EU Horizon 2020 Research & Innovation Program
Digital transformation in Health and Care
SC1-DTH-06-2020
Grant Agreement No. 101016496

SimCardioTest - Simulation of Cardiac Devices & Drugs for in-silico Testing and Certification

Technical Report

D8.7- Data Management Plan v2

Work Package 8 (WP 8)

Project Coordination & Management

Task Lead: Inria, France

WP Lead: Inria, France

PUBLIC

DELIVERABLE INFORMATION

Deliverable number	D8.7 Data Management Plan
Deliverable title	Data Management Plan
Description	Document to detail the data management plan, in particular personal data protection.
Lead authors	Michèle Barbier
Contributors	Maxime Sermesant, Beatriz Trenor, Delphine Feuerstein
Due date	M6
Submission date	28 June 2021
Comments	

Document history			
Date	**Version**	**Author(s)**	**Comments**
10 May 2021	V0.1	M. Barbier	Draft
25 May 2021	V1	M. Barbier	Version 1
22 June 2021	V1.1	M. Barbier	Revised version
17 November 2022	V2	M. Barbier	Format edition
30 April 2024	V3	M. Barbier	Update p5, p12-13
26 June 2024	V3.1	ExCom members	Validation

Table of Contents

EXECUTIVE SUMMARY

This deliverable is the second version of the Data Management Plan (M42). It describes the nature of the data collected and used during the SimCardioTest project.

This report covers data management procedures applicable to all Consortium members of the SimCardioTest Project. Adherence to this plan will be enforced by the SimCardioTest Program Management Office.

This report summarizes the data management plan under three focus areas:

1 Data Types & Classes: There is ownership and liability markings associated with any data content used by the SimCardioTest Consortium. Certain data contents are owned and generated by the Consortium while others are collected from the outside by Consortium members. When such data is released to the public, we must address our approach for compliance and conformance to data confidentiality and privacy laws. Section 2 addresses our approach for proper labelling of each data content used in the project. Specific attention is paid to data embedded with Personally Identifiable Information.

2 Data Repositories: The SimCardioTest Consortium is composed of ten partner Consortiums that are geographically dispersed across Europe and USA. Therefore, SimCardioTest data is stored and used at each partner site. Data may be stored at a local, physical site or virtually in the cloud. Certain SimCardioTest data is also exchanged among partners using available communication and networking resources. Section 3 covers SimCardioTest data storage facilities and the structure of data management authority for access control and safeguarding of stored data.

3 Data Management Guidelines: As each SimCardioTest Consortium member has full knowledge of appropriate marking for each data content and locations where it is stored, the only remaining topics to address for data management are rights, responsibilities and dissemination procedures. Section 4 provides explicit, step-by-step guidelines for authority and rights for proper labelling of each SimCardioTest data content as well as distributed responsibilities for data management at each partner site. It also stipulates communication protocols for secure exchange of SimCardioTest data across partner sites when sensitive data such as Personally Identifiable Information is embedded. Finally, a flowchart describes procedural steps for public release of SimCardioTest data.

1. Introduction

This report provides details of protocols and procedures used by the SimCardioTest Project Consortium in managing all data types and classes associated with activities undertaken by all Consortium members. An additional document closely related to this report is D 9.1: Protection of Personal Data (POPD) concerns ethical issues associated with handling data content collected by the Consortium.

This report on Data Management Plan provides the general framework for data handling (data collection, storage, security and protection, retention, destruction and dissemination). Specific topics being addressed are:

- Characterization of data types and classes associated with the project (Section 2)
- Details of physical sites for data repository (Section 3)
- Role of the Data Manager and execution plan for data management (Section 4)

1.1. Updated version

This report is an updated version of the deliverable D8.7 submitted in June 2021. It is based on a questionnaire sent to partners to assess the changes in two years. No major changes have been noted, the precise location of ORDP data is being partially updated ($2.2, Table 3, p12-13)

- **For UC1:** MPC designed a Database consisting of 3D heart models derived from scans and a Demonstrator of the 3D printed phantom based on human scans (note that both are confidential and not shared with the consortium).
- **For UC3:** the major change is the release of an *open access* database of cardiac models with drug interactions. Users will be able to investigate the effects of different drugs at different concentrations on the electrophysiology and mechanical properties of ventricular tissue. This open database is available on our website and on GitHub.
- The major challenge is for **UC2**: discussion are ongoing on the rights to publish a database on meshes (UPF). Another issue is the size of this database, for which Zenodo presents some limits. We are currently waiting to resolve this issue before releasing the updated version of the deliverable.

