Model Driven Paediatric European Digital Repository

Call identifier: FP7-ICT-2011-9 - **Grant agreement no**: 600932

Thematic Priority: ICT - ICT-2011.5.2: Virtual Physiological Human

Deliverable 1.1

Kick-off Meeting Report

Due date of delivery: 30th April 2013

Actual submission date: May 2nd

Start of the project: 1st March 2013
Ending Date: 28th February 2017

Partner responsible for this deliverable: OPBG

Version: 1.2

Dissemination Level:

Document Classification

Title	Kick-Off Meeting Report
Deliverable	1.1
Reporting Period	March 2013 – February 2014
Authors	OPBG
Work Package	WP1 – Coordination and Project Management
Security	Pu
Nature	Report
Keyword(s)	Kick-Off Meeting

Document History

Name	**Remark**	**Version**	**Date**
M. De Maldè	First Draft	0.1	19th April 2013
M. De Maldè	Second Draft	0.2	23th April 2013

List of Contributors

Name	**Affiliation**
A. Sattanino	Lynkeus
A. Bartoli	Lynkeus
M. De Maldè	Lynkeus
Ludovica Durst	Lynkeus
Callum MacGregor	Lynkeus

List of reviewers

Name	**Affiliation**
E. Morley-Fletcher	Lynkeus
Harry Dimitropoulos	Athena
Tommaso Mansi	SCR
Enrico Bertini	OPBG
Enrico Castelli	OPBG
Clara Malattia	IGG
Melania Manco	OPBG
Sebastien Gaspard	Maat
Frans Steenbrink	Motek

Table of Contents

Introduction

The MD-Paedigree Kick-Off meeting was hosted at the OPBG hospital in Rome on 13th and 14th March 2013.

More than 60 people attended the meeting as representatives of MD-Paedigree's partners.

The Kick-off Meeting has marked a successful first step for effectively starting to implement the project.

Its main objectives were the following:

- To provide an overview of the work that needs to be undertaken in every disease area
- To determine key priorities for the 1st year of the project
- To set the action plan for the next six months period

The non-plenary sessions were split into 5 dedicated working groups, one for each area of the project (Cardiomyopathies, Cardiovascular Disease Risk in Obese Children, Juvenile Idiopathic Arthritis, Neurological and Neuromuscular Diseases, and Infostructure). Each working group was provided with a specific work plan and they all reported back to the Kick-Off Meeting as a whole in the final plenary session. The work plans were delivered to all the participants at the beginning of each separate working session, and are now included in Appendix 2. The same applies to the drafts of the clinical protocols, discussed and edited during the kick-off meeting, and now included in Appendix 1.

Kick Off Meeting Agenda

Wednesday, March 13th		
Morning Session		
Time	**Object**	
10.30-11.00	Welcome address: Giuseppe Profiti (President), Bruno Dallapiccola (Scientific Director), Jacopo De Angelis (Italian National Agency for Promotion of European Research (APRE) - National Contact Point for ICT)	
11.00-11.15	Partners' presentation round: all partners	
11.15-11.30	*The way we were*: Video recollection wrap up of the Health-e-Child and Sim-e-Child projects and of the development of the VPH community: A. Sattanino	
11.30-11.45	**Coffee break**	
11.45-12.15	MD-Paedigree presentation and overview: B. Dallapiccola	
12.15-12.45	Governance structure and calendar of meetings: E. Morley-Fletcher • per Organ • per Disease and Infostructure Areas	
12.45-13.00	Activities' scheduling subdivided by 4 Disease and 1 Infostructure Areas: E. Morley-Fletcher	
13.00-14.00	**Lunch Break**	
Afternoon session		
14.00-14.30	Demonstration of the state-of-the-art of the OPBG's Paediatric Cardiac Digital Repository (PCDR): D. Manset	
	Dedicated Working Groups (Parallel Sessions)	**Lead**
14.30-16.30	1) Cardiomyopathies (WP3, WP8) • Clinical protocols • Clinical use-cases and validation strategies • Data collection goals	OPBG - SAG
14.30-16.30	2) CVD in Obese Children (WP4, WP9) • Clinical protocols • Clinical use-cases and validation strategies • Data collection goals	UCL – SAG
14.30-16.30	3) JIA (WP5, WP10) • Clinical protocols • Clinical use-cases and validation strategies • Data collection goals	IGG – Fraunhofer
14.30-16.30	4)NND (WP6, WP11) • Clinical protocols • Clinical use-cases and validation strategies • Data collection goals	VUmc – MOTEK
14.30-16.30	5) Infostructure (WP12 - WP17) • Prior reusable work • Data input • Access and Interoperability	MAAT
16.30 -17.00	**Coffee Break**	
17.00- 17.30	Project Communication Infrastructure, Web-site and Documents Repository presentation	LYNKEUS
17.30-18.00	Advisory Committees: appointment of members and work programmes • Scientific Committee (Chair D. Comaniciu) • Ethical and Legal Committee (Chair M. Lopez Barahona) • Interoperability Steering Committee (Chair R. Hose) • Users' Board (Chair M. Viceconti)	OPBG

Social Dinner (20.00)

Thursday, March 14th		
Morning Session		
	Dedicated Working Groups (Parallel Sessions)	**Lead**
8.30-10.30	1) Cardiomyopathies (WP3, WP8) • Prior reusable work • User requirements • Planned deliverables • Scheduling and allocation of tasks	OPBG - SAG
8.30-10.30	2) CVD in Obese Children (WP4, WP9) • Prior reusable work • User requirements • Planned deliverables • Scheduling and allocation of task	UCL - SAG
8.30-10.30	3) JIA (WP5, WP10) • Prior reusable work • User requirements • Planned deliverables • Scheduling and allocation of tasks	IGG – Fraunhofer-Utrecht
8.30-10.30	4)NND (WP6, WP11) • Prior reusable work • User requirements • Planned deliverables • Scheduling and allocation of tasks	VUmc - MOTEK
8.30-10.30	5) Infostructure(WP12 - WP17) • Users requirements • User-friendliness • GPUs accelerated computing • Planned deliverables • Scheduling and allocation of tasks Assigned rapporteurs to split and attend the last hour conclusions of the clinical sessions listed above	MAAT
10.30-11.00	**Coffee break**	
11.00-12.00	Continuation of the Five Dedicated Working Groups (Wrap-up, commitment statements, and Infostructure requirements gathering)	
12.00-13.30	Plenary presentation of the Dedicated Working Groups' Conclusions	
13.30-14.30	**Lunch Break**	
Afternoon Session		
14.30-15.00	Action plan for the next 6 months and review of reporting responsibilities: S. Martin - E. Morley-Fletcher	
15.00 -15.15	Plan of dissemination events for 2013 and choice of MD-Paedigree's logo: E. Morley-Fletcher	
15.15-15.45	Clinical Workflows, HTA, Dissemination, Exploitation: E. Morley-Fletcher, G. Pongiglione, K. Stroetmann	
15.45-16.00	Demonstration of the On-line Management Platform functionalities: Lynkeus	
16.00-16.15	Any other business	
16.15-16.30	Wrapping up	

List of participants

N.	Partner	Country	Name	Surname
1	**OPBG**	**Italy**	Bruno	Dallapiccola
			Sonya	Martin
			Nicola	Bergonzi
			Riccardo	Bosco
			Fabrizio	De Benedetti
			Paolo	Tomà
			Giacomo	Pongiglione
			Gabriele	Rinelli
			Anwar	Baban
			Enrico	Castelli
			Enrico	Bertini
			Melania	Manco
			Silvia	Magni Manzoni
			Lorenza	Putignani
			Marcello	Chinali
			Aurelio	Secinaro
			Francesca	Romana Lepri
			Laura	Tanturri
			Maurizio	Petrarca
			Frascarelli	Flaminia
			Benedetta	Leonardi
			Marco	Cirillo
2	**UCL**	**UK**	Andrew	Taylor
			Alex	Jones
			Giovanni	Biglino
3	**IGG**	**Italy**	Clara	Malattia
			Gianmichele	Magnano
			Stefano	Lanni
4	**JHU**	**USA**		
5	**KU Leuven**	**Belgium**	Kaat	Desloovere
			Nathalie	Goemans
6	**VUmc**	**Netherlands**	Jaap	Harlaar
			Marjolen	Van Der Krogt

7	**UMCU (Utrecht)**	**Netherlands**	Berent	Prakken
8	**SAG**	**Germany**	Michel	Suehling
			Tobias	Heimann
			Martin	Kramer
			Maria	Costa
9	**BMR genomics**	**Italy**	Barbara	Simionati
10	**FhG (Fraunhofer)**	**Germany**	Stefan	Wesarg
			Cristina	Oyarzun Laura
11	**INRIA**	**France**	Xavier	Pennec
			Maxime	Sermesant
12	**MOTEK**	**Netherlands**	Fraans	Steembrink
			Ben	Van Basten
13	**SCR**	**USA**		
14	**DUT (DELFT)**	**Netherlands**	Dirkjan	Veeger
15	**URLS**	**Italy**	Fabrizio	Patanè
			Roberto	Di Marco
			Paolo	Cappa
			Stefano	Rossi
16	**USFD (Sheffield)**	**United Kingdom**	Miguel Martin	Fernandez
			Steven	Wood
			Marco	Viceconti
			Claudia	Mazzà
			Rod	Hose
17	**MAAT**	**France**	David	Manset
			Sebastien	Gaspard
18	**HES-SO**	**Switzerland**	Patrick	Ruch
			Emilie	Pasche
			Ranveer	Joyseree

19	**TBV (Transylvania)**	**Romania**	Costantin	Suciu
20	**ATHENA**	**Greece**	Harry	Dimitropoulos
			Omiros	Metaxas
21	**EMPIRICA**	**Germany**	Karl	Stroetmann
			Rainer	Thiel
			Masha	Smirnova
22	**LYN**	**Italy**	Edwin	Morley-Fletcher
			Alessandro	Sattanino
			Callum	MacGregor
			Almerico	Bartoli
			Mirko	De Maldè
			Ludovica	Durst

Working Groups' Minutes

Cardiomyopathies Working Group Minutes

Participants*:

- Gabriele Rinelli - OPBG
- Giacomo Pongiglione –OPBG
- Giovanni Biglino – UCL/GOSH
- Tobias Heimann - SIEMENS
- Michael Sueling - SIEMENS
- Tommaso Mansi (Via TC) – SCR
- Maxime Sermesant

Discussion:

The goal of the discussion was to gain a mutual understanding of each partner's background, strengths, and interests and to develop clinical use cases which can offer a high potential impact for medical and technical partners alike, thus collecting the necessary requirements for use cases and set up a first action plan.

Modelling:

- Modelling activities will start by Siemens Anatomical models: functional/haemodynamical models.
- The Objective is to develop a personalized heart model.
- The general approach is to develop a physiological model (anatomical+dynamic) that matches and adapts to the patient and how it will develop. For computer models in general depending on how you set the variables, blood flow simulation, valves.
    - The modelling activities will encompass different stages: from endocardium to complex valve images. The use of 3D echocardiography will be crucial in this project
    - Towards solutions for landmark detection, body labeling, segmentation, motion estimation and abnormality detection.
    - Proposals for data: geometrical values can be calculated automatically using these kind of models
    - Extraction of models from computer tomography, MRI data (very important in paediatrics).
    - Fast and robust model of estimation methods

- **Genetics:** Isolated DCMP will be enrolled for genetic investigations and screening of 20 DCMP candidate genes. Exome sequencing will be applied for selected familial and complex phenotypes. Research agreement GOSH/ OPBG: Lot of relatively novel genetic investigations will be further continued at Great Ormond Street.

During the discussion T. Mansi from SCR joined the Working Group via teleconference to present the Haemodynamic model.

The working group focused on two proposed clinical scenarios:

- Clinical scenario 1: Advanced Diagnostics
    - Compute novel biomarkers for cardiomyopathy sub-type identification

    - Challenges for patient specific Clinical management
        - Etiology of the disease?
        - Disease prognosis?
        - Optimal therapy?
- Clinical Scenario 2: Model-based Therapy Planning (LVAD implant)

Issues about prediction:

- Endpoint is listed for transplantation.
    - Usually after a transplant you have a group of chronic heart failure patients thoroughly assessed, after 15 years
    - Where these heart models (hemodynamic models) can fit in? From simulations it is difficult to extract numbers to use for prediction. The model gives you something you can't measure very easily.
    - Q: from the data that we had a year ago, can we predict the outcomes 8 months later?

Discussion based on INRIA's Presentation:

- Statistical and biophysical modelling of the heart
- Electromechanical model of the heart, 4 elements:
    - anatomy: how to personalize the fibres;
    - electrophysiology: Parameters ECG (electrical conductivity) Literature (Anisotropy)
    - biomechanics: adjustment of parameters: mass, anisotropy>literature •Elasticity, active contraction (MRI, Trial and errors)
    - Haemodynamics

Discussion follows with a focus on Clinical Protocols

- Study Setting: 33-month longitudinal cohort study
- Study Population: Inclusion criteria, Exclusion criteria
- General study Design: Clinical assessment, Diagnostic testing, Imaging
- Different protocols between chronic and acute. (OPBG almost all chronic). Acute are much younger, once they translate to dilated cardiomyopathy.
- The majority of patients in the study will be chronic: 2 decisions if transplant or stable (majority). Acute patients will be very few.
- The group pointed out that it is necessary to start to upload data in the repository very early, and that for the imaging data it's crucial that everybody uses the same protocols.

The Group continued the session with a discussion about Ethical implication and Informed consent.

Working Group Conclusions

Clinical question / use case:

How will CDM patients evolve in the next 20 months: will they be in stable condition, or will they need a transplant?

Hypotheses:

- High septal strain predicts patients who will require early heart transplants?
- Correlation between size of heart and outcome?

Commitments:

- Beginning of April 2013: Clinical protocols ready
- September 2013: Ethical clearance, start of image data exchange to fine-tune protocols, check processing capabilities
- October 2013: Start of study data acquisition at UCL
- February 2014: Start of study data acquisition at OPBG
- November 2014: End of baseline acquisition

Requirements for MD-Paedigree's Infostructure:

- Upload of images usable in September 2013
- Automatic pseudonymization usable in October 2013
- Solution for uploading and accessing clinical variables usable in October 2013

CVD Risk in Obese Children and Adolescent Minutes

Participants

- Andrew Taylor - UCL
- Marcello Chinali - OPBG
- Alexander Jones - UCL
- Melania Manco - OPBG
- Lorenza Putignani - OPBG
- Xavier Pennec - INRIA
- Cristina Oyarzun Laura - FRAUNHOFER
- Alessandro Sattanino - LYNKEUS
- Michael Suehling - SIEMENS
- Paolo Tomà - OPBG

The problem of the Integration of data has been deeply discussed, considering the following **data**:

- Clinical
- Metabolomics
- Imaging
  - MRI
  - US

    - IMT
- Genetics
- Microbioma/metagenome
- Process of assessment and analysis automated

Questions emerged and discussed in the meeting:

Questions

- Should we have control group?
- Should we acquire other blood sample during stress?
- Should we take other circulating adipokines (Ghrelin)?
- Can we get accurate info about fat distribution from US?
- CMR - Why LGE, or should we do ECV?
- Vascular assessment - ? IMT, ? PWV (MR)? endotelial function (how), ? MR stress, other measures of systemic vasculature
- Will we get meaningful data from our genetics analysis?
- What will Microbioma/metagenome analysis show?

Issues raised when discussing the protocol preparation (in bold the open questions needing decision soon)

- Metabolic meal - assess IR, lipid and vascular response: Originally we wanted to do Glucose Test but we'll probably do a metabolic meal because that might give us more dynamic (liquid response) information.
- T1 mapping of the liver (T1, Dixon, Elastography)
- Genetics - extracted DNA to BMR Genomics in Padova
- MRI protocol to include assessment after meal

- **Should we carry out Mendelian randomisation first?**
- **When to start? - 1.5T at OPBG** (in January 2014 a new machine is delivered)
- **Where is all the data going to go? Interim solution?**

Some specific discussions

- What we do know about liquid load is that it causes vascular changes in a way that glucose doesn't, so you see endotheliuma changes. If we do it in a dynamic way we can look at the changes at heart rate blood pressure, vascular resistance, endothelium changes etc.
- Will the Ethical Committees give us permission to study children with mixed meals?
- Necessary to deliver some imaging data to Siemens, from which to extract parameters, image analysis system. About the fat content, we should try to automate the measurements of fat, and to detect slight variations/remodelling of the heart.
- Define a list of key variables for Siemens.
- For the clinicians, what background data they want to obtain on these (180) people?
- What questionnaires? Smoking/alcohol, exercise data, social class.

Next steps schedule

- Clinical protocols completed - Month 1

- Data proformas completed - Month 1
- Ethics to be submitted - Month 2
- Ethics completed - Month 4/5
- Trial data set to the technical partners - Month 3-6
- Define research hypotheses - Months 3-6

JIA Working Group Minutes

Participants:

- Fabrizio De Benedetti - OPBG
- Flaminia Frascarelli - OPBG
- Stefano Lanni - IGG
- Gian Michele Magnano - IGG
- Silvia Magni Manzoni - OPBG
- Claudia Mazzà - USFD
- Lorenza Putignani -OPBG
- Stefano Rossi - URLS
- Laura Tanturri - OPBG
- Paolo Tomà - OPBG
- Marco Viceconti - USFD
- Stefan Wesarg – FRAUNHOFER
- Reiner Thiel – EMPIRICA
- Ludovica Durst -LYNKEUS

Discussion:

We revised the study protocol in terms of:

1) Patients selection: all JIA patients with a disease duration ≤ 6 months will be enrolled in the project
2) Clinical data to be collected:
   - Case Reports Forms (CRF) will be amended according to Fabrizio De Benedetti and Berent Prakken suggestions.
   - Protocols for collection and storage of biologic samples will be sent by Lorenza Putignani (stools for microbiote analysis) and Berent Prakken (blood and synovial fluid) to be included in the study protocol.
3) Imaging data:
   - Ultrasound (US) will be performed in all patients enrolled. The US protocol has been discussed and a training session will be organized within the next two months (Stefan Wesarg will be invited to participate for a potential post-processing analysis of US assessment). Each centre is asked to verify whether the US machine stores images using DICOM. 3D US will not be applied for the project since it is not available in all centres.

- MRI assessment of both ankles (clinically affected and non-affected side) is required at baseline. A second MRI assessment will be performed at follow-up (month 12 and/or month 24- it still needs to be determined) only in patients with persistent active disease (or in all patients with baseline ankle involvement-this still needs to be determined), in order to pursue a more accurate assessment of disease progression. A standard position of foot during MRI assessment needs to be yet established.
- DXA: We need to know the DXA equipment used in each centre and whether data are stored using DICOM (please communicate it as soon as possible).

4) Gait analysis:
   - Claudia Mazzà will revise all literature of the topic in order to propose a protocol for gait analysis in JIA, with a special focus on ankle/foot.
   - A training/calibration session for gait analysis involving all operators in the field of the centres will be scheduled by next month. Afterwards gait analysis will be done in healthy controls (around 10-15 patients, age 5-8 years), before starting recruitment of JIA patients.
   - An initial problem was raised by UMCU in consideration of the fact that they do not routinely perform CGA in their patients care. A few days later, Berent Prakken was however capable of guaranteeing the availability of Dr. Jan Jaap van der Net, a renowned expert in biomechanical problems in children with arthritis, who will be able to perform CGA within UMCU, where he is head of the Department of physical therapy & biomechanics in children.
5) Timing:
   - Baseline: Clinical data and US assessment. Ankle MRI and gait analysis will be performed only in patients with ankle involvement.
   - Month 6: Clinical data and US assessment. DXA and Gait analysis will be performed only in patients with ankle involvement at baseline.
   - Month 12: Clinical data and US assessment. Ankle MRI will be performed in patients with persistent ankle disease activity. To be discussed whether to repeat ankle MRI in all patients with baseline ankle involvement.
   - Month 18: Clinical data and US assessment.
   - Month 24: Clinical data and US assessment. MRI: to be discussed

Questions to Infostructure Partners about data collection:

Q1. Can you characterize the data that you want to exploit in your system? Can you provide data samples?

- JIA will have the same data & forms as we had in HeC with only some small differences. For example, the Microbiota and cytokine information will now be added (via Excel/CSV or PDF files).
- We also have imaging:
  - MRI (DICOM)
  - Ultra Sound (fake DICOM support that just exports JPEG images and it gets much more complex with 3D ultrasound)
  - DXA (probably DICOM) but not integrated with the central PACS system.
- 8-digit patient ID numbers will be used internally at each hospital to link patient data from different modalities (e.g. imaging with lab results). This ID will be unique within each hospital. This will be easy for what is supported by PACS, but not sure what will happen for things like DXA for example (i.e. the association between a patient's DXA results and other data for the same patient must be created by clinicians somehow).

- The Microbiota output is typically a table/Excel/CSV file, where each line/entry: Coded Species (ID name), % of concentration. Again identity association must be done for this.
- Immunological data is also a table/CSV file, where each line/entry: Name of cytokine, amount of concentration.
- In addition we will have Gait analysis data… see below Q2.

Q2. What do you want from the simulation? How will you interpret the simulation data? What do you want to search?

- Image processing models, geometries, 3D: add some fields for predictions from the models for each patient, which will probably be numbers. Can be stored as DICOM comments? The models considered are from offline transactions producing just biomarkers – not interactive models.
- Modellers must exchange data between themselves – it must be secure and part of the MD-Paedigree system (they need a secure repository). They usually work with shared folders.

BIOLOGICAL SAMPLES WILL BE COLLECTED AT DISEASE ONSET (BASELINE VISIT), WHEN PATIENT WILL ACHIEVE CLINICAL REMISSION STATE (according to the Wallace criteria for remission), **AND DURING FLARE OF DISEASE.**

All centres are asked to provide/confirm the names of the Rheumatologists, Radiologists and Physiatrists involved in the project as well as the name of the persons dedicated to data collection.

NND Working Group Minutes

Participants:

- Jaap Harlaar – Vumc
- Kaat Desloovere – KU Leuven
- Enrico Bertini - OPBG
- Paolo Cappa - URLS
- Enrico Castelli - OPBG
- Maria Jimena Costa – Siemens
- Roberto Di Marco – URLS
- Miguel Martin Fernandez - USFD
- Nathalie Goesman – KU Leuven
- Fabrizio Patanè – URLS
- Franz Steenbrink - MOTEK
- Ben Van Basten - MOTEK
- Marjolen Van der Krogt - VUmc
- Dirkjan Veeger – TU Delft

Discussion:

The Group discussed about the advanced use of Clinical Gait Analysis and about the high heterogeneity in the results of this kind of tool.

Two strategies were individuated:

- enhance the understanding of diseases' etiology;
- Identify patterns enabling statistical based prediction

The Group clearly identified and discussed the main objectives of WPs 6 and 11.

At the beginning of the WP6 meeting, there were some different expectations from the various group members, which, however, could eventually be aligned, allowing to settle most of the issues initially raised, and to define a top priorities list.

The Group discussed about

- Data collection:
  - The Leuven Group (Kaat Desloovere) will provide database for CGA and Data collection.
  - The CGA Data are on CP Patients, since for Duchenne and SMA 3 no CGA is foreseen neither before or after intervention.
  - Another issues emerged is that Leuven could not provide any MRI data.
- Patient enrolment (both for retrospective database and new data collection) :
  - Enrolment criteria: decide a common diagnosis to select patients
  - Furthermore has been foreseen 40 CGA of healthy children, to provide a benchmark for the CGA of pathological conditions.
- Clinical protocols (data collection):
  - Decision to define within a month protocols for scans (MRI and DXA) and acquisitions. Siemens and Sheffield will advise on how to get the best quality (essential to develop the protocol)
  - Siemens will perform MRI segmentation but many MRI Data set are needed to perform automatic extraction and statistical analysis.
  - Not entirely clear what kind of information was needed to feed the probabilistic modelling: there have been preliminary discussion with people from Athena but group must define exactly inputs required, especially for the retrospective data.
  - DTI are also needed for evaluating muscle fibre orientation.
- Models: discussion about what models have to be used as basis for further developments: MOTEK HBM or OpenSym? Group decided to exploit both models.

Discussion with Infostructure Partners about data collection:

- The NND data is a combination of a set of discrete values and a set of continuous waveforms (time series from which clinical relevant parameters are extracted). The data is Patient-centric. For each patient they have a full set of gain analysis data, with multiple trials, etc. Bayesian classification of gait patterns has already been investigated in one centre, but it's a complex task.
- Clinicians disagree from country to country on treatment options (e.g. Botox treatments are not acceptable in some countries, whereas in other countries they have no problem).
- Generally a difficult task. We need to find homogenous groups of patients using baseline/pre-data and post-data & maybe a control group.

- Maybe not focus on classification, as it is a very hard problem, but instead focus on finding *similar patients* which is very important.
- We need to visit at least one lab to understand fully the data.
- Pressure data could also be included…. But maybe best approach is to simplify, else it can become very complex.
- Search criteria: for example use continuous waveforms to find similar patients.

The persons responsible per specific tasks (needed in the forms to be filled for the ethical approval) have been almost completely defined and the WP leaders will complete it within 2 weeks . A top priority for WP11 is defining the appropriate sequences of development (who is doing what?)as well as the requirements and parameters.

Plan of the year: the Group plans to have defined , within January 2014, the first stages of the physical model they plan to develop and to have completed the data set for the retrospective data.

Data collection:

- 10 patients not only MRI but also DXA
- Healthy children: 20 MRI without CGA;
- 10 for CP (for each clinical centre);
- 10 for Duchenne (for each clinical centre);
- 10 for SMA (for each clinical centre);

Timing:

- Marker protocols within month 3
- Protocols for MRI and DXA ready by the end of month 2

Infostructure Working group Minutes

Participants:

- CS - Constantin Suciu –TBV
- PR - Patrick Ruch, HES-SO
- EP - Emilie Pasche, HES-SO
- RJ - Ranveer Joyseeree, HES-SO
- HD - Harry Dimitropoulos, ATHENA
- OM - Omiros Metaxas, ATHENA
- FP - Fabrizio Patane, URLS
- RH - Rod Hose, UCFD/VPH-Share
- SW - Steven Wood, UCFD/VPH-Share
- MK - Martin Kramer, SIEMENS
- GP - Giacomo Pongiglione, OPBG
- PK - Paolo Kappa, OPBG
- AS - Alberto Sanna, HSR

- CM - Callum McGregor, LYNKEUS
- KS - Karl Stroetmann - EMPIRICA/HTA
- SG - Sebastien Gaspard, MAAT/GNUBILA
- DM - David Manset, MAAT/GNUBILA
- MS - Maria Smirnova, EMPIRICA

The Group started debating the legacy from project partners + external projects' reusable assets:

- AITION, PAROS, ADP (ATHENA)
  - Presentation by Harry Dimitropoulos and Omiros Metaxas from ATHENA
  - Athina Distributed processing and querying engine
    - Distributed querying over federated heterogeneous sources
    - Distributed processing algorithms
  - Data information and validation
    - Data curation and validation techniques
    - Enhanced DCV on top of open source madIS
  - Semantic modeling and representation
    - Tools for mapping ontologies and data, large scale
    - Ontology-based querying
  - Personalization, adaption and representation
    - Provides user behaviors, interests, preferences, attitudes identification
    - Graph-based user-actor representation
    - PAROS application suite
  - Biomedical knowledge discovery KDD
    - AITION platform, now quite advanced platform
      - Identify disease signatures
      - Deliver highly accurate and reusable predictive statistical simulation models
      - Data driven and model driven
  - Data policy definition and implementation
    - Compliance with OpenAIRE etc
    -
      - How does ATHENA articulate with other lifescience initiatives' contents?
      - ATHENA mainly concerned with mechanisms for relating publications to datasets
  - Not everything open sourced as of yet (excepted madIS)
    - Project specific IPR sharing is put in place in general

- PCDR, HeC, SeC (MAAT/GNUBILA)
  - Presentation online: http://prezi.com/tqyqxmfxqmij
- XIP, THESEUS (SIEMENS)

  - Presentation given by Martin Kramer from SIEMENS

- Theseus Medico project (German-funded)
  - Improving clinical workflow of radiologists using semantics
  - Dealing with different types of images
    - Image analysis as basement
    - + Text parser / mining systems
  - Semantic annotation, ontology mapping and knowledge creation
    - Applicable to medical images, reports, treatment plans and guidelines, Internet databases and expert knowledge
    - On top of this, many applications: semantic reading, search, img and text linking, etc
  - Architecture made of semantic server managing data commands and importers
  - Everything saved in RDF format
  - Image parsing in Windows Azure
    - Worker roles (DLLs) processing in the Cloud
  - Web-based DICOM viewer
    - Open source XIP
- Potential contributions
  - Web-based DICOM viewer
  - Medical annotation ontology
  - User-friendly interfaces
  - Windows Azure interface
- Open issues
  - Anonymization tools
  - Security

- VPH-Share (USFD)

  - VPH-Share
  -
    - Presentation by Steven Wood and Rod Hose from UCFD
      - Project is ongoing since 2 years already
    - Architecture looks similar to that of MD-Paedigree
      - Access mechanisms to HPC infrastructures
        - Compute services such as workflow models
        - Separated storage from databases
      - Knowledge Management
      - Knowledge Discovery, Data Inference
      - Semantic Services
      - Decided to provide access to Amazon and Rackspace
        - Based on OpenStack
          - Have documented SOTA 18 months old on this area

        - Using Taverna workflow system
    - Semantics
        - Terminology repository and search service
        - LinkedData model for data publication
        - Light weight annotation model
    - Potential areas of collaboration between VPH-Share and MD-Paedigree
        - Would like to share tools between systems
        - No data validation processes (curation?), interested in our approaches
        - Lot of experience in data integration in hospitals (EPR extraction, PACS integration etc)
        - Easy to use semantic annotation tools, publication to SPARQL and SQL endpoints
        - Data publication suite
            - No fine-grained security for access controls. Republish datasets based on access rights to be respected
            - Anonymization is "data owners' problem"
                - Clinical research targeted so anonymization isn't main concern
                - No ethical approval required in VPH-Share
                - Just providing recommendations to users

- KRESHMOI, EHR4CR (HES-SO)
    - Presentation given by Patrick Ruch from HES-SO
    - Involved in WP13, WP15 and WP16
    - Strong expertise in Content-based image retrieval
        - Questions answering
        - Support to biocuration (2 million queries a year at HES-SO)
    - Participation to KRESHMOI project
        - Multimodal search
        - Clinical practice guidelines (normal values)
        - Genetic curation etc
    - May contribute something like a search engine, perhaps natural language based
        - Advanced search using image, sequence, ontology etc
        - Data-driven design of clinical trials
            - Rely on EHR4CR developments
                - Local ontologies to build global view, DebugIT-like
                - Local mappings to access data on the fly and when central repository is updated
- epSOS, p-medicine
    - DM reports on epSOS Patient Summary and discussion with French NCP at ASIP Santé

- MD-Paedigree may introduce a Patient Summary like concept for research
  - Uniform view of diagnosis, prescription etc
  - We should not reinvent the wheel and thus use existing ontologies
- 2 different approaches possible
  - epSOS has 38 value sets to enable mapping to patient data at European level
    - Straight forward as focussing on value sets instead of full ontologies
    - Mappings available for most EU countries + USA
    - HL7 CDA, CEN 13606 standards
    - All these aspects expected to be covered in WP13
  - p-medecine tried to integrate all oncology ontologies into one single
    - More complex as aligning different ontologies
    - More into the business of clinical trials

➢ Open questions
- MD-Paedigree aims to store different types of data
  - Clinical, kinematic, kinetic, imaging, reports, etc
  - Data controlled will need to know who's downloaded/accessed what data etc
    - Will also need consent forms?
  - CUSTODIX company cited as expert in the area
    - Can act as a Trusted Third Party Service (TTPS)
    - Involved in EHR4CR and EMIF
  - What is the role of LinkedData in MD-Paedigree?
- User Board --> involve GEANT?
- Interoperability Steering Committee --> involve epSOS?
- Common project software repository + licencing conditions from partners contributing assets

➢ Users requirements - HES-SO lead
- DM summarizes yesterday's discussions
  - Correction: VPH-Share provides anonymization tools, does not enforce their use but rather encourage it
- GP kick starts requirements brainstorming, ideal vision of the MD-Paedigree system:
  - Heterogeneous data integration
    - Entering data should be the same as entering routine data
      - Maybe we can adopt a form for the reports
        - Forms should not be too "categorized"
        - Data should be transferred to the system automatically
      - User-friendly interfaces with just the variables to be entered
        - List of variables we want to use as such criteria
        - +- Z-values
      - Data should be curated once imported into the system
  - Should be possible to deal with text reports

- Informed consent is the problem of clinical partners not infostructure
- Anonymization is an issue for users
  - May have 2 levels
    - (1) Open public, strictly compliant to HIPAA like recommendations
    - (2) Private reserved for research
  - Images/DICOM is an issue
    - How to ensure anonymization of complex and heterogeneous DICOM files?
  - How to come back to patient in case needed?
- Data curation
- Similarity search
  - List of meaningful variables for selection
    - Clinical structured representation
      - E.g. should make measurements homogeneous (Z-values) and comparable
  - What does "most similar" mean?

o Identified major use-cases

- (0) User-friendly, simple and secure access to the system
  - No change to clinical practice
- (1) Pseudonymize/anonymize all types of collected data
  - Patient informed consent
    - Not practical for every patient, especially of retrospective data
    - Informed consent may be imposed right before exam (prospective data)
      - COAST database in the US exploited retrospective data without asking for consents. Just ensured anonymity of data
  - HIPAA too constraining for MD-Paedigree research purposes
    - (1) Open public, strictly compliant to HIPAA like recommendations
    - (2) Private, less HIPAA compliant, reserved for research
      - Access rights to be defined
      - HIPAA alterations to be documented and explained
  - Architecturally implies servers to be located within the hospital firewall/security as is the case of PCDR
  - DICOM images anonymzation
    - Ultra-sound images pose a problem as identifying information is written on the image itself
      - Solutions exist to blank specific areas in the image
    - Anonymization profile(s) to be defined for the various DICOM modalities
- (2) Automatically collect, integrate, curate and represent data in the right way
  - List of diagnosis/intervention codes from source systems
- (3) Similarity search throughout population stored in the system
  - 2 sub scenarios
    - (a) Individual versus population

- (b) Similarity groups
- Define similarity in terms of clinical parameters
  - (i) Simplest is form-based profiled from user and specific measurements/variables
    - Number of templates referring to different categories etc
    - May also be able to construct new personalized templates
    - May choose your own clinical parameters' weights
    - Questions
      - How do we manage these templates?
      - How complex has the template to be?
      - How do you update the templates?
        - The template is your query in the end
      - Are their existing systems?
        - Archimedes (Kaiser Permanente) system to be checked
          - California-based, clinicians + insurances working together
          - Integration of data from 8 EPIC systems into a data warehouse
          - Outcomes improvement (50% better in the US)
        - KS to circulate information about it
      - How many medical events per patient typically?
        - 10 to 15 medical events
        - Prescription is included in the visit
  - Self learning-based, simple distance + user feedback
    - Euclidian distance does not work as certain parameters may vary in different ways
  - Could define similarity based on user feedback, from simple euclidian distance search
  - Experts based weighting of clinical variables
- (4) Modeling and simulation
  - What will be the flow of simulation parameters etc
  - Simulation workflows, need to know inputs and outputs
    - Clear definition input and operational output
    - Can models from previous project/groups be shared/accessed?
  - In particular what was done wrt GPUs
  - To be further discussed next round
- (5) Connect with external literature
  - Case reports, etc
  - Medline, Cokran(?) instances in HeS-So

        - RSNA full text as well
- List of user requirements questions for other groups
    - Can we characterize data per disease groups?
        - Can we obtain a data sample asap?
    - What do you want from the simulations?
        - How will you interpret outcome of such simulations?
- Scheduling and allocation of tasks
    - Review of area scheduler
    - Planned deliverables
        - D13.1
            - Draft August 2013
            - Consolidated September 2013
            - Due November 2013
        - D14.1
            - Draft August 2013
            - Consolidated September 2013
            - Due November 2013
- Self-Assessment Criteria Definition
    - To be defined on a WP/Task basis by WP leaders for the next monthly TC
- Data samples for testing purposes
    - To be asked and gathered centrally somewhere (FTP server?)
- Assigned rapporteurs to split and attend the last hour conclusions of the clinical sessions listed above
    - Ranveer to centralize requirements / main contact person at HES-SO
        - (1) Tobias Heimann/SIEMENS - Cardiomyopathies (CMP)
            - ==> GP/OPBG + TH/SIEMENS
            - Data acquisition not before January 2014
            - MRI protocol to be defined by AT
            - Metadata collected in an Excell form and appended in civis in rome and to the pt file in london.
            - Echo data to be added on individual base and on local preference.
            - No need for cardiac biopsy except selected cases.
            - Gene analysis out a pre defined list.
            - Easy usable upload for images latest by September
            - Automatic pseudonymization mechanism by October
            - Solution for uploading and accessing clinical variables usable in November
        - (2) Andrew Taylor/GOSH - CVD in Obese Children (CVD)
            - ==> MK/SIEMENS
            - 1. Which type of data will be used? Will test data be available?
                - CVD has following medical data:
                    - DICOM (3D Images)

        - Excel sheets/CSV with parameters of patients
    - List of parameters will be set up asap
    - Test data (one dataset of each type) will be provided asap
    - Original patient IDs will be used internally at each hospital to link different patient data. In case of exports for MD Paedigree partners this patient ID will be replaced by a generated unique ID which will be saved in a central mapping file. Only one responsible person will store and update this mapping file between internal patient ID and anonymized UID.
  - 2. What your requirement and wishes about an infostructure system?
    - Quick solution (in 1-2 months available) for sharing test data
    - Upload CSVs and images
    - Are there anonymization tools available? Tools should automatically generate mapping files between original patient ID and anonymized UID
    - Filling online forms with all parameters of the patient
- (3) Alberto Martin/IGG - JIA / (JIA)
  - ==> Fabrizio Patane/URLS + OM/ATHENA
  - Q1. Can you characterize the data that you want to exploit in your system? Can you provide data samples?
  - JIA will have the same data & forms as we had in HeC with only some small differences. For example, the Microbiota and cytokine information will now be added (via Excel/CSV or PDF files).
  - We also have imaging:
    - MRI (DICOM)
    - Ultra Sound (fake DICOM support that just exports JPEG images and it gets much more complex with 3D ultrasound)
    - DXA (probably DICOM) but not integrated with the central PACS system.
  - 8-digit patient ID numbers will be used internally at each hospital to link patient data from different modalities (e.g. imaging with lab results). This ID will be unique within each hospital. This will be easy for what is supported by PACS, but not sure what will happen for things like DXA for example (i.e. the association between a patient's DXA results and other data for the same patient must be created by clinicians somehow).
  - The Microbiota output is typically a table/Excel/CSV file, where each line/entry: Coded Species (ID name), % of concentration. Again identity association must be done for this.
  - Immunological data is also a table/CSV file, where each line/entry: Name of cytokine, amount of concentration.
  - In addition we will have Gait analysis data… see below Q2.

