PRIMAGE

Medical imaging
Artificial intelligence
Childhood cancer research

D10.2 – Ethics review plan

Project Full Title: *PRedictive In-silico Multiscale Analytics to support cancer personalized diaGnosis and prognosis, Empowered by imaging biomarkers*
Project acronym: *PRIMAGE*
Project type: Horizon 2020 | RIA (Topic SC1-DTH-07-2018)
Grant agreement no: 826494

Document information:

Deliverable no.	**D10.2**
Title	Ethics review plan
Work Package	WP10
Dissemination level	Public
Nature	Report
Responsible partner	HULAFE
Main Contact person	Dr. Luis Martí-Bonmati (marti_lui@gva.es)
Deliverable due submission date	28/02/2019
Deliverable actual submission date	28/02/2019

History of changes:

Change explanation	Pages affected	Change made by (Name & surname)	Date

Table of Contents

Table of figures

Disclaimer

The opinions stated in this report reflects the opinions of the authors and not the opinion of the European Commission.

Acknowledgement

This document is a deliverable of the PRIMAGE project. This project has received funding from the European Union´s Horizon2020 research and innovation programme under grant agreement no.: 826494.

Introduction

The scope of this deliverable is to detail the ethical aspects of the procedures to follow for the development of the project PRIMAGE: PRedictive In-silico Multiscale Analytics to support cancer (GA 826494).

Relevant Acronyms

AEPD	Agencia Española de Protección de Datos
CCRI	St. Anna kinderkrebsforschung - Children Cancer Research Institute
DIPG	Diffuse Intrinsic Pontine Glioma
EC	European Commission
GDPR	General Data Protection Regulation
HULAFE	Hospital Universitario y Politécnico La Fe
LOPD	Ley de Ordenación y Protección de Datos
NB	Neuroblastoma
PI	Principal Investigator
SIOPE	European Society for Paediatric Oncology
UKOELN	Klinikum der Universitaet Zu Koeln
UNIPI	Universita di Pisa

1. General Considerations

The key aspect of PRIMAGE project is to develop an open cloud-based platform with the integration of data and in-silico models to support decision making in the clinical management of two paediatric cancers, Neuroblastoma (NB), the most frequent solid cancer of early childhood, and the Diffuse Intrinsic Pontine Glioma (DIPG) the leading cause of brain tumour-related death in children.

The research conducted will respect fundamental ethical principles, including those reflected in the Charter of Fundamental Rights of the European Union (7th December 2000) and worldwide: Dignity, Freedom, Equality, Solidarity, Citizens' Rights and Justice. Additionally, the project will be in accordance with the European Human Rights Convention, especially with regard to privacy and autonomy.

The present document summarizes the general Ethical Plan of the PRIMAGE project, which will be the basis for the deliverables required in WP11 (Ethics Requirements). In the deliverables D11.1, D11.2, D11.3 and D11.4, the three clinical partners of the project (HULAFE, CCRI and KOELN) will reflect the assurance of compliance with the respective national policies as well as those within the European Union regarding ethics, data protection and good clinical practices.

1.1 Project Objectives

PRIMAGE objectives are classified in 8 major outcomes, each one requiring of the achievement of specific scientific and technical objectives.

Objective1: To design a Decision Support System (DSS) for cancer management with advanced functionality and usability, under a user-centric approach guided by our clinical partners (European Key Opinion Leaders in Paediatrics Oncology), aligned with their current work flows and aimed to conquer trust and gain acceptability by the clinical practitioners.

Objective 2: To implement PRIMAGE DDS as a cloud-based platform, with an hybrid model of use of (i) open public cloud (based on EOSC services) and (ii) private clouds in order to deliver a platform to (i) be used by the scientific community to promote use of de-identified clinical curated data in Open Science and also (ii) to be suitable for future commercial exploitation under PaaS business models, as an scalable, safe and cost-effective cloud infrastructure.

Objective 3: To establish a symbiosis with major European initiatives for shared clinical data repositories for Neuroblastoma (c.a. 2600 patients) and DIPG (c.a. 700 patients), both managed by SIOPE, where (i) retrospective clinical Big Data is used for biomarkers and in-silico model training and validation and (ii) PRIMAGE delivers methodologies to facilitate the extraction, de-identification and quality control of imaging and clinical data from hospital databases, thus contributing to feeding new clinical cases into the existing shared repositories. Additionally, this project will develop a strategy for embedding imaging biobanks into wider biobanks networks (e.g. BBMRI-ERIC) and facilitate data cross-linking with other biorepositories.