1.2. Legal framework

The objective of this section is to describe the legal framework of the project linked with the data policy. In this sense this section reviews the IPR framework established by the SimCardioTest Grant Agreement (GA), the SimCardioTest Consortium Agreement (CA) and the overarching European legislation regarding data protection.

Grant agreement and Consortium agreement

SimCardioTest data include different types of data hence this legal framework on data management derives from the articles of the GA and the CA. Of them, main points to take into consideration are the following:

- **The GA defines in article 26 the ownership of the project results**. These results are any tangible or intangible output of the actions including data and information and are owned by

the institution that generates them. In case the results are generated by two or more institution rules defined in article 26.2 of the GA and article 8.2 of SimCardioTest CA must be applied.

- Unless it goes against the legitimate interest of the beneficiaries **the results must be disseminated by disclosing them**. This means that the beneficiaries have the right to protect the results in case the institution plans to protect or exploit the results.

- As defined in the article 8.4 of the CA 30 calendar days prior to any publication notice must be given to the other parties. Objections must be raised in writing within 15 calendar days after the receipt of the notice. The publication will be permitted if no objection is made within this time limit.

Personal data protection in SimCardioTest - EU regulations

- The project is committed to observe the EU Data Protection Directive 94/46/EC when dealing with data and the forthcoming European General Data Protection Regulation.

- SimCardioTest has a whole Work Package dedicated to Ethical Issues (WP9) to manage all ethical issues that arise during the project lifetime. Personal health data have been collected in pseudonymised way and not shared between partners.
- IPR's and exploitation rights (including data) are extensively covered in D8.3 "Plan for managing Knowledge and Intellectual Property" (M7, M42).

1.3. Open Access in H2020

SimCardioTest is participating in the Research Data Pilot and will open access to datasets with multiple purposes (Ref GA article 29.3)

In addition, the workflow related to publications is clearly stated in the CA (article 8). The same instructions are provided in SimCardioTest Project Quality Plan (D8.2). Most of these datasets can be reused, but the terms of use will be properly specified. It is expected that during the project certain contents can be generated and they can be of interest for researchers. With this purpose, the project identified which datasets can be made public, and which only re-used by SimCardioTest members. This is treated on a case-by-case basis. Data research will be published following the same approach as for publications. This means that data will be accessible within the 6 months after its publication.

1.4. Related document

SimCardioTest Grant Agreement establishes the rights and obligations of beneficiaries towards the European Commission.

SimCardioTest Consortium Agreement sets out the legal basis for the share of rights, obligations and responsibilities related to the implementation of the project among the beneficiaries themselves.

SimCardioTest D8.2: Project Quality Plan defines the general approach for SimCardioTest quality assurance and the procedures to be followed for project governance, partner communication and meetings, reports and deliverable production (confidential).

SimCardioTest D8.3: Intellectual Property Management Plan, describes SimCardioTest Document to detail the IP management.

SimCardioTest D9.1 POPD Ethics requirements (confidential) reports ethical issues associated with handling data content collected by the Consortium.

1.5. Governance

SimCardioTest Consortium has in place an effective data governance which is ensured by four important entities in addition to the Parties of SimCardioTest:

- **The Governing Board (GB)** as the ultimate decision-making body of the consortium; to be consulted for any contractual decision, and patenting/licencing/publishing process.

- **The Executive Comittee (ExCom) is** the supervisory body for the execution of the project: *inter-alia* Day-to-day decision-making; Assessment of the results obtained and possible minor adaptation of the future work; Deciding on technical reviews; Approving and releasing the WP deliverables

- The **Project Coordinator (PC)** as the legal entity acting as the intermediary between the Parties and the Funding Authority. The Project Coordinator shall, in addition to its responsibilities as a Party, perform the tasks assigned to it as described in the Grant Agreement and this Consortium Agreement. The Project Coordinator chairs the Governing Board meetings.

- The **Exploitation Manager (EM**), who is responsible, together with the Parties willing to exploit the Results for (*i*) the tracking of the Results of the Project their potential exploitation and (*ii*) ensuring the authorship and knowledge property during the dissemination and communication obtained during the Project. This role is carried out by the Project Manager with the support of Inria's Innovation Transfer and Partnerships Department (STIP).

○ Local Repository Management Guidelines

All SimCardioTest data stored on a computer, an external storage unit or paper document are interpreted as data at a Local Repository at a partner's site where this information content resides. Data may reside with a researcher, on-site IT facility or an external storage unit via a third-party cloud service. Local management of SimCardioTest data is the sole responsibility of the partner or its designated data management authority. As long as each partner adheres to the industry's best practices for access control, data integrity and security measures, no further action is required from the Consortium.
It is expected that appropriate data maintenance procedures are in place at each local LO-R site such that whereabouts of all stored SimCardioTest data can be tracked and accounted for in case further distribution, retention or destruction actions need to be taken. When partner A needs to share SimCardioTest data (RE marking) with Partner B, the two involved parties agree upon the most suitable and secure option for data delivery, especially when the data is embedded with Personal Identifiable Information.