- Q2. What do you want from the simulation? How will you interpret the simulation data? What do you want to search?
- Image processing models, geometries, 3D: add some fields for predictions from the models for each patient, which will probably be numbers. Can be stored as DICOM comments? The models considered are from offline transactions producing just biomarkers – not interactive models.
- Modellers must exchange data between themselves – it must be secure and part of the MD-Paedigree system (they need a secure repository). They usually work with shared folders.

- (4) Jaap Harlaar/VUMC - NeuroMuscular Diseases (NMD)
  - ==> Fabrizio Patane/URLS
  - Their data is a combination of a set of discrete values and a set of continuous waveforms (time series from which clinical relevant parameters are extracted). The data is Patient-centric. For each patient they have a full set of gain analysis data, with multiple trials, etc. Bayesian classification of gait patterns has already been investigated in one centre, but it's a complex task.
    - Clinicians disagree from country to country on treatment options (e.g. Botox treatments are not acceptable in some countries, whereas in other countries they have no problem).
    - Generally a difficult task. We need to find homogenous groups of patients using baseline/pre-data and post-data & maybe a control group.
    - Maybe not focus on classification, as it is a very hard problem, but instead focus on finding similar patients which is very important.
    - We need to visit at least one lab to understand fully the data.
    - Pressure data could also be included…. But maybe best approach is to simplify, else it can become very complex.
    - Search criteria: for example use continuous waveforms to find similar patients.

- Open issues
  - Data samples
    - DM to provide data repository ASAP
      - Echo and MR data may be provided asap
      - Ethical committee examined protocols yesterday and agreed and approved in principle
        - Informed consent to indicate what, why and where the data will go
          - May include a clause for secondary use in similar projects

        - May need 2 consents, 1 for the project and 1 for the secondary use
- Clinical system vs research platform
    - Completely different task
        - User-friednliness, architecture, policies, etc may differ slightly
        - Whatever we do needs to be unified and data should go to the same place
        - No impact on daily clinical work though
        - Informed consent isn't a valid question, as it will be necessary in both cases
        - May start as if it was a multi-centric clinical study with a centralized CRO (e.g. OPBG)
        - System may be developed as social network involving the patient in the loop
            - Giving some patients an access to the system
                - Mapping their identity to anonymous records in the system
                - Asking patients to sign consents online
                - Paradigm shift concept
    - Is clinical system a wish or a goal in the project lifetime?
    - Need a written policy, a vision statement
- What are the uses cases?
- What are the user workflows for each disease group?
- What are the interfaces between systems of all infostructure partners?

Timing of next step:

- April
    1. 3 self-assessment criteria per task per WP by WP leader
    2. 1st round of User Requirements @ OPBG
- June/July
    1. 2nd round of User Requirements
- August
    1. 1st draft of D13.1 and D14.1
- September
    1. Consolidated drafts of D13.1 and D14.1 for review during bi-annual meeting
- October
    1. D13.1 and D14.1 internal review
- November
    1. D13.1 and D14.1 submission

The group finally highlighted two main issues:

- **Need of data (meaningful data samples)**

    1. Data sources to be characterized (type of data, formats, technologies, standards if known)

- **Need of a list of inputs and operational outputs for simulation workflows**

    1. Access to existing models
    2. **The project needs a policy whether clinical system vs research platform?**
    3. Impacts on data privacy (i.e. anonymization/pseudonymization, informed consent, secondary use of data etc.)

Appendix 1 - Clinical protocols discussed during the Kick-off meeting

Protocol no.:	
Title:	Data acquisition and processing for Cardiomyopathies
Acronym:	MD-Paedigree – WP 3
Multicentric/Monocentric Study	Multicentric
Principal Investigator/Coordinator	Prof. Bruno Dallapiccola

WP 3 - Data acquisition and processing for Cardiomyopathies

Protocol no:
Version: 4 Apr 12, 2013**CONFIDENTIAL**

Sponsor:	European Commission though the Bambino Gesù Children Hospital
Person responsible of the study	Dr. Gabriele Rinelli
Person responsible for OPBG	Dr. Gabriele Rinelli
Data Management/Statistical analysis:	

Protocol approved and signed by:

Principal Investigator:

Prof. Bruno Dallapiccola

Responsible Work Package:
Dr. Gabriele Rinelli.
Responsible Unit:
Dr. Gabriele Rinelli.
Acronym List

AEs	Adverse Events
EC	Ethical Committee
CRF	Case Report Form
GCP	Good Clinical Practice
ICH	International Conference on Harmonisation
IRB/IEC	Institutional Review Board/Independent Ethics Committee
NSAEs	Non Serious Adverse Events
SAEs	Serious Adverse Events
SOPs	Standard Operating Procedures
……	……..

INDEX

1 INTRODUCTION

1.1 BACKGROUND OF THE MD-PAEDIGREE PROJECT

MD – PAEDIGREE *Model-Driven European Paediatric Repository* is a research project funded by the European Commission under the Virtual Physiological and Human Area (VPH) of the ICT Theme of the Seventh Framework Programme (contract no. 600932). The project foresees 7 world-renowned clinical centres of excellence pursuing improved interoperability of paediatric biomedical information, data and knowledge, by developing a set of reusable and adaptable multi-scale models for more predictive, individualised, effective and safer paediatric healthcare. The project is scientifically and technologically supported by one of the leading industrial actors in medical applications in Europe, and operating in conjunction with highly qualified SMEs and some of the most experienced research partners in the Virtual Phisiological Human (VPH) community.

MD-Paedigree validates and brings to maturity patient-specific computer-based predictive models of various paediatric diseases (cardiomyopathy, risk of cardiovascular disease in obese children and adolescents, Juvenile Idiopathic Arthritis (JIA) and neurological and neuromuscular diseases (NND), thus increasing their potential acceptance in the clinical and biomedical research environment by making them readily available not only in the form of sustainable models and simulations, but also as newly-defined workflows for personalised predictive medicine at the point of care.

MD-Paedigree aims to advance the state-of-the-art of patient-specific computational modelling of different paediatric diseases and translate the latest advances into clinics to improve disease understanding, therapy outcome, and provide an infostructure platform for assessing new therapies at the point of care.

MD-Paedigree's goals therefore are to integrate and share highly heterogeneous biomedical information, data and knowledge, using best, jointly develop reusable, adaptable and composable multi-scale VPH workflow models and to support evidence-based translational medicine at the point of care practices from the biomedical semantic Web.

Illness in infants, children, and adolescents are a large and under-appreciated public health problem. Because paediatric patients now have a much longer life expectancy, the burden and costs are substantial for families and society. For instance, recent studies have shown that the number of adult patients with congenital heart disease is already similar to that of the paediatric population and will continue to grow. It can therefore be stated that in the longer term the VPH scientific approach requires a fundamental research investment in paediatrics, where the conceptual revolution which is underway, transforming the nature of healthcare from reactive to preventive, can best be applied, moving towards an approach based on personalised, predictive, preventive, and participatory (P4) medicine, increasingly focused on wellness.

MD-Paedigree represents a major step towards personalised paediatric e-Health, based on data-driven models, patient-specific simulations and a sustainable data and model repository. The project can in fact impact the way healthcare is practiced in the future. MD-Paedigree's impact on dealing with very concrete and exemplary clinical cases is demonstrated through the following applications in the diseases scenarios. MD-Paedigree encompasses complete services for storage,

similarity search, outcome analysis, risk stratification, and personalised decision support in paediatrics within its innovative model-driven data and workflow-based models repository, leveraging on service and knowledge utilities. It fosters the state-of-the-art of patient-specific computational modelling of the selected diseases and translates the latest advances into clinics to improve disease understanding and provide a platform for testing new therapies at the point of care. In fact, MD-Paedigree demonstrates how a dedicated VPH repository can provide full accessibility to existing and further developing knowledge in paediatrics to all interested bio-medical researchers precisely through linking data with models, thus proving the large scale benefits of having both the data and models readily available at the point of care.

The data collection performed within Health-e-Child and Sim-e-Child projects (already funded by the European Commission in previous calls for proposals) is still available to the MD-Paedigree consortium thanks to the continuity in the eHealth platform which is shared by all three projects.

Pathology	No of patients/Time	
	Spinal Muscular Atrophy (SMA) 20 ambulant patients (severity grade type 3); 10 patients for each centre for biophysical modeling; 10 patients among the 3a subgroup (symptoms of weakness appearing before age 3 years); 10 patients among the 3b subgroup (weakness appearing after the age of 3 years.	
	Duchenne Muscular Dystrophy (	

1.2 BACKGROUND OF THE DATA ACQUISITION AND PROCESSING FOR CARDIOMYOPATHIES STUDY

In paediatric cardiovascular disease, predicting how patients will respond to treatments (operations, catheter interventions, pharmacology), which treatments to use, and when to treat can be difficult to define due to small patient numbers and limited outcome data. When children present with new onset heart failure, there are five possible outcomes: full recovery, dilated cardiomyopathy (DCM) requiring drug therapy, DCM requiring transplantation or mechanical support, another diagnosis (other forms of cardiomyopathy, metabolic disease) or death. At

presentation, however, it is very difficult to predict which group any patient will end up in. Data suggests that good systolic function and younger age are good prognostic indicators for survival [Andrews RE et al., 2008]18, but better prognosticators are necessary. FP7- ICT-2011.5.2 600932 - MD-Paedigree –Part B 7.

Over the last decade, there has been a huge investment into information technology and computer modelling to build models of the heart that are able to gather any kind of clinical information and produce realistic representations of the cardiovascular system. Modelling of patient bioinformatic data may provide better insight into prognosis of cardiomyopathies, which would help in patient management and in telling families how their child will progress. Would he/she recover completely or would he/she require heart transplant? These models have now reached high levels of reproducibility, opening new avenues for more efficient, safer, and cost-effective patient management. However, their comprehensive validation is still limited.

2 AIMS OF STUDY

2.1 MAIN GOAL

Main objective of the "data acquisition and processing for cardiomyopathies" study is to evaluate predictors of cardiac failure in children and adolescents with cardiomyopathy (CMD) and, by doing so, to provide clinical, and cardiac structural, geometrical and functional data to build a "VPH Infostructure" data repository, to be readily available to clinicians and researchers on CMD in the youth.

MD-Paedigree will re-use the models developed in Health-e-Child and Sim-e-Child (research projects funded by the European Commission in previous calls for proposals, and participated by several MD-Paedigree project partners) and extend them to cardiomyopathies. The objective is to capture the main features of the cardiovascular system, including the heart, arteries and peripheral circulation, to predict cardiomyopathy progression an plan therapies like heart transplant and ventricular assist devices. Investigative data provided by imaging, pressure monitoring, clinical observations and exercise will be used to build these models and to validate them, by comparing model prediction with actual outcome. By merging all scattered information obtained from different diagnostic tools in clinical practice, and obtaining a generative model of heart function in children, our model will provide cardiologists the tools to deliver patients the best possible medical care.

2.2 PRIMARY END POINT

The primary end point of the study is the successful collection of clinical, laboratory and diagnostic data to be subsequently modelled and simulated by the IT experts. The primary end point envisages:

- **Enrolment of 180 DCM patients**: Enrolment of 180 patients, at baseline, with clinical, laboratory and diagnostic tool analysis achieved from month 4 to month 20, including echocardiographic, MRI and exercise test parameters.

- **Re-evaluation of all patients**: All 180 patients enrolled during D3.2 re-evaluated at follow up (month 21 to 36) to evaluate changes in clinical, laboratory and cardiac geometry and functional parameters.

3 STUDY DESCRIPTION

3.1 STUDY DESIGN

The data acquisition and processing for cardiomyopathies study is a 33-month observational longitudinal cohort study. The study will be performed in two cohorts of CMD children (total N=180). Ninety patients (approximately 45 girls) for each clinical Centre will be consecutively enrolled. Patients will be evaluated at the baseline (month 4 to month 20) and re-evaluated between month 21 and month 36.
The study will include clinical evaluation, laboratory testing, genetic testing, and diagnostic testing including functional class assessment tests (6-minute walk test and cardiopulmonary test), as well as imaging modalities (echocardiography and cardiac MRI). Study protocol details are described below.

SUBJECTS SELECTION
Inclusion Criteria:

Study cohort will include children and adolescents (age 2-11 and 12-18 years old) of both genders with established diagnosis of acute or chronic DCM (including both primary and secondary DCMs). In details inclusion criteria for the present analysis will be: presence of biventricular heart physiology, LV ejection fraction <50% and/or fractional shortening <25%, diagnosed by echocardiogram, and increased left ventricular end-diastolic diameter >2 standard deviations from the expected normal limit. Patients will be enrolled in the two clinical study centers (Ospedale Pediatrico Bambino Gesù, University College London).

Exclusion Criteria:
Patients will be excluded from the study, in the presence of one or more of the following: systemic hypertension (>95th percentile for age and height), persistent high rate supraventricular arrhythmias, pericardial disease (including restrictive and constrictive pericarditis), univentricular heart physiology, cor pulmonale.

STUDY DESCRIPTION

Patients will undergo clinical evaluation, laboratory testing, genetic testing, and diagnostic testing. All data will be collected at baseline and at one follow-up visit which will take place after 16 months (± 2 weeks), with the exclusion of genetic testing, which will be performed only once during the study in a subgroup of selected patients (see exclusion criteria in the genetic testing paragraph for details).
The study will include clinical evaluation, laboratory testing, genetic testing, and diagnostic testing, including functional class assessment tests (6-minute walk test and cardiopulmonary test) as well

as imaging modalities (echocardiography and/or cardiac MRI). Study protocol details are described below.

CLINICAL EVALUATION PROTOCOL

Clinical evaluation will be performed at baseline and during the follow-up visit it will include past medical history interview, clinical evaluation, standard diagnostic testing (X-rays and ECG) and laboratory testing. Clinical protocol details are specified below.

Medical history interview and clinical symptoms

- *ONSET*:
  - acute or chronic presentation (Y/N categorical variable)
    - IF CHRONIC: time from onset of the disease (months, continuous variable)
    - IF ACUTE: reason for referral to cardiologist (e.g. cardiomegaly at Chest RX and/or arrhythmias at ECG) (descriptive variable).
- *SYMPTOMS*:
  - cough (Y/N categorical variable),
  - poor feeding (Y/N categorical variable),
  - irritability (Y/N categorical variable),
  - pallor (Y/N categorical variable),
  - sweating (Y/N categorical variable).

- *COMORBIDITIES*:
  - neurological (Y/N categorical variable)
  - muscular disorder (Y/N categorical variable)
  - renal disease (Y/N categorical variable).

In all cases, known factors of myocardial damage will be evaluated. In details the three major causes known in paediatric population with DCMP: viral myocarditis, autoimmunity and genetic predisposition.

- *DEMOGRAPHICS:*
  - age (years, moths)
  - ethnic origin (categorical variable)
  - previous hospitalization (Y/N categorical variable, if yes also include number of episodes and dates when available).

- *TREATMENT:*
  - ongoing treatments (type [categorical variable] and dosage [continuous variable])
  - treatments taken and stopped within 3 months before type [categorical variable]

CLINICAL EXAMINATION:

Clinical parameters will be collected. In details:

- Symptoms severity (System will be applied to assess patients throughout the entire study period):

- NYHA or Ross classification as appropriate [categorical variable], according to the following tables:

Table 1. NYHA Class. Used for patients older than 6 years [Rosenthal et al, Journal of Heart and Lung transplantation, 2004]

Class I	Patients with cardiac disease but without resulting limitation of physical activity. Ordinary physical activities does not acute undue fatigue, palpitation, dyspnoea, anginal pain.
Class II	Patients with cardiac disease with slight limitation of physical activity. They are comfortable at rest. Ordinary physical activities results in fatigue, palpitation, dyspnoea, anginal pain.
Class III	Patients with cardiac disease resulting in marked limitation of physical activity. They are comfortable at rest. Less than ordinary physical activities results in fatigue, palpitation, dyspnoea, anginal pain.
Class IV	Patients with cardiac disease resulting in inability to carry on any physical activity without discomfort. Symptoms are present at rest. If any physical activity is undertaken, discomfort is increased .

Table 2. Ross Classification. [Ross et al, American Journal of Cardiology, 1987]

Class I	Patients with cardiac disease but without resulting limitation of physical activity. No limitation or symptoms, during feeding.
Class II	Mild tachypnoea or diaphoresis with feeding in infants. Dyspnoea on exertion in older children. No growth failure.
Class III	Marked tachypnoea or diaphoresis with feeds or on exertion. Prolonged feeding times. Growth failure form congestive heart failure.
Class IV	Tachypnoea, retractions, grunting or diaphoresis at rest.

- Assess stage severity, according to Guidelines (on the basis of the history, symptoms and ventricular function).

- Physical examination including:
  - Presence of enforced S3,S4, S2 (Y/N categorical variable)
  - Murmurs:
    - Mitral (Y/N categorical variable, if yes also add grading)
    - Tricuspid ((Y/N categorical variable, if yes also add grading)
  - Liver enlargement and tenderness (Y/N categorical variable, if yes also add grading)
  - Presence of pulmonary rales (Y/N categorical variable).

- CLINICAL PARAMETERS INCLUDING:
  - Systolic (SBP) and diastolic blood pressure (DBP) will be measured, using appropriate cuff sizes, three times while the subjects are seated after resting for 5 minutes. Measurements will be averaged for the analysis, according to the Fourth report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents (to the nearest mm of Hg, continuous variable)
  - Weight (in Kg, to the nearest 50g [continuous variable])
  - Height (in m, to the nearest 0.5cm [continuous variable])
  - Heart rate (bpm', [continuous variable])

- Respiratory rate (bpm', [continuous variable])
- Saturation (% [continuous variable]).

LABORATORY TESTING:

- Perform baseline laboratory blood samplings for:
  - Haematology:
    - red blood cell count,
    - haemoglobin
    - haematocrit
    - mean cell volume
    - white blood cell count
    - differential white blood cell count
    - platelet count
  - biochemistry:
    - sodium
    - potassium
    - chloride
    - creatinine
    - total proteins
    - albumin
    - alkaline phospatases
    - AST
    - ALT
    - glycemia
    - LDH
    - creatinin phosphokinase (CPK)
  - EGA
  - Metabolic disease screening:
    - Acylcarnitine
    - Aminoacidemia
    - Ammonemia
    - acid posphatases
    - IEF sialotransferrin
  - Urine analysis and organic aciduria
  - BNP dosing 0-

- Diagnostic testing:
  - 12-lead ECG
    - cardiac rhythm
    - heart rate (bpm')
    - PR interval (msec)
    - QRS duration (msec)
    - uncorrected QT duration (msec)
    - corrected QT duration applying the Bazett's formula (msec)
    - Presence of LVH (by either Sokolow-Lyon voltage criteria, or Cornell voltage-duration measurement).
  - Chest Xray for pulmonary congestion (yes/no)

- Holter ECG 24h (mean heart rate, mean daytime heart rate, mean nighttime heart rate, analysis of arrhythmias: supraventricular ectopic beats (yes/no); number (numeric variable), ventricular ectopic beats (yes/no; number (numeric variable), supraventricular tachyarrhythmia (Yes/no), ventricular tachyarrhythmia (yes/no)

ESTIMATION OF FUNCTIONAL CLASS AND CARDIOPULMONARY TEST

- Six minutes walking test (6MWT): the test will measure the distance that the patient can walk on a flat, hard surface in a period of 6 minutes. It evaluates the global and integrated responses of all the systems involved during exercise, including the pulmonary and cardiovascular systems, systemic circulation, peripheral circulation, blood, neuromuscular units, and muscle metabolism. Most patients do not achieve maximal exercise capacity during the 6MWT; instead, they choose their own intensity of exercise and are allowed to stop and rest during the test. Since most activities of daily living are performed at submaximal levels of exertion, the 6MWT reflects with accuracy the functional exercise level for daily physical activities.
  - Parameters recorded during the test:
    - The longer walked distance (meters)
    - Heart rate at the beginning of the test (rest) and maximal heart rate (bpm)
    - Blood pressure at rest and maximal blood pressure (mmHg)
    - Maximal speed (Km/h)
    - Presence/absence of symptoms, arrhythmias and electrocardiography findings.

- Cardiopulmonary test (CPX). CPX will be performed to analyze gas exchange at rest, during exercise, and during recovery and yield breath-by-breath (BBB) measures of oxygen uptake (VO2), carbon dioxide output (VCO2), and ventilation (VE). Advanced computerized systems will provide both simple (direct) and complex (integrated) analysis of data. Data will be integrated with standard variables measured during exercise testing, including heart rate, blood pressure, work rate, electrocardiography findings, and symptoms, to provide a comprehensive assessment of exercise tolerance and exercise responses. Exercise testing is performed on a treadmill following the Bruce Protocol.
  - Parameters recorded during the test:
    - Time of exercise (minutes) and work rate (METs)
    - Heart rate at the beginning of the test (rest) and maximal heart rate (bpm')
    - Blood pressure at rest and maximal blood pressure (mmHg)
    - Presence/absence of symptoms, arrhythmias and electrocardiography findings
    - Maximal oxygen uptake (VO2) and maximal carbon dioxide output (VCO2) in L/min and expressed as mL/Kg/min and Gas Exchange Rate (VO2/VCO2)
    - Ventilation (VE) in L/min.

GENETIC TESTING

Patients will undergo specific genetic investigations as DCM can be primary due to an underlying genetic defect.

Genetically based DCM is sub-classified into isolated DCM and DCM secondary to multisystemic genetic disorders, mostly muscular dystrophies. Isolated primary DCMP is rare in paediatric patients and recently numerous genes causative of familial DCMP have been identified. However, these gene mutations account for about 20 to 40% of CMPs. After exclusion of secondary causes and apparent absence of secondary multisystemic disorder, the majority of these patients are yet

underdiagnosed and difficult to characterise. DCMP secondary to muscular dystrophies can be familial and accurate history and first degree relatives physical examination and investigation might reveal more than one affected individual in the same family.
In the present study all patients will be evaluated by trained clinical geneticists. Metabolic investigations will be performed including baseline blood tests (as specified in the laboratory testing section). Apparently isolated DCM patients will be clinically evaluated by a neurologist, and, when indicated, specifically tested to exclude systemic neuromuscular disorders, as Duchenne and Becker muscular dystrophies and Barth syndrome. Clinical assessment will include: family history based on three generations, with specific enquiry about heart failure, sudden death, conduction disorders, stroke, muscular dystrophy and related anomalies, sensorineural deafness, muscle weakness; parental cardiovascular assessment, evaluation of muscle bulk and joint contractures for ruling out multisystemic muscular dystrophies.
The following conditions will be excluded from genetic study: myocarditis investigated by endomyocardial biopsy (viral, bacterial or fungal infections), nutritional deficiency (Kwashiorkor, pellagra, thiamine deficiency, selenium deficiency), collagen diseases (rheumatic fever, rheumatoid arthritis, systemic lupus erythematosus, dermatomyositis, Kawasaki disease), hematologic diseases (thalassemia, sickle cell disease, iron deficiency anemia), exposure to cardiotoxic drugs (anthracycline, cyclophosphamide, chloroquine, iron overload), endocrine disorders (hypothyroidism, hyperthyroidism, hypoparathyroidism, pheochromocytoma, hypoglycemia), and metabolic disorders except for Barth syndrome. Exclusion will include anomalous origin of left coronary artery from pulmonary artery.

Panel of genes that will be screened in these patients include the followings:

Gene Symbol	Protein Name	OMIM	% of FDC Caused by Mutations in This Gene[1]	Allelic Disorders
TNNC1	Troponin C, slow skeletal and cardiac muscles	191040	?	
PSEN1	Presenilin-1	104311	<1%	Early-onset Alzheimer disease
PSEN2	Presenilin-2	600759	<1%	Early- and late-onset Alzheimer disease
MYH6	Myosin-6	160710	?	FHC
PLN	Cardiac phospholamban	172405	?	
VCL	Vinculin	193065	?	
SGCD	Delta-sarcoglycan	601411	?	Delta sarcoglycanopathy (LGMD2F)[6]
ACTC1	Actin, alpha cardiac muscle 1	102540	<1%	FHC[5]
TAZ	Tafazzin	30094	? XL	Barth syndrome, endocardial fibroelastosis type 2, familial isolated non-compaction of the left ventricular myocardium

LDB3	LIM domain-binding protein 3	605906		?
DES	Desmin	125660	<1%	Desminopathy, Myofibrillar myopathy
TPM1	Tropomyosin alpha-1 chain	191010	?	FHC
TNNI3	Troponin I, cardiac muscle	191044	? AR	FHC, restrictive cardiomyopathy
TNNT2	Troponin T, cardiac muscle	191045	**2%-4%**	FHC
MYBPC3	Myosin-binding protein C, cardiac-type	600958	?	FHC
MYH7	Myosin-7	160760	**5%-8%**	Laing distal myopathy, FHC
LMNA	Lamin-A/C	150330	**7%-8%**	Partial lipodystrophy, CMT2B1, Emery-Dreifuss muscular dystrophy, Hutchinson-Gilford progeria syndrome, LGMD1B[6]
SCN5A	Sodium channel protein type 5 subunit alpha	600163	**2%-4%**	Long QT syndrome type 3, Brugada syndrome, idiopathic ventricular fibrillation, sick sinus syndrome, cardiac conduction system disease

Sample collection:
Samples will be collected among the 180 patients included in those with apparently idiopathic DCM, and after exclusion of secondary acquired causes of DCM.
After obtaining informed consent from patients or their legal representatives, blood samples will be collected from patients and their parents. Blood samples (in EDTA tubes) of 3-10ml will be conserved according to standard protocols at each participating hospital.
DNA will be extracted from blood with the QIAamp DNA Blood Mini/Midi Kit. Shipment of the samples according to standardized international rules will be carried out by the participating centers.

Next Generation Sequencing
This is divided in two parts. First screening of candidate genes for all patients with apparently isolated DCMP. Second in selected complex and familial cases exome sequencing will be applied.

NGS will use an opportunely Truseq custom panel designed to sequence simultaneously the CDS of the gene chosen for DCMP. Library samples will be prepared using the Truseq custom Amplicon Kit (Illumina). Sequencing will be performed on an Illumina multiplexed MiSeq sequencing platform. For each run, equimolar pools of libraries, integrated with indices to support up to 96 samples per run, will be sequenced using a paired-end 250 bp read length protocol to obtain high coverage (>98%) and depth (>100x). Data analysis from alignment to variant calling will be implemented using the MiSeq Reporter software (Illumina) and the Integrative Genomics Viewer (Broad

Institute). All identified mutations will be confirmed by sanger sequencing, same for unanalyzed coding regions.
Exome sequencing will be performed in selected unresolved cases, mainly familial and complex multiorgan phenotypes. Since 2005, next-generation DNA sequencing platforms have become widely available. This approach enables rapid sequencing of all exome sections of the genome, commonly referred to as exome sequencing, in a single reasonably priced experiment. The total size of the human exome is approximately 30 Mb which comprises approximately 1% of the entire human genome. The number of variants that are identified in exome sequencing studies varies greatly. Typically, between 20 000 and 50 000 variants are identified per sequenced exome. In order to reduce the number of false-positive calls, variants are first filtered based on quality criteria (e.g. variation can be excluded based on the predicted functional consequence of the variant by excluding dbSNP, non-coding and synonymous variants while missense and nonsense variants, coding indels and variants located within consensus splice sequences are maintained since they could potentially be pathogenic). It may also be possible to prioritize variants on the basis of the severity of their predicted effects. These steps usually reduces the number of potential candidate mutations by 90–95%. After this, between 150 and 500 private non-synonymous or splicesite variants are prioritized as potential pathogenic variants. By applying strategies for disease variant prioritization, an increasing number of studies have shown the successful application of exome sequencing to disease gene identification. The primary successes for exome sequencing have been in finding mutations that cause rare, familial forms of disease. It is estimated that application of Exome Sequencing achieves a success rate of 60–80% or Mendelian disorders.

CARDIAC IMAGING

In the presence of early DCM, the heart can preserve normal heart chamber systolic function undergoing a number of adaptive mechanisms, which include change in cardiac geometry as well as in subclinical cardiac systolic and diastolic parameters. These subtle abnormalities can often be detected in the early stages of the disease through diagnostic imaging, before the evidence of overt reduction in the ejection fraction and can be used to predict the development of the disease. Cardiac imaging will include Echocardiography and cardiac MRI.

- Echocardiography will be performed in all patients (including both acute and chronic DCM) and will be used to derive advanced measures of cardiac function including cardiac geometry, systolic and diastolic parameters. Three-dimensional echocardiography will be also used to evaluate systolic function and systolic synchronicity.
- Cardiac MRI will be performed in patients with chronic DCM (i.e. symptoms>three months) and will be used to evaluate cardiac volume and mass as well as transvalvular blood flows and cardiac fibrosis.

ECHOCARDIOGRAPHY:
All patients included in the study will undergo a complete echocardiographic examination. Abnormalities in cardiac shape and geometry will be assessed through measurements of left ventricular chamber dimensions and wall thickness. Analysis of cardiac function of both cardiac systole and diastole are routinely performed in clinical practice and monitored over time in order to evaluate the development o f the disease as well as the effect of therapy. Traditional indices of cardiac systolic function include fractional shortening, ejection fraction and midwall fractional

shortening. However, to more accurately evaluate the efficiency of cardiac contraction, traditional indices of systolic function can be integrated with the evaluation of cardiac wall stress derived from the analysis of both volume and pressure heart load. Furthermore analysis of contractile vectors, provide information to understand the interaction between changes in geometry and systolic dysfunction. Diastolic function is usually evaluated though the analysis of transmitral inflow velocities and myocardial relaxation velocities, which has also been shown to provide accurate estimates of cardiac filling pressure. Thus, cardiac function results from a complex interaction between heart geometry, ejection function, heart relaxation and interventricular dependence. In the presence of cardiomyopathies, abnormalities in one or more of these cardiac parameters can be observed.

Echocardiograms will be performed by expert sonographers on high-quality fully equipped commercially available echocardiograms. Exams will be performed in a dimly lit room with subjects examined with the head of the examining table elevated approximately 30° in a partial decubitus position maintained by using standard pillows. Recordings will be made from the subcostal view, parasternal acoustic window in both long- and short-axis views, and apical acoustic window to record two-, four-and five-chamber images, and color Doppler recordings. Exams will follow a standardized protocol (see below) and reviewed off-line by two independent readers in each study center using ad-hoc computerized review station with monitor screen overlay for performance of needed measurements.

ECHOCARDIOGRAPHIC IMAGES ACQUISITION PROTOCOL:

- **Parasternal long axis view:**
    - 2D harmonic (4 beats)
    - Color flow Doppler for evaluation of AI (4 beats)
    - Color flow Doppler for evaluation of MR (4 beats)
    - M-Mode of LV for measurements (50-75 mm/s display, at least 3 beats)
    - M-mode of left atrium at the level of aortic valve (50-100 mm/s display) (3 beats).

- **Parasternal short axis**
    - 2D harmonic at aortic level (4 beats)
    - 2D harmonic at mitral level (4 beats)
    - 2D harmonic at LV papillary muscle level (4 beats)
    - 2D harmonic at LV apical level (4 beats)
    - Color flow Doppler for evaluation of AI (4 beats)
    - Color flow Doppler for evaluation of MR (4 beats)
    - Color flow Doppler for evaluation of pulmonic valve (4 beats)
    - PW spectral Doppler at the level of the right ventricular out flow tract (50-75 mm/s display, at least 3 beats)
    - CW spectral Doppler for evaluation of PI (50-75 mm/s display, at least 3 beats)
    - Color flow Doppler of tricuspid valve (4 beats)
    - CW spectral Doppler for evaluation of TR (50-75 mm/s display, at least 3 beats).

- **Apical 4-chamber view**

- 2D harmonic acquired at held expiration (4 beats)
  - Avoid 2D foreshortening obtaining the longest and widest LV cavity
  - Optimal visualization of the endocardium
  - Exclude the papillary muscle
- M-mode of the tricuspid lateral wall plane (TAPSE)
- M-mode of the mitral lateral wall plane (MAPSE)
- Color flow Doppler for the evaluation of MR (4 beats)
- PW spectral Doppler at the trans-mitral flow velocity with sample volume at the mitral tips (50-75 mm/s display, at least 3 beats)
- CW spectral Doppler of mitral valve for evaluation of MR (4 beats)
- Color flow Doppler for the evaluation of TR (4 beats)
- CW spectral Doppler for evaluation of TR (50-75 mm/s display, at least 3 beats)
- Doppler tissue imaging myocardial velocity mitral annulus (Inter-ventricular septum, lateral wall, right ventricular free wall at the tricuspid annular level) freeze spectral (50-75 mm/s display, at least 3 beats).

- **Apical 3 or 5-chamber view**

  - 2D harmonic (4 beats)
  - PW spectral Doppler of the LVOT (50-75 mm/s display, at least 3 beats)
  - CW spectral Doppler of the aortic valve (50-75 mm/s display, at least 3 beats)
  - Color flow Doppler for evaluation of AI (4 beats)

- **Apical 2-chamber view**

  - 2D harmonic (4 beats)
  - Color flow Doppler for evaluation of MR (3 beats)

- **Subcostal view**

  - 2D harmonic or M-mode of the inferior vena cava asking the subject to breathe (4 beats)

- **Three dimensional imaging:**

  - From the apical 4-chamber view one ECG-triggered real time complete full volume acquisition including LV and LA.

- **Speckle tracking:**

  - For speckle tracking 4 beats ECG-triggered image acquisition of the apical view (4-chamber, 2- chamber, 5- chamber view) and parasternal short axis (mitral level, LV papillary muscle level, apical level).

Detailed echocardiographic variable list:

- Left ventricular geometry:
    - LV diameter (mm)
    - LV wall thickness (mm)
    - LV mass (g)
    - Sphericity index (ratio).

- Left atrial geometry:
    - LA diameters (mm)
    - LA volume (mL)

- Mitral valve geometry and function:
    - MV diameter (mm)
    - Trans-valvular velocities, peak and mean gradients (cm/s; mmHg)
    - Mitral valve regurgitation (categorical variable)

- Left ventricular function (including measures of cardiac preload and afterload):
    - Systolic:
        - End-systolic and end-diastolic volumes (ml)
        - Endocardial shortening (%)
        - Ejection fraction (%)
        - Stroke volume and cardiac output (ml)
        - Valvular annulus displacement for tricuspid valve (mm)
        - Right ventricular fractional area change (%)

    - Global and segmental contractility:
        - Cardiac global longitudinal, radial and circumferential strain (%).
        - Cardiac rotation of LV base and apex (degrees).
    - Diastolic:
        - Transmitral Doppler velocities, time and ratio. (cm/s)
        - Myocardial tissue Doppler velocities (cm/s)

- Ventricular systolic synchronicity:
    - Interventricular:
        - Interventricular mechanical delay (IVMD, msec)
    - Intraventricular (LV)
        - 3D Volume synchronicity (3DSDI, %)

Cardiac Magnetic Resonance imaging (CMR):

CMR sequences will be performed in all patients with chronic DCM (i.e. symtoms > three months). Contrast-enhanced MR images will allow depicting and quantifying myocardial inflammation, infiltration, and fibrosis. Black-blood fast spin-echo MR images will be used for the morphologic assessment of the heart with high spatial resolution and T2-weighted MR images for the evaluation of the acute myocardial edema. Flow mapping technique will allow assessing qualitatively and quantitatively flow volumes, velocities, and flow fractions in any oblique cardiac plane of any valvular heart disease and calculation of the stroke volumes from aortic and pulmonary arteries. CMR exam will be focused on cine images, used for qualitative evaluation of

regional and global systolic function in two-chamber, four-chamber and short-axis and for a quantification of chamber volumes and myocardial function (obtained by a stack of short-axis sections from the mitral annulus to the apex). Short-axis sections will be analyzed for measurements of end diastolic and systolic volumes and cardiac mass. Black-blood fast spin-echo MR images will also be obtained for the morphologic assessment of the heart and T2-weighted MR images for the evaluation of the acute myocardial edema. Late-gadolinium-enhanced images will show the difference between viable and nonviable myocardium with the overall and predominantly spatial distribution of the enhancement (subepicardial, midwall, or subendocardial).

CMR IMAGES ACQUISITION PROTOCOL:

- ***Equipment***
    1. *MR scanner at 1.5 T Static Magnetic Field*
    2. *Cardiac package including post-processing tools*
    3. *Monitoring equipment (blood pressure, electrocardiogram for monitoring of cardiac rhythm, intercom to communicate with patient)*
    4. *Defibrillator*
    5. *Drugs for emergency treatment*
    6. *Anesthetic Room for scan performed in General Anesthetic*
- ***Preparation***
    1. *Obtain metal check list and consent form for Godolinium iv administration*
    2. *Place Peripheral iv line access possibly at antecubital vein*
    3. *Place ECG electrodes according to MR scanner properties and recommendations*
    4. Fast and secure the patient in the scanner

- ***Acquisition CMR Protocol (based on SCMR Standards protocol*** *http://www.scmr.org/navigation/CMR-in-specific-circumstances.html) also available as Philips Exam Card*
    1. *Scout imaging – transtransaxial, coronal, sagittal*
    2. *Cardiac planes planned with Interactive Real Time sequence (if available)*
    3. *Single-shot Axial stacks PD or T1 Black Blood TSE*
        a. *8 mm slice thickness*
        b. *2 mm gap or distant factor 25%*
        c. *triggered in end-diastole*
    4. *Ventricle Structure and Function module*
        a. Balanced SSFP cine or single shot
            i. Vertical long axis prescribed orthogonal to transaxial scouts aligned through the apex and center of the mitral valve
            ii. Horizontal long axis aligned orthogonal to the vertical long axis, passing through the apex and center of the mitral valve
        b. Balanced SSFP short axis (SA) cine images, from the mitral valve plane through the apex. The basal most short axis slice should be located immediately on the myocardial side of the atrioventricular junction at enddiastole prescribed from the previously acquired long axis cines.
            i. Slice thickness 5-6 mm, with 2 mm interslice gaps to equal 7-8 mm.
            ii. Temporal resolution ≤ 45 ms between phases
            iii. TE shortest

iv. Parallel imaging used as available

c. Balanced-SSFP long axis cine images 7 mm slice thickness.

i. The 4 chamber (4CH) long axis is prescribed from the vertical long axis through the apex and center of the mitral and tricuspid valves. This can be cross-checked on basal short axis cines, using the costophrenic angle (margin) of the RV free wall.

ii. Vertical long axis or 2CH, prescribed from the scout already acquired

iii. LV outflow tract (LVOT) long axis, passing through the apex, the center of the mitral valve and aligned with the center of LVOT to aortic valve, as seen on a basal short axis cine.

iv. LVOT cross-cut in coronal (optional)

d. RV Long axis images should include an RV vertical long axis view aligned with tricuspid inflow and a RV outflow tract view (sagittal plane through the pulmonary valve).

e. PD or T1 weighted short-axis BB TSE 8 mm ECG triggered in diastole, same views as planned for cine imaging at least 10 slices

5. Tagging (optional)

a. echo-planar imaging sequence GRE with spatial modulation of magnetization (CSPAMM) with systolic and diastolic grid-tags within a single breath-hold
b. 3 short axis stacks (apex, mid-ventricular, base)
c. VLA and 4CH views

6. T1 Mapping pre contrast (optional)

Mid-ventricular Short axis view TI Scout (possibly modified look-locker inversion recovery sequence, if available)

1. Perfusion (optional)

a. Saturation-recovery imaging with gradient echo-echo planar (GRE-EPI) hybrid, GRE, or SSFP readout
b. Short-axis view imaging (at least 3 slices per heart beat)
   i. Slice thickness 8 mm c. Parallel imaging, 2-fold acceleration;
   ii. In-plane resolution < 3 mm;
   iii. Readout temporal resolution ~100 – 125 ms or shorter as available;
   iv. Contrast is given (0.05 – 0.1 mmol/kg, 3–7 ml/s) followed by at least 30 ml saline flush (3–7 ml/sec);
   v. Breathhold starts during early phases of contrast infusion before contrast reaches the LV cavity;
   vi. Image for 40–50 heart beats by which time contrast has passed through the LV myocardium.