Objective 4: To validate new imaging biomarkers of MR, PET/CT and MIBG for NB and DIPG and develop diagnostic models based on imaging and cross-linked to established biological biomarker panels for each disease for supporting decision by clinical practitioners.

Objective 5: To deliver in-silico models of solid tumour growth for NB and DIPG using multiscale simulation frameworks to couple model at subcellular scale, to cell, to tissue and to complete organ models, enabling assessment of a radiotherapy and chemotherapy treatment on tumour progression for a given patient under.

Objective 6: To promote the usability of the new generated knowledge from the in-silico models for virtual diagnosis (O4) and tumour growth (O5), by (i) delivering visualisation methods for high dimensional data analysis, and (ii) implementing Artificial Intelligence methodologies to generate responses to the most relevant 5 Clinical End Points for NB and DIPG (as preliminary defined in page 19).

Objective 7: To integrate a functional prototype of PRIMAGE cloud-based platform, offering predictive tools to assist management of NB and DIPG paediatric cancers, from diagnosis to prognosis, therapies choice and

treatment follow up, based on the use of novel imaging biomarkers (O4), tumour growth models (O5), advanced visualisation of predictions with weighted confidence scores and responses to a set of CEPs, obtained from the use of patient-specific in-silico models in combination with AI analytics on relevant patient clusters.

Objective 8: To validate PRIMAGE platform performance in multicentre non-interventional studies for NB (approx. 150 patients) and for DIPG (approx. 75 patients) in Spain, Germany and Austria, involving clinical multidisciplinary clinical teams in the use of the Decision Support System and, its assessment according to Key Performance Indicators.

PRIMAGE project involves an observational study on different types of data such as clinical information, medical imaging, pathology, histology and genomics from NB and DIPG patients. All data to be studied in the project will be properly anonymized and codified in the participating hospital with a new id specific to PRIMAGE project. Dissociated data will be securely stored in the central cloud platform of PRIMAGE project, to which only project participants will have access.

This project will undertake early **non-interventional clinical in-silico validation** of the cloud-based decision support platform system, in the application context of data from newly diagnosed NB and DIPG patients. A detailed plan of each clinical study will be prepared as part of in T8.1 to follow the standard validation procedure by the Ethical Committee at the participant hospitals in Spain, Germany and Austria. These validation studies are aimed to a preliminary assessment of the performance, relevance and usability of the proposed in-silico tools.

The ethical-sensitivity of this project is related to the proposed research objectives, which imply R&D works where clinical data needs to be gathered by the main clinical partners (Hospitals) involved in the project (HULAFE, CCRI and UKOELN) and later shared through the PRIMAGE platforms, where other partners will have access for data analysis.

As with regards to the potential impact of PRIMAGE's research, we have evaluated the ethical implications of the intended use or potential use of our research outcomes and have determined that the proposed research outcomes do not contravene ethical fundamental rights. Our research objectives are related **to a novel in-silico-based diagnosis and treatment system for oncology diseases**, which will have to be in compliance with ethical and safety requirements in order to achieve the CE marking as a medical device class I in the future. Thus, the ethical compliance of the PRIMAGE system is guaranteed. PRIMAGE system design and testing is guided by GLP and GCP criteria, is subject of review by ethical experts from the Research Ethics Committee of HULAFE.

More precisely, the ethics issues involve:

- The research with human beings.
- The identification, handling and storage of information of persons sampled by population/medical characteristics;
- The management, storage, analysis and distribution of individual sensitive and/or private data.
- The conditions to make such results widely available while continuing protecting individual's information by data anonymization.
- The issues of possible misuse of results produced.

This leads to consider a number of ethical issues and relevant legislations and international texts to be followed.

First, this project confirms that the proposed research does not involve:

- Research activity aimed to study human response to a specific treatment.
- Research activity aimed to collect biological samples.
- Research activity aimed at human cloning for reproductive purposes.
- Research activity intended to modify the genetic heritage of human beings, which could make such changes heritable.
- Research activity intended to create human embryos solely for the purpose of research or for the purpose of stem cell procurement, including by means of somatic cell nuclear transfer.

-

Even though PRIMAGE project is an observational study and an in-silico clinical validation simulation, with regards to humans protection, the studies from which the data is extracted will be conducted in compliance with the rules of the ICH-GCP (International Conference on Harmonization for Good Clinical Practice and with the Council of Europe Convention on Human Rights and Biomedicine (Additional Protocols on Biomedical Research-2005).