2. Data Types & Classes

2.1 Data Types

Every SimCardioTest data are categorized into different subtypes: *i)* Collected Data (C-Data), *ii)* Internal Model Data (M-Data) and *iii)* Generated Data (G-Data).

i. Collected Data (C-Data) (internal or external)
   a. *External* collected data originates from a source outside of the Consortium such as public media, a private source, a governmental or regulatory body or prior work product from a Consortium member.
   b. *Internal* collected data encompasses all data collected by the Consortium such as test data

ii. Internal Model Data (M-Data) applies to all non-collected data such as documents, technical diagrams and descriptions, software code, algorithms and models developed by any Consortium member as a work product of the SimCardioTest project. Any prior work is labelled as External Collected Data.

iii. Generated Data (G-Data) is any Collected or Model Data that has undergone processing by any Consortium member. For example, when a collected data item is used in a SimCardioTest report, the resulting report document becomes a generated data item. Likewise, when an open-source software code developed by a third party (i.e. external collected data) is modified and used by a Consortium member to generate a new software tool, this tool and its embedded code are identified as generated data.

Three sub-categories exist for generated data (GD):

   a. *Collected* GD: Collected data is minimally processed or used as-is
   b. *Derived* GD: Collected data is extensively processed
   c. *Non-derived* GD: Generated by processing Internal Model data

All data types and their associations are shown in Figure 1. Data Types is to associate each data item with ownership and liability. Table 1 shows the assigned ownership and liability level for each data type.

Since no clinical trial involving humans is planned within this project, no personal data will be generated. All the data used for training and testing are already pseudonymised with no access to the corresponding table.

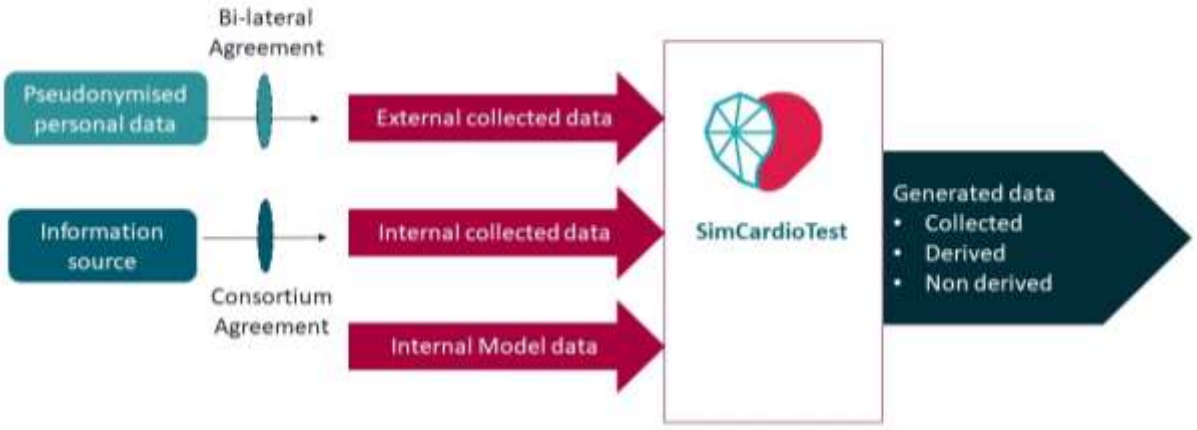


***Figure1**: Summary of Data Type Categories (Generated, Collected and Internal Model Data)*

As evident from the short descriptions of collected and generated data, the objective of categorizing Data Types is to associate each data item with ownership and liability.

- If entity A is the original source of data content X, then it is the primary owner of X.
- If entity B acquires data X from entity A, then B is a secondary owner.
- If data X is further transferred to entity C, it is also classified as a secondary owner
- Whether B has the right to use data X or further dissemination to C is governed by the data rights agreement ("contract") negotiated between A and B.

SimCardioTest Consortium is composed of researchers located within Europe. The European Union General Data Protection Regulation (EU GDPR) policies are applicable to all members and external participants that exchange data with the Consortium.