2. Flow (2D Phase Contrast Velocity Encoding)

a. Through-plane perpendicular to the ascending aorta (AO)

b. Through-plane perpendicular to main pulmonary artery (MPA)

c. Through-plane aligned to mitral valve annulus (AVV)

d. Free breathing (at least 3 NSA)

e. VENC set at 150-200

3. Late gadolinium enhancement (LGE)

a. Need at least 8-10 minute wait after gadolinium injection (0.1–0.2 mmol/kg). Note – The delay may be shorter than 10 minutes if lower doses are used as blood pool signal falls below that of late enhanced myocardium.

b. 2D segmented inversion recovery GRE imaging during diastolic stand-still

c. Same views as planned for cine imaging (short- and long-axis views)
d. Slice thickness, same as for cine imaging
e. In-plane resolution, ~1.4–1.8 mm
f. Acquisition duration per R-R interval below 200 ms but should be less in the setting of tachycardia.
g. Inversion time set to null normal myocardium. Alternative is to use fixed TI with a phase-sensitive sequence.
h. Read-out is usually every other heart beat but should be modified to every heart beat in the setting of bradycardia, and every third heart beat in the setting of tachycardia or arrhythmia.

4. T1 Mapping (diffuse fibrosis)

Mid-ventricular short axis view TI Scout (possibly modified look-locker inversion recovery sequence, if available at 15 min post-contrast (optional)

- ***Standard Analysis***
  - ***1.*** *Ventricle analysis*
    a. All short axis images are evaluated with analysis packages for countouring endocardial and epicardial borders at end-diastole and end-systole.
    b. The inclusion or exclusion of papillary muscles in the LV mass should be the same as that used in normal reference ranges used for comparison.
    c. Care must be used at the 1 or 2 most basal slices. Due to systolic movement of the base towards the apex in normally contractile ventricles, the end-systolic phase will include only left atrium. This may not be the case in a severely dysfunctional LV. Either way, this slice at enddiastole will include LV mass and volume.
  - **2.** Flow analysis
    a. Contouring of vessel edges (AO,MPA) over the entire cardiac cycle from the Magnitudo Sequence.
    b. Carefully contour AV inflow and outflow edges to exclude amount of flow from the LVOT and RVOT
  - **3.** Late enhancement
    a. Interpret visually using AHA 17-segment model.
    b. Estimate area (mean transmural extent) of enhancement within each segment (0%, 1–25%, 26–50%, 51–75%, 76–100%).
    c. Quantify LGE in g

<u>Detailed CMR variable list:</u>

A) Left ventricular (LV) analysis:
   a. LV EDV (ml)
   b. LV ESV (ml)
   c. LV SV (ml)
   d. LV CO (l/min)
   e. LV EF (%)
   f. LV diastolic mass (g)
   g. LV diastolic longitudinal diameter (mm)
   h. LV diastolic transverse diameter (mm)
   i. Mitral valve annulus diameters 4CH 2CH (mm)

j. LV Fibrosis (g)
k. LV Fibrosis/myocardial mass (%)

B) Right ventricular (RV) analysis:
   a. RV EDV (ml)
   b. RV ESV (ml)
   c. RV SV (ml)
   d. RV CO (l/min)
   e. RV (EF %)

C) Left atrium analysis:
   a. End-Systolic (ES) atrial area 4CH (mm)
   b. ES atrial area 2CH (mm)
   c. ES atrial orthogonal diameter 4CH 2CH (mm)

D) Transvalvular flow analysis:
   a. Aortic (AO) FF (ml)
   b. AO BF (ml)
   c. AO NF (ml)
   d. AO Vmax (m/s)
   e. AO RF (%)
   f. Mean Pulmonary artery (MPA) FF (ml)
   g. MPA BF (ml)
   h. MPA NF (ml)
   i. MPA Vmax (m/s)
   j. MPA RF (%)
   k. Mitral valve (MV) FF (ml)
   l. MV BF (ml)
   m. MV NF (ml)
   n. MV Vmax (m/s)
   o. MV RF (%)

SUBGROUP OF PATIENTS ENTERING CARDIAC TRANSPLANT LIST

It can be predicted that a number of patients with DCM (expected 10%), during the study project timeframe will enter the cardiac transplant list due to worsening of their medical condition. In these patients standard clinical procedure requires a number of medical testing, including cardiac catheterization for the evaluation of intracardiac pressures. For the purpose of the present study no change in the standard practice of care for these patients will be made. To guarantee a complete set of information for these patients, data derived from cardiac catheterization performed for clinical purposes in the patients included in the study, will be merged to the patients' records.

4. WITHDRAWAL FROM THE STUDY

Participation in this study is entirely voluntary. The parents may decide, at any time, that they do not wish their child to take part in the study without altering current or future treatment in any way.
If at any stage of the project the parents wish to withdraw their child from the study, the researcher responsible for the study will arrange for their child's information to be immediately removed from computer records and for any remaining samples we hold belonging to him/her to be destroyed.

6. STUDY PLANNING

6.1 EFFICACY PARAMETERS

Clinical, immunological genetic and imaging data will be recorded at each study point. This data will be gathered and stored in a standardized manner building upon the Health-e-Child software tools which will be extended for the purpose of integrating model information related to a wider range of diseases. The tools to be developed will also include the aspect of a multidimensional longitudinal analysis that yields the opportunity to identify potential new outcome measures (imaging or biological biomarkers) for the assessment of treatment efficacy. Furthermore, the prognostic value on an individual level of multidimensional data, including modern imaging modalities, genetic and meta-genetic data will be explored through the development and integration of appropriate data clustering methods.

6.2 EXPERIMENTAL DESIGN

The study will last 4 years. It is designed as a prospective longitudinal study.

6.3 DATA PROTECTION

This kind of this project requires that a substantial amount of personal data, including genetic information, be collected from the participants and shared across a network.The project will be carried out in accordance with the applicable European and National data privacy protection laws and regulations. All data will be gathered in an anonymous form so that no data may be traceable to a patient other than by the local treating clinicians. Only the respective hospitals will have access to the key of re-identification. Therefore, no project partner or other third party outside the respective hospitals involved, will have access to the identifiable patient data. Furthermore, only anonimyzed data will be processed or used in the project.
This information, handled in an anonymous manner, will be granted to regulatory authorities for regular reviews of clinical study procedures and/or data, in order to protect child's privacy.

7. SECURITY EVALUATION

7.1 DEFINITIONS

No adverse effects are foreseen as consequence of the clinical study.

8. SAMPLE DIMENSION AND STATISTIC METHODOLOGY

8.1 STATISTIC DESIGN

As the main goal of whole project is to establish a data repository for pediatric diseases, the sample size has been set by taking into account primary endpoints and study power but also available resources at each center, and study feasibility. In particular for the genetic analysis no study power is foreseen.

Cardiomyipathy modeling will be done by the following partners: Siemens AG, Institut Nationale de Recherche en Informatique et en Automatique (INRIA) and Siemens Corporate Research.

8.2 MANAGEMENT OF MISSING DATA

Patients whose information is missing or incorrectly transcribed will not be used for the analysis, but will be marked appropriately in the electronic database by inserting a variable "Flag".

9. AMINISTRATIVE AND ETHICAL PROCEDURES

Confidentiality

Clinical data will be acquired as required by each partner's national law.
At each clinical center patient's data will be collected and stored as electronic files and will be accessible by the responsible research personnel. Access to data will be granted using their personal credentials. Access to the file will be protected and the log of the user who performed the operation will be required at regular intervals. The data manager will perform regularly a data backup.

Data publication and final report

The ownership of scientific data will be shared between all the partners involved in the Project. The WP leader, Dr. Gabriele Rinelli, and researchers who will conduct the study, will endeavor to promote the dissemination of the results through the project website, communications in national and international scientific meetings, publication in international journals of high scientific profile. The dissemination and publication of the results by the experimenters will be promoted in accordance with the provisions in force concerning the confidentiality of sensitive data. In all scientific publications the efforts of all researchers will be recognized.

All health professionals involved in the project will seek to minimize the physical and psychological discomfort caused to patients and parents from participating in this study. In order to ensure the well-being, they will not be notified in any way about the personal results of genetic investigations.

9.1 AUTORISATIONS

The protocol will undergo the approval of the Ethical Committee for the study implementation before the enrollment of the patients.

9.2 INFORMED CONSENT

Each parent/patient who will be asked for his/her enrollment in the study will be informed on every aspect concerning the study itself and, before the recruitment begins, he/she must sign the informed consent model. The date of the signature of the consent will be recorded on the CRF. A Copy of the informed consent model must be handed to the parent/patient.

9.3 INSURANCE COVERAGE

Insurance coverage used is as foreseen by each research structure for clinical and research activities. ***9.4 USE OF THE INFORMATION AND DATA PUBBLICATION***

Researchers must have the right to disseminate/publish the results of the study, with respect to the current legislation on sensible data protection and Intellectual property Rights. European Commission does not have any right in the publication/dissemination of study results

9.5 CLINICAL PROTOCOL AMENDMENTS

Clinical protocol amendment is intended to be any modifications to the final version of the clinical protocol that, after being approved by the Researcher and by the Ethical Committee, cannot be informally modified.

9.6 DATA MANAGEMENT AND DOCUMENT PRESERVATION

Researchers will manage the preservation of patients identification codes for at least 15 years after the end of the study. Patients clinical data and other original data will be kept for the maximum period allowed by the Hospital (min. 15 years).

9.7 BUDGET

The project costs are covered by the specific Grant by the European Commission. Such grant will cover 75% of staff costs, of consumables and equipment used for the project, and of all other expenses involved. Annex 1 (GPF) contains the project's entire budget, including details of the Institution's budget.

10. RESEARCHER RESPONSIBILITY

Researcher, as principle investigator of the study declares to be aware of all the duties committed to him/her and his/her collaborators for the conduction of the study. Except where explicitly edited, the term "researcher" on this Protocol and on the CRF is referred to the researcher or to any other person appointed by him/her, who will have power to sign documents instead of him/her.

The researcher will conduct the study according to the principles of the Helsinki declaration (1964) and Amendments (Appendix 1), and to the European Guidelines on Clinical Good Practices.

11. Annexes

Annex 1 GPF
Annex 2 DOW

12. REFERENCES

1. Andrews RE, Fenton MJ, Ridout DA, Burch M; British Congenital Cardiac Association. New-onset heart failure due to heart muscle disease in childhood: a prospective study in the United Kingdom and Ireland. Circulation 2008; 117(1):79-84
2. B. Georgescu, S. Zhou, D. Comaniciu, A. Gupta: Database-Guided Segmentation of Anatomical Structures with Complex Appearance, IEEE Int'l Conf. Computer Vision and Pattern Recognition, San Diego, CA, Vol. II, 429-436, 2005.

3. S.K. Zhou, F. Guo, J. Park, G. Carneiro, J. Jackson, M. Brendel, C. Simopoulos, J. Otsuki, D. Comaniciu: A Probabilistic, Hierarchical and Discriminant Framework for Rapid and Accurate Detection of Deformable Anatomic Structures, IEEE Int'l Conf. Computer Vision, Rio de Janeiro, Brazil, 2007
4. S.K. Zhou, D. Comaniciu: Shape Regression Machine. In: N. Karssemeijer and B. Lelieveldt (Eds.): Information Processing in Medical Imaging (IPMI) 2007, Springer LNCS 4584, pp. 13–25, 2007.
5. J. Park, S.K. Zhou, J. Otsuki, C. Simopoulos, D. Comaniciu: Automatic Cardiac View Classification of Echocardiogram, IEEE Int'l Conf. Computer Vision, Rio de Janeiro, Brazil, 2007
6. L. Yang, B. Georgescu, Y. Zheng, D.J. Foran, D. Comaniciu: A Fast and Accurate Tracking Algorithm of Left Ventricles in 3D Echocardiograpy, IEEE Int'l Symposium on Biomedical Imaging, Paris, France, 2008.
7. R. Socher, A. Barbu, D. Comaniciu: A Learning-Based Hierarchical Model for Vessel Segmentation, IEEE Int'l Symposium on Biomedical Imaging, Paris, France, 2008.
8. J. Zhang, S.k. Zhou, L. McMillan, D. Comaniciu: Discriminative Learning for Deformable Shape Segmentation: A Comparative Study, European Conf. Computer Vision, Marseille, France, 2008
9. J.H. Park, S.K. Zhou, D. Comaniciu: Automatic Mitral Valve Inflow Measurements from Doppler Echocardiography, Int'l Conf. Medical Image Computing and Computer Assisted Intervention, (MICCAI'08), New York, NY 2008
10. Y. Zheng, B. Georgescu, H. Lingm S.K. Zhou, M. Scheuering, D. Comaniciu: Constrained Marginal Space Learning for Efficient 3D Anatomical Structure Detection in Medical Images, IEEE Int'l Conf. Computer Vision and Pattern Recognition, Miami, Florida, 2009
11. Y. Zheng, B. Georgescu, D. Comaniciu: Marginal Space Learning for Efficient Detection of 2D/3D Anatomical Structures in Medical Images, Information Processing in Medical Imaging (IPMI), Williamsburg, VA, 2009. Oral
12. Nordsletten, DA and Niederer, SA and Nash, MP and Hunter, PJ and Smith, NP. Coupling multiphysics models to cardiac mechanics. Progress in Biophysics and Molecular Biology. 104(1-3), 77-88, 2011
13. R. Clayton and A. Panfilov. A guide to modelling cardiac electrical activity in anatomically detailed ventricles. Progress in biophysics and molecular biology, 96 (1-3):19–43, 2008.
14. K. Costa, J. Holmes, and D. McCulloch. Modelling cardiac mechanical properties in three dimensions. Philosophical Transactions A, 359(1783):1233, 2001.
15. S. Niederer and N. Smith. An improved numerical method for strong coupling of excitation and contraction models in the heart. Progress in biophysics and molecular biology, 96(1-3):90–111, 2008.
16. Chapelle, D. and Le Tallec, P. and Moireau, P. and Sorine, M., An energy-preserving muscle tissue model: formulation and compatible discretizations, International Journal of Multiscale Computational Engineering, 2011
17. T. Mansi, B. André, M. Lynch, M. Sermesant, H. Delingette, Y. Boudjemline, N. Ayache : Virtual Pulmonary Valve Replacement Interventions with a Personalised Cardiac Electromechanical Model. In Recent Advances in the 3D Physiological Human. Springer, 2009
18. M. Sermesant, R. Chabiniok, P. Chinchapatnam, T. Mansi, F. Billet, P. Moireau, J.M. Peyrat, K. Wong, J. Relan, K. Rhode, M. Ginks, P. Lambiase, H. Delingette, M. Sorine, C.A. Rinaldi, D. Chapelle, R. Razavi, and N. Ayache. Patient-specific electromechanical models of the heart for the prediction of pacing acute effects in CRT: A preliminary clinical validation. Medical Image Analysis, 2012. Note: In press.
19. Delingette H, Billet F, Wong KC, Sermesant M, Rhode K, Ginks M, Rinaldi CA, Razavi R, Ayache N. Personalization of cardiac motion and contractility from images using variational data assimilation. IEEE Trans Biomed Eng. 2012 Jan;59(1):20-4. Epub 2011 Jun 27. PMID: 21712158

20. Chabiniok R, Moireau P, Lesault PF, Rahmouni A, Deux JF, Chapelle D. Estimation of tissue contractility from cardiac cine-MRI using a biomechanical heart model. Biomech Model Mechanobiol. 2011 Jul 28. [Epub ahead of print] PMID 21796413

21. M. Sermesant, H. Delingette, and N. Ayache. An electromechanical model of the heart for image analysis and simulation. IEEE TMI, 25(5):612–625, 2006.

22. T. Mansi, V. Mihalef, P. Sharma, B. Georgescu, X. Zheng, S. Rapaka, A. Kamen, D. Mereles, H. Steen, B. Meder, H. Katus, D. Comaniciu. Data-Driven Computational Models of Heart Anatomy, Mechanics and Hemodynamics: an Integrated Framework. In IEEE ISBI conference, 2012.

23. P. Reymond, Y. Bohraus, F. Perren, F. Lazeyras, and N. Stergiopulos, "Validation of a patientspecific 1-D model of the systemic arterial tree", American Journal of Physiology, published online, doi: 10.1152/ajpheart.00821.2010

24. M. Olufsen, C. Peskin, W.Y. Kim, and E. Pedersen, "Numerical Simulation and Experimental Validation of Blood Flow in Arteries with Structured-Tree Outflow Conditions", Annals of Biomedical Engineering, vol. 28, pp. 1281–1299, 2000

25. Michael Markl PhD, Julia Geiger MD, Bernd Jung PhD, Daniel Hirtler MD, Raoul Arnold MD, Noninvasive evaluation of 3D hemodynamics in a complex case of single ventricle physiology, JMRI, 2012

26. Griffith B., Luo X., McQueen D., Peskin C.. Simulating The Fluid Dynamics Of Natural And Prosthetic Heart Valves Using The Immersed Boundary Method, 2009, International Journal of Applied Mechanics, vol. 1, no. 1, pp. 137-177.

27. Schenkel, T, Mauro M, Reik M, Markl M, Jung B, Oertel H. MRI-Based CFD Analysis of Flow in a Human Left Ventricle: Methodology and Application to a Healthy Heart, 2009, Annals of Biomedical Engineering, vol 37. No. 3, pp.503-515.

28. S. Kulp, M. Gao, S. Zhang, Z. Qian, S. Voros, D. Metaxas, and L. Axel. "Using High Resolution Cardiac CT Data to Model and Visualize Patient-Specific Interactions Between Trabeculae and Blood Flow." Proceedings of MICCAI 2011.

29. Saber N., Wood N., Gosman A. D., Merrifield R., Yang G.-Z., Charrier C., Gatehouse P., Firmin D. - Progress Towards Patient-Specific Computational Flow Modeling of the Left Heart via Combination of Magnetic Resonance Imaging with Computational Fluid Dynamics, Annals of Biomedical Engineering, 2003, 31(1), 42-52.

30. Mihalef2011: V. Mihalef, R. Ionasec, P. Sharma, B. Georgescu, M. Suehling, D. Comaniciu: Patient-Specific Modeling of the Whole Heart Anatomy, Dynamics and Hemodynamics from 4D Cardiac CT Images, Interface Focus Journal of the Royal Society, 2011

31. Ralovich K., Mihalef V., Sharma P. , Itu L., Vitanovski D., Ionasec R., Suehling M., Everett A.,Pongiglione G., Navab N., and Comaniciu D., Modeling and Simulation Framework for Hemodynamic Assessment of Aortic Coarctation Patients, ISMRM 2012

32. Itu, L., Sharma, P., Gulsun, M., Mihalef, V., Kamen, A., and Greiser, A., "Determination of timevarying pressure field from phase contrast MRI data", Journal of Cardiovascular Magnetic Resonance, 14 (Suppl 1), 2012.

33. Itu, L., Sharma, P., Mihalef, V., Kamen, A., Suciu, C., Comaniciu, D., "A Patient-Specific Reduced- Order Model For Coronary Circulation", IEEE International Symposium on Biomedical Imaging, Barcelona, Spain, May 2012.

34. T. Mansi, V. Mihalef, P. Sharma, B. Georgescu, X. Zheng, S. Rapaka, A. Kamen, D. Mereles, H. Steen, B. Meder, H. Katus, D. Comaniciu. Data-Driven Computational Models of Heart Anatomy, Mechanics and Hemodynamics: an Integrated Framework. In IEEE ISBI conference, 2012.

35. Benson, D. Computational methods in lagrangian and eulerian hydrocodes. Comput. Meth. In Appl. Mech. and Eng. 99, 235–394, 1992.

36. Kristin McLeod, Alfonso Caiazzo, Miguel A. Fernández, Tommaso Mansi, Irene E. Vignon- Clementel, Maxime Sermesant, Xavier Pennec, Younes Boudjemline, and Jean-Frederic Gerbeau. Atlas-Based Reduced Models of Blood Flows for Fast Patient-Specific Simulations. In Proc. MICCAI Workshop on Statistical Atlases and Computational Models of the Heart: Mapping Structure and Function + a Cardiac Electrophysiological Simulation Challenge (STACOM+CESC'10), volume 6364 of LNCS, Beijing, pages 95-104, September 2010. Springer.

37. Tommaso Mansi, Ingmar Voigt, Benedetta Leonardi, Xavier Pennec, Stanley Durrleman, Maxime Sermesant, Hervé Delingette, Andrew M. Taylor, Younes Boudjemline, Giacomo Pongiglione, and Nicholas Ayache. A Statistical Model for Quantification and Prediction of Cardiac Remodelling: Application to Tetralogy of Fallot. IEEE Transactions on Medical Imaging, 9(30):1605-1616, September 2011

38. Kristin McLeod, Tommaso Mansi, Maxime Sermesant, Giacomo Pongiglione and Xavier Pennec. Statistical Shape Analysis of Surfaces in Medical Images Applied to the Tetralogy of Fallot Heart. In Mathematical Problems in Computational Biology and Biomedicine, F. Cazals and P. Kornprobst (Eds), Springer 2012. To appear.

39. Stefan Sommer, Mads Nielsen, and Xavier Pennec. Sparsity and Scale: Compact Representations of Deformation for Diffeomorphic Registration. In IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 2012), Breckenridge, Colorado, USA, January 2012.

WP 4: Risk of cardiovascular disease in obese children and adolescents

Protocol no:

MD PAEDIGREE **WP 4Version 2:** April 15, 2013**CONFIDENTIAL**

Protocol no.:	MD PAEDIGREE **WP 4**
Title:	Data acquisition and processing for the estimation of CVD risk in obese children
Acronym:	MD-Paedigree – WP 4
Multicentric/Monocentric Study	Multicentric
Principal Investigator	Dr. Melania Manco
Sponsor:	Bambino Gesù Children'S Hospital (BGCH)
Responsible Work Package 4 :	Prof. Prof. Andrew Taylor (University College London)
Scientific Coordinator of the Project	Prof. Bruno Dalla Piccola(Ospedale Pediatrico Bambino Gesù)
Data Management/Statistical analysis:	Dr. Michael Suehling - Siemens

Protocol approved and signed by:

Scientific Coordinator of the Project

Prof. Bruno Dallapiccola

Responsible Work Package:
Prof. Andrew Taylor

Principal Investigator:
Dr. Melania Manco

Acronym List

AEs	Adverse Events
EC	Ethical Committee
CRF	Case Report Form
GCP	Good Clinical Practice
ICH	International Conference on Harmonisation
IRB/IEC	Institutional Review Board/Independent Ethics Committee
NSAEs	Non Serious Adverse Events
SAEs	Serious Adverse EventsSOPs Standard Operating Procedures
……	……..

INDEX

1 INTRODUCTION

1.1 GENERAL BACKGROUND OF THE PROJECT

MD – PAEDIGREE *Model-Driven European Paediatric Repository* is a research project funded by the European Commission under the Virtual Physiological and Human Area (VPH) of the ICT Theme of the Seventh Framework Programme (contract no. 600932). The project foresees 7 world-renowned clinical centres of excellence pursuing improved interoperability of paediatric biomedical information, data and knowledge, by developing a set of reusable and adaptable multi-scale models for more predictive, individualised, effective and safer paediatric healthcare. The project is scientifically and technologically supported by one of the leading industrial actors in medical applications in Europe, and operating in conjunction with highly qualified SMEs and some of the most experienced research partners in the Virtual Phisiological Human (VPH) community.

MD-Paedigree validates and brings to maturity patient-specific computer-based predictive models of various paediatric diseases (cardiomyopathy, risk of cardiovascular disease in obese children and adolescents, Juvenile Idiopathic Arthritis (JIA) and neurological and neuro-muscular diseases (NND), thus increasing their potential acceptance in the clinical and biomedical research environment by making them readily available not only in the form of sustainable models and simulations, but also as newly-defined workflows for personalised predictive medicine at the point of care.

MD-Paedigree aims to advance the state-of-the-art of patient-specific computational modelling of different paediatric diseases and translate the latest advances into clinics to improve disease understanding, therapy outcome, and provide an infostructure platform for assessing new therapies at the point of care.

MD-Paedigree's goals, therefore, are to integrate and share highly heterogeneous biomedical information, data and knowledge, using best, jointly develop reusable, adaptable and composable multi-scale VPH workflow models and to support evidence-based translational medicine at the point of care practices from the biomedical semantic Web.

Ilness in infants, children, and adolescents are a large and under-appreciated public health problem. Because paediatric patients now have a much longer life expectancy, the burden and costs are substantial for families and society. For instance, recent studies have shown that the number of adult patients with congenital heart disease is already similar to that of the paediatric population and will continue to grow. It can therefore be stated that in the longer term the VPH scientific approach requires a fundamental research investment in paediatrics, where the conceptual revolution which is underway, transforming the nature of healthcare from reactive to preventive, can best be applied, moving towards an approach based on personalised, predictive, preventive, and participatory (P4) medicine, increasingly focused on wellness.

MD-Paedigree represents a major step towards personalised paediatric e-Health, based on data-driven models, patient-specific simulations and a sustainable data and model repository. The

project can in fact impact the way healthcare is practiced in the future. MD-Paedigree's impact on dealing with very concrete and exemplary clinical cases is demonstrated through the following applications in the diseases scenarios. MD-Paedigree encompasses complete services for storage, similarity search, outcome analysis, risk stratification, and personalised decision support in paediatrics within its innovative model-driven data and workflow-based models repository, leveraging on service and knowledge utilities. It fosters the state-of-the-art of patient-specific computational modelling of the selected diseases and translates the latest advances into clinics to improve disease understanding and provide a platform for testing new therapies at the point of care. In fact, MD-Paedigree demonstrates how a dedicated VPH repository can provide full accessibility to existing and further developing knowledge in paediatrics to all interested bio-medical researchers precisely through linking data with models, thus proving the large scale benefits of having both the data and models readily available at the point of care.

The data collection performed within Health-e-Child and Sim-e-Child projects (already funded by the European Commission in previous calls for proposals) is still available to the MD-Paedigree consortium thanks to the continuity in the eHealth platform which is shared by all three projects. In addition, the new patients' recruitment to be performed within MD-Paedigree consists of:

<table>
<tr><th>Pathology</th><th colspan="2">No of patients/Time</th></tr>
<tr><td>Cardiomyopathies</td><td>180 children, by month 33: 60 patients (among which 30 girls) for each clinical centre,90 for BGCH</td><td rowspan="3">Genetic and meta-genomic:

180 patients with cardiomiopathies, 180 with CVD risk in obesity, 200 with JIA, and 100 unaffected subjects (control group).</td></tr>
<tr><td>CVD risk in obese children</td><td>180 patients , by month 36: 60 (among which 30 girls) for each clinical centre, 90 for BGCH.</td></tr>
<tr><td>Juvenile Idiopathic Arthitis (JIA)</td><td>Altogether 200 patients by month 28.</td></tr>
<tr><td rowspan="3">NND</td><td colspan="2">Cerebral Palsy: 50 patients for each clinical centre for probabilistic modelling, as well as 600 retrospective patients from KU Leuven and OPBG.</td></tr>
<tr><td colspan="2">Spinal Muscular Atrophy (SMA)
20 ambulant patients (severity grade type 3);
10 patients for each centre for biophysical modeling;
10 patients among the 3a subgroup (symptoms of weakness appearing before age 3 years);
10 patients among the 3b subgroup (weakness appearing after the age of 3 years.</td></tr>
<tr><td colspan="2">Duchenne Muscular Dystrophy (DMD)</td></tr>
</table>

	Clinical data will be collected by OPBG, KU Leuven and VUA from 20 ambulant genetically confirmed DMD Patients. 10 patients with an age ranging between 5 and 6 years, additional 10 patients with an age ranging between 7 and 8 years.

1.2 BACKGROUND OF WP 4 - "DATA ACQUISITION AND PROCESSING FOR THE ESTIMATION OF CVD RISK IN OBESE CHILDREN"

The World Health Report 2002 revealed that, in developed countries, approximately one third of all coronary heart diseases and ischaemic strokes and almost 60% of hypertensive diseases can be directly attributed to obesity [WHR2002]. These figures confirm obesity as one of the primary risk factors for cardiovascular disease (CVD), a risk factor that originates early in life. As autopsy studies have shown, the levels of lipids, blood pressure, and obesity in the young are directly associated with the extent of early atherosclerosis of the aorta and coronary arteries [Berenson et al.,1998]20. For this reason, it is of particular concern that there has been a significant increase in childhood and adolescent obesity over the last decade. In the United States, 32% of children and adolescents are now at or above the eightyfifth percentile of the 2000 BMI-for-age growth charts [Ogden et al., 2008; Kuczmarski et al., 2000], but also in the United Kingdom, the prevalence of obesity in children is approaching one third [BHF2008]. One of the challenges concerning the study of childhood obesity and its influence on CVD risk is the required time span for longitudinal studies: cardiovascular events occur mostly later in adulthood, which means that longitudinal studies have to comprise several decades. Nonetheless, cross-sectional studies are able to show correlation between childhood obesity and established surrogate markers for CVD, such as atherosclerosis and cardiac hypertrophy. The Strong Heart Study [Chinali et al., 2006; Chinali et al., 2008], which analysed data from over 450 adolescents, demonstrated that in patients with obesity and/or metabolic syndrome a significantly higher prevalence of left ventricular hypertrophy and left atrial dilation paired with impairment in both systolic and diastolic function is observed. Insulin resistance (IR) is an established determinant in the pathogenesis of CVD; it is constantly observed in patients with hypertension, dyslipidemia and atherosclerosis. Evidence supports firmly that body fat distribution (subcutaneous, visceral, muscle and hepatic fat) modulates IR and cardiovascular risk more than total body adiposity, thus explaining why some individuals who are seemingly equally obese and share common lifestyle and dietary habits tend to have higher IR and CVD risk than others.

MD-Paedigree will integrate the variety of known biomarkers for CVD risk assessment into one common framework, enhance body fat distribution biomarker measurement, and analyse interdependencies between the biomarkers. In addition, MD-Paedigree will develop computational models with high predictive power to better understand the mechanism of CVD development. These models will also allow the simulation of interventions to make personalised predictions for the optimal therapy.

Obesity is commonly acknowledged as a major risk factor for cardiovascular disease (CVD). However, the precise mechanism leading to the development of cardiovascular risk in obesity from childhood to adolescence to adulthood remains largely unsolved [Lloyd et al., 2010; Cornier et al., 2011]. In particular, it is still unclear whether childhood obesity increases CVD risk simply because of the tracking of obesity from childhood to adulthood or via the development of CVD risk factors already present in childhood and adolescence. Many structural and functional changes in the adolescent heart, such as left ventricular (LV) hypertrophy, left atrial (LA) enlargement, and

subclinical impairment of LV systolic and diastolic function are believed to be precursors to more overt forms of cardiac dysfunction and heart failure [Abel et al., 2008].

In order to rate the degree of obesity for clinical diagnostics and studies, the body mass index (BMI) is still the primary measure, also in children [Lloyd et al., 2012]. However, BMI only estimates the general adiposity of a subject, while it does not take into account the distribution of adipose tissue within the body. Specifically, visceral adipose tissue (VAT), the fat between the abdominal organs, has shown to correlate highly with CVD [vanGaal et al., 2006]. In addition, subjects with normal BMI may still have high body fat content, which has proved to be a significant CVD risk factor for adults [Romero-Corral et al., 2010].

Complementary to BMI, imaging modalities such as computed tomography (CT) or magnetic resonance imaging (MRI) allow measuring specific adipose tissue types and have established themselves as important tools for diagnosis. While CT and MRI are the current gold standard for adipose tissue quantification, high costs (and the radiation exposure of CT) restrict these modalities to large-scale studies, and ultrasound (US) is becoming an affordable, non-invasive alternative [deLucia2010]. In particular, echocardiography allows to measure epicardial adipose tissue (EAT) and has emerged as a novel approach to accurately estimate VAT [Iacobellis2003]. However, the quantification of adipose tissue from image data is still mostly performed manually; a tedious and time-consuming process prone to subjective bias. In particular, for the analysis of EAT from MRI, the adipose tissue has to be measured and contoured manually, which leads to noticeable discrepancies between different observers [Flüchter2007]. Measuring the thickness of EAT from US is even more challenging, which is why commonly several manual measurements are performed with electronic callipers and averaged [Iacobellis2008].

In order to decrease the manual workload of the operators, several methods have been proposed for semi- or completely automated image-based quantification of adiposity. The extraction of adipose tissue from MRI has been studied extensively, either for selected body regions [Poll et al. 2002, Peng et al. 2007, Positano et al. 2009] or for whole-body scans [Kullberg et al. 200985, Würslin et al. 201086, Wald et al. 201287]. Since adipose tissue features high intensities in MRI, many authors use thresholding to separate it from the surrounding tissue. Although an automatic selection of thresholds has been proposed [Gronemeyer et al. 2000], different adipose tissue types (VAT and subcutaneous adipose tissue, SAT) still have to be separated manually. An automatic algorithm for this problem was developed [Positano et al. 2004] 88, based on an active contour algorithm. Liou et al. [2006] 89 proposed to use morphological operations, edge detection, and knowledge-based curvature fitting. In all these approaches, bone marrow is often misclassified as adipose tissue, because it features similar intensities in MRI. Thomas et al. [1998] 90 excluded bone marrow by user interaction, while Shen et al. [2003] eliminated the paravertebral adiposity tissue automatically. [Kullberg et al. 2007] used geometrical models of the pelvis and vertebra to exclude these structures and thresholding and morphological operations to automatically separate VAT and SAT. Zhou et al. [2011] employed fuzzy c-means clustering and thresholding to quantify VAT and SAT in both water-saturated and non-water saturated MR images. While automated ultrasound segmentation is feasible for a variety of anatomical structures [Noble et al. 2006], it has rarely been used on adipose tissue. One of the few approaches was proposed by Ng et al. [2009] who used US radiofrequency signals from different locations and beam angles and calculated the spectrum dispersion within the image. Pixels which represent adipose tissue change faster than other areas. To the best of our knowledge, there are no automatic algorithms quantifying intraabdominal fat from US.

In MD-Paedigree, we will re-use our proven anatomical organ models developed in Health-e-Child and Sim-e-Child to add prior knowledge to image analysis. This will enable us to assess different adipose tissue types automatically from image data and use this information in our further analysis. In addition to the fat distribution data from imaging, we will also use established biomarkers such as blood pressure, metabolic and haemodynamic data to estimate the CVD risk. Currently, most studies that analyse different factors of CVD risk employ univariate or, at best, multivariate but linear models, which represent a major limitation. Univariate models can only identify independent contributors to the risk, while they do not shed much light on the interplay between the factors. As demonstrated by [Colombet et al., 2000], cardiovascular risk can be modelled by multivariate machine learning models with only ten clinical variables (representing commonly acknowledged markers of CVD risk). In a similar study, Kurt et al. [2008] successfully modelled the risk of coronary artery disease with a multi-layer perceptron (MLP) and a comparable set of 8 clinical variables. Sumathi and Santhakumaran [2011] trained an Artificial Neural Network (ANN) on a set of 15 clinical variables and claimed to use it successfully for early diagnosis of hypertension. In MD-Paedigree, we will construct multivariate nonlinear models of CVD risk involving state-of-the-art statistical and machine learning techniques. This will not only help to build more accurate models of CVD risk, but also to better understand the mechanism of CVD development via the identification of important risk factors and understanding of their interrelation. Such personalised risk models may become a more reliable alternative or at least a useful complement to the CVD risk prediction charts of WHO [Prevention, 2007], especially since these charts are available for adults only.

A common drawback of the existing works of multivariate modelling is that the underlying techniques like Multi-layer-Perceptron (MLP) or Artificial Neuron Networks (ANN) are basically “black box” models, i.e. the reasons for their results cannot be conveyed to their human users, which leads to low acceptance rates among clinicians.

In our modelling, we will focus on case-based reasoning and discriminative distance learning instead Tsymbal et al, 2009; Zhou et al., 2006]. Since these systems base their decisions on concrete patient cases and are able to present the relevant cases (i.e. the ones utilised for decision making) to the user, they provide easy and intuitive decision support and a possibility for personalised therapy planning, based on the clinical history of retrieved similar patients.

Our work will be centred on the similarity search based decision support system HeC CaseReasoner [Manset et al., 2009] developed in the Health-e-Child project. It features recently suggested techniques for discriminative distance learning, including learning from equivalence constraints and the intrinsic random forest similarity.

The basic philosophy behind the design of CaseReasoner is to provide clinicians with a flexible and interactive tool to enable operations such as data filtering and similarity search over a grid of clinical centres, and to facilitate the exploration of the resulting data sets. The major aim is to let clinicians explore and compare the patients’ records, regardless of geographical location, and to visualize their place in the distribution of both the whole population of patients, as well as in the distribution of its semantic subsets (Figure 8). The search platform can then be used for several tasks such as case-based retrieval [Depeursinge et al. 2010], support for curation [Ruch et. al. 2006] and ultimately decision support [Ruch et al. 2008; Pasche et al. 2011]. HeC CaseReasoner employs a domain-independent technology, and has been applied within Health-e-Child for decision support in three domains: cardiology, neurooncology, and rheumatology. With MD-Paedigree, HeC CaseReasoner will be further extended and applied to decision support in the domain of modelling cardiovascular risk in obese children and adolescents.

2 AIMS OF STUDY

2.1 MAIN GOALS

Our major objectives with modelling the cardiovascular risk in the obese child and adolescent are:

a) automated, objective quantification of different adipose tissue types and their distribution from MRI and ultrasound data;

b) collection of a large number of additional factors contributing to the risk, including metabolic and haemodynamic factors, clinical and family histories, and their interrelation;

c) construction of personalised multivariate retrieval-based models for the assessment of cardiovascular risk using state-of-the-art machine learning techniques, both for cross-sectional and longitudinal studies;

d) interpretation of the models with the purpose of better understanding the mechanism of cardiovascular dysfunction from childhood to adolescence and adulthood, and quantitative evaluation of their predictive performance with cross-validation and sensitivity analysis, and with evaluation on unseen subsequently acquired cases.

Main goal of the study is therefore to collect clinical, biochemical and imaging data to estimate cardiovascular risk associated with obesity in adolescents and to identify significant predictors of increased risk as estimated by changes in arterial stiffness over the time.
On the basis of the outcomes of the data collected and subsequent analysis MD-Paedigree will integrate the variety of known biomarkers for CVD risk assessment into one common framework, enhance body fat distribution biomarker measurement, and analyse interdependencies between the biomarkers. In addition, MD-Paedigree will develop computational models with high predictive power to better understand the mechanism of CVD development. These models will also allow the simulation of interventions to make personalised predictions for the optimal therapy.