Important EU directives relevant with the patient recruitment and clinical trial having a significant impact in the project are the following:

a) The Charter of fundamental Rights of the EU.

b) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

c) International Ethical Guidelines for Health-related Research Involving Humans Prepared by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO).

The project is also aware with relevant International conventions and declarations such as:

a) The principals and precepts of the Declaration of Helsinki (Ethical Principles for Medical Research Involving Human Subjects, 1964 and later amendments).
b) Convention of the Council of Europe on Human Rights and Biomedicine signed in Oviedo on 4 April 1997.
c) Universal Declaration on the human genome and human rights adopted by UNESCO.

In addition, the project will comply with the internal legislation in the different host institutions and their relevant national legislation. Clinical data will be provided by the following partners Hospital Universitario y Politécnico La Fe (Spain); St. Anna kinderkrebsforschung-Children Cancer Research Institute (Austria) and Klinikum Der Universitaet Zu Koeln (Germany). As stated in Legal Basis for Ethics in H2020 Reg. No 1290/2013: article 23, “Participants shall comply with national legislation, regulations and ethical rules in the countries where the action will be carried out. Where appropriate, participants shall seek the approval of the relevant national or local ethics committees prior to the start of the action.” The proposal will be submitted to obtain the approval of the Ethics committee. The follow up of these committees will be made available to the European Commission as a demonstration that all levels of safety, ethics and security have been considered in the study.

2. Main Ethical Issues

According to the EC ethical self-assessment checklist (European Commission, 2015a) (See Section 2: Main Ethical Issues), we have identified the main ethical aspects involving sensitive ethical issues. We have grouped the potential issues accordingly to different security level of patient’s information:

a.- Research on human beings: Children.

b.- Handling of human cells / tissues.

c.- Handling of human personal data.

European Commission, 2015a aims at assisting applicants at getting their proposal “Ethics-Ready”—i.e. to identify and deal correctly with any ethical issues that may arise in the course of a project in compliance with H2020 Ethical Standards and Guidelines

2.1 Research on human beings: children

This project will undertake studies for the training and validation of PRIMAGE platform. Data will be derived from cases affected by Neuroblastoma and DIPG, both diseases affecting minors.

Physical interventions on the study participants (non-invasive, not collecting biological samples)

Clinical validation of PRIMAGE platform will run in parallel with patient´s way and will not imply any interventional measure on the participant patients. The available data will be used to construct, fine-tuning and validate the diagnosis and treatment functionalities of PRIMAGE for these diseases. PRIMAGE will act as a parallel method to deeply process all the information and enhance the clinical management of future patients with more information. For the registry, patients whose data will be used in an anonymized way for the training and validation of the platform will undergo the standard of care physical interventions for diagnosis and treatments of DIPG and NB as in normal clinical practice (such as imaging, histology, and genetic tests). Minors (with age-adapted information) and their legal guardians are usually informed about the diagnostic and therapeutic procedures to follow and their consent is asked according to GCP.

2.2 Human cells / tissues

PRIMAGE does not involve any collection of biological samples within the project. Only information such as digital data from biology/pathology studies included in the normal clinical practice will be used in the study.

2.3 Protection of personal data

PRIMAGE aims at R&D and validate new tools for a better diagnosis and treatment management of oncology diseases, especially and within the scope of the project: Neuroblastoma and DIPG diseases. For this purpose, we will have to process and analyse different types of personal data (imaging, pathology, histology, genomics) during the project.

Legal framework

The procedures in PRIMAGE for data collection, storage, protection, retention and destruction will comply with national and EU legislation. The procedures will be implemented according to law: The General Data Protection Regulation 2016/679 (GDPR) is a European Union (EU) regulation, which replaces the Data Protection Directive 95/46/EC.

According the 2016/679 GDPR on the protection of individuals with regard to the processing of personal data and on the free movement of such data, specify that "any person whose identity can be determined directly or indirectly, in particular by reference to an identification number or one or more specific elements characteristic of his physical, physiological, psychological, economic, cultural or social identity shall be identifiable".

This project is being performed by experienced groups of researchers who already have a proven track record in performing this type of studies with the highest standards. Their ability to comply with applicable ethical rules and to undertake such studies across national boundaries has been demonstrated by the successful participation of in many international studies. These experienced groups will continue to support and promote the adoption of the current ethical conventions amongst groups that are less familiar with the implications and practical application of these regulations.