- If entity B uses data X without prior approval from or agreement with entity A, then B may be liable for violating EU or other national data usage laws and policies.
    - For SimCardioTest, it is a secondary owner for External Collected Data but the primary owner for Internal Collected Data.
    - For internal Model Data, the Consortium member that generated the original content is the primary owner.
    - Care should be taken when dealing with either External or Internal Collected Data since Personally Identifiable Information may be embedded in such data. Protection of Personal Data by the SimCardioTest member is discussed thoroughly in D9.1 and will be expanded further in this report in regards to sharing Personal data internally across partner sites.

***Table 1**: Data Type Categories and associated Ownership & Liability Levels*

Data Type	Ownership	Liability to Data Law Violation
External Collected Data	Secondary	Low
Internal Collected Data	Primary	Low
Internal Model Data	Primary	Low
Generated Data - Collected	Secondary	Low (w/ contract), Medium (w/o contract)
Generated Data - Derived	Primary	Low (w/ contract), Medium (w/o contract)
Generated Data – Non-Derived	Primary	Low

Liability level "Medium" implies the resulting violation penalty may be a fine or some other form of compensation.

In Section 4, we outline our data management plan to adhere to policies and best practices that prevent and protect any Consortium member from handling data content with High liability level, if any.

2.2 Data Classes

Just as Data Types are embedded with *ownership* and *liability* watermarks, Data Classes are associated with read, write and distribution *privileges* and *responsibilities*. The Consortium use the following Data Class labels:

1) PU — public or ORDP (Open Research Data Pilot) - Data is public, releasable (no restriction for distribution or dissemination)
2) CO — confidential - Data is private, restricted to internal use by Consortium members only
3) RE — restricted - Data is private, restricted to a subset of Consortium members
4) PP — project partner - Data is private, restricted to Consortium members and external participants such as the Advisory Board members.

An illustration of SimCardioTest data types and associated privilege levels is presented as a Venn diagram in Figure 2 and examples of classes and types of SimCardioTest data are provided in Table 2. Note that additional privileges can be stipulated for each data class. For example, Data Class RE ("restricted") can be further classified with additional restrictions. A data content (marked RE) between partners A & B may stipulate Read-Only privilege or limited to person X from A and persons Y from B will almost always have a marking stating "No Further Dissemination or Distribution" without prior approval from the Consortium.

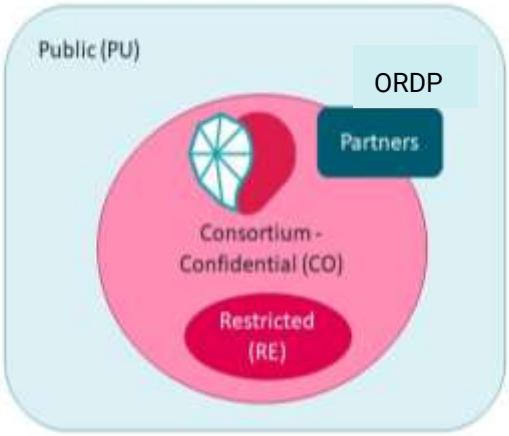


Figure 2: *Venn diagram illustrating various Data Classes*

SimCardioTest will release as ***Open Access (PU / ORDP)*** the non-confidential project developments and outcomes to ensure impact and fostering further research on cardiac CM&S for in-silico clinical trials. At the same time, this will be balanced by keeping some developments private to enhance competitiveness of the companies involved in the consortium and potential further commercial exploitation.

The **Data Class RE** option allows collaboration among partners and protection of intellectual property that may result from such activity. Since all SimCardioTest data content must belong to any of the four classes listed above, there exists no data content that is shared with external participants but restricted to a subset of Consortium members. That is, any data shared with an external entity must be visible to all Consortium members. This clause does not prevent any Consortium member from collaborating with an external entity on project-directed or related research areas but the information must be disclosed to all Consortium members. The concern raised here is protection of data privilege rights and data usage liability. In summary, no data class exists between a subset of SimCardioTest Consortium members and external participants. Likewise, all Public releases (PU) are consented by all Consortium members and all Open Research Data Pilot.

Table 2: *SimCardioTest example of data type and classes*

Label (SCT: SimCardioTest)	**Example**	**Type**	**Data detail**	**Data classes**	**Sensitivity**
SCT- Report	Data Management Plan	C- internal Data	Information on data, storage & sharing	ORDP	None
SCT-Database	Database of fluid simulations in virtual population	G-Data (non-derived)	Database of meshes and associated quantities (velocities, pressures) from virtual population	ORDP	None

SCT- Demo	Demonstrator of simulation of drug effects in cellular and 3D models	M-Data	Demonstrator of simulation of drug effects in cellular and 3D models	PU	low
SCT-RE-Demo	Demonstrator 3D printed Phantom based on data of coronary and ventricular meshes	G-Data (derived	Database of coronary tree and cardiac ventricles meshes	RE	low
SCT-RE-Report	Verification and validation reports on existing pacing leads	M-Data C-Data	Bench and pre-clinical V&V reports to abide by all applicable standards	RE	High
SCT- CO-Database	Database of atrial meshes	G-Data (derived)	Database of meshes derived from the atrial images	CO	low

Open Research Data Pilot Access

The Open Research Data Pilot of the European Commission enables open access and reuse of research data generated by Horizon 2020 projects.