In a sub-sample of obese patients, the study will also assess the metabolic and cardiovascular response to a lipid reach liquid meal (ancillary study). Indeed, a mixed meal (carbohydrates and lipids) is able to elicit a powerful metabolic and cardiovascular response. The ancillary study may involve also offsprings of morbidly obese individuals and patients with type 2 diabetes since it has been firmly demonstrated that they have a significantly higher cardiovascular risk of age matched peers.

2.2 PRIMARY END POINT

The primary end point of the study is the successful collection of clinical, laboratory and diagnostic data to be subsequently modelled and simulated by the IT experts. The primary end point envisages estimation of the cardiovascular risk associated with obesity in 180 adolescents (60 for each clinical partner involved, 90 for Bambino Gesu' Children's Hospital) will include evaluation of insulin resistance based on modelling of glucose and insulin values following a high energy liquid mixed meal , lipid profile, liver function tests, white blood cell count, circulating adipokines and markers of low-grade inflammation; and estimation of glucose tolerance. Adiposity at the abdominal (visceral adipose tissue, VAT; subcutaneous adipose tissue, SAT; hepatic and pancreatic fat fractions, HFF & PFF, respectively) and epicardic (epicardic adipose tissue, EAT) sites will be

estimated by ultrasonography (US) and magnetic resonance imaging (MRI). Stratification of the risk will include information of left ventricular (LV) morphology and haemodynamics by US and Cardiac Magnetic resonance (CMR) following the mixed meal..

3 STUDY DESCRIPTION

3.1 STUDY DESIGN:

MAIN STUDY

The study is designed as longitudinal observation of 180 obese adolescents (60 for each clinical partner involved, 90 males, 90 for BGCH). They will be evaluated at the baseline (month 1 to month 18) and re-evaluated between 19 and 36 months.
At ages 14 to 16 years (baseline study) and 15.5 to 17.5 years (follow-up study) , participants will visit a clinical research facility, where they will be asked to provide informed, written consent, to complete questionnaires and undergo the research investigations consisting of clinical and laboratory evaluations, ultrasound evaluation of heart morphology and function, ultrasound evaluation of liver brightness which are routinely performed in overweight and obese individuals [i.e. lipid profile, liver function tests, white blood cell count, circulating adipokines and markers of low-grade inflammation; and estimation of glucose tolerance]. Adiposity at the abdominal (visceral adipose tissue, VAT; subcutaneous adipose tissue, SAT; hepatic and pancreatic fat fractions, HFF & PFF, respectively) and epicardic (epicardic adipose tissue, EAT) sites will be estimated also by magnetic resonance imaging (MRI).

ANCILLARY STUDY

In a sub-sample of obese patients and eventually in offsprings of severely obese and diabetic patients an oral metabolic tolerance test, oMTT, containing 75 g glucose and 75g of fat per m^2 body surface area (prevalently saturated fatty acids) will be performed during RMI scanning (ancillary study). The lipid meal is expected to boost both the insulin and the cardiovascular response. Offsprings of obese and diabetic patients are suitable candidates to the study since they may present with metabolic responses not different from obese patients.

No sedation will be required for RMI studies of adolescent patients.

Hence, stratification of the risk will include evaluation of insulin resistance based on modelling of glucose and insulin values following the oral glucose tolerance test and/or the high energy liquid mixed meal (in patients undergoing the ancillary study), information of left ventricular (LV) morphology and haemodynamics by US and Cardiac Magnetic resonance (CMR). Reduced elasticity, distensibility, and stiffness at the follow-up will be used as end-point estimate of cardiovascular disease.

MAIN STUDY

Questionnaires (attached to the protocol)
A paper copy of a basic questionnaire will be sent by post or given in person to the participants prior to their attendance at the clinical research facility. The full questionnaire will be completed at the facility. An accompanying letter will ask them to complete as much of the questionnaire as possible at home. Particular attention will be drawn to data that might require help from family

members to obtain eg. Family history. A trained professional will then take the participants through their answers when they attend the clinical research facility to ensure complete and accurate responses and to address any questions or uncertainty that the participants may have. The questionnaires will address the following:

- ✓ Name, sex, date of birth, contact details (address, email, telephone)
- ✓ Ethnic group
- ✓ Educational attainment (grades)
- ✓ Maternal & paternal social class [The National Statistics Socio-Economic Classification 2001]
- ✓ Mother's age, weight & height
- ✓ Father's age, weight & height
- ✓ Birth weight and length of gestation
- ✓ Family history of medical conditions, including hypertension, diabetes, angina, myocardial infarction, stroke, peripheral vascular disease, and hypercholesterolaemia
- ✓ medical history including history of cardiac disease or cardiac surgery, any endocrine abnormality, including diabetes, familial hypercholesterolaemia, renal disease
- ✓ menstrual history& contraceptive use in the girls
- ✓ medication history, including steroid use
- ✓ smoking, alcohol, caffeine and recreational drug use history, including time of last consumption
- ✓ time of last consumption of food / drink
- ✓ activity / exercise [Baecke et al. 1982]

Self-assessment of pubertal status. Participants will be asked to compare themselves to a series of images representing Tanner stage to estimate their progress through puberty [Carel and Leger 2008]. This method has been shown to have good levels of agreement with physician examination in a large population of obese and non-obese boys and girls [Sun et al. 2012].

Perceived stress. Participants' sense of being under stress over the preceding 4 weeks will be assessed with the ten-item version of the Perceived Stress Scale [Cohen et al. 1983]. Each item (e.g., "Over the past week how often have you felt that you were unable to control the important things in your life?") is rated on a five-point scale. Total scores may range from 0 to 40, with higher scores indicating greater perceived stress. The Perceived Stress Scale has been widely used in research on stress and health, and has high internal consistency.

Emotional well-being. The 28-item version of the General Health Questionnaire (GHQ) [Richard et al. 2004] will be used to assess emotional wellbeing over the preceding 4 weeks. The GHQ is widely used in clinical and population studies, and its validity has been established against psychiatric interviews. It has high internal consistency. The hospital anxiety and depression scale (HADS) [Zigmond and Snaith 1983]will be used to augment this.

Eating style. Four scales of eating style will be assessed (routine restraint, compensatory restraint, susceptibility to external cues, and emotional eating) using the Weight-Related Eating Questionnaire (WREQ) [Schembre et al. 2009]. This 16-item questionnaire has been shown to have good psychometric properties and construct validity in an ethnically diverse population of young adult to elderly men and women [Schembre and Geller 2011].

Anthropometrics

Height will be measured to the nearest 1 mm, without shoes and with the Frankfurt plane of the participant's head aligned by eye to be parallel to the ground. A calibrated stadiometer or height board will be used. Weight will be determined to the nearest 10 grams, using calibrated scales, with the participant wearing only light clothing and no shoes. Waist and hip circumferences will be determined according to standard practice [Molarius et al. 1999], using a flexible measuring tape.

Routine laboratory tests will include evaluation of fasting glucose, insulin, c-peptide, lipid profile (total and HDL cholesterol, triglycerides), liver function tests (alanine-aminotransferase, aspartate amino transferase, γ-glutamyl transferase), glycated haemoglobin, white blood cell count; glucose tolerance by a standard OGTT (1.75 g/kg body weight up to a maximum of 75 g). Glucose, insulin and c-peptide will be measured at baseline and 30, 60, 90 and 120 min. Systolic (SBP) and diastolic blood pressure (DBP) will be measured three times while the subjects are seated, and the measurements will be averaged for the analysis.

All the examinations will not require withdraw of additional amount of blood respect to analyses routinely performed in obese patients (average 20 ml of blood).

Estimation of adipokines, low-grade inflammation and insulin resistance

Measurements of adipokines and markers of inflammation. Blood samples will be withdrawn to measure fasting plasma adipokines (leptin, adiponectin), circulating markers of inflammation (C-reactive protein, CRP; Tumor-Necrosis Factor-alpha, TNF-alpha; Interleukin 6, IL6) and endothelium dysfunction (e-Selectin, Intercellular Adhesion Molecule 1, ICAM-1).

Assessment of the renin-angiotensin-aldosterone axis. With the patient in the supine position, blood will be obtained for measuring plasma renin activity, aldosterone, cortisol, serum sodium and potassium. After being upright and ambulating for 2 hours, repeated blood samples will be obtained. Dietary sodium intake will be assessed by measuring 24 hour urinary sodium excretion.

Assays of adipokines will be centralized with anonymized samples sent to Department of Metabolic Diseases, University Medical Center Utrecht, Utrecht, the Netherlands, responsible person Dr. Hank Shipper.

On residual blood/plasma genetic analyses and assay of cytokines will be performed.

Body fat assessment

We will use a graphics processing unit (GPU) implementation of the T2*-IDEAL algorithm [Yu et al. 2007; Kowalik et al. 2011] to measure body fat content. This iteratively separates MR images into fat and water components, which can then be used to measure the proportion of fat in each 3x3x10 mm voxel. Data will be acquired in a continuous stack of 10 mm thick slices from the neck to the knees. To prevent motion artefact, we will use breath holding for the thorax and abdomen and cardiac gating for slices containing the heart. Fat quantification in the head, arms and below the knees is impractical due to the need for participant re-positioning or specialized coils. Due to their low fat content, we will exclude these body parts using anatomical landmarks to ensure consistency between participants.

Estimation of Insulin Resistance and secretion

Insulin resistance will be estimated in fasting condition, after the glucose load and patients undergoing the oMTT also after the mixed meal. It will be computed by means of the following methods: [QUICKI= 1/[log plasma fasting insulin(mIU/l) + log plasma fasting glucose (mg/dl)]; HOMA-IR = (insulin x glucose)/22.5); WBISI = (10,000/square root of [fasting glucose x fasting insulin] x [mean OGTT glucose x mean OGTT insulin]. Insulin secretion will be estimated by the deconvolution method. Beta-Cell ability to adapt insulin secretion to changes in insulin sensitivity (glucose disposition indexes) will be assessed.

Genetic Analysis

Sample collection, storage and DNA extraction

Blood samples will be collected from 180 obese patients at the baseline for DNA analysis. Faecal samples will be collected, at baseline and after 18 months for micorbioma/metagenome analyses. A database of patients, including name, age, disease, laboratory data but also specific indications about antibiotic, prebiotics, and probiotics administration will accomplish faecal sample datasheets for appropriate later description of gut microbiota enterotypes.

Blood and faecal samples will be stored at 4°C for at maximum of 24 hours or, alternatively, at -80°C until shipment to genomic/metagenomic facilities for automatic DNA extraction and targeted-sequencing. The samples will be sent every two months or when suitable for the laboratories included in the study in dry ice by express courier.

In detail, DNA will be extracted from faecal samples at the Bambino Gesù Children's Hospital, laboratory of mIcrobiology (responsible person: Dr. Lorenza Putignani).

DNA extracted from blood from genetic analysis will be sent to BMR Genetics s.r.l., via Redipuglia, 22 - 35131 PADOVA, Italia Codice fiscale e Partita IVA: 03888370289 (responsible person Dr. Barbara Simionati).

All samples will be anonymzed prior to be sent.

After DNA analysis, residual biological material will be destroyed within 12 months from the end of the follow up study. Indeed no future studies are foreseen.

Patients and legal representative will be informed of all results except for results of genetic analysis.

Genetic analysis

Genetic analysis will be performed on blood samples withdrawn at baseline, in order to build a genetic score of cardiovascular disease (CVD) risk. Genomic DNA will be extracted by GeneCatcher gDNA Blood Kit (Invitrogen). Analyses will be performed by using the Illumina technology. Candidate Single Nucleotide Polymorphisms (SNPs) for estimation of CVD risk in the MD-Paedigree study. DNA analysis. Analysis (DNA extraction and SNPs analysis) of a custom of SNPs in 180 patients plus the statistical analysis in order to build a genetic score of CVD risk. SNPs will be selected among SNPs identified in previous Genome Wide Association (GWAS studies). Selection will be based on either statistical significance threshold of the genetic association with the investigated variable (dyslipidemia, left ventricular hypertrophy, hypertension, type 2 diabetes, increased visceral adiposity and fatty liver) and/or clinical significance in a customized metabochip . Two genetic risk scores will be constructed on an a priori basis. Genetic risk scores will be the sum of all cardiovascular risk alleles from all SNPs, both those associated with CVD (increased

stiffness/IMT) and those associated with risk factors as done previously (Raynter NP; JAMA 2010; 303: 631-7; Peterson RE, Hum Genet 2011; 129: 221-30). The SNPs affecting more than one phenotype will be included once.

Associated variable/phenotype Single Nucleotide Polymorphisms (SNPs)

Associated variable/phenotype	**SNP**	**Reference**
Blood lipids	91 SNPS to be selected based upon significance	Teslovich TM Nature 2010; 466: 707-713 Aulchenko YS Nature genetics 2009; 41: 47-55
Blood pressure	rs3918226 NOS3 rs4846049 MTHFR-NPPB rs2004776 AGT rs661348 LSP1/TNNT3 rs11105354 ATP2B1 rs2014408 SOX6 rs1799945 HFE rs1421811 NPR3 rs9930333 FTO rs16933812 PAX5 rs7638110 MRPS22 rs17773430 MCR4	Johnson T, AJHG 2011 89: 688-700; Melka MG, JCEM 2012; 97:E145-E150
Type 2 diabetes/fasting glucose	rs560887 G6PC2 rs10830963 MTNR1B rs2191349 DGKB-TMEM195 rs780094 GCKR rs11708067 ADCY5 rs7944584 MADD rs10885122 ADRA2A rs174550 FADS1 rs11605924 CRY2 rs11558471 SLC30A8 rs4506565 TCF7L2 rs4607517 GCK rs7034200 GLIS3 rs340874 PROX1 rs11920090 SLC2A2 rs11071657 C2CD4B rs10923931 NOTCH2 rs11899863 THADA rs243021 BCL11A rs7578326 IRS1 rs13081389 PPARG rs6795735 ADAMTS9	

	rs1470579 IGF2BP2 rs1801214 WFS1 rs4457053 ZBED3 rs10440833 CDKAL1 rs849134 JAZF1 rs972283 KLF14 rs896854 TP53INP1 rs10965250 CDKN2A/B rs13292136 CHCHD9 rs12779790 CDC123/CAMK1D rs5015480 HHEX/IDE rs2334499 HCCA2 rs231362 KCNQ1 (a) rs163184 KCNQ1 (b) rs5215 KCNJ11 rs1552224 CENTD2 rs1531343 HMGA2 rs4760790 TSPAN8/LGR5 rs7957197 HNF1A rs11634397 ZFAND6 rs8042680 PRC1 rs11642841 FTO rs4430796 HNF1B (TCF2) rs5945326 DUSP9	
Left ventricular dimension	rs756529 KCNB1	Arnnet DK, BMC Medical Genetics 2009, 10:43 Vasan RS JAMA 2009; 302: 168-78
Fatty liver	rs738409 PNPLA3 rs2854116 APOC3 rs12979860 IL28B rs1260326 GCKR rs4986790 TLR4	
Visceral adiposity	CYP17A1 rs1004467 NT5C2 rs11191548 SH2B1 rs7498665	
Levels of adiponectin	ADIPOQ rs17366653	
Levels of CRP	rs2847281 rs6901250 rs4705952	Dehghan A, Circulation 2011; 123: 731-8
Genetic score (Hypertension+left ventricular wall thickness+stroke+CAD)	29SNPs	Ehret GB Nature 2011; 478:103-109

Microbioma/Metagenome analysis

Metagenome data analysis will be carried out on fecal samples from obese patients collected at baseline and at 18 months, and re-evaluated at the follow-up to investigate the risk to develop CVD associated with specific gut taxa at the baseline and during observational time-course (enterotype) (Vulevic et al., J Nutr. 2013 Mar;143(3):324-31. Epub 2013 Jan 9; Di Girolamo et al., 2012 Dec;6(6):759-73).

Microbioma analysis will provide an opportunity to understand how gut microbiota taxa distribution may possibly correlate with CVD risk (Vulevic et al., J Nutr. 2013 Mar;143(3):324-31. Epub 2013 Jan 9; Di Girolamo et al., 2012 Dec;6(6):759-73). Stool samples which will be collected at baseline and after 18 months. The results of gut microbiota analysis will be integrated with clinical data to assess how they correlate with obesity indexes, and in particular to explore the prognostic value of the presence of major gut taxa patterns in conditioning disease susceptibility as well as the immune response.

In order to analyse the taxonomic gut content of obese patients, a targeted approach based on sequencing of the variable regions V1 and V3 of 16S rRNA locus will be used (Aagaard *et al*., 2012. PLoS One 7(6):e36466. Epub Jun 13). Fecal samples will be collected and analysed at onset of disease, at time of clinical remission, and during disease flares, with a prediction of approximately 400 samples.

Analysis of microbiome of fecal samples will be carried out following DNA extraction (automatic EZ1 Biorobot, Qiagen), and further pyrosequencing using a 454 Junior apparatus and sequence analysis; comparison will be performed with the recently developed MEGAN 4 software (available at http://www-ab.informatik.unituebingen.de/software/megan) (Mitra et al., BMC Genomics 2011), or with the PhylOTU software (https://github.com/sharpton/PhylOTU) (Wylie et al., 2012. PLoS One 7(6):e35294. Epub Jun 13.), in order to identify the microbiota operational taxonomic units (OTUs).

ANCILLARY STUDY

MD-paedigree gives us the opportunity to provide addition information on the cardiovascular risk elicited by a mixed meal which resembles daily diet by performing an ancillary study. Such as not all the patients will undergo the study, but 60 patients are sufficient to achieve statistically significant results. Ideally, each centre will contribute with 20 patients, but one centre can replace an other in case of failure in the recruitment.
The amount of blood required for the ancillary study cannot exceed 100 ml per patient. By considering a medium body weight for each patient of 60 kilos, such amount is below the threshold suggested by the WHO as a safe (WHO guidelines).

oMTT
The oMTT will be performed simultaneously to an MRI scan. While the complete metabolic test will last 240 min, the MRI scan lasting up to 1 hour 30 minutes.
Prior to the MRI scan, an intravenous cannula will be placed in a peripheral vein and blood for a baseline metabolic profile will be obtained. To obtain useful basal metabolic measures, participants will be asked to fast overnight and drink nothing but water on the day of their assessment. They will be asked to consume their last meal prior to fasting at 9pm. They will also be asked to abstain from smoking, alcohol, recreational drugs or caffeine consumption and from formal physical exercise for the preceding 24 hours.

After resting in the MRI scanner for 15 minutes, during which time planning (scout) scans will be carried out, resting haemodynamic parameters will be measured. The participants will then be asked to ingest a lipid and glucose rich meal as describe din detail below. The haemodynamic and metabolic responses to this meal will then be measured with repeated MR assessments, blood samples and saliva samples. Prior to completion of this assessment, a complete scan of the body will be carried out to accurately quantify body fat and its distribution.

Intravenous cannulation and venous blood sampling

This will be carried out according to standard hospital practice, using the largest suitable cannula up to gauge 18 in a peripheral vein. The following protocol will be used:

Hand hygiene:

- ✓ Decontaminate hands before and after each participant contact and before applying and after removing gloves.
- ✓ Use correct hand hygiene procedure.

Personal protective equipment:

10 Wear gloves.

11 Remove and discard gloves immediately after the exposure-prone activity.

Skin preparation:

6 Use 2% chlorhexidine/70% alcohol applicator (ChloraprepSepp®) and **allow to dry** (NB – this is important to avoid contamination of sample as well as for hygiene reasons).

7 Do not re-palpate the vein after the skin preparation.

Dressing:

✓ Use a sterile, semi-permeable, transparent dressing to allow observation of insertion site.

Safe maximum total blood draw will be defined according to body weight (3 mL/kg per 24 hours), in accordance with safe practice guidelines (WHO guidelines) and this limit will not be exceeded. In any case, the amount of blood withdrawn will not exceed 16 ml.

High energy liquid mixed meal

Studies will be carried out after a minimum 12h overnight fast. After fasting blood has been drawn, participants will consume a standard liquid meal (oral metabolic tolerance test, oMTT). This will contain heavy whipping cream and 75g glucose. The quantity of cream will be varied to deliver 75g of fat per m^2 body surface area. The drink will have a volume of approximately 500 mL and will be consumed within 10 minutes. This regimen has been shown to stimulate significant responses in vascular inflammatory markers [Ceriello et al. 2004] but also it can also boost the glucose induced insulin response to a different degree in normal-weight and obese individuals [Manco M et al 2004].

Assessments of insulin resistance using a similar meal were found to correlate well with standard oral glucose tolerance testing in children [Chandler-Laney et al. 2013].

Haemodynamic response to meal (MR)

To control for the effects of acute stress, participants will be asked to rate their level of anxiety or stress on a visual analog scale (1-100) just prior to entry to the MR scanner, just prior to the meal, 30 mins after the meal, and just prior to completion of the scan. Room temperature, which is known to affect the vasculature, will be recorded to the nearest 0.1°C.

Oscillometric BP will be measured in the non-dominant arm at one-minute intervals. All imaging will be performed on a 1.5T MR scanner (Avanto or Aera, Siemens Medical Solutions, Erlangen, Germany).

Flow quantification will be performed through-plane in a cross-section of the ascending aorta as it passes the bifurcation of the pulmonary arteries using an ECG-gated spiral phase-contrast MR sequence, as described previously [Steeden et al. 2011]. This technique will allow images to be acquired within a short breath-hold (~5 seconds) with a spatial resolution of 1.6x1.6 mm and a temporal resolution of 30 milliseconds.

Flow images will be processed to derive stroke volume (SV) and CO. Total peripheral resistance (TPR) will be calculated by dividing the mean BP (MBP) by CO. Total arterial compliance (TAC) will be calculated by optimization of a two-element windkessel model, as previously described [Stergiopulos et al. 1994].

Peak mesenteric artery flow assessed with phase-contrast MRI occurs 20 minutes following meal ingestion [Masui et al. 1994] as do changes in internal carotid and vertebral artery flow [Totman et al. 2009]. Cardiac index peaks 30-60 minutes after a meal in healthy young people and at 30 minutes for the elderly or individuals with dysautonomia [Lipsitz et al. 1993]. Forearm vascular resistance peaks at 45 minutes in all groups. There is minimal change in mean arterial blood pressure (MABP) in healthy young subjects but substantial reductions in MABP at 30 minutes in dysautonomic patients. Heart rate (HR) peaks at 30 minutes in elderly or dysautonomic patients, with a concomitant peak in noradrenaline levels in these subjects, but HR rises gently to a maximum at 1 hour for the healthy young. Taken together, these data suggest that the peak vascular load after meal ingestion is at some time between 20 and 40 minutes but that this load is well regulated in healthy young patients, where precise autonomic regulation of the circulation is preserved.

To achieve optimal characterisation of the cardiovascular response to a meal, HR will be monitored continuously, blood pressure (BP) will be measured every 1 minute, stroke volume will be assessed every 5 minutes for the first 40 minutes and then at 50 minutes and 1 hour after the meal. Flow into the carotids, vertebral arteries, superior mesenteric artery and flow at the descending aorta just proximal to the junction of the iliac arteries will be assessed at the same time to assess proportional vascular changes in the head, lower limbs and mesentery. These data and the vascular parameters derived from them will be compared to the same measures taken just prior to meal ingestion after participants had rested for 15 minutes.

Participants will spend a total of 1 hour and 30 minutes in MRI scanner, during which time, they will be asked to watch a restful and un-stimulating movie (Winged Migration).

<u>Appetite / satiety response to meal</u>

Prior to the meal and at every blood draw, participants will be asked to rate their appetite / satiety on four visual analog scales (scored from 0 to 100 mm)[Blundell et al. 1993]:

2 “How hungry do you feel right now?” (0 = not at all, 100 = extremely)
3 “Rate your desire to eat at this moment.” (0 = none, 100 = very strong)
4 “How full do you feel right now?” (0 = not at all, 100 = completely)
5 “How much food will you eat at your next opportunity?” (0 = none, 100 = a very large amount)

Blood and saliva measures of metabolic response to meal

Blood will be drawn just prior to consumption of the meal (T0). Further samples will be taken at 15, 30, 45, 60, 90, 120, 180, and 240 minutes following the meal (T15-T240). Saliva samples for cortisol assay will be taken at T0 and then at T15, T30, T45, T60, T120, T180 and T240 minutes following the meal to assess differences in hypothalamic-pituitary-adrenal axis response, which are known to differ according to fat distribution (Vicennati et al. 2002).Saliva samples will be collected using standard equipment (Salivette® Cortisol - Sarstedt, Nümbrecht, Germany). Concentration of salivary free cortisol will be measured using a commercially available chemiluminescence-immuno-assay (IBL, Hamburg, Germany).

The schedule for the measurement of each compound from blood will be as follows:

Glucose metabolism / insulin resistance will be assessed with these measures:
HbA_1C [van 't Riet et al. 2012]: T0
Glucose: T0, T15, T30, T45, T60, T90, T120
Insulin: T0, T15, T30, T45, T60, T90, T120, T240
C-peptide: T0, T15, T30, T60, T120

Hormonal response to meal:
GLP-1[Carroll et al. 2007; Baggio and Drucker 2007]: T0, T15, T30, T60, T90
GIP[Baggio and Drucker 2007]: T0, T15, T30, T60, T90
Leptin[Carroll et al. 2007]: T0, T240
Adiponectin: T0, T30, T60, T120, T240
Acylation-stimulating protein (ASP)[van Oostrom et al. 2004]: T0
Complement 3 (C3)[van Oostrom et al. 2004]: T0, T120
Ghrelin [Carroll et al. 2007]: T0, T15, T30, T45, T90, T180
ACTH[Vicennati et al. 2002]: T0, T15, T30, T45, T60, T120
Noradrenaline: T0, T15, T30, T60, T120
Adrenaline: T0, T15, T30, T60, T120

Lipid response to meal:
TG: T0, T60, T120, T240
FFA: T0, 60, 120, 240
HDL: T0, 60, 120, 240
Total cholesterol: T0, 60, 120, 240

Liver function
ALT: T0
AST: T0
GGT: T0

Inflammatory response to meal:
FBC (WBC, differential WCC, platelets)[Raz et al. 2013]: T0 T240
hsCRP[Raz et al. 2013]:T0 T120T240
TNF alpha[Nappo et al. 2002]:T0 T240
IL6[Nappo et al. 2002]:T0 T240
Chemerin:T0 T120 T240
Cathepsin-S:T0 T60 T120 T240

Nitrotyrosine[Ceriello et al. 2004]: T0 T30 T60 T90 T120 T240
sICAM-1[Ceriello et al. 2004; Nappo et al. 2002]: T0 T30 T60 T90 T120 T240
sVCAM-1[Ceriello et al. 2004; Nappo et al. 2002]: T0 T30 T60 T90 T120 T240

3.2 SUBJECTS SELECTION

For the main study, three clinical units will be involved for the enrollment of 180 individuals, 60 at each center.

University College London – Great Ormond Street Hospital (GOSH) for Children NHS Foundation Trust - Centre for Cardiovascular Imaging, UCL Institute of Cardiovascular Sciences. Responsible for the coordination of activities under WP4 and for the study coordination in UCL will be Prof. Andrew Taylor, Professor of Cardiovascular Imaging, Head – Centre for Cardiovascular Imaging, UCL Institute of Cardiovascular Sciences; Director – Centre for Cardiovascular MR.

Ospedale Pediatrico Bambino Gesù: Scientific Directorate, Research Area for Preventive and Predictive Medicine; Unit for Multifactorial Diseases, Bambino Gesù Pediatric Hospital, Rome. Responsible for the study coordination in OPBG will be Dr. Melania Manco, MD PhD.

Johns Hopkins University Hospital: Paediatric Cardiology - Helen Taussig Congenital Heart Center. Responsible for the study coordination in JHU: Dr. Allen Everett, paediatric cardiologist.

At each center, 90 obese adolescents (30 males; age 14-16.5) will be enrolled among patients admitted to clinic units. Patients will be selected from amongst those consecutively referred from November 2013 to October 2014. Inclusion criteria will be obesity (percentile of Body Mass Index ≥95° which equals to 1.645 SDS according to Kuczmarski RJ) with no systemic, endocrine and genetic disease. Exclusion criteria will be use of medication; alcohol and recreational drug. Enrolled patients will be restudied between November 2014 and April 2015.

For the ancillary study 20 offsprings of morbidly obese or patients affected by type 2 diabetes or family history of severe obesity or type 2 diabetes (at least one parent with a BMI >40 kg/m2 or type 2 diabetes under medication) per each centre can be enrolled instead of obese patients.

4 WITHDRAWAL FROM THE STUDY

Participation in this study is entirely voluntary. The parents may decide, at any time, that they do not wish their child to take part in the study without altering current or future treatment in any way.
If at any stage of the project the parents wish to withdraw their child from the study or the adolescent to retire the absent, the researcher responsible for the study will arrange for their child's information to be immediately removed from computer records and for any remaining samples we hold belonging to him/her to be destroyed.

5 PATIENT'S STUDY

5.1 STUDY TO BE PERFORMED

History will be recalled and clinical data collected as described in the paragraph 3.1. Patients will undergo routine laboratory tests (assay of fasting glucose, insulin, lipid profile, liver function tests, blood cell count and oral glucose tolerance test).
In addition, as required by the research protocol, they will undergo ultrasound and RMI estimation of abdominal and heart adiposity, ultrasound estimation of heart morphology and function, ultrasound estimation of intima media thickness.
Biological sample residual from routine laboratory assays will be used for the assay of circulating levels of adipokines and DNA analysis. Hence, no extra blood will be withdrawn for the purposes of this research protocol.
A database of patients, including name, age, disease, laboratory data but also specific indications about antibiotic, prebiotics, and probiotics administration will accomplish faecal sample datasheets for appropriate later description of gut microbiota enterotypes. Samples will be stored at 4°C for at maximum of 24 hours or, alternatively, at -80°C until shipment to metagenomic facilities for automatic DNA extraction and targeted-sequencing. In the latter case, the samples will be sent every two months or when suitable for the laboratories included in the study in dry ice by express courier.

6 STUDY PLANNING

6.1 EFFICACY PARAMETERS

Patients will be enrolled in the first 18 months of the study and re-evaluated 18 months later.
Reduced elasticity, distensibility, and stiffness at the follow-up will be used as end-point estimate of cardiovascular disease. Indeed, arterials stiffness is a marker of early atherosclerosis with good sensitivity and specificity.
Stiffness, Ultrasound and RMI assessment will be performed by trained project personnel to reduce inter-individual variability.

6.2 EXPERIMENTAL DESIGN

The study will last 4 years. It is designed as a prospective longitudinal study. The timeframe for patient recruitment spans the first 28 months. Follow up data for each data (clinical, imaging, immunologial etc) will be collected at follow-up visit as indicated in details in patient study session.

6.3 DATA PROTECTION

All collected data will be anonymised. Clinical data and biological samples will be coded and stored as such. The code will be generated by software using a system of 128-bit encryption. The code will be stored in a close drawer by Prof. Andrew Taylor. At the end of the study, the key code will be destroyed and, hence, data anonymised. From this moment on, it will not be possible for anyone to discover the patient's identity.
All clinical data will be communicated to participants and/or legal representative except for genetic testing.

7 SECURITY EVALUATION

7.1 DEFINITIONS

No adverse effects are foreseen as consequence of the clinical study.

8 SAMPLE DIMENSION AND STATISTIC METHODOLOGY

8.1 STATISTIC DESIGN

Main study

As the main goal of whole project is to establish a data repository for pediatric diseases, the sample size (180 patients, 60 for each clinical center, 90 for BGCH) has been set by taking into account primary endpoints and study power but also available resources at each center, and study feasibility. In particular for the genetic analysis no study power is foreseen.

Modeling of cardiovascular risk as estimated by the arterial stiffness will be done by partners Siemens AG, Fraunhofer Research Institute (Fraunhofer Gesellschaft zur Foerderung)and Institut Nationale de Recherche en Informatique et en Automatique (INRIA), using a multivariate approach.

Indeed, most studies that analyse different factors of CVD risk employ univariate or, at best, multivariate but linear models, which represent a major limitation. Univariate models can only identify independent contributors to the risk, while they do not shed much light on the interplay between the factors. As demonstrated by [Colombet et al., 2000] 96, cardiovascular risk can be modelled by multivariate machine learning models with only ten clinical variables (representing commonly acknowledged markers of CVD risk). In a similar study, Kurt et al. [2008] 97 successfully modelled the risk of coronary artery disease with a multi-layer perceptron (MLP) and a comparable set of 8 clinical variables. Sumathi and Santhakumaran [2011]98 trained an Artificial Neural Network (ANN) on a set of 15 clinical variables and claimed to use it successfully for early diagnosis of hypertension. In MD-Paedigree, we will construct multivariate nonlinear models of CVD risk involving state-of-the-art statistical and machine learning techniques. This will not only help to build more accurate models of CVD risk, but also to better understand the mechanism of CVD development via the identification of important risk factors and understanding of their interrelation. Such personalised risk models may become a more reliable alternative or at least a useful complement to the CVD risk prediction charts of WHO [Prevention, 2007] 99, especially since these charts are available for adults only.

A common drawback of the existing works of multivariate modelling is that the underlying techniques like Multi-layer-Perceptron (MLP) or Artificial Neuron Networks (ANN) are basically “black box” models, i.e. the reasons for their results cannot be conveyed to their human users, which leads to low acceptance rates among clinicians.

In our modelling, we will focus on case-based reasoning and discriminative distance learning instead Tsymbal et al, 2009 100; Zhou et al., 2006 101]. Since these systems base their decisions on concrete patient cases and are able to present the relevant cases (i.e. the ones utilised for decision making) to the user, they provide easy and intuitive decision support and a possibility for personalised therapy planning, based on the clinical history of retrieved similar patients.

Ancillary study

Table 1. Power calculations at the 80% level for samples sizes from N=60 to N=180.

Metabolic / vascular parameter	SD	Mean response to mixed meal in normals	Sample Size (subjects)			
			60	90	120	180
Fasting glucose (mmol/L) ENREF_1 [Ceriello 2004]	0.89	-	0.65	0.53	0.46	0.38
HbA_1C (%)	0.89	-	0.65	0.53	0.46	0.38
Resting systolic BP (mmHg) ENREF_2 [Gray L, 2011]	12.7	-	9.2	7.6	6.5	5.4
Heart rate response to meal (bpm) ENREF_3 [Lipsitz LA 1993]	4.2	4.5	3.1	2.5	2.2	1.8
Change in peak systolic flow velocity in superior mesenteric artery in response to meal (cm/s) ENREF_4 [Masui 1994]	19.0	29.6	13.8	11.3	9.8	8.0
Glucose response to meal (mmol/L) ENREF_1 [Ceriello, 2004]	3.3	4.7	2.4	2.0	1.7	1.4
Triglyceride response to meal (mmol/L) ENREF_1 [Ceriello, 2004]	0.59	1.03	0.43	0.35	0.31	0.25

Calculations were performed assuming two-sample comparisons of means between two groups (eg. Obese versus offspring of obese and diabetic patients) that are equally sized (50% of N) at a significance level of 0.05. Values are differences between the means in two groups that could be detected for each parameter, at each sample size. Estimates of variance (SD) and typical mean response to a mixed meal, where appropriate, are given as drawn from the literature, which is referenced.

8.2 MANAGEMENT OF MISSING DATA

Patients whose information is missing or incorrectly transcribed will not be used for the analysis, but will be marked appropriately in the electronic database by inserting a variable "Flag".

9. AMINISTRATIVE AND ETHICAL PROCEDURES

Confidentiality

Clinical data will be acquired as required by each partner's national law.

At each clinical center patient's data will be collected and stored as electronic files and will be accessible by the responsible research personnel. Access to data will be granted using their personal credentials. Access to the file will be protected and the log of the user who performed the operation will be required at regular intervals. The data manager will perform regularly a data backup.

Data publication and final report

The ownership of scientific data will be shared between all the partners involved in the Project. The WP leader, Dr. Andrew Taylor, and researchers who will conduct the study, will endeavor to promote the dissemination of the results through the project website, communications in national and international scientific meetings, publication in international journals of high scientific profile. The dissemination and publication of the results by the experimenters will be promoted in accordance with the provisions in force concerning the confidentiality of sensitive data. In all scientific publications the efforts of all researchers will be recognized.

All health professionals involved in the project will seek to minimize the physical and psychological discomfort caused to patients and parents from participating in this study. In order to ensure the well-being, they will not be notified in any way about the personal results of genetic investigations.

9.1 AUTORIZATIONS

Study implementation and patient enrollment will undergo prior approval of the present Protocol by the Local Ethical Committee of each partner involved in patient enrollment.

9.2 INFORMED CONSENT

Each parent/patient who will be asked for his/her enrollment in the study will be informed on every aspect concerning the study itself and, before the recruitment begins, he/she must sign the informed consent model. The date of the signature of the consent will be recorded. A Copy of the informed consent model must be handed to the parent/patient.

9.3 INSURANCE COVERAGE

Insurance coverage used is as foreseen by each research structure for clinical and research activities. ***9.4 USE OF THE INFORMATION AND DATA PUBBLICATION***

Researchers must have the right to disseminate/publish the results of the study, with respect to the current legislation on sensible data protection and Intellectual property Rights. European Commission does not have any right in the publication/dissemination of study results

9.5 CLINICAL PROTOCOL AMENDMENTS

Clinical protocol amendment is intended to be any modifications to the final version of the clinical protocol that, after being approved by the Researcher and by the Ethical Committee, cannot be informally modified.

9.6 DATA MANAGEMENT AND DOCUMENT PRESERVATION

Researchers will manage the preservation of patients identification codes for at least 15 years after the end of the study. Patients clinical data and other original data will be kept for the maximum period allowed by the Hospital (min. 15 years).

9.7 BUDGET

The project costs are covered by the specific Grant by the European Commission. Such grant will cover 75% of staff costs, of consumables and equipment used for the project, and of all other expenses involved. Annex 1 (GPF) contains the project's entire budget, including details of the Institution's budget.

10. RESEARCHER RESPONSIBILITY

The researcher, as principle investigator of the study declares to be aware of all the duties committed to him/her and his/her collaborators for the conduction of the study. Except where explicitly edited, the term "researcher" on this Protocol and on the CRF is referred to the researcher or to any other person appointed by him/her, who will have power to sign documents instead of him/her.

The researcher will conduct the study according to the principles of the Helsinki declaration (1964) and Amendments (Appendix 1), and to the European Guidelines on Clinical Good Practices .