PRIMAGE approach: Dissociated data

Based on the experience and know-how of members of our consortium already dealing with this sort of data, we have prepared a clear approach for data management during the PRIMAGE project, based on the principle of working with “dissociated data”, instead of “personal data”.

In this sense, the provisions governing the protection of personal data come to the conclusion that the person affected will not be determinable when their identification requires a disproportionate effort that is sufficient to dissuade the person who accesses the data from the identification of the person to whom the same is concerned.

As the cloud servers of PRIMAGE will be located in the European Union, we have studied in details the Spanish regulation: The Spanish Data Protection Agency (Agencia Española de Protección de Datos -AEPD-) considers that in order for a dissociation procedure to be considered sufficient for the purposes of the LOPD, "it will be necessary for the application of such procedure to be impossible to associate the data that is available to a specific subject". The AEPD has been pointing out that this will require that there is no possibility, even remote, that, through the use, prior, contemporaneous or later of any means (computer process, program, system tool, etc.), the information concerning those affected by the processing of data, that is in possession of the consultant, can reveal its identity.

In this regard, PRIMAGE procedures will follow the following specifications related to the management of dissociated data:

- The software incorporated in PRIMAGE’s platform to eliminate the metadata of DICOM images and any other sample or biomarker does not allow any PRIMAGE’s member at any time and in any way to have access to personal information. Specifically, the fields ‘Patient Name’, ‘Patient Id’, ‘Study Id’, ‘Patient Birthdate’ are removed from the all the imaging headers in the user computer before uploading the data to the PRIMAGE DB.

- The medical professional should not provide for any reason to PRIMAGE patient information that allows PRIMAGE to know the patient's identity. In particular, the only data that PRIMAGE and the medical professional must share, must be the data itself (without metadata containing personal data of the patient) and the image identification code (not the patient) that PRIMAGE must generate before sending the image to the DB. This code should only allow the identification of the data (i.e. an image), but not the patient (it is the responsibility of the medical professional to relate the code of the image to the patient).

All the data processed in PRIMAGE database is anonymised. Therefore, no personal data is stored nor processed in the PRIMAGE system.

PRIMAGE project will use as main sources the retrospective data at a European level, for NB and for DIPG, coordinated by their respective European networks. The formal procedure to request access to data has been initiated and the responsible entities for the management of the respective registries are partners of PRIMAGE:

- DIPG Registry: This cases and imaging dataset has clinical, biological and centrally reviewed radiology data of patients with DIPG, both in- and outside clinical trials, for c.a. 700 anonymised patients.
- SIOPEN-r-net Registry: NB specific. NB specific registry, managed by SIOPEN (International society of paediatric oncology European neuroblastoma research group). This registry has clinical, biological and imaging data of high risk NB patients, for c.a. 2.500 patients, participants in academia-promoted clinical trials.

This project will also use stored data on DIPG and NB from our three clinical partners, which have not yet been incorporated in SIOPE-DIPG (and will be as a result of this project) and for NB patients that were not participant in the SIOPEN clinical trials.

The three Paediatric Oncology Units participants in PRIMAGE are national reference centres for NB in Spain, Germany and Austria, managing the clinical files for these nations. Finally, also the incorporation of data from the Pisa oncological imaging biobank (managed by partner UNIPI) and from the Valencia imaging biobank (managed by partner HULAFE-GIBI) will be considered as part of this task, as both resources are made available to this project.

3. Conflict of Interest

There are not conflicts of interest to be declared.

4. Ethical Management

The PI and his team are a group of senior health care professionals with large experience and proven competence in the management of ethical research issues, so all of them are committed to carry out in the correct way all ethical issues that are contemplate in the project.

5. Ethical Plan

This section addresses all the aspects related to ethical assurance activities that will be undertaken in the context of the PRIMAGE project. In other words, this section describes the provisions established in the project to ensure that any PRIMAGE-related research activity will comply with the applicable ethical requirements/principles. The section is organized as such:

• **Section 5.1 Introduction: Ethical assurance strategy for the PRIMAGE project**. This section introduces the reader to the overall ethical assurance strategy chosen for the PRIMAGE project.

• **Section 5.2 PRIMAGE data collection plan & Handling procedures**.

• **Section 5.3 Relevant Ethical Requirements for the PRIMAGE Project.** This section determines the relevant ethical issues that need to be defined for the PRIMAGE project. The section does so by comparing the data collection plan against the EC ethical guidance (European Commission, 2015a).