Non-confidential data and models of drugs/devices will be released as Open Access, including the virtual populations generated for each case. The developed CM&S for each case will have cloud-based APIs integrated within the in-silico clinical platform of SimCardioTest (Table 3). SimCardioTest Consortium will

- Release Open Source tools for in-silico models for each use case;
- Release reusable virtual populations generated for each use case;
- Cloud-based APIs for each use case, integrated within the in-silico clinical trial platform.

The sensitivity of these data is null.

Table 3: *SimCardioTest open access Data - Open Research Data Pilot access and metadata, update 2024*

Title	details	
Data Management Plan	– To detail the data management plan, Information on data, storage & sharing in particular personal data protection – data from all partners	– Released in 18 March 2021 by Inria – open access – located on Inria server (SimcardioTest website), Sophia-Antipolis
Database of virtual human cardiac models with drug interaction	– Database of detailed cardiac models for drug interaction simulation (Cardiac meshes) – data from UPV, SRL	– released in January 2023, by SRL – open access – located on GitHub and on our website

Database of fluid simulations in virtual population	– Database of meshes and associated quantities (velocities, pressures) from virtual population – Demonstrator of computer models for device planning. – data from UPF, SRL, UBx	– under discussion regarding the issue of the size of this database, for which Zenodo presents some limits – open access
Report on the standardised models for drugs	– Description of models for drug interaction – Data from SRL, UPV, EXC	– Available in D4.1- Report on the standardised models for use case 3 on drugs – located on Inria server (SimCardioTest website), Sophia-Antipolis
Description of biomarkers for drug safety, efficacy & cardiotoxicity	– Description of defined biomarkers. – data from EXC, SRL, UPV	– Available in D4.4 - Description of biomarkers for drug safety, efficacy & cardiotoxicity – To be released in July 2024 located on Inria server (SimcardioTest website), Sophia-Antipolis
Report on the in-silico trial of drug efficacy & cardiotoxicity	– Description of simulation results from in-silico trials on drugs. – UPV, Inria, IST	– Available in D5.3 - Report on the in-silico trial of drug efficacy & cardiotoxicity – To be released in July 2024 on Inria server (SimcardioTest website), Sophia-Antipolis
Database of device configurations	– Virtual library atrial meshes and device configurations – data from UBx, Inria, BSC, UPF	– To be released – under discussion regarding the issue of the size of this database, for which Zenodo presents some limits
Report on the standardized models for LAAO	– Description of models for LAAO devices – data from UPF – January 2022 by Oscar	- Available in July 2024 on Inria server (SimcardioTest website), Sophia-Antipolis

2.3 Size of the data

We do not envision very large datasets in SimCardioTest but the exact size is difficult to predict. Data on models can be kilobytes, data on meshes Mbytes, and output from simulation Mbytes to Gbytes. Finally, research papers and code files are usually not exceeding few megabytes. The exception is the size of the database of atrial meshes which is several terabytes. Discussion are ongoing to find the best location for storage.

3. Data Repository Sites

SimCardioTest consortium is composed of ten research Consortiums headquartered in the United States (1), Spain (2), Italy (1), Belgium (1), Norway (1) and France (4). Therefore, SimCardioTest data exist at each partner facility. To simplify data management responsibilities across sites, we define the terms

- Local Repository (LO-R): Physical data storage at a partner site
- Consortium Repository (CO-R): Physical data storage at the Program Management site
- External Repository (EX-R): Physical data storage at a third-party site

1. **SimCardioTest data at Local Repository** (LO-R) at site A is managed by the partner responsible for site A. For example, SimCardioTest data at the local repository by Simula is managed by its Information Technology (IT) or other data management authority. Whether SimCardioTest data at Simula is stored physically at a local machine or online via a Cloud/Web service is outside the scope of this report.

2. **SimCardioTest data at Consortium Repository** (CO-R) is managed by Inria in Sophia Antipolis, France. The Consortium Repository can be viewed as the umbrella site for the majority of SimCardioTest data. Since INRIA is both a partner/Coordinator and the headquarters of the Project Management site, it also houses a Local Repository (LO-R). Both the public website and internal CO-R of the SimCardioTest Consortium are maintained by the Project Manager. After project completion, maintenance of the public website and CO-R will be carried out by a SimCardioTest Consortium member from Inria or delegated to an IT staff member at Inria. In essence, this staff member becomes a SimCardioTest Consortium member and must abide by all data privacy, usage and distribution policies stipulated by the Consortium.