11. Annexes

Annex 1 GPF
Annex 2 DOW
Annex3 Questionnaires

12 REFERENCES

1. World Health Report 2002. Reducing risks, promoting healthy life. World Health Organisation: Geneva, 2002.
2. Berenson GS, Srinivasan SR, Bao W, et al. Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults: the Bogalusa Heart Study. N Engl J Med. 1998; 338: 1650–1656.
3. Ogden CL, Carroll MD, Flegal KM. High body mass index for age among US children and adolescents, 2003–2006. J Am Med Assoc 2008; 299: 2401–2405.
4. Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, Flegal KM, Guo SS, Wei R et al. CDC growth charts: United States. Adv Data 2000; 314: 1–27.
5. British Heart Foundation. European Cardiovascular Disease Statistics, 2008.
6. Chinali M, de Simone G, Roman MJ, Lee ET, Best LG, Howard BV, Devereux RB. Impact of obesity on cardiac geometry and function in a population of adolescents: the Strong Heart Study. J Am Coll Cardiol. 2006;47(11):2267-73.
7. Chinali M, de Simone G, Roman MJ, et al. Cardiac markers of pre-clinical disease in adolescents with the metabolic syndrome: the Strong Heart Study. J Am Coll Cardiol 2008; 52: 932-938
8. Manco M. Metabolic syndrome in childhood from impaired carbohydrate metabolism to nonalcoholic fatty liver disease., J Am Coll Nutr. 2011 Oct;30(5):295-303.
9. L. Lloyd, S. Langley-Evans, S. McMullen, Childhood obesity and adult cardiovascular disease risk: a systematic review, Int J Obesity 34, 2010, 18-28.
10. M. Cornier, J. Marshall, J. Hill, D. Maahs, R. Eckel, Prevention of overweight/obesity as a strategy to optimize cardiovascular health, Circulation 124, 2011, 840-850.
11. E. Abel, S. Litwin, G. Sweeney, Cardiac remodeling in obesity, Physiol Rev 88, 2008, 389-419.
12. L. Lloyd, S. Langley-Evans, S. McMullen, Childhood obesity and adult cardiovascular disease risk: a systematic review, Int J Obesity 34, 2010, 18-28.
13. van Gaal LF, Mertens IL, de Block CE. Mechanisms linking obesity with cardiovascular disease. Nature 2006; 444:875–880.

14. Romero-Corral, V. Somers, J. Sierra-Johnson, Y. Korenfeld, S. Boarin, J. Korinek, M. Jensen, G. Parati, F. Lopez-Jimenez, Normal weight obesity: a risk factor for cardiometabolic dysregulation and cardiovascular mortality, Eur Heart J. 31(6), 2010, 737-746.
15. de Lucia Rolfe E, Sleigh A, Finucane FM, Brage S, Stolk RP, Cooper C, Sharp SJ, Wareham NJ, Ong, KK. Ultrasound Measurements of Visceral and Subcutaneous Abdominal Thickness to Predict Abdominal Adiposity Among Older Men and Women. Obesity 2010; 18:625–631.
16. Iacobellis G, Assael F, Ribaudo MC, Zappaterreno A, Alessi G, di Mario U, Leonetti F. Epicardial Fat from Echocardiography: A New Method for Visceral Adipose Tissue Prediction. Obesity Research 2003; 11:304–310
17. Flüchter S, Haghi D, Dinter D, Heberlein W, Kühl HP, Neff W, Sueselbeck T, Borggrefe M, Papavassiliu T. Volumetric Assessment of Epicardial Adipose Tissue With Cardiovascular Magnetic Resonance Imaging. Obesity 2007; 15:870–878.
18. Iacobellis G, Willens HJ, Barbaro G, Sharma AM: Threshold values of high-risk echocardiographic epicardial fat thickness. Obesity 16, 887-892, 2008.
19. L.W. Poll, H. J. Wittsack, J. A. Koch, R. Willers, A. Scherer, C. Kapitza, L. Heinemann and U. Mödder. Quantification of total abdominal fat volumes using magnetic resonance imaging. European journal of medical research 7(8): 347-352, 2002 83. Q. Peng, R. W. McColl, Y. Ding, J. Wang, J. M. Chia, and P. T. Weatherall. Automated method for accurate abdominal fat quantification on water saturated magnetic resonance images. Journal of Magnetic Resonance Imaging, 26:238–746, 2007.
20. V. Positano, T. Christiansen, M. F. Santarelli, S. Ringgaard, L. Landini, and A. Gastaldelli. Accurate segmentation of subcutaneous and intermuscular adipose tissue from mr images of the thigh. Journal of Magnetic Resonance Imaging, 29(3):677–684, 2009
21. J. Kullberg, H. Ahlstroem, L. Johansson, and H. Frimmel. Automated and reproducible segmentation of visceral and subcutaneous adipose tissue from abdominal mri. International Journal of Obesity, 31(12):1806-1817, 2007
22. C. Würslin, J. Machann, H. Rempp, C. Claussen, B. Yang, and F. Schick. Topography mapping of whole body adipose tissue using a fully automated and standardised procedure. Journal of Magnetic Resonance Imaging, 31:430-439, 2010
23. D. Wald, B. Teucher, J. Dinkel, R. Kaaks, S. Delorme, H.-P. Meinzer, T. Heimann. Automated quantification of adipose and skeletal muscle tissue in whole-body MRI data for epidemiological studies. In SPIE Medical Imaging 2012: Computer-Aided Diagnosis, 2012.
24. V. Positano, A. Gastaldelli, A. M. Sironi, M. F. Santarelli, M. Lombardi, and L. Landini. An accurate and robust method for unsupervised assessment of abdominal fat by mri. Journal of Magnetic Resonance, 20:684–689, 2004.
25. T.-H. Liou, W. P. Chan, L.-C. Pan, P.-W. Lin, P. Chou, and C.-H. Chen. Fully automated large-scale assessment of visceral and subcutaneous abdominal adipose tissue by magnetic resonance imaging. International Journal of Obesity, 30:844–852, 2006.
26. Thomas E.L., Saeed N., Hajnal J.V., Brynes A., Goldstone A.P., Frost G., and Bell J.D. Magnetic resonance imaging of total body fat. Journal of Applied Physiology 85, 1998, 1778–1785.
27. W. Shen, Z. Wang, M. Punyanita, J. Lei, A. Sinav, J. G. Kral, C. Imielinska, R. Ross, and S. B. Heymsfield. Adipose tissue quantification by imaging methods: A proposed classification. Obesity Research, 11:5–16, 2003

28. J. Kullberg, H. Ahlstroem, L. Johansson, and H. Frimmel. Automated and reproducible segmentation of visceral and subcutaneous adipose tissue from abdominal mri. International Journal of Obesity, 31:1806–1817, 2007.
29. Zhou, H. Murillo, and Q. Peng. Novel segmentation method for abdominal fat quantification by MRI. Journal of Magnetic Resonance Imaging, 2011.
30. J.A. Noble and D. Boukerroui. Ultrasound image segmentation: A survey. IEEE Transactions on Medical Imaging, 25(8):987-1010, 2006.
31. J. Ng, R. Rohling and P.D. Lawrence. Automatic measurement of human subcutaneous fat with ultrasound. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 56(8):1642- 1653, 2009
32. Colombet I., Ruelland A., Chatellier G., Gueyffier F., Degoulet P., Jaulent M., Models to predict cardiovascular risk: comparison of CART, multilayer perceptron and logistic regression, AMIA 2000.
33. Kurt I., Ture M., Kurum T.A., Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease, Expert Systems with Applications, 34(1), 2008, 366-374.
34. Sumathi B., Santhakumaran A., Pre-diagnosis of hypertension using artificial neural network, Global J. of Computer Science and Technology 11(2), 2011.
35. Prevention of Cardiovascular Disease, Pocket Guidelines for Assessment and Management of Cardiovascular Risk, WHO, Geneva, 2007.
36. Tsymbal A, Zhou SK, Huber M. Neighborhood graph and learning discriminative distance functions for clinical decision support. Conf Proc IEEE Eng Med Biol Soc. 2009;2009:5617-20. PMID: 19964399
37. S. K. Zhou, J. Shao, B. Georgescu, D. Comaniciu, Boostmotion: Boosting a discriminative similarity function for motion estimation, in: Proc. of Int. Conf. on Computer Vision and Pattern Recognition, CVPR'06, Vol. 2, IEEE CS Press, 2006, 1761-1768.
38. D. Manset, F. Pourraz, A. Tsymbal, J. Revillard, K. Skaburskas, R. McClatchey, A. Anjum, A. Rios, M. Huber, Gridifying biomedical applications: experiences of the Health-e-Child project, In: Handbook of Research on Computational Grid Technologies for Life Sciences, Biomedicine and Healthcare, IGI Global Publishers, 2009, 469-493.
39. Depeursinge, A. Vargas, A. Platon, A. Geissbuhler, P.-A. Poletti, H. Müller, 3D Case-Based Retrieval for Interstitial Lung Diseases, MICCAI workshop on Medical Content-Based Retrieval for Clinical Decision Support, Springer, LNCS 5853, 2010, 39-48.
40. Ruch P: Automatic assignment of biomedical categories: toward a generic approach. Bioinformatics 22(6): 658-664 (2006)
41. Ruch P, Gobeill J, Tbahriti I,Geissbühler A. From Episodes of Care to Diagnosis Codes: Automatic Text Categorization for Medico-Economic Encoding. AMIA. 2008.
42. Pasche E, Gobeill J, Teodoro D, Vishnyakova D, Gaudinat A, Ruch P, Lovis C. Using multimodal mining to drive clinical guidelines development. Stud Health Technol Inform. 2011;169:477-81.
43. Vulevic J, Juric A, Tzortzis G, Gibson GR. A Mixture of trans-Galactooligosaccharides Reduces Markers of Metabolic Syndrome and Modulates the Fecal Microbiota and Immune Function of Overweight Adults. J Nutr. 2013 Mar;143(3):324-31. doi: 10.3945/jn.112.166132. Epub 2013 Jan 9.

44. Di Girolamo F, Del Chierico F, Caenaro G, Lante I, Muraca M, Putignani L. Human serum proteome analysis: new source of markers in metabolic disorders. Biomark Med. 2012 Dec;6(6):759-73. doi: 10.2217/bmm.12.92.
45. Aagaard K, Riehle K, Ma J, Segata N, Mistretta TA, Coarfa C, Raza S, Rosenbaum S, Van den Veyver I, Milosavljevic A, Gevers D, Huttenhower C, Petrosino J, Versalovic J. 2012. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. PLoS One 7(6):e36466. Epub Jun 13.
46. Mitra S, Stärk M, Huson DH. Analysis of 16S rRNA environmental sequences using MEGAN. BMC Genomics. 2011 Nov 30;12 Suppl 3:S17. Epub 2011 Nov 30.
47. Wylie KM, Truty RM, Sharpton TJ, Mihindukulasuriya KA, Zhou Y, Gao H, Sodergren E, Weinstock GM, Pollard KS. 2012. Novel bacterial taxa in the human microbiome. PLoS One 7(6):e35294. Epub Jun 13.
48. Baecke JA, Burema J, Frijters JE. 1982. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr 36(5):936-942.
49. Baggio LL, Drucker DJ. 2007. Biology of incretins: GLP-1 and GIP. Gastroenterology 132(6):2131-2157.
50. Blundell JE, Burley VJ, Cotton JR, Lawton CL. 1993. Dietary fat and the control of energy intake: evaluating the effects of fat on meal size and postmeal satiety. Am J Clin Nutr 57(5 Suppl):772S-777S; discussion 777S-778S.
51. Carel JC, Leger J. 2008. Clinical practice. Precocious puberty. N Engl J Med 358(22):2366-2377.
52. Carroll JF, Kaiser KA, Franks SF, Deere C, Caffrey JL. 2007. Influence of BMI and gender on postprandial hormone responses. Obesity (Silver Spring) 15(12):2974-2983.
53. Ceriello A, Quagliaro L, Piconi L, Assaloni R, Da Ros R, Maier A, Esposito K, Giugliano D. 2004. Effect of postprandial hypertriglyceridemia and hyperglycemia on circulating adhesion molecules and oxidative stress generation and the possible role of simvastatin treatment. Diabetes 53(3):701-710.
54. Chandler-Laney PC, Higgins PB, Granger W, Alvarez J, Casazza K, Fernandez JR, Dalla Man C, Cobelli C, Gower BA. 2013. Use of a simple liquid meal test to evaluate insulin sensitivity and beta-cell function in children. Pediatric obesity.
55. Kowalik G, Steeden JA, Atkinson D, Muthurangu V. 2011. A networked GPU reconstructor within the clinical workflow for rapid fat quantification. Proceedings of the 19th Annual Meeting of ISMRM (Abstract 2708).
56. Molarius A, Seidell JC, Sans S, Tuomilehto J, Kuulasmaa K. 1999. Waist and hip circumferences, and waist-hip ratio in 19 populations of the WHO MONICA Project. Int J Obes Relat Metab Disord 23(2):116-125.
57. Nappo F, Esposito K, Cioffi M, Giugliano G, Molinari AM, Paolisso G, Marfella R, Giugliano D. 2002. Postprandial endothelial activation in healthy subjects and in type 2 diabetic patients: role of fat and carbohydrate meals. J Am Coll Cardiol 39(7):1145-1150.
58. The National Statistics Socio-Economic Classification. UK2001.
59. Raz O, Steinvil A, Berliner S, Rosenzweig T, Justo D, Shapira I. 2013. The effect of two iso-caloric meals containing equal amounts of fats with a different fat composition on the inflammatory and metabolic markers in apparently healthy volunteers. J Inflamm (Lond) 10(1):3.
60. Richard C, Lussier MT, Gagnon R, Lamarche L. 2004. GHQ-28 and cGHQ-28: implications of two scoring methods for the GHQ in a primary care setting. Soc Psychiatry Psychiatr Epidemiol 39(3):235—243.

61. Schembre S, Greene G, Melanson K. 2009. Development and validation of a weight-related eating questionnaire. Eating behaviors 10(2):119-124.
62. Schembre SM, Geller KS. 2011. Psychometric properties and construct validity of the Weight-Related Eating Questionnaire in a diverse population. Obesity (Silver Spring) 19(12):2336-2344.
63. Steeden JA, Atkinson D, Hansen MS, Taylor AM, Muthurangu V. 2011. Rapid flow assessment of congenital heart disease with high-spatiotemporal-resolution gated spiral phase-contrast MR imaging. Radiology 260(1):79-87.
64. Stergiopulos N, Meister JJ, Westerhof N. 1994. Simple and accurate way for estimating total and segmental arterial compliance: the pulse pressure method. Ann Biomed Eng 22(4):392-397.
65. Sun Y, Tao FB, Su PY, China Puberty Research C. 2012. Self-assessment of pubertal Tanner stage by realistic colour images in representative Chinese obese and non-obese children and adolescents. Acta Paediatr 101(4):e163-166.
66. Tsai WC, Li YH, Lin CC, Chao TH, Chen JH. 2004. Effects of oxidative stress on endothelial function after a high-fat meal. Clin Sci (Lond) 106(3):315-319.
67. van 't Riet E, Rijkelijkhuizen JM, Alssema M, Nijpels G, Stehouwer CD, Heine RJ, Dekker JM. 2012. HbA1c is an independent predictor of non-fatal cardiovascular disease in a Caucasian population without diabetes: a 10-year follow-up of the Hoorn Study. European journal of preventive cardiology 19(1):23-31.
68. van Oostrom AJ, van Dijk H, Verseyden C, Sniderman AD, Cianflone K, Rabelink TJ, Castro Cabezas M. 2004. Addition of glucose to an oral fat load reduces postprandial free fatty acids and prevents the postprandial increase in complement component 3. Am J Clin Nutr 79(3):510-515.
69. Vicennati V, Ceroni L, Gagliardi L, Gambineri A, Pasquali R. 2002. Comment: response of the hypothalamic-pituitary-adrenocortical axis to high-protein/fat and high-carbohydrate meals in women with different obesity phenotypes. J Clin Endocrinol Metab 87(8):3984-3988.
70. Yu H, McKenzie CA, Shimakawa A, Vu AT, Brau AC, Beatty PJ, Pineda AR, Brittain JH, Reeder SB. 2007. Multiecho reconstruction for simultaneous water-fat decomposition and T2* estimation. J Magn Reson Imaging 26(4):1153-1161.
71. Ceriello A, Quagliaro L, Piconi L, Assaloni R, Da Ros R, Maier A, Esposito K, Giugliano D. Effect of postprandial hypertriglyceridemia and hyperglycemia on circulating adhesion molecules and oxidative stress generation and the possible role of simvastatin treatment. *Diabetes*. 2004;53:701-710.
72. Gray L, Lee IM, Sesso HD, Batty GD. Blood pressure in early adulthood, hypertension in middle age, and future cardiovascular disease mortality: HAHS (Harvard Alumni Health Study). *J Am Coll Cardiol*. 2011;58:2396-2403.
73. Lipsitz LA, Ryan SM, Parker JA, Freeman R, Wei JY, Goldberger AL. Hemodynamic and autonomic nervous system responses to mixed meal ingestion in healthy young and old subjects and dysautonomic patients with postprandial hypotension. *Circulation*. 1993;87:391-400.
74. Masui T, Isoda H, Mochizuki T, Takahashi M, Kaneko M, Shirakawa T, Ota A. Effects of meal intake on the flow velocity in the superior mesenteric artery: evaluation with 2D phase mapping MRI. *J Comput Assist Tomogr*. 1994;18:590-595.

WP 5: Data acquisition and processing for Juvenile Idiopathic Arthritis

Protocol no:
MD-Paedigree – WP 5**Version 2:** Apr 15, 2013**CONFIDENTIAL**

Protocol no.:	
Title:	WP 5 – Data acquisition and processing for Juvenile Idiopathic Arthritis
Acronym:	MD-Paedigree – WP 5
Multicentric/Monocentric Study	Multicentric
Scientific Coordinator of the Project	Prof. Bruno Dallapiccola (Ospedale Pediatrico Bmbino Gesù)
Sponsor:	Bambino Gesù Children'S Hospital(European Commission)
Responsible Work Package 5:	Prof. Alberto Martini (Istituto Giannina Gaslini)
Principal Investigator	Dr. Fabrizio De Benedetti (Ospedale Pediatrico Bambino Gesù)
Data Management/Statistical analysis:	

Protocol approved and signed by:

Scientific Coordinator of the Project

Prof. Bruno Dallapiccola

Responsible Work Package 5:
Prof. Alberto Martini.

Principal Investigator

Dr. Fabrizio De Benedetti
Acronym List

AEs	Adverse Events
EC	Ethical Committee
CRF	Case Report Form
GCP	Good Clinical Practice
ICH	International Conference on Harmonisation
IRB/IEC	Institutional Review Board/Independent Ethics Committee
NSAEs	Non Serious Adverse Events
SAEs	Serious Adverse Events
SOPs	Standard Operating Procedures
……	……..

INDEX

1 INTRODUCTION

1.1 GENERAL BACKGROUND OF THE PROJECT

MD – PAEDIGREE *Model-Driven European Paediatric Repository* is a research project funded by the European Commission under the Virtual Physiological and Human Area (VPH) of the ICT Theme

of the Seventh Framework Programme (contract no. 600932). The project foresees 7 world-renowned clinical centres of excellence pursuing improved interoperability of paediatric biomedical information, data and knowledge, by developing a set of reusable and adaptable multi-scale models for more predictive, individualised, effective and safer paediatric healthcare. The project is scientifically and technologically supported by one of the leading industrial actors in medical applications in Europe, and operating in conjunction with highly qualified SMEs and some of the most experienced research partners in the Virtual Phisiological Human (VPH) community.

MD-Paedigree validates and brings to maturity patient-specific computer-based predictive models of various paediatric diseases (cardiomyopathy, risk of cardiovascular disease in obese children and adolescents, Juvenile Idiopathic Arthritis (JIA) and neurological and neuro-muscular diseases (NND), thus increasing their potential acceptance in the clinical and biomedical research environment by making them readily available not only in the form of sustainable models and simulations, but also as newly-defined workflows for personalized predictive medicine at the point of care.

MD-Paedigree aims to advance the state-of-the-art of patient-specific computational modelling of different paediatric diseases and translate the latest advances into clinics to improve disease understanding, therapy outcome, and provide an infostructure platform for assessing new therapies at the point of care.

MD-Paedigree's goals therefore are to integrate and share highly heterogeneous biomedical information, data and knowledge, using best, jointly develop reusable, adaptable and composable multi-scale VPH workflow models and to support evidence-based translational medicine at the point of care practices from the biomedical semantic Web.

Illness in infants, children, and adolescents are a large and under-appreciated public health problem. Because paediatric patients now have a much longer life expectancy, the burden and costs are substantial for families and society. For instance, recent studies have shown that the number of adult patients with congenital heart disease is already similar to that of the paediatric population and will continue to grow. It can therefore be stated that in the longer term the VPH scientific approach requires a fundamental research investment in paediatrics, where the conceptual revolution which is underway, transforming the nature of healthcare from reactive to preventive, can best be applied, moving towards an approach based on personalized, predictive, preventive, and participatory (P4) medicine, increasingly focused on wellness.

MD-Paedigree represents a major step towards personalized paediatric e-Health, based on data-driven models, patient-specific simulations and a sustainable data and model repository. The project can in fact impact the way healthcare is practiced in the future. MD-Paedigree's impact on dealing with very concrete and exemplary clinical cases is demonstrated through the following applications in the diseases scenarios. MD-Paedigree encompasses complete services for storage, similarity search, outcome analysis, risk stratification, and personalized decision support in paediatrics within its innovative model-driven data and workflow-based models repository, leveraging on service and knowledge utilities. It fosters the state-of-the-art of patient-specific computational modelling of the selected diseases and translates the latest advances into clinics to improve disease understanding and provide a platform for testing new therapies at the point of care. In fact, MD-Paedigree demonstrates how a dedicated VPH repository can provide full accessibility to existing and further developing knowledge in paediatrics to all interested bio-medical researchers precisely through linking data with models, thus proving the large scale benefits of having both the data and models readily available at the point of care.

The data collection performed within Health-e-Child and Sim-e-Child projects (already funded by the European Commission in previous calls for proposals) is still available to the MD-Paedigree consortium thanks to the continuity in the eHealth platform which is shared by all three projects. In addition, the new patients' recruitment to be performed within MD-Paedigree consists of:

<table>
<tr><th>Pathology</th><th>No of patients/Time</th><th></th></tr>
<tr><td>Cardiomyopathies</td><td>180 children, by month 33: 60 patients (among which 30 girls) for each clinical centre.</td><td rowspan="3">Genetic and meta-genomic:

180 patients with cardiomiopathies, 180 with CVD risk in obesity, 200 with JIA, and 100 unaffected subjects (control group).</td></tr>
<tr><td>CVD risk in obese children</td><td>180 patients , by month 36: 60 (among which 30 girls) for each clinical centre.</td></tr>
<tr><td>Juvenile Idiopathic Arthitis (JIA)</td><td>Altogether 200 patients by month 28.</td></tr>
<tr><td rowspan="3">NND</td><td colspan="2">Cerebral Palsy: 50 patients for each clinical centre for probabilistic modelling, as well as 600 retrospective patients from KU Leuven and OPBG.</td></tr>
<tr><td colspan="2">Spinal Muscular Atrophy (SMA)
20 ambulant patients (severity grade type 3);
10 patients for each centre for biophysical modeling;
10 patients among the 3a subgroup (symptoms of weakness appearing before age 3 years);
10 patients among the 3b subgroup (weakness appearing after the age of 3 years.</td></tr>
<tr><td colspan="2">Duchenne Muscular Dystrophy (DMD)
Clinical data will be collected by OPBG, KU Leuven and VUA from 20 ambulant genetically confirmed DMD Patients. 10 patients with an age ranging between 5 and 6 years, additional 10 patients with an age ranging between 7 and 8 years.</td></tr>
</table>

1.2 BACKGROUND OF THE DATA ACQUISITION AND PROCESSING FOR JUVENILE IDIOPATHIC ARTHRITIS STUDY

Juvenile idiopathic arthritis (JIA) is a broad term that describes a clinically heterogeneous group of arthritis which has an onset before age of 16 years, lasts more than 6 weeks and is of unknown origin. The cause and pathogenesis of JIA are still poorly understood, but likely they include both genetic and environmental components. Moreover, disease heterogeneity implies that different factors probably contribute to its pathogenesis and causes [Prakken B et al., 2011]27.

Affected joints develop synovial proliferation and infiltration by inflammatory cells which may ultimately lead to destructive lesions of joint structures, disability and high disease-related costs. Indeed, JIA which affects approximately one in 1,000 children represents the leading cause of childhood disability from a

musculoskeletal disorder. Current classification, which is based on clinical criteria, is still unsatisfactory: considerable heterogeneity in disease course and treatment response exists, both between and within subtypes of JIA.

Unfortunately, the present ability to predict the disease course and outcome is limited. Within the FP6 Health-e-Child project, ICT tools for diagnosis and scoring of JIA, based on image data of the wrist, have been developed.

This framework is the basis for the developments planned for MD-Paedigree. Comprehensive and accurate computer models derived from patient-specific data across multiple scales covering body, organs, tissues, and molecular levels are developed.

This data is gathered and stored in a standardized manner building upon the Health-e-Child software tools developed for wrist analysis in the context of JIA. These tools are extended for the purpose of integrating model information related to a wider range of joints, covering morphology, gait analysis, bio/genetic data. The tools to be developed will also include the aspect of a multidimensional longitudinal analysis that yields the opportunity to identify potential new outcome measures (imaging or biological biomarkers) for the assessment of treatment efficacy. Furthermore, the prognostic value on an individual level of multidimensional data, including modern imaging modalities, genetic and meta-genetic data, will be explored through the development and integration of appropriate data clustering methods.

By collecting patient specific multi-scale and multi-dimensional information and automating image and data analysis at the point of care, this project has a strong clinical impact on early diagnosis, prediction of disease and of treatment outcome.

The impact of biomechanical property alterations on subsequent progression of structural damage in patients with chronic inflammatory arthritis is not yet characterized. Personalized joint biomechanical modeling allows critical evaluation of the forces within the joint under physiologic and pathological loading conditions. Evaluation of the impact of joint mechanical abnormalities on disease progression is needed for an accurate outcome prediction. The modelling predictions could have significant implications in early diagnosis and therapeutic intervention. In this perspective, early signs of structural damage will be evaluated also using MR imaging analysis. In the frame of the EU FP6 Health-e-Child project, a great deal of effort had been spent in order to standardize imaging procedures and devise paediatric-targeted scoring systems for the assessment of disease activity and damage in JIA considering the wrist [Malattia C, et al., 2011] 107. The collaboration between clinical and IT partners has enabled the development and validation of computerized quantitative measurements of inflammation and destructive changes that have shown potential value as predictors of future damage [Malattia C et al. ,2012;] 108. In continuity from the work developed in Health-e-Child, which has led to advanced personalized modelling of disease progression, the goal will be to implement a more robust multi-scale, personalized and predictive computer-based model of JIA – this time focusing on a wider range of joints than the wrist joint. It will span body, organ, tissue and molecular level with adequate information fusion and in addition information obtained from gait analysis. This allows for pattern discovery in multimodal data through correlations between clinical data, imaging, immunological, metagenomic data (gut microbiota), and a biomechanical gait model. The driving force behind this project stems from the integration of data coming from a new cohort of patients (approximately 200 patients) into the framework developed within the Health-e-Child project that will be further extended and adapted to the needs of MD-Paedigree.

Initial imaging will be performed at disease onset and followed for 2 years at least, in order to expand predictive multi-scale models in JIA. The longitudinal design of the study will allow a dynamic process of testing multi-scale disease models for each patient at follow-up visits to further personalize treatment strategies.

Imaging of the Affected Joints

By fusing the information on the anatomy and the physical properties of the tissues provided by the imaging technologies, with the functional information provided by the CGA, it will be possible to personalise a whole body-level model of the musculoskeletal dynamics capable of predicting the forces acting on a given joint during the patient movements [Martelli et al, 2011]109 [Taddei et al. 2011]110. These forces will then be applied to an organ-level finite element model of the joint, where the mechanical properties of the tissues will be informed as much as possible from the imaging data [Viceconti et al, 2012]111. Among the other things we shall explore the possibility to derive cancellous bone anisotropy from DTI-like MRI imaging, mechanical properties of the cartilage from information obtained by MRI, etc. We shall also correlate the biomechanical predictions with the signatures of the disease that can be quantified, such as the extension and the location of the cartilage erosion, or the alteration of the subchondral bone, to the predictions of stress and strains obtained by the organ-level model. As shown in literature [Magni-Manzoni S et al., 2012]112, a combination of MRI and US imaging is beneficial for the assessment of JIA. High-resolution US will be performed in order to better define the extent of the disease. The severity of joint involvement will be judged sonographically by a variety of parameters such as joint effusion, synovial thickening and hyperaemia, cartilage integrity and bone erosions. Quantitative assessments of these parameters will be extracted from the US equipment based on standardized scanning planes by means of 2D imaging. In addition to MRI and US, whole-body Dual X-ray Absorptiometry (DXA) will be performed. Total body DXA provides an accurate measurement of the areal body density over the frontal plane, separating the bone mass, the lean mass (muscles), and the fat mass with good accuracy. This imaging modality will be used to personalise multi-scale models of the musculoskeletal system capable of predicting the forces transmitted at the joints during a given movement. DXA images will be used not only to FP7- ICT-2011.5.2 600932 - MD-Paedigree –Part B 16 personalise these generic models anatomically: total body density will be used to define the inertial properties ofthe model; lean mass will be used to estimate the muscles cross-section in the musculoskeletal model; bone density will be used to personalise the bone stiffness in the joint models. All these personalised models will be composed in an integrative multiscale representation of the patient's musculoskeletal system, capable of predicting, for example, the forces being transferred to the joint cartilage during a given movement as captured during the gait analysis.

Articulated Modelling of the Affected Joints for Automated Biomarker Extraction

The progress beyond the Health-e-Child project is defined by clinical as well as technical aspects. The wrist MRI scores, as well as the automated software for the quantitative assessment of disease activity and damage, developed in the frame of Health-e-Child, will be adapted to investigate the ankle. Focusing on the locomotory system, especially the juvenile ankle, enables the physician to study the effects of JIA on the joint motion, which form another scale in the patient-specific model. MD-Peadigree aims to automate and extend the multimodal image analysis [Malattia C, et al., 2008] 113, 114, and therefore, standardise the derived biomarkers by means of model-based segmentation of MRI images. For this purpose, an articulated model of the juvenile ankle will be developed and used. It includes the bones' shape, the spatial relation between the bones and their appearance in MRI images. By simulating the joint articulation, it will allow for the adaption to a specific MRI-scan, resulting in patient-specific models. In order to generate a personalised morphological model for JIA, an articulated joint model – consisting of bones, cartilage and ligaments representing the variation in shape, image appearance and spatial relations trained using machine learning methods – will be developed. It will be built from manual annotations by experts on morphological MRI datasets of patients suffering from JIA. Data from US evaluation will be also included.

Musculoskeletal Modelling of the Joint Kinetics

Furthermore, the role of the musculoskeletal dynamics and of the mechanical properties of the joint tissues in conditioning disease progression or in response to treatment will be investigated. The integration of image based patient-specific models with gait cycle analysis will allow the generation of highly personalised multiscale models of the musculoskeletal system capable of elucidating the role of biomechanical properties in onset and/or progression of structural damages. Three-dimensional clinical gait analysis (CGA)

is a well-established method enabling, when a strict analysis of causes of errors is carried out and periodical validation procedures are implemented (see for more details the paragraph Neurological and Neuro-muscular Diseases - NND) - Progress beyond the State-of-the-Art) highly objective and reliable evaluation of gait in both healthy and diseased populations. CGA including kinematics and kinetics, provide more information about gait changes, such as joint angles and moments, which cannot be quantified in a standard clinical setting. The kinematics shows the joint movement, while the kinetics describes the forces involved in movement (e.g. ground reaction forces, joint moments, and joint powers). By examining kinetics, the mechanisms of gait deviation can be described and the early use of gait analysis can be instrumental in discovering developments of potentially destructive gait deviations. Patients will be dressed with skin-attached markers that are both visible in MRI imaging, radiopaque (so they appear also in the DXA image) and, successively, reflective markers will be reapplied in the same anatomical positions, so they can be tracked during gait analysis. Whole body imaging and gait analysis will be performed one after the other with the patient dressed with the markers. This will provide a fiducial registration framework between anatomical and functional data. The imaging protocol will be agreed with the modellers, in order to ensure that the highest amount of information is transferred to the predictive models. Each patient will be examined using three-dimensional clinical gait analysis (CGA), ground force platform, and cutaneous electromyography (EMG). Depending on the joint of interest, the patient will FP7- ICT-2011.5.2 600932 - MD-Paedigree – Part B 17 be asked to repeat a few times a given movement, selected among those most common in daily life (i.e. for lower limb, level walking, stair climbing, sit to stand, etc.), and the relative motions and muscle activation signals will be recorded. An expert physiatrist will examine the gait analysis data to exclude specific gait abnormalities. Using the fiducial marker set, the motion data will be fused with the imaging data, and with the internal, musculoskeletal, and joint models fitted to the imaging data. This will result in a body-organ multi-scale model capable of predicting the forces being transferred to the joint during each of the recorded movements. EMG data will not be used to inform the model, but will be compared with the activation patterns predicted by the models, so as to verify that the model is operating consistently with the patient's neuromuscular activation strategy. The body model will use inverse kinematics to find the optimal registration framework between the model and the recorded kinematics, so as to reduce as much as possible the so-called skin artefacts. Then, inverse dynamics will be used to compute the joints torque that is required to generate the recorded movement. An optimisation scheme will be used to compute muscle activations and joint forces. This time-varying system of musculo-articular forces will be applied as boundary condition to a finite element model of the joint being investigated. The individualised finite element model will predict the mechanical stresses and strains induced in the various joint tissues by the given movement, and information to be used as an additional "biomarker" in the evaluation of the individual clinical case.

Immunological and Genetic Analysis

Imaging data will be integrated with immunological and metagenomic data in order to try to identify surrogate parameters for disease activity, disease severity, risk of side effects and treatment outcomes. New particle-based multiplex immunoassay, such as the Luminex technology [de Jager W et al., 2007]115, allowing the measurement of multiple circulating and/or synovial cytokines, as well as of other immune mediators, will be used to define the individual immunological profile for each patient. Furthermore, paired peripheral blood and synovial fluid mononuclear cells subpopulations (naive and effectors T cells, B cells, monocytes, etc.) will be evaluated by cytofluorimetric analysis. We will also look at phenotypic markers, mRNA, epigenetic markers (methylation FOXP3) and functionality (in vitro suppression assays).

Analysis of gut microbiota (the genome of microbes present in the gastrointestinal tract) will provide new insight into the environmental factors which regulate innate and adaptive immune homeostasis and affect the development of systemic autoimmune diseases. The gastrointestinal tract is the largest human immune organ and home to a complex community of trillions of bacteria that are engaged in a dynamic interaction with the host immune system. (The human body contains over 10 times more microbial cells than human cells). Communication between the microbiota and the host establishes and maintains immune homeostasis, enabling protective immune responses against pathogens while preventing adverse inflammatory responses to harmless commensal microbes. Correlations have been found between the

composition of gut microbiota and some preferential immune responses (i.e. Th17 response). By analysing the gut microbiota of JIA patients collected in specific disease states (at the onset, when patient will achieve clinical remission state, and during flare of the disease) we aim to explore its potential role in conditioning disease susceptibility as well as immune response in the different stages of disease, thus adding a further important dimension to multiscale analysis. Investigating the interaction of gut microbes and the host immune system will improve the understanding of the pathogenesis of this autoimmune disease, and provide innovative foundations for the design of novel immuno- or microbe-based therapies.

Prediction of the Disease Course

The impact of biomechanical property alterations on subsequent progression of structural damage in patients with chronic inflammatory arthritis is not yet characterised. Personalised joint biomechanical modeling allows critical evaluation of the forces within the joint under physiologic and pathological loading conditions. Evaluation of the impact of joint mechanical abnormalities on disease progression is needed for an accurate outcome prediction. The potential of the multi-scale modeling methods proposed, is to make the exploration of complex systemic interactions between the neuromuscular control, the musculoskeletal functional anatomy, and the local biomechanical determinants acting in the joint space at the tissue level, possible. The modelling predictions could have significant implications in early diagnosis and therapeutic intervention. In this perspective, early signs of structural damage will be evaluated using MRI. Demographic clinical imaging and laboratory data in the form of text, images, annotations, videos, biomarkers and articulated models will be entered in the MD-Paedigree digital repository and will be continuously analysed providing potentially more accurate disease model tools. The combination of different assessment techniques will enable to enhance the value of a multidisciplinary management of JIA. The multidimensionality of the human and microbial phenotypes (and the dynamic, nonlinear interactions) will be explored by means of improved informatics tools, including new approaches for understanding the complexity of the metadata, in order to better understand the implications of gut microbiota variations in human health and disease.

The prognostic value on an individual level of multidimensional data, including modern imaging modalities, immunological, metagenomic data, as well as articulated models and biomechanical models will be explored. JIA constitutes an ideal domain for assessing the merits of simulators and predictors based on data generated across different scales. The validity and effectiveness of the proposed solutions will be assessed by using the model to address several open issues in JIA with a strong clinical impact on early diagnosis, prediction of disease and of treatment outcome.

2 AIMS OF THE STUDY

2.1 MAIN GOAL

The goal of the study is to collect clinical, immunological, metagenomic and imaging data for the subsequent integrated analysis of JIA. Data collection is set-up as a prospective longitudinal study. The timeframe for patient recruitment spans the first 28 months. The objective is to acquire data from about 200 patients within the first 28 months (baseline acquisitions). For each patient, follow-up data will be collected for monitoring disease course and to identify outcome predictors.

2.2 PRIMARY END POINT

To provide potentially more accurate disease model tools through the collection of demographic clinical imaging and laboratory data in the form of text, images, annotations, videos, biomarkers and articulated models to be entered and continuously analysed in the MD-Paedigree digital repository. The combination of different assessment techniques will enable to enhance the value of a multidisciplinary management of JIA.

2.3 SECONDARY END POINT

To clinically assess multidisciplinary derived models, in order to improve the identification of markers of outcome prediction and risk stratification, and thus to derive and evaluate personalised treatment models.

3 STUDY DESCRIPTION

3.1 STUDY DESIGN

The study will last 4 years. It is designed as a prospective longitudinal study. The expected patient sample is 200 JIA patients. All consecutive patients with JIA, disease duration < 6 months and active arthritis will be enrolled in the study.

The following **clinical data** will be acquired at 6 months follow up intervals for the first two years from patient enrolment.

- demographic data such as gender, age at disease onset, JIA subtype according to ILAR classification, etc.
- clinical variables including standardised and validated measures of disease activity and disease damage (e.g. number and site of inflamed joints, presence of systemic feature, functional ability, the Juvenile Arthritis Disease Activity Score, the Juvenile Arthritis Damage Index etc.) will be collected at enrolment and every 6 months.

Information concerning previous and ongoing treatment will be recorded.

Routine laboratory tests to extract markers of inflammation such as ESR, CRP, antinuclear antibodies, and rheumatoid factor will be performed at enrolment and every 6 months.

Immunological and Genetic Analysis

Sample collection, **storage and DNA extraction**: samples will be collected from 200 patients for rheumatology. A database of patients, including name, age, disease, laboratory data but also specific indications about antibiotic, prebiotics, and probiotics administration will accomplish, faecal sample datasheets for appropriate later description of gut microbiota enterotypes. Samples will be stored at 4°C for at maximum of 24 hours or, alternatively, at -80°C until shipment to metagenomic facilities for automatic DNA extraction and targeted-sequencing. In the latter case, the samples will be sent every two months or when suitable for the laboratories included in the study in dry ice by express courier.