5.1 Introduction: Ethical assurance strategy for the PRIMAGE project

The use of Big Data is core to PRIMAGE project´s aims of generating new knowledge from advance in-silico tools. The use of clinical data in PRIMAGE will be undertaken under the strictest administrative and contractual procedures to ensure legal and ethical compliance.

Herein is described PRIMAGE ethical assurance strategy:

1. INITIAL PROJECT-LEVEL ETHICAL APPROVAL

PRIMAGE deals with Real World Data (RWD) from NB and DIPG patients. To start the project, it is mandatory the acquisition of consent from the Ethical Committees from the different clinical partners providing clinical information to extract relevant patients' data from their local repositories. This consent provides an

initial check of the ethical worthiness of the project, and authorize PRIMAGE to start its data collection undertakings.

The PRIMAGE partners' institutions providing data on patients' clinical information are:

- Hospital Universitario y Politécnico La Fe (HULAFE)
- St. Anna Kinderkrebsforschung- Childres´s Cancer Research Institute (CCRI) (as hosting partner of SIOPEN registry),
- Klinikum der Universitaet Zu Koeln (UKOELN)
- Universita di Pisa (UNIPI).

The procedure followed for the initial ethical approval of the PRIMAGE data collection plan and the data handling procedures presented in this deliverable is as following:

1. A Project is submitted to the Ethical Committee of PRIMAGE project Coordinator, HULAFE. In Annex I is available a template of this document.

2. Once the approval is granted to PRIMAGE Coordinator, the final approved Project Plan will be submitted to the respective Ethical Committees of the other institutions. Currently, the project has been submitted to the Ethical Committee of HULAFE and approval is pending.

3. To add more evidence to the classification of the PRIMAGE project as a non-interventional study, the Project has also been submitted to The Spanish Agency of Medicines and Medical Devices.

4. Release of the approvals submission to the Ethical Committee. After having the approval of the research Ethics Committee, data collection activities in the context of the PRIMAGE project will be initiated.

To be compliance with the General Data Protection Regulation (GDPR), a Project Plan will be also submitted to the Data Protection Officer (DPO).

We expected the Ethics Committee to waive the requirement of an **informed consent** from the patients based on the International Ethical Guidelines for Health-related Research Involving Humans Prepared by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO) (2016) in which it is stated that:

> "*A research ethics committee may approve a modification or waiver of informed consent to research if:*
> - *the research would not be feasible or practicable to carry out without the waiver or modification;*
> - *the research has important social value; and*
> - *the research poses no more than minimal risks to participants.",*

PRIMAGE clinical partners and related institutions should be waived of the informed consent as the PRIMAGE project:

- could not be carried out in most NB and DIPG patients as they are not alive;
- has important social value as its outcome will be a decision support tool which will help on a more accurate diagnosis, prognosis and therapies follow up in these diseases.
- only implies in-silico data management and analysis, the research having no interventional activity and no risks to participants.

According to these requirements, informed consent forms and information sheet should not be compulsory in projects based on data collection for quality issues from healthcare reports, such as HULAFE and other health research institutions. A certification that the project accomplishes with the National Law on Data Protection (LOPD) will be provided instead.

2. CONTINUOUS ETHICAL ASSURANCE

Ethical assurance means have been defined specifically for the PRIMAGE project based on consideration of relevant EC ethics guidelines (European Commission, 2015a).

In the course of the project, the continuous application of the PRIMAGE protocol will be monitored by the PRIMAGE Ethical Manager, who has been designated by the PRIMAGE Steering Committee to monitor the ethical worthiness of the project.

It will be also designated a Data Manager in each Institution providing patient data to ensure that the data collection procedures meet the Good Clinical Practice (GCP) standards.

5.2 PRIMAGE Data Collection Plan & Handling procedures

PRIMAGE envisages the collection of Real World Data (RWD) from NB and DIPG patients. To do so, previously the respective institutions providing the data will have to have the approval from their ethics Committees.

Relevant RWD will be accessed from the different data repositories, PACS and biobanks available for the project (Pisa University Biobank, HULAFE's Biobank, UKOELN, DIPG Registry, SIOPEN Neuroblastoma Registry). All this data is already anonymised. Furthermore, all the clinical partners and biobanks owners (HULAFE, CCRI, UKOELN and UNIPI) have protocols in place for full compliance with the EU and their national Data Protection Laws, as part of their daily routine.