3. Certain SimCardioTest data may be stored at a **third-party External Repository** (EX-R) such as a web-based collaboration tool, data editing tool, electronic mail or a Cloud data storage service such as GitHub. All SimCardioTest data stored on a computer, an external storage unit or paper document is interpreted as data at a Local Repository at the site where this information content resides.

4. Data Management Plan

In Section 3, we cover SimCardioTest data in terms of Type (ownership and liability) and Class (privilege and responsibility). In Section 4 we provide physical and virtual locations for hosting and storage of SimCardioTest data. In this section, we cover the following topics:

- Procedures & Safeguards against handling High liability data sets
- Authority for Data Class markings
- Local Repository Management Guidelines
- Consortium Repository Management Guidelines
- External Repository Management Guidelines

We also address communication methods and security measures used in the transfer of SimCardioTest data from one site to the other, such as from a local site to an external web host and vice versa.

4.1. Procedures & Safeguards against handling of High liability data sets:

Any SimCardioTest data with High liability pertains to collected or derived-generated data with embedded personally identifiable information, with contract —i.e. usage with approval from the primary owner(s). The Coordinator and the Project Management Office are not gatekeepers for data collected by a member of a partner Consortium. It is the responsibility of each Partner or the designated Principal Investigator to comply with EU GDPR and national data privacy laws and policies in collecting both external and internal data. Whenever there is a personal information concern for data collection methods used or from examination of the data content itself, the data set in question will be investigated and resolved collectively by all partners of the Consortium.

4.2. Data sets with Medium liability,

This category applies to non-personal data collected and generated data without an agreement. For example, an image or video from a website may not contain any personal identifiable information but it is the property of the primary owner. Utilization of such data for internal research activity does not violate any data privacy law but public release of generated data requires approval or contract with the primary owner. This topic is addressed further in the next section. Any public release of SimCardioTest data with PU marking requires approval from all partners.

4.3. Authority for Data Class Markings

Any SimCardioTest **data with PU** (Public) marking must be examined and approved for public release by all partners. This is not an issue for official deliverables with PU markings to the European Commission which are already well scrutinized by the entire Consortium. Here, the main concern is other types of releases: data content on the Consortium public website, presentations at public venues, technical manuscripts and publications at conferences and journals, etc. In short, PU markings are under the authority of the Consortium (i.e. all partners) as a whole. No partner or Consortium member will mark a SimCardioTest data content with a PU marking unless authorized/ cleared by the Consortium.

Any SimCardioTest data shared internally within the Consortium is already accepted as **data with CO** (Confidential) marking. No explicit marking of such data is required.

Data with PP (Project Partners marking) applies to all SimCardioTest data exchanged with external partners who are affiliated or associated with the SimCardioTest project. The partner that interacts with external partners will tag a PP marking to all SimCardioTest private data being shared. If such data includes collected or generated data owned by another partner, the first partner will request approval for PP release of such information. All PP releases should have a marking that stipulates against further dissemination without prior approval. In general, as a common courtesy, all PP data releases will be peer-reviewed by the Governing Board.

Staff and associates at partner sites —who are not Consortium members of SimCardioTest— are not generally considered as external partners. Three options exist for dissemination of private SimCardioTest information to such entities: 1) they join the Consortium as members, 2) they join as external participants or 3) case-by-case approval by the Consortium.

Any **data with RE** (Restricted) marking refers to data that is not shared with the entire Consortium. This does not imply that certain partners are withholding information from the rest of the Consortium. In certain situations, such as joint collaboration during technical manuscript preparation for a

conference or information disclosure for a patent application, it is likely that not all partners will contribute to such an effort. For RE class data, no explicit marking is expected and it is left to the collaborating partners to find suitable options for data exchange.

4.4. Consortium Repository Management Guidelines

The purpose of the Consortium/confidential Repository (CO-R) is not to serve as a backup storage for LO-R, nor as a single source for all SimCardioTest data. Instead, CO-R is used to host the Consortium's public website as well as bookkeeping of data classes with markings PU (Public), CO (Consortium) and PP (Project Partners)**. It does not host RE-marked data**.

It also stores program management information (i.e. private SimCardioTest data, P-Data) such as Consortium member contact information, meeting schedules, meeting notes, milestone and delivery date charts, etc. In short, it is the one-stop internal information portal for all Consortium members. The Consortium Repository is managed by a local IT authority at INRIA in Sophia Antipolis, France under the supervision of the Coordinator and the Project Manager.