Synovial and blood Cytokine and inflammatory mediators profile

Biological samples (blood, and synovial fluid from patients with clinical indication to perform local steroid injection) will be collected at disease onset, when patient will achieve clinical remission state (according to Wallace criteria for remission in JIA) and during flare of the disease.
For biomarkers a high throughput methodology will be used, namely the multiplex immuno assay or Luminex®. This is a bead-based assay that allows the detection of more than 100 soluble mediators in a single sample of 50 microliter of body fluid, such as plasma or synovial fluid. Partner UMCU is an international expertise centre for this technology and has developed a "home-brew" assay for the determination of over 100 soluble factors, mostly cytokines and all directly related to inflammation, and thus potential co-determining risk factors for inflammation. A set of markers related to inflammation and/or cardiovascular risk, mostly adipokines and cytokines. We will perform pilot experiments in small proof-of principle cohorts (max 20 patients) will be measured to determine the panel that will be measured in a large validation cohort. These markers will be measured in peripheral blood plasma, and, if available in synovial fluid. In a smaller subpopulation of patients , based on the results from the previous studies, we will perform T cell characterisation in paired peripheral blood and synovial fluid derived mononuclear cells focusing on regulatory T cells (natural and induced regulatory T cells expressing FOXP3, Tr1 cells) and effector T cells (Th17, Th1 cells). Phenotypic markers, mRNA, epigenetic markers (methylation FOXP3) and functionality (in vitro suppression assays) will be observed.

Microbioma analysis will provide an opportunity to understand how the gut microbiota regulates innate and adaptive immune homeostasis and affects the development of systemic autoimmune diseases. Dysregulation of host responses as a consequence of dysbiosis in the gut lumen could affect distant anatomical sites through the activation of host immune responses.
Stool samples which will be collected at disease onset, when patient will achieve clinical remission state (according to Wallace criteria for remission in JIA) and during flare of the disease. The results of gut microbiota analysis will be integrated with clinical immunological and imaging data to assess how it does affect human health, and in particular to explore the prognostic value of the presence of major clustering patterns at the gastrointestinal tract in conditioning disease susceptibility as well as the immune response in the various phases of the disease.

In order to analyse the taxonomic gut content of JIA patients, a targeted approach based on sequencing of the variable regions V1 and V3 of 16S rRNA locus will be used (Aagaard *et al*., 2012. PLoS One 7(6):e36466. Epub Jun 13). Fecal samples will be collected and analysed at onset of disease, at time of clinical remission, and during disease flares, with a prediction of approximately 400 samples.

Analysis of microbiome of fecal samples will be carried out following DNA extraction (automatic EZ1 Biorobot, Qiagen), and further pyrosequencing using a 454 Junior apparatus and sequence analysis; comparison will be performed with the recently developed MEGAN 4 software (available at http://www-ab.informatik.unituebingen.de/software/megan) (Mitra et al., BMC Genomics 2011), or with the PhylOTU software (https://github.com/sharpton/PhylOTU) (Wylie et al., 2012. PLoS One 7(6):e35294. Epub Jun 13.), in order to identify the microbiota operational taxonomic units (OTUs).

The following imaging techniques will be performed:

Ultrasound imaging: high-resolution ultrasound (U/S) evaluation of joints will be performed using a machine, equipped with broadband linear-array transducers. At the patient enrolment in addition to all clinically affected jointselbows, knees and ankles will be also investigated with U/S for a more accurate assessment of disease extension. U/S follow-up data will then be acquired at 6 months follow-up intervals for the first two years from patients enrolment in the baseline affected and newly affected joints . The U/S scanning protocols will be based on the standardized technical guidelines issued by the European Society of Musculoskeletal Radiology and the OMERACT US group. The severity of joint involvement may be judged sonographically by a variety of gray-scale parameters, including the amount of joint effusion, the presence of synovial thickening, the degree and duration of synovial hyperemia, the occurrence of cartilage abnormalities and bone erosions. In our protocols, quantitative assessments of these parameters will be extracted from the U/S equipment based on standardized scanning planes by means of 2D imaging. Correlation between the site of gray-scale damage and the site of hyperemia will be performed in order to identify patterns of hyperemia that may be predictive of disease progression.
In patients with ankle involvement the following investigations will be also performed:

- Digital plain radiography: will be performed at the enrolment and after 2 years to assess the presence and the degree of local growth disturbances, abnormal joint alignment (i.e. joint subluxation, dislocation and flexion/extension defects) and structural damage progression.Magnetic Resonance (MR): will be performed on a 1.5 Tesla MR scanner at the time of patient enrolment and after 2 year. The following image sequences will be used: TSE T1 3D; TSE T2 fat sat; GRE 3D fat sat. MRI detectable pathological findings will be extracted using both a semi-quantitative and a quantitative approaches.
  In case of unilateral involvement, the controlateral ankle will be also scanned at the same time to be used as a control.
- Dual X-ray Absorptiometry (DXA): will be performed at 6 month follow-up visit. Total body DXA provides an accurate measurement of the areal body density over the frontal plane, separating with good accuracy the bone mass, the lean mass (muscles), and the fat mass. Being a radiological image, it provides a fairly accurate spatial location of the joint centres, and of other skeletal landmarks.

Gait cycle analysis

Quantitative gait assessment will be carried out at the enrolment with CGAs installed at the labs and the reflective markers will be attached bilaterally on the participant's skin at the shoulders, trunk, pelvis, legs and feet. Children will be revaluated at 6 month follow-up visit.. The same examiner will perform the clinical measurement and marker placement in these children. To evaluate kinetic and kinematic variables in all three anatomical planes we plan to calculate from five gait cycles beginning with the left foot strike and five gait cycles beginning with the right foot strike. Differences will be evaluated in children with JIA between pre- and post-treatment gait analyses using a Repeated Measures Analysis of Variance (ANOVA). Non-parametric statistical (Mann–Whitney) tests will be used to determine differences between children with JIA before treatment and controls, and between children with JIA after treatment and controls.

3.2 SUBJECTS SELECTION

Data collection will be performed in three leading European Pediatric Rheumatologic Centers (IRCCS Istituto Giannina Gaslini, Genova, Italy; IRCCS Ospedale Pediatrico Bambino Gesù, Rome; Universitair MediscCentrum Utrecht, the Netherlands). The following patient selection criteria will be applied.

Inclusion criteria:

- Children and adolescents with JIA according to ILAR criteria and disease duration < 6
- months.
- Parents or legal guardian (and the subject when age is appropriate) must be willing to sign the consent/assent forms.

Exclusion criteria

- Patients requiring general anesthesia or with contraindication to MRI will be excluded from the study.

The timeframe for patient recruitment spans the first 28 months. The objective is to acquire data from about 200 patients within the first 28 months. For each patient, follow-up data will be collected for at least 24 months for monitoring disease course and to identify outcome predictors.

4. WITHDRAWAL FROM THE STUDY

Participation in this study is entirely voluntary. The parents may decide, at any time, that they do not wish their child to take part in the study without altering current or future treatment in any way.
If at any stage of the project the parents wish to withdraw their child from the study, the researcher responsible for the study will arrange for their child's information to be immediately removed from computer records and for any remaining samples we hold belonging to him/her to be destroyed.

5. PATIENT'S STUDY

5.1 STUDY TO BE PERFORMED

Clinical assessment (physical and rheumatologic examinations using standardised and validated measures of disease activity and damage) will be performed every six months for each patient. The examining rheumatologists will be blinded to results of imaging assessment. The following data will be collected through standardised paper case report forms

Patient data: abbreviation for centre of origin (IGG, OPBG, UMCU) as given in the supplemental files, initials, date of birth. For instance patient Mario Rossi, date of birth 15 October 1990 from Istituto G.Gaslini will be reported as: IGG (for Istituto G.Gaslini) MR 15 10 1990. The full code will thus be IGG -MR 15 10 1990.

- Patient characteristics recorded at baseline will include: age at onset, sex, disease duration, JIA subtype (according to ILAR criteria), date of inclusion into the study, pharmacologic treatment (dose and duration) including joint injections.
- Physician's global assessment of overall disease activity, measured on a 10 cm visual analog scale (VAS) (0= no activity; 10 = maximum activity).
- Parent's global assessment of the child overall well-being, measured on a 10 cm VAS (0= very good; 10 = very poor).
- Functional ability assessed by C-HAQ (Childhood Health Assessment Questionnaire; grade 0-3 for 8 criteria;).
- Number of joints with active arthritis. Joint with active arthritis is a joint with swelling not due to bony enlargement or, if no swelling is present, limitation of motion accompanied either by pain on motion and/or tenderness .
- Number of joints with limited range of motion.
- Juvenile Arthritis Disease Activity Score (JADAS)
- Juvenile Arthritis Damage Index (JADI) will be performed once a year for the assessment of the long-term damage in patients with JIA.
- Information concerning previous and ongoing treatment will be recorded.

Laboratory assessment included: Westergren erythrocyte sedimentation rate (ESR), C reactive protein (CRP), WBC differential count, Haemoglobin level, platelet count, anti-nuclear antibodies (ANA), rheumatoid factor (FR).

Immunological and genetic assessment

Biological samples (blood, and synovial fluid from patients with clinical indication to perform local steroid injection) will be collected at disease onset, when patient will achieve clinical remission state (according to Wallace criteria for remission in JIA) and during flare of the disease.

For biomarkers we will use a high throughput methodology, namely the multiplex immunoassay or Luminex®. This is a bead-based assay that allows the detection of more than 100 soluble mediators in a single sample of 50 microliter of body fluid, such as plasma or synovial fluid.

Panel for measurements in Multiplex Immuno Assay:

- IL-1RA IL-23 CCL7/MCP-3 RBP4 TREM-1
- IL-1a IL-25 CCL11/Eotaxin TPO KIM-1/TIM-1
- IL-1b IL-27 CC17/Tarc SAA-1 Cathepsin B
- IL-2 IL-33 CCL18/PARC G-CSF Cathepsin L
- CCL19/MIP-3beta
- IL-3 TNF-alpha M-CSF Cathepsin S
- IL-4 TNF-beta CCL22/MDC GM-CSF sPD-1
- IL-5 IFN-alpha CCL27/C-TACK SCF Granzyme-B
- IL-6 IFN-beta CXCL-5/ENA-78 HGF sIL-1RI
- IL-7 IFN-gamm CXCL8/IL-8 EGF sIL-1RII
- IL-9 MIF CXCL9/MIG FGF basic sTNF-RI
- IL-10 LIF CXCL10/IP-10 NGF sTNF-RII
- IL-11 OSM CXCL13/BLC BDNF sIL-2R
- IL-12 TSLP XCL-1 VEGF sIL-6R
- IL-13 OPG Adiponectin sICAM sSCF-R
- IL-15 OPN Adipsin sVCAM

- IL-16 CCL1/I-309 Leptin sCD14
- IL-17 CCL2/MCP-1 Chemerin sCD163
- IL-18 CCL3/MIP- Omentin MMP-8 1alpha
- IL-21 CCL4/MIP-1beta Resistin TIMP-1
- IL-22 CCL5/RANTES PAI-1 S100A12.

In a smaller subpopulation of patients we will perform T cell characterisation in paired peripheral blood and synovial fluid derived mononuclear cells focusing on regulatory T cells (natural and induced regulatory T cells expressing FOXP3, Tr1 cells), and effector T cells (Th17, Th1 cells). We will both look at phenotypic markers, mRNA, epigenetic markers (methylation FOXP3) and functionality (in vitro suppression assays).

Meta-genomic data analysis (Microbiome) will be performed on stool samples which will be collected at disease onset, when patient will achieve clinical remission state and during flare of the disease. Faecal samples, will be accomplished by sample datasheets which will include specific indications on antibiotic, prebiotics, and probiotics administration, appropriate for later description of gut microbiota enterotypes. Samples will be stored at 4°C for at maximum of 24 hours or, alternatively, at -80°C until shipment from different laboratories to metagenomic facilities for DNA extraction and sequencing processing. In the latter case, the samples will be sent every two months or when suitable for the laboratories included in the study in dry ice by express courier.

Imaging assessments:

Ultrasound imaging: high-resolution ultrasound (U/S) evaluation of joints will be performed using a machine, equipped with broadband linear-array transducers At the patient enrolment in addition to all clinically affected joints, wrists, 2nd and 3rd MCP and IP joints, elbows, knees and ankles will be also investigated with U/S for a more accurate assessment of disease extension. U/S follow-up data will then be acquired at 6 months follow-up intervals for the first two years from patients enrolment. The U/S scanning protocols will be based on the standardised technical guidelines issued by the European Society of Musculoskeletal Radiology and the OMERACT US group. The severity of joint involvement may be judged sonographically by a variety of gray-scale parameters, including the amount of joint effusion, the presence of synovial thickening, the degree and duration of synovial hyperemia, the occurrence of cartilage abnormalities and bone erosions.

In patients with ankle involvement the following investigations will be also performed:

- Ankle digital plain radiography will be performed at the enrolment and after 2 years as standard routine practice in order to assess the presence and the degree of local growth disturbances, abnormal joint alignment and disease progression.
- Magnetic Resonance (MR): ankle MRI will be performed on a 1.5 Tesla MR scanner at the time of patient enrolment and after 2 year from baseline evaluation. The MRI will be performed only in cooperating patients who do not require general anesthesia. MRI do not expose to ionizing radiation. The following image sequences will be used in the study protocol: Morphological study: TSE T1 3D; TSE T2 fat sat; GRE 3D fat sat. MRI detectable pathological findings will be extracted using both a semi-quantitative and a quantitative

approaches. In case of unilateral involvement, the controlateral ankle will be also scanned at the same time to be used as a control.

Dual X-ray Absorptiometry (DXA) : will be performed at 6 month follow-up visit.

Gait cycle analysis

Quantitative gait assessment will be carried out with CGAs installed at the labs and the reflective markers will be attached bilaterally on the participant's skin at the shoulders, trunk, pelvis, legs and feet. Children with JIA will be evaluated the second time after treatment. The same examiner will perform the clinical measurement and marker placement in children with JIA. To evaluate kinetic and kinematic variables in all three anatomical planes we plan to calculate from five gait cycles beginning with the left foot strike and five gait cycles beginning with the right foot strike.

6. STUDY PLANNING

6.1 EFFICACY PARAMETERS

Clinical, immunological genetic and imaging data will be gathered and stored in a standardized manner building upon the Health-e-Child software tools which will be extended for the purpose of integrating model information related to a wider range of joints, covering morphology, gait analysis, bio/genetic data. The tools to be developed will also include the aspect of a multidimensional longitudinal analysis that yields the opportunity to identify potential new outcome measures (imaging or biological biomarkers) for the assessment of treatment efficacy. Furthermore, the prognostic value on an individual level of multidimensional data, including modern imaging modalities, genetic and meta-genetic data will be explored through the development and integration of appropriate data clustering methods.

6.2 EXPERIMENTAL DESIGN

The study will last 4 years. It is designed as a prospective longitudinal study. The timeframe for patient recruitment spans the first 28 months. Follow up data for each data (clinical, imaging, immunological, etc.) will be collected at follow-up visit as indicated in details in patient study session.

6.4 DATA PROTECTION

This kind of project requires that a substantial amount of personal data, including genetic information, is collected from the participants and shared across a network.The project will be carried out in accordance with the applicable European and National data privacy protection laws and regulations. All data will be gathered in an anonymous form so that no data may be traceable to a patient other than by the local treating clinicians. Only the respective hospitals will have access to the key of re-identification. Therefore, no project partner or other third party outside the respective hospitals involved, will have access to the identifiable patient data. Furthermore, only anonimyzed data will be processed or used in the project.

This information, handled in an anonymous manner, will be granted to regulatory authorities for regular reviews of clinical study procedures and/or data, in order to protect child's privacy .

7. SECURITY EVALUATION

7.1 DEFINITIONS

No adverse effects are foreseen as consequence of the clinical study.

8. SAMPLE DIMENSION AND STATISTIC METHODOLOGY

8.1 STATISTIC DESIGN

As the main goal of whole project is to establish a data repository for pediatric diseases, the sample size (180 patients, 60 for each clinical center) has been set by taking into account primary endpoints and study power but also available resources at each center, and study feasibility. In particular for the genetic analysis no study power is foreseen.

Modeling of JIA will be done by the following partners: University of Sheffield, La Sapienza University of Rome, Fraunhofer Research Institute (Fraunhofer Gesellschaft zur Foerderung), MOTEK SrL.

Differences will be evaluated in children with JIA between pre- and post-treatment gait analyses using a Repeated Measures Analysis of Variance (ANOVA). Non-parametric statistical (Mann–Whitney) tests will be used to determine differences between children with JIA before treatment and controls, and between children with JIA after treatment and controls.

Data processing will include the following workflow:

- Data preprocessing: data validation, discretization, null & outlier removal [Data Curator & Validator (DCV) related]
- Normalization: descriptors will be cross-mapped to standard data dictionaries (e.g. epSOS value sets, ICD-10 diagnosis, LOINC labs...);
- Normalization: source numeric data will be normalized to generate z-values;
- Data enrichment: normal values will be added from legacy guidelines;
- Data inferences: aggregated scores will be computed out of various fields (e.g. Body mass index out of weight/size);
- Data analytics: standard statistical tests will be applied (e.g. standard deviations);
- Data analytics: distance measures will be applied to generate case-based retrieval application (e.g. given a set of values, a ranked list of similar profiles will be returned);
- Data analytics: clustering of instances using statistical & visualization algorithms such as C4.5, Rocchio.
- For imaging data: we will extract visual features from the images that describe image regions, so local image content; we will try to aggregate data across cases in terms of visual data but also textual or structured data to be able to differentiate normal form abnormal visual data.
- Data analytics: Statistical modeling & simulation based on probabilistic techiques (e.g. graphical probabilistic networks) [AITION related]
-

8.2 MANAGEMENT OF MISSING DATA

Patients whose information is missing or incorrectly transcribed will not be used for the analysis, but will be marked appropriately in the electronic database by inserting a variable "Flag".

9. AMINISTRATIVE AND ETHICAL PROCEDURES

Confidentiality

Clinical data will be acquired as required by each partner's national law.

At each clinical centre patient's data will be collected and stored as electronic files and will be accessible by the responsible research personnel. Access to data will be granted using their personal credentials. Access to the file will be protected and the log of the user who performed the operation will be required at regular intervals. The data manager will perform regularly a data backup.

Data publication and final report

The ownership of scientific data will be shared between all the partners involved in the Project. The WP leader, Prof. Alberto Martini, and researchers who will conduct the study, will endeavor to promote the dissemination of the results through the project website, communications in national and international scientific meetings, publication in international journals of high scientific profile. The dissemination and publication of the results by the experimenters will be promoted in accordance with the provisions in force concerning the confidentiality of sensitive data. In all scientific publications the efforts of all researchers will be recognized.

All health professionals involved in the project will seek to minimize the physical and psychological discomfort caused to patients and parents from participating in this study. In order to ensure the well-being, they will not be notified in any way about the personal results of genetic investigations.

9.1 AUTORISATIONS

The protocol will undergo the approval of the Ethical Committee for the study implementation before the enrolment of the patients.

9.2 INFORMED CONSENT

Each parent/patient who will be asked for his/her enrolment in the study will be informed on every aspect concerning the study itself and, before the recruitment begins, he/she must sign the informed consent model. The date of the signature of the consent will be recorded on the CRF. A Copy of the informed consent model must be handed to the parent/patient.

9.3 INSURANCE COVERAGE

Insurance coverage used is as foreseen by each research structure for clinical and research activities. ***9.4 USE OF THE INFORMATION AND DATA PUBBLICATION***

Researchers must have the right to disseminate/publish the results of the study, with respect to the current legislation on sensible data protection and Intellectual property Rights. European Commission does not have any right in the publication/dissemination of study results

9.5 CLINICAL PROTOCOL AMENDMENTS

Clinical protocol amendment is intended to be any modifications to the final version of the clinical protocol that, after being approved by the Researcher and by the Ethical Committee, cannot be informally modified.

9.*6 DATA MANAGEMENT AND DOCUMENT PRESERVATION*

Researchers will manage the preservation of patients identification codes for at least 15 years after the end of the study. Patients clinical data and other original data will be kept for the maximum period allowed by the Hospital (min. 15 years).

9.7 BUDGET

The project costs are covered by the specific Grant by the European Commission. Such grant will cover 75% of staff costs, of consumables and equipment used for the project, and of all other expenses involved. Annex 1 (GPF) contains the project's entire budget, including details of the Institution's budget.

10. RESEARCHER RESPONSIBILITY

Researcher, as principle investigator of the study declares to be aware of all the duties committed to him/her and his/her collaborators for the conduction of the study. Except where explicitly edited, the term "researcher" on this Protocol and on the CRF is referred to the researcher or to any other person appointed by him/her, who will have power to sign documents instead of him/her.
The researcher will conduct the study according to the principles of the Helsinki declaration (1964) and Amendments (Appendix 1), and to the European Guidelines on Clinical Good Practices .

11. Annexes

Annex 1 GPF
Annex 2 DOW

12. REFERENCES

1. D. Manset, F. Pourraz, A. Tsymbal, J. Revillard, K. Skaburskas, R. McClatchey, A. Anjum, A. Rios & M. Huber, Gridifying Biomedical Applications in the Health-e-Child Project, Chapter XXIV of the Handbook of Research on Com- putational Grid Technologies for Life Sciences, Biomedicine and Healthcare. ISBN 978-1-60566-374-6 IGI Global Publishers, May 2009.
2. N. Wilkins-Diehr, D. Gannon, G. Klimeck, S. Oster & S. Pamidighantam Univ. of California, San Diego, CA. "TeraGrid Science Gateways and Their Impact on Science". In Computer (Nov. 2008). Volume: 41 Issue:11. On page(s): 32 – 41. ISSN: 0018-9162
3. Future for European Grids: Grids and Service Oriented Knowledge Utilities - Vision and Research Directions 2010 and Beyond, European Communities, 2006. FP7- ICT-2011.5.2 600932 - MD-Paedigree –Part B 113
4. Andrews RE, Fenton MJ, Ridout DA, Burch M; British Congenital Cardiac Association. New-onset heart failure due to heart muscle disease in childhood: a prospective study in the United Kingdom and Ireland. Circulation 2008; 117(1):79-84
5. Malattia C, Damasio MB, Pistorio A, Ioseliani M, Vilca I, Valle M, Ruperto N, Viola S, Buoncompagni A, Magnano GM, Ravelli A, Tomà P, Martini A. Development and preliminary validation of a paediatric-targeted MRI scoring system for the assessment of disease activity and damage in juvenile idiopathic arthritis. Ann Rheum Dis. 2011; 70(3):440-6.

6. Malattia C, Damasio MB, Basso C, Verri A, Magnaguagno F, Viola S, Gattorno M, Ravelli A, Tomà P, Martini A. Dynamic contrast-enhanced magnetic resonance imaging in the assessment of disease activity in patients with juvenile idiopathic arthritis.Rheumatology (Oxford). 2010;49(1):178-85

7. Martelli, S., F. Taddei, A. Cappello, S. van Sint Jan, A. Leardini and M. Viceconti. “Effect of sub-optimal neuromotor control on the hip joint load during level walking." J Biomech 2011; 44: 1716-1721.

8. Taddei, F., S. Martelli, G. Valente, A. Leardini, M. G. Benedetti, M. Manfrini and M. Viceconti. "Femoral loads during gait in a patient with massive skeletal reconstruction." Clin Biomech 2012;27:273-280.Viceconti, M., F. Taddei, L. Cristofolini, S. Martelli, C. Falcinelli and E. Schileo (2012). "Are spontaneous fractures possible? An example of clinical application for personalised, multiscale neuro-musculo-skeletal modelling." J Biomech 45(3): 421-426.

9. Magni-Manzoni S, Malattia C, Lanni S, Ravelli A. Advances and challenges in imaging in juvenile idiopathic arthritis.Nat Rev Rheumatol. 2012;8(6):329-36. Buoncompagni A, Magnano GM, Ravelli A, Tomà P, Martini A. Development and preliminary validation of a paediatric-targeted MRI scoring system for the assessment of disease activity and damage in juvenile idiopathic arthritis. Ann Rheum Dis. 2011; 70(3):440-6.

10. Malattia C, Damasio MB, Magnaguagno F, Pistorio A, Valle M, Martinoli C, Viola S, Buoncompagni A, Loy A, Ravelli A, Tomà P, Martini A. Magnetic resonance imaging, ultrasonography, and conventional radiography in the assessment of bone erosions in juvenile idiopathic arthritis. Arthritis Rheum. 2008; 59(12):1764-72.

11. de Jager W, Hoppenreijs EP, Wulffraat NM, Wedderburn LR, Kuis W, Prakken BJ Blood and synovial fluid cytokine signatures in patients with juvenile idiopathic arthritis: a crosssectional study. Ann Rheum Dis. 2007;66 (5):589-98.

WP 6 - Data acquisition and processing for Neurological and Neuromuscular Diseases (1)

Protocol no:

MD-Paedigree

WP 6**Version 1:** Apr 18, 2013**CONFIDENTIAL**

Protocol no.:	
Title:	WP 6 - Data acquisition and processing for NND
Acronym:	MD-Paedigree – WP 6
Multicentric/Monocentric Study	Multicentric
Principal Investigator	Dr.Enrico Castelli
Sponsor	Bambino Gesù Children's Hospital (European Commission)
Person responsible for the study WP6	Prof. Jaap Harlaar (Vrije Universiteit Amstrerdam)

Scientific Coordinator of the Project	Prof. Bruno Dallapiccola
Data Management/Statistical analysis:	

Protocol approved and signed by:

Scientific Coordinator of the Project:

Prof. Bruno Dallapiccola

Responsible Work Package 6:
Prof. Jaap Harlaar
Principal Investigator:
Dr. Enrico Castelli

Acronym List

AEs	Adverse Events
EC	Ethical Committee
CRF	Case Report Form
GCP	Good Clinical Practice
ICH	International Conference on Harmonisation
IRB/IEC	Institutional Review Board/Independent Ethics Committee
NSAEs	Non Serious Adverse Events
SAEs	Serious Adverse Events
SOPs	Standard Operating Procedures

INDEX

1 INTRODUCTION

1.1 BACKGROUND OF THE MD-PAEDIGREE PROJECT

MD – PAEDIGREE *Model-Driven European Paediatric Repository* is a research project funded by the European Commission under the Virtual Physiological and Human Area (VPH) of the ICT Theme of the Seventh Framework Programme (contract no. 600932). The project foresees 7 world-renowned clinical centres of excellence pursuing improved interoperability of paediatric biomedical information, data and knowledge, by developing a set of reusable and adaptable multi-scale models for more predictive, individualised, effective and safer paediatric healthcare. The project is scientifically and technologically supported by one of the leading industrial actors in medical applications in Europe, and operating in conjunction with highly qualified SMEs and some of the most experienced research partners in the Virtual Phisiological Human (VPH) community.

MD-Paedigree validates and brings to maturity patient-specific computer-based predictive models of various paediatric diseases (cardiomyopathy, risk of cardiovascular disease in obese children and adolescents, Juvenile Idiopathic Arthritis (JIA) and neurological and neuro-muscular diseases (NND), thus increasing their potential acceptance in the clinical and biomedical research environment by making them readily available not only in the form of sustainable models and simulations, but also as newly-defined workflows for personalised predictive medicine at the point of care.

MD-Paedigree aims to advance the state-of-the-art of patient-specific computational modelling of different paediatric diseases and translate the latest advances into clinics to improve disease understanding, therapy outcome, and provide an infostructure platform for assessing new therapies at the point of care.

MD-Paedigree's goals therefore are to integrate and share highly heterogeneous biomedical information, data and knowledge, using best, jointly develop reusable, adaptable and composable multi-scale VPH workflow models and to support evidence-based translational medicine at the point of care practices from the biomedical semantic Web.

Ilness in infants, children, and adolescents are a large and under-appreciated public health problem. Because paediatric patients now have a much longer life expectancy, the burden and costs are substantial for families and society. For instance, recent studies have shown that the number of adult patients with congenital heart disease is already similar to that of the paediatric population and will continue to grow. It can therefore be stated that in the longer term the VPH scientific approach requires a fundamental research investment in paediatrics, where the conceptual revolution which is underway, transforming the nature of healthcare from reactive to preventive, can best be applied, moving towards an approach based on personalised, predictive, preventive, and participatory (P4) medicine, increasingly focused on wellness.

MD-Paedigree represents a major step towards personalised paediatric e-Health, based on data-driven models, patient-specific simulations and a sustainable data and model repository. The project can in fact impact the way healthcare is practiced in the future. MD-Paedigree's impact on dealing with very concrete and exemplary clinical cases is demonstrated through the following applications in the diseases scenarios. MD-Paedigree encompasses complete services for storage, similarity search, outcome analysis, risk stratification, and personalised decision support in paediatrics within its innovative model-driven data and workflow-based models repository,

leveraging on service and knowledge utilities. It fosters the state-of-the-art of patient-specific computational modelling of the selected diseases and translates the latest advances into clinics to improve disease understanding and provide a platform for testing new therapies at the point of care. In fact, MD-Paedigree demonstrates how a dedicated VPH repository can provide full accessibility to existing and further developing knowledge in paediatrics to all interested bio-medical researchers precisely through linking data with models, thus proving the large scale benefits of having both the data and models readily available at the point of care.

The data collection performed within Health-e-Child and Sim-e-Child projects (already funded by the European Commission in previous calls for proposals) is still available to the MD-Paedigree consortium thanks to the continuity in the eHealth platform which is shared by all three projects. In addition, the new patients' recruitment to be performed within MD-Paedigree consists of:

<table>
<tr><th>Pathology</th><th colspan="2">No of patients/Time</th></tr>
<tr><td>Cardiomyopathies</td><td>180 children, by month 33: 60 patients (among which 30 girls) for each clinical centre.</td><td rowspan="3">Genetic and meta-genomic:

180 patients with cardiomiopathies, 180 with CVD risk in obesity, 200 with JIA, and 100 unaffected subjects (control group).</td></tr>
<tr><td>CVD risk in obese children</td><td>180 patients , by month 36: 60 (among which 30 girls) for each clinical centre.</td></tr>
<tr><td>Juvenile Idiopathic Arthitis (JIA)</td><td>Altogether 200 patients by month 28.</td></tr>
<tr><td rowspan="3">NND</td><td colspan="2">Cerebral Palsy: 50 patients for each clinical centre for probabilistic modelling, as well as 600 retrospective patients from KU Leuven and OPBG.</td></tr>
<tr><td colspan="2">Spinal Muscular Atrophy (SMA)
Data will be collected by OPBG, KU Leuven and VUA from 20 ambulant patients (severity grade type 3).</td></tr>
<tr><td colspan="2">Duchenne Muscular dystrophy (DMD)
Clinical data will be collected by OPBG, KU Leuven and VUA from 20 ambulant genetically confirmed DMD Patients.</td></tr>
</table>

1.2 BACKGROUND OF THE DATA ACQUISITION AND PROCESSING FOR NND STUDY

In Neurological and Neuromuscular Diseases (NND) as well as in certain chronic paediatric diseases of the musculoskeletal system, treatments are strongly guided by maximising the walking function of the human movement system, because walking is considered as clinically meaningful by patients. This generalises to

most mobility-related functions. The most common paediatric disorder within the NND disease area is Cerebral Palsy (CP) whose incidence ranges between 2 to 3.6 per 1,000 live births [Odding E. et al., 2006]28. CP includes a group of non-progressive, often changing, motor impairment disorders, secondary to lesions in the sensory-motor cortex and corticospinal tract, arising in the early stages of the child's development. Conventional clinical gait analysis (CGA) is already an important tool in the treatment of children with CP that aims to improve or sustain walking performance, but its potential is under-utilised and recent developments need full exploration.

The second important disorder is Spinal Muscular Atrophy (SMA), an autosomal recessive disease characterised by degeneration of motoneurons in the spinal cord. SMA is caused by mutations of the survival motor neuron 1 gene (SMN1). Estimated incidence is 1 in 6,000-8,000 live births. This disease is characterised by progressive generalised muscle weakness and atrophy predominating in proximal limb muscles. For ambulant SMA patients, new methods for functional motor evaluation based on gait modelling would allow to increase sensitivity to change in assessing weakness and fatigability.

The third disorder, Duchenne Muscular Dystrophy (DMD) is the most common and severe form of muscular dystrophy, with an FP7- ICT-2011.5.2 600932 - MD-Paedigree –Part B 9 incidence around 1 in 3,600 juveniles. This disorder is caused by a mutation in the dystrophin gene, that codes for a protein which is a major structural component of the muscle. The absence of dystrophin results in muscle degeneration, difficulty in walking (resulting in wheelchair use from 14 years of age), followed by loss of arms and hands function. In the last few years, following a rapidly increasing number of potentially effective therapeutic approaches for DMD, the request for validated and sensitive outcome measures to be used in clinical trials has increased. Although walking is a common task executed by a healthy individual in a seemingly effortless manner, it implies a complex involvement of inputs from several senses (visual, vestibular, proprioceptive, somatosensory), partly automated by the so called spinal central pattern generator (CPG) [Spardy LE et al, 2011]29. These inputs are known to interact with each other, but the way in which this is performed is not fully exploited at present [Baker R , 2006]30. Nevertheless, the current insights are certainly at an advanced state that allows for meaningful application towards pathological walking, where decision support is needed [FM Chang et al., 2010] 31. In the clinical practice of specialised centres, CGA is used to evaluate the joint and muscle functions in their functional context, i.e. during gait [Cappozzo A. et al., 2005]32. Common CGA measures 3D kinematics (by 3D optoelectronic registration of skin mounted markers). Each relevant degree of freedom (DOF) is expressed as a function of the gait cycle. Moreover, using a mass distribution model and measuring ground reaction forces, the net moments for each DOF are calculated using inverse dynamics analysis.

Muscle activation patterns, for all relevant muscles, are measured using electromyography (EMG) for each targeted muscle. Finally, the energy cost of walking can be evaluated using metabolic measurements. CGA is a special form of personalised computer-aided medicine that supports clinical decision making [Novacheck TF et al., 2010]33. Unfortunately, the output of CGA is not yet in a format that permits clear, unambiguous interpretation, because of the redundancy of the Neuro-Musculo-Skeletal System (NMSS) which obstructs distinguishing cause from compensation. Even though recent developments in modelling the NMS Physiome as a part of EU funded Virtual Physiological Human efforts are at an advanced state, their results have not yet been implemented in clinical practice, and the full potential of CGA still needs to be reaped. A combination of standard protocols of gait analysis, biophysical modelling and large scale statistical analysis can therefore be expected to provide a powerful framework for meaningful interpretation.

Protocols and personalised models in Advanced Clinical Gait Analysis

To reiterate the conclusion of the NDD clinical background section: the potential of gait analysis to serve clinical decision making in NDD is generally under-used for several reasons. These will be taken up within the MD-Paedigree project.

Protocol definitions for clinical gait analysis

Three levels of protocol definitions are needed to assure multicentre reliable data for the repository:

Technical Quality assurance for CGA laboratories

It is important to realise that for accurate data from the experimental systems a strict analysis of causes of errors and periodical validation procedures needs to be implemented in the gait labs [Cedraro, A.et al., 2009116, Chiari L. et al.,2005 117]. If the adopted experimental procedure permits the gathering of valid data, the first important prerequisite for reliable and accurate results from a particular subject is fulfilled. Within MD-Paedigree these quality assurance (QA) procedures will therefore be formalised between laboratories for clinical gait analysis. MD-Paedigree will constitute a European standard for technical QA and have this approved by the important European bodies on clinical gait analysis, i.e. the ESMAC. A consensus meeting will be part of this.

Standardisations of gait analysis protocols: Marker placements

One of the main non-technical sources of error in CGA using OptoElectronic Movement Analysis systems is caused by marker artifacts, resulting from skin movement relative to the bone [Leardini et al., 2005] 118. Recently it has been shown that, in the case of well-trained staff, errors due to marker misplacements and skin movement artifacts will stay within a few degrees of error of the joint kinematics graphs [J.L. McGinley et al. 2009] 119. This error level is considered to be just clinically acceptable. This means that all gait labs should fulfill the requirements to be qualified for MD-Paedigree graded gait analysis. In analogy with the Technical Quality Assurance (TQA), MD-Paedigree will strongly promote interoperability and constitute a protocol for standardised marker placement, as well as standard procedures to evaluate this within and between laboratories. In parallel, we shall explore the possibility to use imaging/gait analysis protocols, where patients are dressed with radiopaque/MRI opaque and reflective markers attached to the skin as used in gait analysis protocols, while the imaging protocol is conducted. These data will make possible to use sophisticated inverse kinematics modelling methods to minimise the skin artifacts, and to obtain accurate estimations of the skeletal kinematics.

Standardisations of gait analysis protocols: operational protocols

The results of kinematics and kinetics of CGA are also dependent on the use of standard protocols for instruction on walking targets. In particular, the enforcement of a precise walking speed is of major influence on the output [Schwartz MH et al. 2008] 120. As such, instructions should be carefully standardised and protocols developed that use multiple walking speeds. It has been suggested and shown by previous studies [Bovi G. et al., 2001]121, that these protocols are necessary to detect important pathological features of the NMSS of the subject, especially in patients with CP [van der Krogt MM. et al., 2009]122. EMG recordings and oxygen consumption will be part of the overall assessment procedures. Moreover, in order to feed the development of probabilistic models a standardised description of therapies will be completed. This description will be used to longitudinally describe the applied clinical workflows that are currently used to improve gait performance in children with NND.

Conclusion: the established and clinically authorised protocols (technical, marker and procedures) of CGA will be an important step forward for the NND paediatric care in the EU, along with the establishment of a reliable MD-Paedigree database for typically developing children.

Application of computational biophysical models of the NMSS in CGA

For clinical gait analysis the use of Neuro-Musculo-Skeletal (NMS) models is an important step forward in the interpretation of its results, aiming to inform the clinical decision-making. Because of the modelling based interpretation, the physician no longer needs to interpret the results of clinical gait analysis, within his own informal frame of interpretation. Using NMS models the results of CGA are quantitatively "translated" into the function and performance of the underlying structures, i.e. muscle activation, muscle forces, and joint loads that make possible to unravel the aetiology of the pathological gait pattern of the subject under study. The EU project “Personalised models of the Neuro-Musculo-Skeletal Physiome” (NMS Physiome 123) is moving towards the development of PPI (Predictive, Personalised and Integrative) musculoskeletal medicine. NMS Physiome is a part of the European Union’s Virtual Physiological Human initiative.