Therefore, patients' information are completely anonymized and dissociated before any manipulation by the members of PRIMAGE consortium. Therefore, any compilation of Patient Identifications, Names or Study identifiers will remain at the hospital-institution side and will be blinded to all the PRIMAGE's team at the centralized repository. No personal sensitive data will be accessed and/or manipulated by PRIMAGE's personnel.

Hospital centres have in place all necessary measurements to ensure a responsible data management approach is taken for data collection and processing, and that appropriate control mechanisms are implemented and closely supervised. All the Hospital centres in the project have a lot of experience in running clinical trials. They use protocols and command chains are in full compliance with their national Data Protection Laws. A Data Manager will be nominated on each clinical partner to ensure that the data is ethically handled.

The procedures for data collection, storage, protection, retention and destruction will comply with national and EU legislation (Regulation 2016/679 GDPR, General Data Protection Regulation). The procedures will be implemented according to law.

Particularly, in Spain, the Spanish Data Protection Agency (*Agencia Española de Protección de Datos,* AEPD) considers that in order for a dissociation procedure to be considered sufficient for the purposes of the LOPD, *"it will be necessary for the application of such procedure to be impossible to associate the data that is available to a specific subject*". The AEPD has been pointing out that this will require that there is no possibility, even remote, that, through the use, prior, contemporaneous or later of any means (computer process, program, system tool, etc.), the information concerning those affected by the processing of data, that is in possession of the consultant, can reveal its identity.

Accordingly, in order to understand that the dissociation has been performed correctly, it is necessary that it is not possible to identify the patient. The AEPD has stated that, in order for a dissociation procedure to be considered sufficient for the purpose of the LOPD to understand that personal data are not being processed, "it will be necessary for the application of said procedure to be impossible to associate a certain data with a specific subject".

In this sense, the provisions governing the protection of personal data come to the conclusion that the person affected will not be determinable when their identification requires a disproportionate effort that is sufficient to dissuade the person who accesses the data from the identification of the person to whom the same is concerned.

Datasets used for PRIMAGE platform testing:

- **Imaging Data**

Imaging data represents the highest challenge in terms of storage and processing. In PRIMAGE data repositories, for each patient, imaging data is linkable to their available anonymized biological, pathological and genetics. The use of common metadata frameworks and image analysis techniques for automated data annotation for each image is proposed to generate common repositories.

- **Genetics and other molecular data**

This project uses exiting knowledge on biological biomarkers, currently on clinical use or at advanced clinical validation stage. This type of data will be used in combination with related imaging and clinical data, facilitating the multidisciplinary Big Data analytics.

- **Clinical Data**

Use of natural language processing tools for automated extraction of relevant pathological data, including data on patient characteristics and phenotypes will be extracted from the EHRs. Data will then be structured, curated and stored.

Currently, repositories for clinical data are very fragmented. This is owing to the lack of use of common data filing structures for the different data types. Aiming to ovoid this, in PRIMAGE project, partners will implement:

(i) Protocols for image acquisition and processing, for standardisation of data formats for different data types;
(ii) Tools to streamline the processes of data extraction, field mapping, pseudomization, annotation, curation and secured routing;
(iii) A common repository, based on open-cloud infrastructure with data querying based on semantic annotation.

Clinical data repositories

The use of clinical data (imaging, clinical, genetic data) will be done in two phases:

Phase 1: Compilation of clinical data for PRIMAGE training and Knowlegde Extraction and Building (in the Big Data domain)

▪ For computational simulation developments: including testing of in-silico models for tumour growth, testing of advanced visualization solutions, identification and testing of imaging biomarkers and training of predictive models for CEPs.

▪ For testing of the integrated PRIMAGE decision support platform: extensive lab testing of the platform will be undertaken using datasets from 2010 to December 2020, initially split from the curated clinical datasets, to ensure these data are only used in the platform testing (and not in the developments).