It is anticipated that CO-R will remain operational after the completion of the project. The designated IT authority can be a SimCardioTest Consortium member, a non-member staff or an external service provider as long as it adheres to data usage and privacy policies stipulated by the Consortium. Similar to the data manager at LO-R, the CO-R IT authority must adhere to the industry's best practices for access control, data integrity and security measures such that no further action is required from the Consortium. It is expected that appropriate data maintenance procedures are in place at the CO-R site such that whereabouts of all stored SimCardioTest data can be tracked and accounted for in case further distribution, retention or destruction actions need to be taken. Under this model, all SimCardioTest members are provided Read-Write access to all SimCardioTest data (with PU, CO and PP markings) stored at CO-R.

4.5. External Repository Management Guidelines

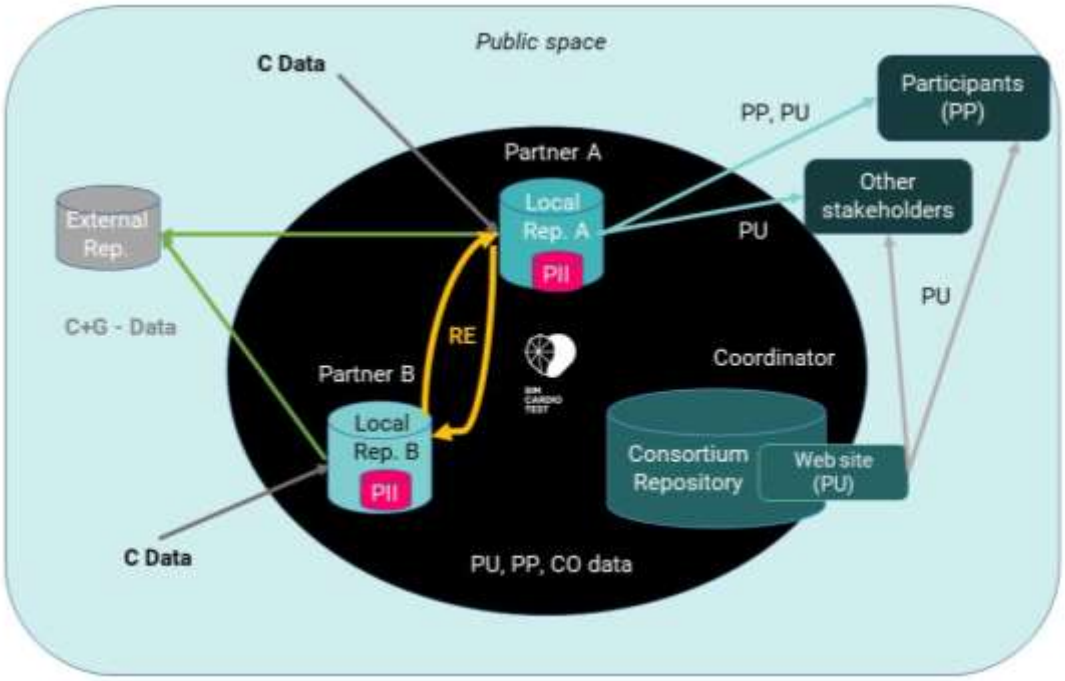
Compared to LO-R and CO-R, an external repository is a third-party site (typically web-based) with lax or no mechanisms for access control, data integrity and information security. However, these sites may be useful for coordination and collaboration purposes among Consortium members. They can be used to exchange messages, collaborate software code development, etc. In general, usage of such third-party sites will not be managed or overseen by the Consortium as long as the following data management policy is adhered to: no uploading or sharing of Personal identifiable information - embedded data or private SimCardioTest data.

4.6. Communication Methods and Security Measures for Data Transfer

The only SimCardioTest data types with high sensitivity and liability concerns are those with embedded Personal Identifiable Information. Safeguarding of private SimCardioTest data such as Intellectual Property is also important at the partner level but not managed by the Consortium.
There is no transfer of personal identifiable information data between partners.

Figure 3 provides a look-quick guide to the overall data management plan.