A key result of this project, conducted by Prof. Viceconti, at MD-Paedigree partner USFD, is the integration of an advanced software application for the pre-processing of imaging and gait analysis data into a full musculoskeletal model (NMS Builder) and the OpenSIM musculoskeletal modelling environment developed by Stanford University. NMS Builder is already available in prototypical form to all partners of the MD-Paedigree consortium. Although NMS computational models are thus well known in the biomechanical research community, as yet only one company, MOTEK, has incorporated gait analysis and model based interpretation of gait for market delivery. Their model (the HBM model) is computationally very efficient: even without high performance computers it can run in real time. More complex modelling activities can be conducted using the NMS Physiome tools. The actual problem of accuracy of NMS models is that all models currently used in paediatric gait analysis are based on data scaled from a single cadaver in a simple way. Sensitivity studies have shown that such a gross simplification in applying generic models is too inaccurate, and, especially in the case of children, dedicated and validated models, fused with medical imaging data, should be developed in order to yield reasonable accuracy for clinical application in this population. The first level of MS models in CGA is the mass distribution model of body segments. Mass distribution means that the masses, centre of mass and inertial properties of each segment need to be known for accurate calculation of inverse dynamics resulting in valid joint kinetics. What is needed is a method for scaling that allows application, in clinical workflows, to enable personalised medicine. MD-Paedigree will develop and evaluate a scaling method for the NMSS of children, to be applied in existing NMS models that are used in CGA. Validation will be based on MRI measures. Next to anthropometrics scaling is the alternative to use a 2D image, generated by a whole body DXA image, morphed to a generic 3D skin model of a child. The advantage is that DXA provides accurate measurement of the areal density of the bone, fat, and lean tissues the inertial properties of each segment. The second level of personalised MS models in CGA are to account for the subject specific bony deformities. The bony deformities that should be accounted for can be limited to the clinically well known deformities in CP. These deformities have significant influence on the output of NMS model calculations (i.e. femoral anteversion and tibial torsion). These effects could primary be modelled by morphing the generalised bony structures towards the actual morphology of the bone. The most important effects of bony deformities should be parameterised by the effects on axis alignment: (a) introducing a skewness of the principal axes of rotation of the joints in the kinematic chain of linked segments, and (b) the altered lever arms of muscles with respect to these principal axes of rotation of the joint. Again anthropometric measures and DXA will be explored. The third level of personalised modelling is to account for pathology specific muscle parameters. These models should focus on the parameters that are known to be of large influence on the second step in inverse dynamics, i.e. the estimation of muscle forces based on optimisation criteria on how to explain the net joints moments from CGA. This means that especially muscle contractures, altered muscle structure and hypertonia (in CP), as well as muscle weakening (in DMD and SMA), must be targeted. US measures of the muscle belly, along with fibre directions will enable estimates of the muscle Physiological Cross sectional Area (PSCA), while dynamometric evaluations will yield measures of muscle belly length and optimal fibre length.

Supporting probabilistic models, despite the strong potential of biophysical models of the NMSS, will only hold a certain amount of predictive value, i.e. as far as their assumed accuracy will allow. However, in clinical practice, even if the pathology cannot be fully explained by biophysical modelling, the use of probabilistic models is still extremely powerful in supporting clinical decision making. Until now only two gait laboratories in the world (Gillette Children's, Minneapolis, US and Pellenberg, Leuven, Belgium) have explored the possibilities of generating decision rules from their dataset [van Gestel et al., 2011 124]. These laboratories are the only ones that have created a large enough set of reliable data to make such an effort worthwhile. In MD-Paedigree the clinical partners will collect data, according to the dataset and quality protocols defined on the basis of standardised formats, for feeding into the repository.

2 AIMS OF THE STUDY

2.1 MAIN GOAL

Main goal of the study is to acquire sets of data (gait analysis and images) related to Neurological and Neuromuscular Diseases for the repository, and to develop probabilistic modelling and biophysical modelling.

The most common paediatric disorder within the NND disease area is Cerebral Palsy (CP) whose incidence ranges between 2 to 3.6 per 1,000 live births. CP includes a group of non-progressive, often changing, motor impairment syndromes, secondary to lesions in the sensory-motor cortex and corticospinal tract, arising in the early stages of the child's development. Conventional clinical gait analysis (CGA) is already an important tool in the treatment of children with CP that aims to improve or sustain walking performance, but its potential is under-utilised and recent developments need full exploration. The second important disorder is Spinal Muscular Atrophy (SMA), an autosomal recessive disease characterised by degeneration of motoneurons in the spinal cord. SMA is caused by mutations of the survival motor neuron 1 gene (SMN1). Estimated incidence is 1 in 8,000 live births. The third disorder, Duchenne Muscular Dystrophy (DMD) is the most common and severe form of muscular dystrophy, with an incidence around 1 in 3,600 juveniles. This disorder is caused by a mutation in the dystrophin gene, that codes for a protein which is a major structural component of the muscle. The absence of dystrophin results in muscle degeneration, difficulty in walking (resulting in wheelchair use from 14 years of age), followed by loss of arms and hands function. In the last few years, following a rapidly increasing number of potentially effective therapeutic approaches for DMD, the request for validated and sensitive outcome measures to be used in clinical trials has increased.
A combination of standard protocols of gait analysis, biophysical modelling and large scale statistical analysis can therefore be expected to provide a powerful framework for meaningful interpretation.

2.2 PRIMARY END-POINT

Successful collection of 130 CP patients clinical gait dataset: a clinical gait dataset according to defined standards of 130 CP patients reprocessed form existing databases (100) and new measurements within the first 26 months of activity. Moreover a clinical gait dataset will be established according to defined standards of 130 CP patients reprocessed form existing: a comprehensive clinical dataset of gait analysis data for CP, data sets of 30 CP patients.

3 STUDY DESCRIPTION

Spinal Muscular Atrophy (SMA) Data will be collected by OPBG, KU Leuven and VUA from 20 ambulant patients (severity grade type 3); 10 patients for each centre for biophysical modeling; 10 patients among the 3a subgroup (symptoms of weakness appearing before age 3 years); 10 patients among the 3b subgroup (weakness appearing after the age of 3 years.
Duchenne Muscular dystrophy (DMD) Clinical data will be collected by OPBG, KU Leuven and VUA from 20 ambulant genetically confirmed DMD Patients. 10 patients with an age ranging between 5 and 6 years, additional 10 patients with an age ranging between 7 and 8 years.

3.1 STUDY DESIGN

Gait analysis collection for CP

Gait analysis data will be provided to the work packages that are involved in biophysical and probabilistic modelling.
A complete dataset related to clinical gait analysis consists of:
1. A standardised anamnesis
2. Standard clinical testing: Physical Examinations and Tests; Questionnaires
3. Xray s if applicable
4. From gait analysis: Kinematic data; Kinetic data; EMG Data; O2 Data.

Contextual data, like treatments received

Criteria for selection are based on children with CP that are routinely measured in the gait lab: classified as GMFCS 1-3; diplegic or hemiplegic; sufficient cognitive skills; without relevant visual deficit; and older than 6 years.
1. Complete data sets of 10 CP patients for each clinical center will be provided for biophysical modelling.
2. For the probabilistic modelling, as many as the clinical load would allow, can be included, the aim is 50 patients per center before month 36.

Image acquisition

In WP 11 some advanced modeling is developed, that the fusion of multimodal sources of data (MRI, DXA and CGA). As an input to this WP, each clinical center (VUA, OPBG, KU Leuven) will acquire at least 10 subjects with both MRI and DXA, including the markers that are needed for gait analysis. Volume of interest includes pelvis, femur, tibia, foot. The first three subjects should be acquired within the first year of the project. Images will have to be anonymised before making them available for the technical partners.

3.2 SUBJECTS SELECTION

Data collection will be performed in three leading European Centers (Ospedale Pediatrico Bambino Gesù (OPBG, Rome, Italy), Katholieke Univesiteit Leuven (KULeuven, Belgium), Vrije Universiteit Amsterdam (VUMC, The Netherlands).

The subjects will be enrolled between the in- and out-patients coming for assessment and rehabilitation in the involved centers. A MD of the centre will explain to the parents and the child the aims and the features of the research, in order to have their consent.

The following cohorts must be considered:

Control group: healthy patients:

- Patients, with normal neuromuscular or scheletric development and requiring MRI for problems not related to the locomotor system (backbone, hip, lower limbs);
- 10 OPBG patients.

CP patients - Prospective group:

- Children with a diagnosis of Cerebral Palsy following the definition of Rosenbaum: "Cerebral Palsy (CP) describes a group of permanent disorders of the development of movement and posture, causing activity limitation, that are attributed to non progressive disturbances that occurred in the developing fetal or infant brain. The motor disorders of cerebral palsy are often accompanied by disturbances of sensation, perception, cognition, communication, and behaviour, by epilepsy, and by secondary musculoskeletal problems". (Rosenbaum P et Al. A report: the definition and classification of cerebral palsy. Dev Med Child Neurol Suppl. 2007 Feb;109:8-14.)
- 150 CP children (50 from each involved centre) will be selected for this study in the first 36 months, 30 with a neurological diagnosis of diplegia, 10 with right hemiplegia and 10 with left hemiplegia

CP patients - Retrospective group:

- Children with a diagnosis of Cerebral Palsy following the definition of Rosenbaum: "Cerebral Palsy (CP) describes a group of permanent disorders of the development of movement and posture, causing activity limitation, that are attributed to non progressive disturbances that occurred in the developing fetal or infant brain. The motor disorders of cerebral palsy are often accompanied by disturbances of sensation, perception, cognition, communication, and behaviour, by epilepsy, and by secondary musculoskeletal problems". (Rosenbaum P et Al. A report: the definition and classification of cerebral palsy. Dev Med Child Neurol Suppl. 2007 Feb;109:8-14.), who have already performed a Gait analysis in the participants' (OPBG or Leuven) Gait analysis Laboratories.
- 600 patients in total (50% emiplegici e 50% diplegici) 400 form Leuven and 200 form OPBG.

Inclusion criteria are:

All groups:

- Age 6 – 11
- Parent/guardiuna consent

Control group: healthy patients:

- OPBG patients, with normal neuromuscular or scheletric development;
- Requiring MRI for problems not related to the locomotor system (backbone, hip, lower limbs)

CP patients - Prospective group:

- GMFCS 1-2 (Palisano R et Al. Development and validation of a gross motor function classification system for children with cerebral palsy. Dev Med Child Neurol 1997; 39: 214–23);

- sufficient cognitive skills: Total IQ by Wechsler Intelligence Scale: > 60.

CP patients - Retrospective group:

- patients who have already performed a Gait analysis in the participants' (OPBG or Leuven) Gait analysis Laboratories;
- GMFCS 1-2 (Palisano R et Al. Development and validation of a gross motor function classification system for children with cerebral palsy. Dev Med Child Neurol 1997; 39: 214–23);
- sufficient cognitive skills: Total IQ by Wechsler Intelligence Scale: > 60.

Exclusion criteria are:

All groups:

- Patient history of functional surgery on bones and muscles;
- Lower limbs BoNT A injection in the last 6 months.
- Patient history of hip, backbone and/or lower limb fracture;
- relevant visual deficit non correctable by with lenses;
- significant comorbidities.

Control group: healthy patients:

- patient history of current or past problems connected to the locomotor system.

4. WITHDRAWAL FROM THE STUDY

Participation in this study is entirely voluntary. The parents may decide, at any time, that they do not wish their child to take part in the study without altering current or future treatment in any way. If at any stage of the project the parents wish to withdraw their child from the study, the researcher responsible for the study will arrange for their child's information to be immediately removed from computer records and for any remaining samples we hold belonging to him/her to be destroyed.

5 PATIENT'S STUDY

5.1 STUDY TO BE PERFORMED

All groups:

- a standardized anamnesis (Gestational age, birth weight, Apgar Score at V minute, kind and localization of brain abnormalities at MRI, any previous injection of botulinum toxin, previous orthopedic functional surgery);
- a standard clinical examination (PROM at hip, knee and ankle, MRC strength muscle at major muscle groups of lower limb, MAS at flexor/extensor of hip, knee and ankle) and a neurological assessment;

Control group: healthy patients:

- Dual Energy X-Ray Absorptiometry – DXA
- Clinical gait analysis (CGA) with collection of kinematic, kinetic, EMG and O^2 data.
- Lower limb MRI
- Measurement of metabolic consumption during six minutes walking test (6MWT)

CP patients - Prospective group:

- a Pelvic X-Ray, if not performed in the last year;
- Clinical gait analysis (CGA) with collection of kinematic, kinetic, EMG
- 12 (4 per Center) hemiplegic and 18 (6 per Center) diplegic:
    - Lower limb MRI (pelvis, femur,tibia and foot)
    - Measurement of metabolic consumption during six minutes walking test (6MWT)
    - Dual Energy X-Ray Absorptiometry – DXA

CP patients - Retrospective group:

- Previous Gait analysis data retrieval

No biologic samples will be collected.

Medical information (clinical and instrumental evaluation) collected are part of our common clinical practice but we will inform the child and the parent that the data collected from these surveys will be stored anonymously in this study and in future investigations. An additional informed consent will be required if any of these tests will be required with the sole purpose of research.

6. STUDY PLANNING

6.1 EFFICACY PARAMETERS

Data source: all data will be collected by a MD though observations, visit of the subjects and instrumental examinations. They will be recorded in CR and in ad hoc modules.
Data gathering scheduling: children affected by CP will be assessed only.
All scales and gait analysis protocols are internationally standardised and with a high reliability.

6.2 EXPERIMENTAL DESIGN

The study will last 4 years. It is designed as a prospective and retrospective observational study. The timeframe for patient recruitment spans the first 3 years. Follow-up data for each data will be collected at follow-up visit as indicated in details in patient study session.

6.3 DATA PROTECTION

All collected data will be anonymised. Clinical data and biological samples will be coded and stored as such. The code will be generated by software using a system of 128-bit encryption. The code will be stored in a close drawer Prof. Jaap Harlar. At the end of the study, the key code will be destroyed and, hence, data anonymised. From this moment on, it will not be possible for anyone to discover the patient's identity.
All clinical data will be communicated to participants and/or legal representative except for genetic testing.

7 SECURITY EVALUATION

7.1 DEFINITIONS

No adverse effects are foreseen as consequence of the clinical study.

8. SAMPLE DIMENSION AND STATISTIC METHODOLOGY

8.1 STATISTIC DESIGN

As the main goal of whole project is to establish a data repository for pediatric diseases, the sample size has been set by taking into account primary endpoints and study power but also available resources at each center, and study feasibility.

Data processing will include the following workflow:

- Data preprocessing: data validation, discretization, null & outlier removal [Data Curator & Validator (DCV) related]
- Normalization: descriptors will be cross-mapped to standard data dictionaries (e.g. epSOS value sets, ICD-10 diagnosis, LOINC labs...);
- Normalization: source numeric data will be normalized to generate z-values;
- Data enrichment: normal values will be added from legacy guidelines;
- Data inferences: aggregated scores will be computed out of various fields (e.g. Body mass index out of weight/size);
- Data analytics: standard statistical tests will be applied (e.g. standard deviations);
- Data analytics: distance measures will be applied to generate case-based retrieval application (e.g. given a set of values, a ranked list of similar profiles will be returned);
- Data analytics: clustering of instances using statistical & visualization algorithms such as C4.5, Rocchio.
- For imaging data: we will extract visual features from the images that describe image regions, so local image content; we will try to aggregate data across cases in terms of visual data but also textual or structured data to be able to differentiate normal form abnormal visual data.
- Data analytics: Statistical modeling & simulation based on probabilistic techiques (e.g. graphical probabilistic networks) [AITION related]

NND Modeling will be performed by the following partners: Siemens AG, MOTEK SrL; University of Delft, La Sapienza University of Rome, University of Sheffield.

8.2 MANAGEMENT OF MISSING DATA

Patients whose information is missing or incorrectly transcribed will not be used for the analysis, but will be marked appropriately in the electronic database by inserting a variable "Flag".

9. ADMINISTRATIVE AND ETHICAL PROCEDURES

All clinical procedures and instrumental examination are usually performed for the management and the rehabilitation of children affected by CP.

The MRI study of lower limbs (pelvis, femur, tibia and foot) will be performed only in 10 subjects able to collaborate, avoiding the movement artifacts.
Medical information (clinical and instrumental evaluation) collected are part of our common clinical practice but we will inform the child and the parent that the data collected from these surveys will be stored anonymously in this study and in future investigations. An additional informed consent will be required if any of these tests will be required with the sole purpose of research. A MD of the centre will explain to the parents and the child the aims and the features of the research, in order to have their consent.

9.1 AUTHORISATIONS

The protocol will undergo the approval of the Ethical Committee for the study implementation before the enrollment of the patients.

9.2 INFORMED CONSENT

Each parent/patient who will be asked for his/her enrollment in the study will be informed on every aspect concerning the study itself and, before the recruitment begins, he/she must sign the informed consent model. The date of the signature of the consent will be recorded. A Copy of the informed consent model must be handed to the parent/patient.

9.3 INSURANCE COVERAGE

Insurance coverage used is as foreseen by each research structure for clinical and research activities. ***9.4 USE OF THE INFORMATION AND DATA PUBBLICATION***

Researchers must have the right to disseminate/publish the results of the study, with respect to the current legislation on sensible data protection and Intellectual property Rights. European Commission does not have any right in the publication/dissemination of study results.

9.5 CLINICAL PROTOCOL AMENDMENTS

Clinical protocol amendment is intended to be any modifications to the final version of the clinical protocol that, after being approved by the Researcher and by the Ethical Committee, cannot be informally modified.

9.6 DATA MANAGEMENT AND DOCUMENT PRESERVATION

Researchers will manage the preservation of patients identification codes for at least 15 years after the end of the study. Patients clinical data and other original data will be kept for the maximum period allowed by the Hospital (min. 15 years).

9.7 BUDGET

The project costs are covered by the specific Grant by the European Commission. Such grant will cover 75% of staff costs, of consumables and equipment used for the project, and of all other expenses involved. Annex 1 (GPF) contains the project's entire budget, including details of the Institution's budget.

10. RESEARCHER RESPONSIBILITY

Researcher, as principle investigator of the study declares to be aware of all the duties committed to him/her and his/her collaborators for the conduction of the study. Except where explicitly edited, the term "researcher" on this Protocol and on the CRF is referred to the researcher or to any other person appointed by him/her, who will have power to sign documents instead of him/her.
The researcher will conduct the study according to the principles of the Helsinki declaration (1964) and Amendments (Appendix 1), and to the European Guidelines on Clinical Good Practices .

11. Annexes

Annex 1 GPF
Annex 2 DOW

12. REFERENCES

1. Odding E, Roebroeck ME, Stam HJ. The epidemiology of cerebral palsy: incidence, impairments and risk factors. Disabil Rehabil. 2006 Feb 28;28(4):183-91.
2. Spardy LE, Markin SN, Shevtsova NA, Prilutsky BI, Rybak IA, Rubin JE; A dynamical systemsanalysis of afferent control in a neuromechanical model of locomotion: I. Rhythm generation. J Neural Eng. 2011 Dec;8(6)
3. Baker R, Gait analysis methods in rehabilitation, J of NeuroEngineering and Rehab 2006 3(4), 10 31. FM Chang, JT Rhodes, KM Flynn, JJ Carollo, The role of gait analysis in treating gait abnormalities in cerebral palsy, Orthop Clin North Am. 2010 Oct; 41 (4):489-506;
4. Cappozzo A, Della Croce U, Leardini A, Chiari L. Human movement analysis using stereophotogrammetry. Gait Posture. 2005 Feb;21(2):186-96. Review.
5. Novacheck TF, Trost JP, Sohrweide S. Examination of the child with cerebral palsy. Orthop Clin North Am. 2010 Oct;41(4):469-88. Review.
6. Cedraro, A. Cappello, L. Chiari. A portable system for in-situ re-calibration of force platforms: Experimental validation . Gait Posture 29 (2009) 449–453.
7. L. Chiari, U. Della Croce, A. Leardini, A. Cappozzo "Human movement analysis using stereophotogrammetry Part 2: Instrumental errors" Gait and Posture 21 (2005) 197–211.
8. Leardini A, Chiari L, Della Croce U, Cappozzo A. Human movement analysis using stereophotogrammetry. Part 3. Soft tissue artifact assessment and compensation. Gait Posture. 2005 Feb;21(2):212-25. Review
9. J.L. McGinley et al. The reliability of three-dimensional kinematic gait measurements: A systematic review Gait Posture 2009, 29:3, 360-9
10. Schwartz MH, Rozumalski A, Trost JP. The effect of walking speed on the gait of typically developing children. J Biomech. 2008;41(8):1639-50
11. Bovi G, Rabuffetti M, Mazzoleni P, Ferrarin M. A multiple-task gait analysis approach: kinematic, kinetic and EMG reference data for healthy young and adult subjects. Gait Posture. 2011 Jan;33(1):6-13
12. van der Krogt MM, Doorenbosch CA, Becher JG, Harlaar J. Walking speed modifies spasticity effects in gastrocnemius and soleus in cerebral palsy gait. Clin Biomech (Bristol, Avon). 2009 Jun;24(5):422-8.
13. http://www.nmsphysiome.eu
14. Probabilistic gait classification in children with cerebral palsy: a Bayesian approach. Van Gestel L, De Laet T, Di Lello E, Bruyninckx H, Molenaers G, Van Campenhout A, Aertbeliën E, Schwartz M, Wambacq H, De Cock P, Desloovere K. Res Dev Disabil. 2011 Nov-Dec;32(6):2542-52.

WP 6 - Data acquisition and processing for Neurological and Neuromuscular Diseases (2)

Protocol no:

MD-Paedigree

WP 6**Version 2:** Apr 18, 2013**CONFIDENTIAL**

Protocol no.:	
Title:	WP 6 - Data acquisition and processing for NND
Acronym:	MD-Paedigree –
Multicentric/Monocentric Study	Multicentric
Principal Investigator	Dr.Enrico Bertini Prof. Bruno Dallapiccola
Sponsor:	Bambino Gesù Children's Hospital (European Commission)
Person responsible for the study WP6	Prof. Jaap Harlaar (Vrije Universiteit Amstrerdam)

Scientific Coordinator of the Project	Prof. Bruno Dallapiccola
Data Management/Statistical analysis:	

Protocol approved and signed by:

Scientific Coordinator of the Project:

Prof. Bruno Dallapiccola

Responsible Work Package 6:
Prof. Jaap Harlaar
Principal Investigator:
Dr. Enrico Bertini

Acronym List

AEs	Adverse Events
EC	Ethical Committee
CRF	Case Report Form
GCP	Good Clinical Practice
ICH	International Conference on Harmonisation
IRB/IEC	Institutional Review Board/Independent Ethics Committee
NSAEs	Non Serious Adverse Events
SAEs	Serious Adverse Events
SOPs	Standard Operating Procedures
……	……..

INDEX

1 INTRODUCTION

1.1 BACKGROUND OF THE MD-PAEDIGREE PROJECT

MD – PAEDIGREE *Model-Driven European Paediatric Repository* is a research project funded by the European Commission under the Virtual Physiological and Human Area (VPH) of the ICT Theme of the Seventh Framework Programme (contract no. 600932). The project foresees 7 world-renowned clinical centres of excellence pursuing improved interoperability of paediatric biomedical information, data and knowledge, by developing a set of reusable and adaptable multi-scale models for more predictive, individualised, effective and safer paediatric healthcare. The project is scientifically and technologically supported by one of the leading industrial actors in medical applications in Europe, and operating in conjunction with highly qualified SMEs and some of the most experienced research partners in the Virtual Phisiological Human (VPH) community.

MD-Paedigree validates and brings to maturity patient-specific computer-based predictive models of various paediatric diseases (cardiomyopathy, risk of cardiovascular disease in obese children and adolescents, Juvenile Idiopathic Arthritis (JIA) and neurological and neuro-muscular diseases (NND), thus increasing their potential acceptance in the clinical and biomedical research environment by making them readily available not only in the form of sustainable models and simulations, but also as newly-defined workflows for personalised predictive medicine at the point of care.

MD-Paedigree aims to advance the state-of-the-art of patient-specific computational modelling of different paediatric diseases and translate the latest advances into clinics to improve disease understanding, therapy outcome, and provide an infostructure platform for assessing new therapies at the point of care.

MD-Paedigree's goals therefore are to integrate and share highly heterogeneous biomedical information, data and knowledge, using best, jointly develop reusable, adaptable and composable multi-scale VPH workflow models and to support evidence-based translational medicine at the point of care practices from the biomedical semantic Web.

Ilness in infants, children, and adolescents are a large and under-appreciated public health problem. Because paediatric patients now have a much longer life expectancy, the burden and costs are substantial for families and society. For instance, recent studies have shown that the number of adult patients with congenital heart disease is already similar to that of the paediatric population and will continue to grow. It can therefore be stated that in the longer term the VPH scientific approach requires a fundamental research investment in paediatrics, where the conceptual revolution which is underway, transforming the nature of healthcare from reactive to preventive, can best be applied, moving towards an approach based on personalised, predictive, preventive, and participatory (P4) medicine, increasingly focused on wellness.

MD-Paedigree represents a major step towards personalised paediatric e-Health, based on data-driven models, patient-specific simulations and a sustainable data and model repository. The project can in fact impact the way healthcare is practiced in the future. MD-Paedigree's impact on dealing with very concrete and exemplary clinical cases is demonstrated through the following applications in the diseases scenarios. MD-Paedigree encompasses complete services for storage, similarity search, outcome analysis, risk stratification, and personalised decision support in paediatrics within its innovative model-driven data and workflow-based models repository,

leveraging on service and knowledge utilities. It fosters the state-of-the-art of patient-specific computational modelling of the selected diseases and translates the latest advances into clinics to improve disease understanding and provide a platform for testing new therapies at the point of care. In fact, MD-Paedigree demonstrates how a dedicated VPH repository can provide full accessibility to existing and further developing knowledge in paediatrics to all interested bio-medical researchers precisely through linking data with models, thus proving the large scale benefits of having both the data and models readily available at the point of care.

The data collection performed within Health-e-Child and Sim-e-Child projects (already funded by the European Commission in previous calls for proposals) is still available to the MD-Paedigree consortium thanks to the continuity in the eHealth platform which is shared by all three projects. In addition, the new patients' recruitment to be performed within MD-Paedigree consists of:

<table>
<tr><th>Pathology</th><th colspan="2">No of patients/Time</th></tr>
<tr><td>Cardiomyopathies</td><td>180 children, by month 33: 60 patients (among which 30 girls) for each clinical centre.</td><td rowspan="3">Genetic and meta-genomic:

180 patients with cardiomiopathies, 180 with CVD risk in obesity, 200 with JIA, and 100 unaffected subjects (control group).</td></tr>
<tr><td>CVD risk in obese children</td><td>180 patients , by month 36: 60 (among which 30 girls) for each clinical centre.</td></tr>
<tr><td>Juvenile Idiopathic Arthitis (JIA)</td><td>Altogether 200 patients by month 28.</td></tr>
<tr><td rowspan="3">NND</td><td colspan="2">Cerebral Palsy: 50 patients for each clinical centre for probabilistic modelling, as well as 600 retrospective patients from KU Leuven and OPBG.</td></tr>
<tr><td colspan="2">Spinal Muscular Atrophy (SMA)
Data will be collected by OPBG, KU Leuven and VUA from 20 ambulant patients (severity grade type 3)</td></tr>
<tr><td colspan="2">Duchenne Muscular dystrophy (DMD)
Clinical data will be collected by OPBG, KU Leuven and VUA from 20 ambulant genetically confirmed DMD Patients.</td></tr>
</table>

1.2 BACKGROUND OF THE DATA ACQUISITION AND PROCESSING FOR NND STUDY

In Neurological and Neuromuscular Diseases (NND) as well as in certain chronic paediatric diseases of the musculoskeletal system, treatments are strongly guided by maximising the walking function of the human movement system, because walking is considered as clinically meaningful by patients. This generalises to most mobility-related functions. The most common paediatric disorder within the NND disease area is

Cerebral Palsy (CP) whose incidence ranges between 2 to 3.6 per 1,000 live births [Odding E. et al., 2006]28. CP includes a group of non-progressive, often changing, motor impairment disorders, secondary to lesions in the sensory-motor cortex and corticospinal tract, arising in the early stages of the child's development. Conventional clinical gait analysis (CGA) is already an important tool in the treatment of children with CP that aims to improve or sustain walking performance, but its potential is under-utilised and recent developments need full exploration.

The second important disorder is Spinal Muscular Atrophy (SMA), an autosomal recessive disease characterised by degeneration of motoneurons in the spinal cord. SMA is caused by mutations of the survival motor neuron 1 gene (SMN1). Estimated incidence is 1 in 6,000-8,000 live births. This disease is characterised by progressive generalised muscle weakness and atrophy predominating in proximal limb muscles. For ambulant SMA patients, new methods for functional motor evaluation based on gait modelling would allow to increase sensitivity to change in assessing weakness and fatigability.

The third disorder, Duchenne Muscular Dystrophy (DMD) is the most common and severe form of muscular dystrophy, with an FP7- ICT-2011.5.2 600932 - MD-Paedigree –Part B 9 incidence around 1 in 3,600 juveniles. This disorder is caused by a mutation in the dystrophin gene, that codes for a protein which is a major structural component of the muscle. The absence of dystrophin results in muscle degeneration, difficulty in walking (resulting in wheelchair use from 14 years of age), followed by loss of arms and hands function. In the last few years, following a rapidly increasing number of potentially effective therapeutic approaches for DMD, the request for validated and sensitive outcome measures to be used in clinical trials has increased. Although walking is a common task executed by a healthy individual in a seemingly effortless manner, it implies a complex involvement of inputs from several senses (visual, vestibular, proprioceptive, somatosensory), partly automated by the so called spinal central pattern generator (CPG) [Spardy LE et al, 2011]29. These inputs are known to interact with each other, but the way in which this is performed is not fully exploited at present [Baker R , 2006]30. Nevertheless, the current insights are certainly at an advanced state that allows for meaningful application towards pathological walking, where decision support is needed [FM Chang et al., 2010] 31. In the clinical practice of specialised centres, CGA is used to evaluate the joint and muscle functions in their functional context, i.e. during gait [Cappozzo A. et al., 2005]32. Common CGA measures 3D kinematics (by 3D optoelectronic registration of skin mounted markers). Each relevant degree of freedom (DOF) is expressed as a function of the gait cycle. Moreover, using a mass distribution model and measuring ground reaction forces, the net moments for each DOF are calculated using inverse dynamics analysis.

Muscle activation patterns, for all relevant muscles, are measured using electromyography (EMG) for each targeted muscle. Finally, the energy cost of walking can be evaluated using metabolic measurements. CGA is a special form of personalised computer-aided medicine that supports clinical decision making [Novacheck TF et al., 2010]33. Unfortunately, the output of CGA is not yet in a format that permits clear, unambiguous interpretation, because of the redundancy of the Neuro-Musculo-Skeletal System (NMSS) which obstructs distinguishing cause from compensation. Even though recent developments in modelling the NMS Physiome as a part of EU funded Virtual Physiological Human efforts are at an advanced state, their results have not yet been implemented in clinical practice, and the full potential of CGA still needs to be reaped. A combination of standard protocols of gait analysis, biophysical modelling and large scale statistical analysis can therefore be expected to provide a powerful framework for meaningful interpretation.

Protocols and personalised models in Advanced Clinical Gait Analysis

To reiterate the conclusion of the NDD clinical background section: the potential of gait analysis to serve clinical decision making in NDD is generally under-used for several reasons. These will be taken up within the MD-Paedigree project.

Protocol definitions for clinical gait analysis

Three levels of protocol definitions are needed to assure multicentre reliable data for the repository:
Technical Quality assurance for CGA laboratories

It is important to realise that for accurate data from the experimental systems a strict analysis of causes of errors and periodical validation procedures needs to be implemented in the gait labs [Cedraro, A.et al., 2009116, Chiari L. et al.,2005 117]. If the adopted experimental procedure permits the gathering of valid data, the first important prerequisite for reliable and accurate results from a particular subject is fulfilled. Within MD-Paedigree these quality assurance (QA) procedures will therefore be formalised between laboratories for clinical gait analysis. MD-Paedigree will constitute a European standard for technical QA and have this approved by the important European bodies on clinical gait analysis, i.e. the ESMAC. A consensus meeting will be part of this.

Standardisations of gait analysis protocols: Marker placements

One of the main non-technical sources of error in CGA using OptoElectronic Movement Analysis systems is caused by marker artifacts, resulting from skin movement relative to the bone [Leardini et al., 2005] 118. Recently it has been shown that, in the case of well-trained staff, errors due to marker misplacements and skin movement artifacts will stay within a few degrees of error of the joint kinematics graphs [J.L. McGinley et al. 2009] 119. This error level is considered to be just clinically acceptable. This means that all gait labs should fulfill the requirements to be qualified for MD-Paedigree graded gait analysis. In analogy with the Technical Quality Assurance (TQA), MD-Paedigree will strongly promote interoperability and constitute a protocol for standardised marker placement, as well as standard procedures to evaluate this within and between laboratories. In parallel, we shall explore the possibility to use imaging/gait analysis protocols, where patients are dressed with radiopaque/MRI opaque and reflective markers attached to the skin as used in gait analysis protocols, while the imaging protocol is conducted. These data will make possible to use sophisticated inverse kinematics modelling methods to minimise the skin artifacts, and to obtain accurate estimations of the skeletal kinematics.

Standardisations of gait analysis protocols: operational protocols

The results of kinematics and kinetics of CGA are also dependent on the use of standard protocols for instruction on walking targets. In particular, the enforcement of a precise walking speed is of major influence on the output [Schwartz MH et al. 2008] 120. As such, instructions should be carefully standardised and protocols developed that use multiple walking speeds. It has been suggested and shown by previous studies [Bovi G. et al., 2001]121, that these protocols are necessary to detect important pathological features of the NMSS of the subject, especially in patients with CP [van der Krogt MM. et al., 2009]122. EMG recordings and oxygen consumption will be part of the overall assessment procedures. Moreover, in order to feed the development of probabilistic models a standardised description of therapies will be completed. This description will be used to longitudinally describe the applied clinical workflows that are currently used to improve gait performance in children with NND.

Conclusion: the established and clinically authorised protocols (technical, marker and procedures) of CGA will be an important step forward for the NND paediatric care in the EU, along with the establishment of a reliable MD-Paedigree database for typically developing children.

Application of computational biophysical models of the NMSS in CGA

For clinical gait analysis the use of Neuro-Musculo-Skeletal (NMS) models is an important step forward in the interpretation of its results, aiming to inform the clinical decision-making. Because of the modelling based interpretation, the physician no longer needs to interpret the results of clinical gait analysis, within his own informal frame of interpretation. Using NMS models the results of CGA are quantitatively "translated" into the function and performance of the underlying structures, i.e. muscle activation, muscle forces, and joint loads that make possible to unravel the aetiology of the pathological gait pattern of the subject under study. The EU project “Personalised models of the Neuro-Musculo-Skeletal Physiome” (NMS Physiome 123) is moving towards the development of PPI (Predictive, Personalised and Integrative) musculoskeletal medicine. NMS Physiome is a part of the European Union’s Virtual Physiological Human initiative.

A key result of this project, conducted by Prof. Viceconti, at MD-Paedigree partner USFD, is the integration of an advanced software application for the pre-processing of imaging and gait analysis data into a full

musculoskeletal model (NMS Builder) and the OpenSIM musculoskeletal modelling environment developed by Stanford University. NMS Builder is already available in prototypical form to all partners of the MD-Paedigree consortium. Although NMS computational models are thus well known in the biomechanical research community, as yet only one company, MOTEK, has incorporated gait analysis and model based interpretation of gait for market delivery. Their model (the HBM model) is computationally very efficient: even without high performance computers it can run in real time. More complex modelling activities can be conducted using the NMS Physiome tools. The actual problem of accuracy of NMS models is that all models currently used in paediatric gait analysis are based on data scaled from a single cadaver in a simple way. Sensitivity studies have shown that such a gross simplification in applying generic models is too inaccurate, and, especially in the case of children, dedicated and validated models, fused with medical imaging data, should be developed in order to yield reasonable accuracy for clinical application in this population. The first level of MS models in CGA is the mass distribution model of body segments. Mass distribution means that the masses, centre of mass and inertial properties of each segment need to be known for accurate calculation of inverse dynamics resulting in valid joint kinetics. What is needed is a method for scaling that allows application, in clinical workflows, to enable personalised medicine. MD-Paedigree will develop and evaluate a scaling method for the NMSS of children, to be applied in existing NMS models that are used in CGA. Validation will be based on MRI measures. Next to anthropometrics scaling is the alternative to use a 2D image, generated by a whole body DXA image, morphed to a generic 3D skin model of a child. The advantage is that DXA provides accurate measurement of the areal density of the bone, fat, and lean tissues the inertial properties of each segment. The second level of personalised MS models in CGA are to account for the subject specific bony deformities. The bony deformities that should be accounted for can be limited to the clinically well known deformities in CP. These deformities have significant influence on the output of NMS model calculations (i.e. femoral anteversion and tibial torsion). These effects could primary be modelled by morphing the generalised bony structures towards the actual morphology of the bone. The most important effects of bony deformities should be parameterised by the effects on axis alignment: (a) introducing a skewness of the principal axes of rotation of the joints in the kinematic chain of linked segments, and (b) the altered lever arms of muscles with respect to these principal axes of rotation of the joint. Again anthropometric measures and DXA will be explored. The third level of personalised modelling is to account for pathology specific muscle parameters. These models should focus on the parameters that are known to be of large influence on the second step in inverse dynamics, i.e. the estimation of muscle forces based on optimisation criteria on how to explain the net joints moments from CGA. This means that especially muscle contractures, altered muscle structure and hypertonia (in CP), as well as muscle weakening (in DMD and SMA), must be targeted. US measures of the muscle belly, along with fibre directions will enable estimates of the muscle Physiological Cross sectional Area (PSCA), while dynamometric evaluations will yield measures of muscle belly length and optimal fibre length.

Supporting probabilistic models, despite the strong potential of biophysical models of the NMSS, will only hold a certain amount of predictive value, i.e. as far as their assumed accuracy will allow. However, in clinical practice, even if the pathology cannot be fully explained by biophysical modelling, the use of probabilistic models is still extremely powerful in supporting clinical decision making. Until now only two gait laboratories in the world (Gillette Children's, Minneapolis, US and Pellenberg, Leuven, Belgium) have explored the possibilities of generating decision rules from their dataset [van Gestel et al., 2011 124]. These laboratories are the only ones that have created a large enough set of reliable data to make such an effort worthwhile. In MD-Paedigree the clinical partners will collect data, according to the dataset and quality protocols defined on the basis of standardised formats, for feeding into the repository.

2 AIMS OF THE STUDY

2.1 MAIN GOAL

Main goal of the study is to acquire sets of data (gait analysis and images) related to Neurological and Neuromuscular Diseases for the repository, and to develop probabilistic modelling and biophysical modelling.

The most common paediatric disorder within the NND disease area is Cerebral Palsy (CP) whose incidence ranges between 2 to 3.6 per 1,000 live births. CP includes a group of non-progressive, often changing, motor impairment syndromes, secondary to lesions in the sensory-motor cortex and corticospinal tract, arising in the early stages of the child's development. Conventional clinical gait analysis (CGA) is already an important tool in the treatment of children with CP that aims to improve or sustain walking performance, but its potential is under-utilised and recent developments need full exploration. The second important disorder is Spinal Muscular Atrophy (SMA), an autosomal recessive disease characterised by degeneration of motoneurons in the spinal cord. SMA is caused by mutations of the survival motor neuron 1 gene (SMN1). Estimated incidence is 1 in 8,000 live births. The third disorder, Duchenne Muscular Dystrophy (DMD) is the most common and severe form of muscular dystrophy, with an incidence around 1 in 3,600 juveniles. This disorder is caused by a mutation in the dystrophin gene, that codes for a protein which is a major structural component of the muscle. The absence of dystrophin results in muscle degeneration, difficulty in walking (resulting in wheelchair use from 14 years of age), followed by loss of arms and hands function. In the last few years, following a rapidly increasing number of potentially effective therapeutic approaches for DMD, the request for validated and sensitive outcome measures to be used in clinical trials has increased.
A combination of standard protocols of gait analysis, biophysical modelling and large scale statistical analysis can therefore be expected to provide a powerful framework for meaningful interpretation.

