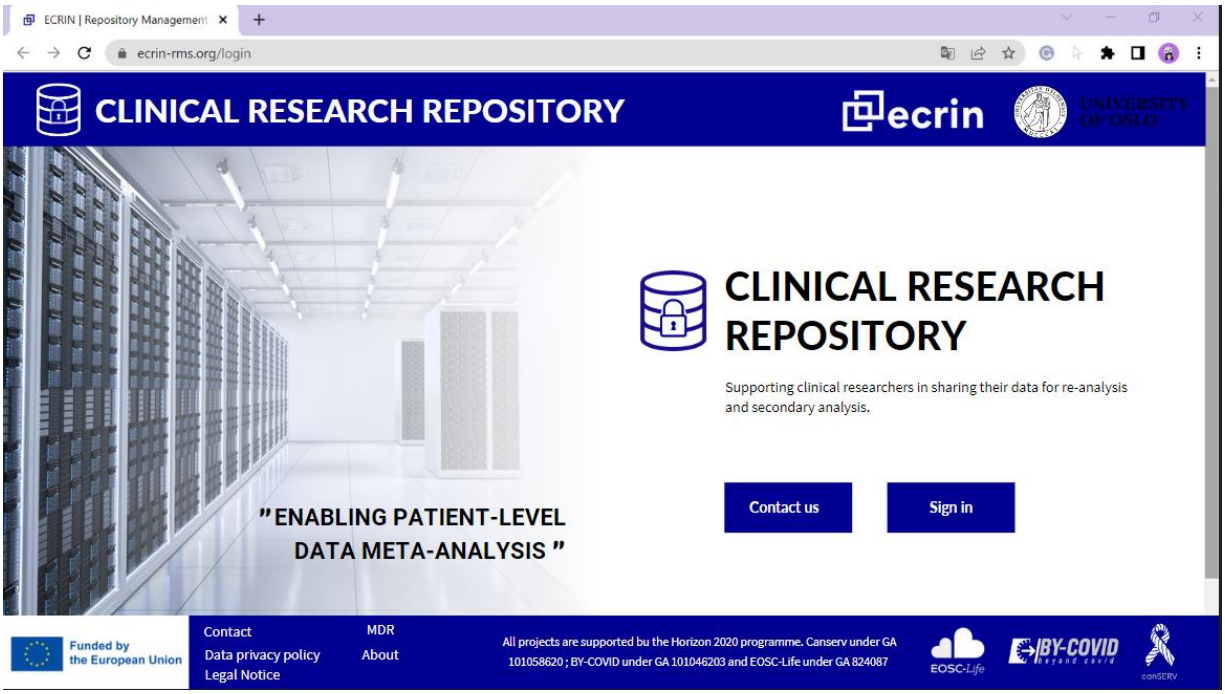
Appendix A: SUS questionnaire



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.

Appendix B: Landing page of the Clinical Research Repository

Design provided by the ECRIN communications team for implementation.



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 824087.