- Any Partner can collect both Internal and External Data and store them at his/her managed LO-R site. If Personal Identifiable Information is embedded in the data, necessary safeguards are used for access control and information security.
- SimCardioTest data resides at different locations: Local Repositories (Partner sites), Consortium Repository (Coordinator site), External Repositories (third-party sites), external Project Partners, Consortium Public Website (hosted at Inria) and Other (any entity not associated with SimCardioTest).
- PU (Public release) marked data only is available at the Public Website and Other. PU data can be delivered from CO-R (to Public Website), Public Website (to Other and Project Partners) or directly from a Partner LO-R site.
- PP (Project Participants) marked data is shared from a Partner to external Participants.
- Two or more Partners can share Collected and Generated Data (C-G-Data) at a third-party, repository (EX-R) as long as no personal information embedded data or private P-Data is released.
- Two or more Partners can exchange RE (Restricted) marked data among themselves.
- The Consortium Repository stores PU, PP and CO marked data as well as Private Data (P-Data) such as Program Management information. Note that P-Data is non-technical. CO-R is also used to maintain information content for the Public Website.
- Personal Identifiable information embedded data can be stored at any Partner LO-R site without access to the identification table.



C-Data = Collected Data (internal or External) (C+G)-Data = Collected & Generated Data
PII = Personally Identifiable Information P-Data = Private Data (e.g. Consortium member info)
PU = Public, **PP** = Project Partner, **CO** = Consortium, **RE** = Restricted (sub-Consortium)

Figure 3: *SimCardioTest Data Recipients, Repositories and Data Flows*

4.7. Summary of SimCardioTest Data Class Markings

Class	Marking (Explicit or Implicit)	Remarks
PU	**Depends**. See note 1) below.	For public release of technical and programming data
PP	**Yes**, explicit. See note 2) below.	For sharing of project data with external participants only
RE	**No**, implicit	For sharing between a subset of partners
CO	**No**, implicit	For sharing among all partners

The figure 4 summarises the flowchart for Clearance Procedures for Public Release of SimCardioTest Data.

Notes: All official public deliverables to the European Commission (and available via Consortium public website) are explicitly marked with PU to avoid any confusion for dissemination rights. Similar approach can be used for SimCardioTest presentations made at public venues such as trade shows and conferences. The PU marking can be included explicitly: "***SimCardioTest Project Data – Public Release***". In general, a lack of marking can lead to follow-on questions and requests from audience members for distribution and dissemination rights. It should also acknowledge the EU funding "This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016496".

1) PP-marking is: "***SimCardioTest Project Confidential – Limited to Designated External Project Participants Only***". No further distribution without approval from the SimCardioTest Consortium.
2) Regarding journal and conference papers, no PU marking is required since there is implicit understanding that they are already cleared internally for public release. The clearance steps are shown in the flowchart below for the most complicated scenario involving collaboration among three or more partners. For the release of a multi-partner collaborated work product, a conformance check by the Consortium is required since it is the Consortium that is liable for potential violation of data privacy laws. The simplest scenario is where a partner prepares a manuscript (with no partner collaboration) for public release. In this case, the sole partner alone assumes liability for violation of data privacy laws. As a common courtesy, CO-marked distribution of a draft manuscript is recommended.

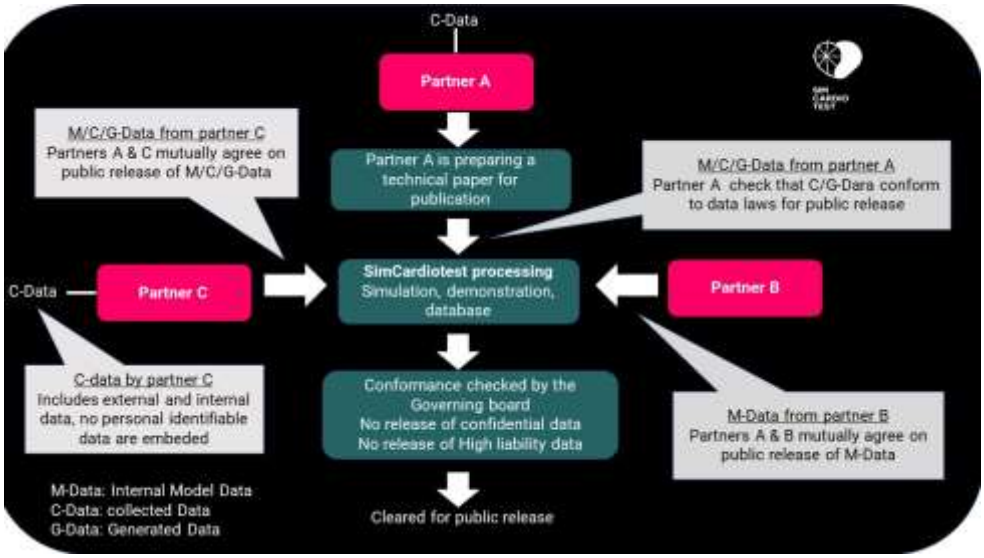


Figure 4*: Flowchart for Clearance Procedures for Public Release of SimCardioTest Data*