The following table summarizes PRIMAGE use of clinical data for Phase 1:

Tabla 1. PRIMAGE Clinical data sources

Source and responsible entity	Characteristics	Use in PRIMAGE
European shared clinical data registries for NB and DIPG		
The SIOPE- DIGPG Registry **Responsible entity:** SIOPE *(PRIMAGE partner)* **Access approval procedure**: Data access authorisation via formal request procedure to be issued by PRIMAGE Project Coordinator	Sample: c.a. **700 anonymised DIPG patients** from all EU nations, both in- and outside clinical trials. Data Type: Diagnostic and follow-up MRI scans linked to e-data transmittal form (eDTF) incl demographics, medical history and physical exam at time of diagnosis, results from radiological, results from pathological review (if available), treatment (incl. radiotherapy, chemotherapy, surgery and supportive), clinical data, and last known status of the patient. Data input current procedure: Local hospitals submit imaging data (centrally reviewed) and fill the eDTF.	PRIMAGE will ·process data for >200 patients for imaging biomarker identification ·data extraction on the eDTFs and use curated data for machine learning based clustering.
The European Neuroblastoma Registry **Responsible entity:** SIOPEN-r-net and CCRI *(PRIMAGE partner)* as coordinating entity **Access approval procedure**: Data access authorisation via formal request procedure to be issued by PRIMAGE Coordinator	Sample: c.a. **2000 high risk** and **c.a. 600 low and Intermediate risk** NB patients participants in academia-promoted clinical trials. Since 2001. Data type: Diagnosis and longitudinal data (clinical, follow-up and biology data for all patients registered in SIOPEN-r-net). Data input procedure: The local hospitals submit imaging data and clinical data in their local forms and languages.	·process data for >200 high risk and >200 low & inter. Risk patients, for imaging biomarker. · use extraction techniques on structured & unstructured clinical data, to be used on machine learning based clustering.
Clinical partners registries at HULAFE-POU, CCRI and UKOELN PRIMAGE clinical partners will make additional retrospective clinical data available, on DIPG (data not yet incorporated in SIOPE-DIPG) and on NB (corresponding to patients that were not participant in the SIOPEN clinical trials).		
HULAFE-POU, CCRI & UKOELN They are the **national reference clinical centres for NB in Spain, Austria and Germany** respectively, with patients´ registries for more than 15 years. **Access approval procedure**: PI for each clinical partner will manage formal approval requests to our hospitals´ legal and ethical departments.	Type of data for NB, includes: ·Imaging: Magnetic Resonance (MR) and 131I-MIBG scintigraphy examinations in DICOM. ·Histological biopsy (if available). ·Complete molecular biology studies according to SIOPEN, in blood, urine and bone marrow. ·Genetic: Complete genetic studies according to SIOPEN (NGS, FISH). ·Electronic Health Records (EHRs): Patient profile. Prescribed treatment. Type of data for DIPG includes: ·Imaging: Magnetic Resonance (MR) and methionine PET/CT examinations in DICOM ·Complete molecular biology studies in blood and urine. CFS liquid biopsy (if available) ·Genetic data (if available) ·EHRs: Patient profile. Prescribed treatment. Complete data is not available for all registries	*Some examples of accessible data:* *For **900 NB** patients HULAFE-POU has imaging data linked to molecular biology, pathology and clinical data.* *For **4000 NB** patients UKOELN has clinical data, for 900 of them also imaging.* *Both have biobank of tumour, bone marrow, blood.*
Imaging biobanks PRIMAGE use of data from Imaging biobanks is mainly devoted to implement and validate protocols to foster their use and contribute to enhance interoperability of imaging biobanks with other biobanks.		
Oncology imaging biobank Managed by **UNIPI**	Imaging data for c.a. 80 NB patients and c.a. 20 DIPG patients from Pisa Univ. hospital	PRIMAGE will use them to test interoperability, their use for imaging biomarker storage, and progress in data linkage
Valencia imaging biobank **BIMCV,** EuBI node for Spain Managed by **HULAFE-GIBI**	The medical imaging data (MRI, CT, PET, MIBG) repository for the healthcare services of the Valencian region (5,1 million population).	

Phase 2: Use of clinical data for PRIMAGE Platform validation (in the Big Data domain)

The clinical data generated from January 2021 to May 2022 will be used for the PRIMAGE platform validation.

For the validation phase it is expected the acquisition of clinical data from 175 NB patients and 75 DIPG patients. Information will be extracted from diagnosis studies for each patient, anonymised and stored, together with other R&D projects in the context of European Open Data research initiatives.

The validation will imply the testing of the platform's usability with the clinicians and ensuring access to complete clinical data, acquired using common protocols, for these cohorts.

The use of data for models´ training, testing and validation will require an extensive curation and quality control. Automated tools will be implemented in this project to streamline the process of extracting, mapping data, controlling the quality, homogenisation / translation / completion of data feeds.

The clinical centres are responsible for fully anonymised data prior to data sharing with the rest of the consortium. Therefore, no transfer of privacy-sensitive data will occur outside the participant organisations responsible for clinical testing. Only fully anonymised data will be shared with the rest of the partners and beyond PRIMAGE consortium.

5.3 Relevant Ethical Requirements for the PRIMAGE Project

In order to identify the main Ethical Requirements that are relevant for the PRIMAGE project, the Data collection plan outlined in the previous section has been checked against the EC ethical self-assessment checklist (European Commission, 2015a) (See Section 2 in this document: Main Ethical Issues).

The Main ethical Issues identified for the PRIMAGE project are related to the following sections of the EC guide:

Section #2: HUMANS

Section #3: HUMAN CELLS / TISSUES

Section #4: PERSONAL DATA

For each section the EC proposes a selection of Ethical Requirements that the research should consider. These requirements are listed in Table 2. The table maps the EC ethical requirement to the initial project-level approvals (on the second column on the table), and the ethical means covered by the PRIMAGE protocol (third column of the table). Such ethical means will be described in the next section (Section 5.4).

Tabla 2. Relevant EC Ethic Requirements for the PRIMAGE project

RELEVANT EC ETHICAL REQUIREMENTS FOR THE PRIMAGE PROJECT	COMMENTS
2.3. Templates of the informed consent/assent forms and information sheets (in language and terms intelligible to the participants) must be kept on file.	N/A The present project will look at NB and DIPG clinical database anonymized information. Hence, only non-interventional studies will be carried out. In this sense and according to Ethical Guidelines for Health-related Research Involving Humans, PRIMAGE clinical partners will request a waiver of the informed consent to their respective Ethics Committees (see 5.1 in this document).
2.8. Details on incidental findings policy must be submitted as a deliverable.	This will be delivered in D11.1
2.9. Copies of opinions/approvals by ethics	This will be delivered in D11.1

committees and/or competent authorities for the research with humans must be submitted as a deliverable.	
2.10. If applicable, for each clinical study, the following documents/information must be submitted as adeliverable (in one package) prior to enrolment of first study subject: (i) Final version of study protocol as submitted to regulators/ethics committee(s), (ii) Registration number of clinical study in a WHO- or ICMJEapproved registry (with the possibility to post results), (iii) Approvals (ethics committees and national competent authority if applicable) required for invitation/enrolment of first subject in at least one clinical centre.	N/A The research does not involve physical intervention on the study's participants.
3.4. In case human cells/tissues are obtained from a biobank, details on the cell/tissue types and on the biobank and access to it must be submitted as a deliverable.	N/A The research does not involve collection of biological samples only patients' data will be extracted and processed.
3.5 . Copies of relevant documents for using, producing or collecting human cells or tissues (e.g., ethics approval, import licence, accreditation/designation/authorisation/licensing) must be kept on file.	N/A The study does not involve collection of biological samples.
4.2 The host institution must confirm that it has appointed a Data Protection Officer (DPO) and the contact details of the DPO are made available to all data subjects involved in the research. For host institutions not required to appoint a DPO under the GDPR a detailed data protection policy for the project must be kept on file.	This will be delivered in D11.4
4.5 The beneficiary must explain why the research data will not be anonymised/ pseudonymised. This must be submitted as a deliverable.	N/A According to PRIMAGE Ethical Plan, all the clinical information used in this project will be anonymized.
4.11 Detailed information on the informed consent procedures in regard to data processing must be kept on file.	The present project will look at NB and DIPG clinical database anonymized information. Hence, only non-interventional studies will be carried out. In this sense, we will request a waiver of the informed consent from the Ethics Committees.
4.12 Templates of the informed consent forms and information sheets (in language and terms intelligible to the participants) must be kept on file.	N/A The present project will look at NB and DIPG clinical database anonymized information. Hence, only non-interventional studies will be carried out. In this sense, we will request a waiver of the informed consent from the Ethics Committees.
4.14 An explicit confirmation that the data used in the project is publicly available and can be freely used for the purposes of the project must be submitted as a deliverable.	This will be delivered in D11.4
4.15 In case of further processing of previously collected personal data, an explicit confirmation	This will be delivered in D11.4

that the beneficiary has lawful basis for the data processing and that the appropriate technical and organizational measures are in place to safeguard the rights of the data subjects must be submitted as a deliverable.	
4.16 The beneficiary must evaluate the ethics risks related to the data processing activities of the project. This includes also an opinion if data protection impact assessment should be conducted under art.35 General Data Protection Regulation 2016/679. The risk evaluation and the opinion must be submitted as a deliverable.	This will be delivered in D11.4