2.2 PRIMARY END-POINT

A comprehensive clinical dataset of gait analysis data for DMD and SMA and MRI and DXA data sets of 30 CP patients.

3 STUDY DESCRIPTION

3.1 STUDY DESIGN

Gait analysis collection for DMD and SMA

Although the problems in DMD and SMA are less complex than in CP (e.g. no spasticity), the protocols developed in T6.1 apply for these populations, to be used in conjunction with modelling, to demonstrate reusability. The clinical problems in DMD and SMA are to trace subtle changes in motor performance during walking, in order to monitor the effects of intervention very quickly. For SMA clinical data will be collected

by OPBG, KU Leuven and VUA from 20 ambulant patients (severity grade type 3); 10 patients will be selected among the 3a subgroup (symptoms of weakness appearing before age 3 years), and 10 patients will belong to the 3b group (weakness appearing after the age of 3 years).
Besides considering type of severity in the selection of patients for data analysis, we will include children having an age range of 5 to 10 years. Particularly, we will recruit :

- children of 5-6 years with the diagnosis of SMA type 3a
- children with age range of 5-10 years with SMA type 3b

Younger children will not be able to fully collaborate during the evaluation process that includes functional motor scales and gait analysis.
All patients will receive a longitudinal full control evaluation at baseline (0), after 12-18 month (1) and 2-3 years(2).

MEASUREMENTS:
1. Functional motor scales:
For SMA patients
Expanded Hammersmith functional motor scale to measure function 6 minutes walk test to measure strength and fatigue, hand held myometer (CITEC) to measure strength (knee flexors and extensors).

2. Gait analysis according to protocols
Compared to SMA, DMD is a rather a homogeneous disorder with well defined natural history endpoints, although the standardised use of steroid treatment and progress in standards of care has changed the natural history of the disease prolonging walking by 2 to 5 years, in relation to natural history data known before systematic steroid treatment, when patients generally lost walking ability between ages of 7-12 years. Clinical data will be collected by OPBG, KU Leuven and VUA from 20 ambulant genetically confirmed DMD patients treated with the same steroid regimen of daily deflazacort 0.75mg/kg/day and with the most common mutations in the dystrophin gene. Age range of patients will be between 5 and 7 years. In particular, we will recruit 10 patients with age between 5 and 6 years, and additional 10 patients with age between 7 and 8 years. In this second DMD group we will observe longitudinally the progression of the disease in the time span of 4 years, because it is known from current natural history data that DMD patients start a downhill progression of function after age of 7-8 years. All patients (10 from OPBG and 10 from KU Leuven) will receive a longitudinal full control evaluation at baseline (0), after 12-18 month (1), and 2-3 years(2).

Measurements:

1. Functional motor scales: the North Star Ambulatory Assessment (NSAA) 6 minutes walk test (6MWT) to measure strength and fatigue, hand held myometer (CITEC) to measure strength (knee flexors and extensors).
2. Gait analysis according to protocols previously identified.
3. In addition OPBG, KU Leuven and VUA will acquire electrocardiographic and echocardiographic data from all
the 20 DMD patients.

Image acquisition

In WP 11 some advanced modeling is developed, that the fusion of multimodal sources of data (MRI, DXA and CGA). As an input to this WP, each clinical center (VUA, OPBG, KU Leuven) will acquire at least 10 subjects with both MRI and DXA, including the markers that are needed for gait analysis. Volume of interest includes pelvis, femur, tibia, foot. The first three subjects should be acquired within the first year of the project. Images will have to be anonymised before making them available for the technical partners.

3.2 SUBJECTS SELECTION

Data collection will be performed in three leading European Centers (Ospedale Pediatrico Bambino Gesù (OPBG, Rome, Italy), Katholieke Univesiteit Leuven (KULeuven, Belgium), Vrije Universiteit Amsterdam (VUA, The Netherlands).

Inclusion criteria are:

- GMFCS 1-2 (Palisano R et Al. Development and validation of a gross motor function classification system for children with cerebral palsy. Dev Med Child Neurol 1997; 39: 214–23);
- sufficient cognitive skills: Total IQ by Wechsler Intelligence Scale: > 60.
-

Exclusion criteria are:

- Patient history of functional surgery on bones and muscles;
- Lower limbs BoNT A injection in the last 6 months;
- Patient history of hip, backbone and/or lower limb fracture;
- relevant visual deficit non correctable by with lenses;
- significant comorbidities.

The subjects will be enrolled between the in- and out-patients coming for assessment and rehabilitation in the involved centers. A MD of the centre will explain to the parents and the child the aims and the features of the research, in order to have their consent.

4. WITHDRAWAL FROM THE STUDY

Participation in this study is entirely voluntary. The parents may decide, at any time, that they do not wish their child to take part in the study without altering current or future treatment in any way. If at any stage of the project the parents wish to withdraw their child from the study, the researcher responsible for the study will arrange for their child's information to be immediately removed from computer records and for any remaining samples we hold belonging to him/her to be destroyed.

5 PATIENT'S STUDY

5.1 STUDY TO BE PERFORMED

For all patients will be carried out:

- a standardized anamnesis (Gestational age, birth weight, Apgar Score at V minute, kind and localization of brain abnormalities at MRI , any previous injection of botulinum toxin, previous orthopedic functional surgery);
- a standard clinical examination (PROM at hip, knee and ankle, MRC strength muscle at major muscle groups of lower limb, MAS at flexor/extensor of hip, knee and ankle) and a neurological assessment;
- a Pelvic X-Ray, if not performed in the last year;
- Clinical gait analysis (CGA) with collection of kinematic, kinetic, EMG.

- Lower limb MRI
- Measurement of metabolic consumption during six minutes walking test (6MWT)
- Dual Energy X-Ray Absorptiometry – DXA

Medical information (clinical and instrumental evaluation) collected are part of our common clinical practice but we will inform the child and the parent that the data collected from these surveys will be stored anonymously in this study and in future investigations. An additional informed consent will be required if any of these tests will be required with the sole purpose of research.

6. STUDY PLANNING

6.1 EFFICACY PARAMETERS

Data source: all data will be collected by a MD though observations, visit of the subjects and instrumental examinations. They will be recorded in CR and in ad hoc modules.
Data gathering scheduling: children affected by CP will be assessed only once while patients with DMD and SMA will be evaluated at baseline (inclusion in the study) and at the 1 and 2 years follow-up.
All scales and gait analysis protocols are internationally standardised and with a high reliability.

6.2 EXPERIMENTAL DESIGN

The study will last 4 years. It is designed as a prospective longitudinal study. The timeframe for patient recruitment spans the first 3 years. Follow-up data for each data will be collected at follow-up visit as indicated in details in patient study session.

6.3 DATA PROTECTION

All collected data will be anonymised. Clinical data and biological samples will be coded and stored as such. The code will be generated by software using a system of 128-bit encryption. The code will be stored in a close drawer Prof. Jaap Harlar. At the end of the study, the key code will be destroyed and, hence, data anonymised. From this moment on, it will not be possible for anyone to discover the patient's identity.
All clinical data will be communicated to participants and/or legal representative except for genetic testing.

7 SECURITY EVALUATION

7.1 DEFINITIONS

No adverse effects are foreseen as consequence of the clinical study.

8. SAMPLE DIMENSION AND STATISTIC METHODOLOGY

8.1 STATISTIC DESIGN

As the main goal of whole project is to establish a data repository for pediatric diseases, the sample size has been set by taking into account primary endpoints and study power but also available resources at each center, and study feasibility. In particular for the genetic analysis no study power is foreseen.

NND Modeling will be done by the following partners: Siemens AG, MOTEK SrL; University of Delft, La Sapienza University of Rome, University of Sheffield.

8.3 MANAGEMENT OF MISSING DATA

Patients whose information is missing or incorrectly transcribed will not be used for the analysis, but will be marked appropriately in the electronic database by inserting a variable "Flag".

9. ADMINISTRATIVE AND ETHICAL PROCEDURES

All clinical procedures and instrumental examination are usually performed for the management and the rehabilitation of children affected by CP, DMD, SMA.
The MRI study of lower limbs (pelvis, femur, tibia and foot) will be performed only in 10 subjects able to collaborate, avoiding the movement artifacts.
Medical information (clinical and instrumental evaluation) collected are part of our common clinical practice but we will inform the child and the parent that the data collected from these surveys will be stored anonymously in this study and in future investigations. An additional informed consent will be required if any of these tests will be required with the sole purpose of research. A MD of the centre will explain to the parents and the child the aims and the features of the research, in order to have their consent.

9.1 AUTHORISATIONS

The protocol will undergo the approval of the Ethical Committee for the study implementation before the enrollment of the patients.

9.2 INFORMED CONSENT

Each parent/patient who will be asked for his/her enrollment in the study will be informed on every aspect concerning the study itself and, before the recruitment begins, he/she must sign the informed consent model. The date of the signature of the consent will be recorded. A Copy of the informed consent model must be handed to the parent/patient.

9.3 INSURANCE COVERAGE

Insurance coverage used is as foreseen by each research structure for clinical and research activities. ***9.4 USE OF THE INFORMATION AND DATA PUBBLICATION***

Researchers must have the right to disseminate/publish the results of the study, with respect to the current legislation on sensible data protection and Intellectual property Rights. European Commission does not have any right in the publication/dissemination of study results.

9.5 CLINICAL PROTOCOL AMENDMENTS

Clinical protocol amendment is intended to be any modifications to the final version of the clinical protocol that, after being approved by the Researcher and by the Ethical Committee, cannot be informally modified.

9.6 DATA MANAGEMENT AND DOCUMENT PRESERVATION

Researchers will manage the preservation of patients identification codes for at least 15 years after the end of the study. Patients clinical data and other original data will be kept for the maximum period allowed by the Hospital (min. 15 years).

9.7 BUDGET

The project costs are covered by the specific Grant by the European Commission. Such grant will cover 75% of staff costs, of consumables and equipment used for the project, and of all other expenses involved. Annex 1 (GPF) contains the project's entire budget, including details of the Institution's budget.

10. RESEARCHER RESPONSIBILITY

Researcher, as principle investigator of the study declares to be aware of all the duties committed to him/her and his/her collaborators for the conduction of the study. Except where explicitly edited, the term "researcher" on this Protocol and on the CRF is referred to the researcher or to any other person appointed by him/her, who will have power to sign documents instead of him/her.

The researcher will conduct the study according to the principles of the Helsinki declaration (1964) and Amendments (Appendix 1), and to the European Guidelines on Clinical Good Practices .

11. Annexes

Annex 1 GPF
Annex 2 DOW

12. REFERENCES

15. Odding E, Roebroeck ME, Stam HJ. The epidemiology of cerebral palsy: incidence, impairments and risk factors. Disabil Rehabil. 2006 Feb 28;28(4):183-91.
16. Spardy LE, Markin SN, Shevtsova NA, Prilutsky BI, Rybak IA, Rubin JE; A dynamical systemsanalysis of afferent control in a neuromechanical model of locomotion: I. Rhythm generation. J Neural Eng. 2011 Dec;8(6)
17. Baker R, Gait analysis methods in rehabilitation, J of NeuroEngineering and Rehab 2006 3(4), 10 31. FM Chang, JT Rhodes, KM Flynn, JJ Carollo, The role of gait analysis in treating gait abnormalities in cerebral palsy, Orthop Clin North Am. 2010 Oct; 41 (4):489-506;
18. Cappozzo A, Della Croce U, Leardini A, Chiari L. Human movement analysis using stereophotogrammetry. Gait Posture. 2005 Feb;21(2):186-96. Review.
19. Novacheck TF, Trost JP, Sohrweide S. Examination of the child with cerebral palsy. Orthop Clin North Am. 2010 Oct;41(4):469-88. Review.
20. Cedraro, A. Cappello, L. Chiari. A portable system for in-situ re-calibration of force platforms: Experimental validation . Gait Posture 29 (2009) 449–453.
21. L. Chiari, U. Della Croce, A. Leardini, A. Cappozzo "Human movement analysis using stereophotogrammetry Part 2: Instrumental errors" Gait and Posture 21 (2005) 197–211.
22. Leardini A, Chiari L, Della Croce U, Cappozzo A. Human movement analysis using stereophotogrammetry. Part 3. Soft tissue artifact assessment and compensation. Gait Posture. 2005 Feb;21(2):212-25. Review
23. J.L. McGinley et al. The reliability of three-dimensional kinematic gait measurements: A systematic review Gait Posture 2009, 29:3, 360-9
24. Schwartz MH, Rozumalski A, Trost JP. The effect of walking speed on the gait of typically developing children. J Biomech. 2008;41(8):1639-50
25. Bovi G, Rabuffetti M, Mazzoleni P, Ferrarin M. A multiple-task gait analysis approach: kinematic, kinetic and EMG reference data for healthy young and adult subjects. Gait Posture. 2011 Jan;33(1):6-13
26. van der Krogt MM, Doorenbosch CA, Becher JG, Harlaar J. Walking speed modifies spasticity effects in gastrocnemius and soleus in cerebral palsy gait. Clin Biomech (Bristol, Avon). 2009 Jun;24(5):422-8.
27. http://www.nmsphysiome.eu
28. Probabilistic gait classification in children with cerebral palsy: a Bayesian approach. Van Gestel L, De Laet T, Di Lello E, Bruyninckx H, Molenaers G, Van Campenhout A, Aertbeliën E, Schwartz M, Wambacq H, De Cock P, Desloovere K. Res Dev Disabil. 2011 Nov-Dec;32(6):2542-52.

Appendix 2 - Working Groups' Papers

Cardiomyopathies

MD-PAEDIGREE KICK OFF MEETING

CARDIOMYOPATHIES WG

Participant's Name	Affiliation

Concept (general)

MD-Paedigree validates and brings to maturity patient-specific computer-based predictive models of various paediatric diseases

- increasing their potential acceptance in the clinical and biomedical research environment
- making them readily available not only in the form of sustainable models and simulations, but also as newly-defined workflows for personalised predictive medicine at the point of care.

These tools can be accessed and used through an innovative model-driven infostructure

- powered by an established digital repository solution
- able to integrate multimodal health data
- entirely focused on paediatrics
- conceived of as a specific implementation of the VPH-Share project, planned to be fully interoperable with it and cooperating, through it, also with p-Medicine.

MD-Paedigree aims at achieving high-level semantic interoperability,

- requiring standards enabling the clinical contents to be interpreted consistently across the different EHR regimes,
- while complete clinical interoperability between systems will require widespread and dependable access to maintained collections of coherent and quality-assured semantic resources,
- including models that provide clinical context,
- mapped to interoperability standards for EHR and PHR and biomedical data,

linked to well specified terminology value sets, derived from high quality ontologies

CONCEPT (SPECIFIC)	**Beyond the state of the art**	**WPs' OBJECTIVES**	**Lead**	**Estimated % realisation**
When children present with new onset heart failure, there are five possible outcomes: full recovery, dilated cardiomyopathy (DCM) requiring drug therapy, DCM requiring transplantation or mechanical support, another diagnosis (other forms of cardiomyopathy, metabolic disease) or death. At presentation, however, it is very difficult to predict which	**1. Anatomy** The first step of the analysis is to compute a detailed model of the cardiac anatomy of a patient. In the course of the Health-e-Child and Sim-e-Child projects, we have been able to extract the anatomy and dynamics of left and right ventricle, and left and right atria, aorta, aortic and mitral valves and pulmonary valve and trunk. Our aim for MD-Paedigree is to integrate these different modules into one robust framework to extract a dynamic anatomical model of the complete	**WP2: Clinical and technical user requirements for disease modelling** • Incorporate into the model the variables that are analysed by the clinicians in their activity. • Ensure that the modeling reflects real clinical needs and is validated against them to assure their robustness and reproducibility. • Provide computational models that can be personalized by adapting the parameters to the integrated data of a patient case • Advance the knowledge about the		

group any patient will end up in. The objective is to: • capture the main features of the cardiovascular system, including the heart, arteries and peripheral circulation, to predict cardiomyopathy progression • plan therapies like heart transplant and ventricular assist devices. • Investigative data provided by imaging, pressure monitoring, clinical observations and exercise will be used to build these models • and to validate them, by comparing model prediction with actual outcome. By merging all scattered information obtained from different diagnostic tools in clinical practice, and obtaining a generative model of heart function in children, our model will provide cardiologists the tools to deliver patients the	heart from MRI and echocardiography data. This will yield a holistic view of the cardiac system, as required by clinicians, especially in the context of cardiomyopathies and their associated complex dysfunctions. **1. Myocardial fibre structure** Fibre architecture plays an important role in the realistic modelling of electrical and mechanical heart activity, but it is not yet possible to acquire in-vivo in-situ images of heart fibres in clinical routine. To cope with this limitation, computational models usually rely on generic fibre orientations. A common approach is to synthesise the variation of fibre orientation using rule-based methods. As a more realistic alternative, we proposed statistical models of heart fibres based on diffusion tensor images. In MD-Paedigree, we will integrate such statistical fibre models into our comprehensive anatomical model. **2. Computational fluid dynamics** With recent advances in patient-specific 4D anatomical modelling and 3D flow measurement techniques, it has become possible to employ computational fluid dynamics (CFD) for haemodynamic assessment and subsequent validation in cardio-vascular applications. Flow patterns and underlying flow parameters obtained from such simulations may be used for early diagnosis, prediction and benchmarking treatment outcomes. While most previous approaches have focused on a single cardiac component, we have recently performed simulations of blood flow in the whole heart using high-quality patient-specific heart models derived from 4D CT, as part of the Sim-e-Child project.	selected diseases by allowing the simulation of different effects on the evolution of the disease • Predict the effect of therapy. • Ensure that MD-Paedigree models have the highest possible impact at the point of care. • Re-use of models between disease areas to leverage synergies where possible. • Existing standards for modelling and tools will be investigated. • The need for new standards will be evaluated and documented.		
		WP3: Data acquisition and processing for Cardiomyopathies • Overall objective: three cohorts of 60 CMD children. • Parents or responsible guardians will be asked for informed consent. • 60 patients (30 girls) for each clinical Centre will be consecutively enrolled. • Inclusion criteria will be age up to 18 years, and established diagnosis of CMD (including both primary and secondary CMDs).		
		WP8: Modelling and simulation for Cardiomyopathies Provide cardiologists the tools to deliver patients the best possible medical care and treatment planning by allowing them to predict and simulate cardiomyopathy progression: 1. Merge all scattered information from different diagnostic tools in clinical practice to: • Capture the main features of the		

best possible medical care.	This was the first time that 4D physiological models of a patient's valves together with the models of the chambers, myocardium, and main vasculature captured from 4D CT have been used to provide patient-specific constraints for the simulations of the blood flow inside the heart. **3. Arterial circulation** In MD-Paedigree, we will re-use models already available in the consortium (developed by INRIA and SCR) to simulate cardiomyopathies and therapies. In particular, aspects of the arterial circulation will be integrated as boundary conditions and modelled using quasi 1D methods with visco-elastic walls. We will also develop a framework to combine the output of these models into a consensus prediction and a variability map, framework that can be enhanced by additional models from the VPH community. In Sim-e-Child, we developed efficient numerical methods for 3D-1D and 3D-0D coupling. We successfully used these methods to couple 3D aortic CFD simulations with both 1D distal vessels and 0D micro-vessel models, and reported excellent agreement between in-vivo and simulated pressure drops across coarctations. More recently, we have also developed estimation algorithms for determining the boundary conditions from routine flow (echo Doppler) and pressure (cuff) measurements. **4. Fluid structure interaction** MD-Paedigree aims to extend our current methodology, which uses a robust one-way interaction to transfer momentum from the moving solid walls to the blood, to a two-way coupled framework that fully models fluid structure	cardiovascular system, including the heart, arteries and peripheral circulation • Obtain a generative model of the heart function in children • Yield a holistic view of the cardiac system 2. Integrate all different modules into one robust framework to extract a dynamic anatomical model of the complete heart from MRI and echocardiography data: • Taking fully into account how haemodynamics play a significant role in determining the progression of cardiomyopathies and their associated complex dysfunctions • Achieving fast and efficient extraction of anatomy and dynamics of left and right ventricles, and left and right atria from US and MR images building on probabilistic, Shape Regression and Trajectory Spectrum Learning techniques already used in the Health-e-Child and Sim-e-Child projects • Re-using models previously developed by INRIA and SCR to model fluid structure interaction (FSI) physics with patient-specific electromechanical models of the heart • Using recent advances in deformation-based shape modelling to model the evolution of the heart over time and also		

	interaction (FSI) physics with patient-specific electromechanical models of the heart. The coupling between fluid and solid will use a previously tested robust algorithm to exchange information between the involved solvers. The blood stress tensor provides traction forces at the endocardium surface, used as boundary conditions by the electromechanical model, while the endocardium velocities are used as boundary conditions for the fluid flow computations. Our FSI model will be also coupled with the models of the systemic and pulmonary arterial circulation for a holistic view of the cardiovascular system at various states (rest, exercise, under vasodilating/vasoconstricting drugs, etc.). **5. Atlas-based techniques** In MD-Paedigree we will explore atlas-based techniques of reduced models to speed up calculations. This allows an order of magnitude of reduction in the number of parameters with reasonable accuracy. Additionally, we will explore the possibility to further regress the common reduced basis not only from the	for biomechanical or haemodynamic simulations leading to potentially stratify the disease • Also integrating statistical models of heart fibres based on diffusion tensor images into comprehensive anatomical models		
		WP7 Genetic and metagenomic analytics To evaluate the role of genetic (assessed by disease-gene or candidate gene analysis) and metagenome (based on gut microbiota profiling) profiles on the development and progress of diseases and on their outcome.		
		WP12: Models validation, outcome analysis and clinical workflows • To clinically validate derived models • To improve prediction of outcome and risk stratification • To establish integrated clinical workflows and personalised treatment models		

	flow but also from additional models or clinical variables, in order to obtain disease/patient-specific reduced flow bases. **6. Statistical shape analysis** In addition to biomechanical or haemodynamic simulations, statistical shape analysis has shown its potential to assess the severity of a disease and predict its evolution. MD-Paedigree aims to use recent advances in deformation-based shape modelling to model the evolution of the heart over time and to potentially stratify the disease. Thanks to an underlying 3D deformation model, such methods can seamlessly integrate not only the shape but also spatial variables such as physical and physiological parameters, flow patterns, etc. **7. Diffeomorphic registration** Recent advances in diffeomorphic registration have shown the feasibility of extracting a sparse multiscale representation of deformations in registration. By regressing these sparse deformation parameters along with the main model parameters with respect to the standard clinical variables, one can create simplified models that are easy to fit to the patient data and provide a clear visual and objective assessment of cardiomyopathies. This information will be integrated with the simulation results for a comprehensive picture of the individual patient.	**WP19: Exploitation, HTA, and Medical Device Conformity** An early evaluation in the form of health technology assessment (HTA) as well as the development of exploitation strategies is essential for the creation of research related services which can prevail in today's highly competitive markets - be they "academic" and RTD markets, be they health services or commercial markets. The workplan is designed to encourage materializing improved disease understanding and therapy outcomes into both clinical routine and translational research, to deploy early prototypes within the developing VPH Infostructure, and to improve in iterative cycles of specifications, refactoring (i.e. improving the design of existing code), and deployment. **Objectives** • Evaluate the MD-Paedigree's models, workflows, and infostructure based on: o its accessibility, usability and effectiveness for the VPH community o the potential of its contributing to personalised healthcare workflows and integration with EHRs/decision support systems, thereby preparing for the transfer into clinical practice o making models and simulations readily available at the points of care and to researchers • Define effectiveness and usability within the context of sharing		

		"developing ICT tools, services and infrastructure to obtain more elaborate and reusable multi-scale models" (call text) as well as developing an appropriate analytical evaluation framework • Explore the health system and business opportunities o to market concrete project outcomes and results o to prevent diseases and contribute to the safety of care o to identify markets and cost models for the effective diffusion of our models, allowing researchers to exploit, share resources and develop new knowledge • Design business plans that prepare pre-market access and that integrate medical device conformity assessment procedures		

Application scenario

Jonathan is a 12 years old boy with Duchenne Muscular Dystrophy (DMD). At clinical evaluation the child reported no dyspnea at rest with some fatigue at mild exercise. Jonathan was free of cardiovascular treatment, heart rate was mildly increased and blood pressure was low-normal. Echocardiography provided information on cardiac geometry and chamber function. A dilated left ventricle with left ventricular hypertrophy was seen. Systolic function was low-normal and diastolic function analysis demonstrated increased left ventricular filling pressure. Additional evaluation with CMR demonstrated a frankly dilated ventricle with mildly reduced ejection fraction and mild diffuse fibrosis of the cardiac muscle. The mitral annular plane was dilated and mitral insufficiency was caused by leaflet tethering. The patient was treated according to current clinical guidelines and a follow-up clinical evaluation and echocardiogram were programmed after three months to evaluate the effect of treatment.

To date, in clinical practice, information on cardiac pathophysiology is based on scattered information derived from different diagnostic techniques. In the present case, relevant information is derived from clinical examination, personal interview, echocardiography and CMR. Clinical outcome highly depends on the physician's experience and ability to manage and exploit the different information sources in a time-consuming process.

MD-Paedigree provides a physician with a robust, multi-scale 4D anatomical, hemodynamic and electromechanical model, in order to integrate all available clinical and diagnostic data. Integrated information on cardiac geometry and volumes (obtained from CMR) is merged to functional information obtained from echocardiography (including filling pressure and cardiac synchrony) and hemodynamic data obtained from clinical examination. Beyond data integration, electromechanical and haemodynamic models of the heart give the possibility to understand the mechanism of muscle dysfunction by integrating information on muscle fibrosis and systolic mechanics and predict the impact of therapy in reducing mitral regurgitation, filling pressure and thus relieve symptoms. Treatment is personalised and tailored to robustly-modelled cardiac morphology and function, integrating all available information on heart geometry, ejection function, heart relaxation, ventricular inter-dependence, valve function and cardiac workload. Prediction of response to drugs helps a physician in prescribing the most effective treatment at the first evidence of cardiac disease. MD-Paedigree's models reduce the timeframe from evidence of disease to optimal medical treatment, thus significantly improving patients' morbidity and mortality

WP2: Clinical and technical user requirements for disease modelling

Tasks	Lead	Deliverables	Deadline
Task 2.1: Conduct interviews with the clinical and technical partners to obtain a complete list of requirements for the disease modelling that will ensure its usefulness within and beyond the project. All WP Leaders will actively contribute to the requirements documentation while they ensure that the respective WP partners are interviewed. 1. Prioritisation criteria: All requirements will be prioritised ensuring that from the start the most important aspects will be implemented to quickly ensure an operational system. 2. Schedule of requirements updating: The requirements list will be continuously updated on a regular basis such that main requirements and system constraints will be released as deliverables.	CHINALI **Estimated % realisation** M3 M6 M9 M12	**D2.1 Initial requirements analysis document including priorities for the implementation.** Initial requirements analysis document including priorities for the implementation: Complete interviews with the clinical and technical partners will be collected to obtain a list of variables and requirements for the disease modelling. Requirements will be prioritized ensuring that from the start the most important aspects will be implemented first.	Month 12 **Lead** CHINALI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP3: Data acquisition and processing for Cardiomyopathies

Tasks	Lead	Deliverables	Deadline
T3.1 Informed Consent & Data Collection Protocol • A 33-month longitudinal study will evaluate predictors of cardiac failure in 180 children with CMD • Clinical parameters including age, gender, reported time from onset of disease, anthropometrics, blood pressure and heart rate, biochemical parameters including NT-proBNP, white blood cell count and markers of low-grade inflammation will be collected. • Cardiac imaging will be used to derive data on cardiac structure, geometry and both systolic and diastolic function as well as cardiac fibrosis, inflammation and infiltration. • Echocardiography will be used to derive advanced measures of cardiac function including diastolic filling physiology, systolic regional strain, papillary muscle function and interventricular dependency. • Three-dimensional echocardiography will be also used to evaluate mitral valve shape and function as well as systolic synchronicity. • Parameters of function will also be merged to parameters on cardiac workload and vascular stiffness. Cardiac MRI will be used to evaluate cardiac volume and mass as well as myocardial inflammation, infiltration, and fibrosis. • Clinical evidence of overt heart failure (defined by hospitalization or reduction in cardiac functional class below or equal to NYHA II) and/or a reduction in ejection fraction by over 10% points, will be considered the endpoint. Patients will be evaluated at the baseline (month 4 to month 20) and re-evaluated between month 21 and month 36.	RINELLI **Estimated % realisation** M3 M6 M9 M12	**D3.1 Form of Informed consent and study protocol for DCM** Approval by the local Ethical Committees. Form of Informed consent and study protocol for DCM: approval by the local Ethical Committees: Study protocol including form of informed consent will be delivered for approval by participating centers' Ethical Committees.	Month 3 **Lead** RINELLI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T3.2 Clinical data & Routine laboratory test data collection **Clinical data:** • personal • history • physical activity • perceived physical functional class • socio-economic status **Anthropometrics:** • Height • weight • BMI • BSA **Routine laboratory tests:** • evaluation of white blood cell count • standard lipid and glucose metabolism • indices of renal function (including serum creatinine, BUN and electrolytes). • Circulating markers of inflammation (C-reactive protein, CRP; Tumor-Necrosis Factor-, TNF-; Interleukin 6 and 10, IL6 and IL10). Systolic (SBP) and diastolic blood pressure (DBP)	RINELLI **Estimated % realisation** M3 M6 M9 M12	**D3.2 Enrolment of 180 DCM patients.** Enrolment of 180 DCM patients: Enrolment of 180 patients, at baseline, with clinical, laboratory and diagnostic tool analysis will be performed from month 4 to month 20 including echocardiographic, MRI and exercise test parameters **D3.3 Re-evaluation of all patients** All 180 patients enrolled during D3.2 will be re-evaluated at follow up (month 21 to 36) to evaluate changes in clinical, laboratory and cardiac geometry and functional parameters	Month 20 **Lead** RINELLI **1st draft ready by:** Month 36 **Lead** CHINALI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Deadline**	
T3.3: Estimation of functional class and cardiopulmonary tests • **Six Minute Walk Test (6MWT)** • **Cardiopulmonary test (CPX)**	CHINALI	
T3.4: Imaging Acquisition and data processing **Echocardiography and MRI**. • **Protocol:** Echocardiograms will be performed by expert sonographers with fully equipped echocardiography machines with tissue Doppler, speckle tracking and 3D capabilities. • Left ventricular (LV) internal dimension, septal and posterior wall thickness will be measured at end-diastole and end-systole following the American Society of Echocardiography recommendations on three cycles. • A necropsy-validated formula will be used to calculate LV mass, which will be normalised for body height in meters to the allometric power of 2.7, in order to linearize the relation between LV mass and height (i.e. body 0growth). • To evaluate the concentricity of LV geometry, myocardial thickness (wall + septum) will be divided by LV minor axis (diameter) to generate a relative wall thickness (RWT). • Traditional indices of LV systolic performance will assess: LV ejection fraction, and LV shortening measured at the midwall level (midwall shortening). • Stroke volume will be determined by and used to calculate cardiac output. LV diastolic properties will be assessed by Doppler interrogation of	**Estimated % realisation**	
	M3	
	M6	
	M9	
	M12	

transmitral peak early (E) and late (A) velocities and by measurement of the deceleration time of peak E velocity.

- Transmitral flow velocities will be merged with mitral annualr velocity (e') derived from tissue Doppler interrogation to derive the E/e' ratio.
- Advanced indices of cardiac geometry and function will include 3D evaluation of LV mass and LV volumes and LA volumes. LV volumes derived from 3DV examination will be used to derive 3D stroke volume, cardiac output and ejection fraction.
- Three dimensional Time-to-minimal-systolic-volume (TMSV) from all 16 LV segments will be recorded and used to calculate cardiac the systolic dyssynchrony index (3D-SDI).
- Two-dimensional images from both parasternal short axis view and apical views will be analysed offline to derive parameters of LV and left atrial speckle strain.
- Speckle tracking strain will be achieved through the combination of speckle tracking, mitral annulus motion, tissue-blood border detection, and the periodicity of the cardiac cycle using R-R intervals. Longitudinal cardiac strain, radial cardiac strain and circumferential cardiac rate, will be performed together with strain rate, cardiac rotation and velocity analysis on the ventricle of echocardiographic images.
- CMR late-enhancement sequences will be performed in all patients.
- Black-blood fast spin-echo MR images will be used for the morphologic assessment of the heart with high spatial resolution and T2-weighted MR images for the evaluation of the acute myocardial edema.
- Flow mapping technique will allow assessing qualitatively and quantitatively flow volumes, velocities, and flow fractions in any oblique cardiac plane of any valvular heart disease and calculation of the stroke volumes from aortic and pulmonary arteries.
- Short-axis sections will be analyzed for measurements of end diastolic and systolic volumes and cardiac mass.
- Black-blood fast spin-echo MR images will also be obtained for the morphologic assessment of the heart and T2-weighted MR images for the evaluation of the acute myocardial edema.
- Late-gadolinium-enhanced images will show the difference between viable and nonviable myocardium with the overall and predominantly spatial distribution of the enhancement (subepicardial, midwall, or subendocardial).

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP8: Modelling and simulation for Cardiomyopathies

Tasks	Lead		Deliverables	Deadline
T8.1: Personalised anatomical and structural heart modelling • anatomical modelling algorithms developed in Sim-e-Child to extract the whole heart from • dynamic imaging data are enhanced to yield robust results on MRI and echo images. • the focus is on estimating these models from "sparse" imaging data such as 2D + t data typically acquired in clinical routine. • a cardiac fibre atlas developed by SCR is registered to the patient datasets. For this purpose, a multi-modal, non-rigid registration algorithm is employed. All methods are validated on the database of images acquired by the clinical partners. **Partners involved: SAG, SCR, INRIA, OPBG, JHU, UCL.**	SUEHLING **Estimated % realisation** M3 M6 M9 M12		D8.1 Personalised anatomical and structural modelling report This report will present the technical advances made in the first 14 months of T8.1. While T8.1 continues with new data until the end of the project, this first period will be the key for the technical feasibility.	Month 18 **Lead** SUEHLING **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead		Deliverables	Deadline
T8.2: Electrophysiological and biomechanical modelling and simulation • the subject-specific models created in T8.1 are used to drive electrophysiological and biomechanical simulations with a simplified blood pressure field. • On a volumetric mesh of the myocardium, a model of cardiac electrophysiology is defined and then personalised using the patient-specific electrophysiology data from T8.1. In addition, the parameters (passive stiffness, regional contractility) of a coupled biomechanical model of the heart are estimated based on the cardiac motion extracted from time series of images. **Partners involved: INRIA, SCR, SAG, OPBG, JHU, UCL.**	PENNEC **Estimated % realisation**		**D8.2 Electrophysiological and biomechanical simulation report** This report will present the technical advances made in the first 20 months of T8.2, which will lay the groundwork for all further processing.	Month 24 **Lead** SUEHLING **1st draft ready by:**
	M3			
	M6			
	M9 M12			

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead		Deliverables	Deadline
T8.3: Hemodynamic modelling and simulation • the haemodynamics blood flow simulation developed in Sim-e-Child is enhanced and adapted for use with cardiomyopathies. • Blood flow velocity fields from MRI are used to define boundary conditions for in- and outflow. • As in T8.2, 0-D and quasi 1-D models of arterial circulation will be employed to simulate the cardiovascular system. • Atlas-based techniques for model reduction are explored to decrease computational complexity for use in clinical routine. Results of the patient-specific simulation of blood flow in the whole heart are validated through qualitative and quantitative comparison with new imaging technologies including 4D PC MRI and 3D Doppler ultrasound. **Partners involved: SCR, INRIA, OPBG, JHU, UCL.**	MANSI **Estimated % realisation** M3 M6 M9 M12		**D8.3 Haemodynamics simulation report** This report will present the technical advances made in the first 20 months of T8.2, which will lay the groundwork for all further processing.	Month 30 **Lead** SUEHLING **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead		Deliverables	Deadline
T8.4: Whole-heart coupled Fluid-Structure-Interaction simulation • cardiac physiology simulation is enhanced by coupling the personalised hemodynamic simulation from T8.4 with the personalised electro-mechanical model from T8.3. • The personalization of the integrated cardiac model is expected to be facilitated by the initial personalization of both electro-mechanical and haemodynamics models. As for T8.2, the FSI model will be validated by simulating therapies, like ventricular assist device, and comparing the predicted outcome with the real one.. **Partners involved: SCR, INRIA, SAG OPBG, JHU, UCL.**	MANSI **Estimated % realisation**		**D8.4 Whole heart, coupled FSI simulation report** This report will present the progress made in the first 14 months of the project in T8.4 and will include a validation of the new FSI model on clinical data.	Month 36 **Lead** SUEHLING **1st draft ready by:**
	M3			
	M6			
	M9 M12			

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead		Deliverables	Deadline
T8.5: Statistical shape, flow and physiological properties modelling and analysis • analyze statistically all the spatially distributed parameters of the previous models in order to extract reduced models which will allow identifying the important disease parameters and modelling the relations between these variables. The statistical shape models developed in Health-e-Child will be extended to allow for geometric sparsity in shape variations and to include other spatial parameters such as fiber orientation, mechanical and electrophysiological parameters, and potentially flow pattern. **Partners involved: INRIA, OPBG, JHU, UCL.**	PENNEC **Estimated % realisation** M3 M6 M9 M12		**D8.5 Statistical shape, flow and physiological properties modelling report** This report will present the results of T8.5, which will include the evaluation of reduced models by means of regression studies.	Month 48 **Lead** SUEHLING **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

<table>
<tr><th colspan="5">WP7 Genetic and metagenomic analytics</th></tr>
<tr><th>Tasks</th><th colspan="2">Lead</th><th>Deliverables</th><th>Deadline</th></tr>
<tr><td rowspan="4">T7.1. Informed consent and data collection protocol.
Informed consent forms and dedicated testing protocols will be prepared for sample collection, DNA extraction and analysis.

T7.2. Sample collection, storage and DNA extraction.
Samples will be collected from 180 patients for cardiology, 200 for rheumatology, 180 for cardiovascular risk in obesity, and from a control group of 100 unaffected subjects
Fecal samples will be stored at -80 °C until further processing and DNA will be extracted according to published methods (Zoetendal, E.G., Heilig, H.G., Klaassens, E.S., Booijink, C.C., Kleerebezem, M., Smidt, H., de Vos, W.M., 2006. Isolation of DNA from bacterial samples of the human gastrointestinal tract. Nat. Protoc. 1, 870–873; Salonen A, Nikilä J, Jalanka-Tuovinen J, Immonen O, Rajilić-Stojanović M, Kekkonen RA, Palva A, de Vos WM., 2010. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. J Microbiol Methods 81(2):127-34), with slight modifications.</td><td colspan="2">BABAN</td><td rowspan="4">D7.1 Recruitment protocol with ethical clearance
Completion of the recruitment protocol, consensus and ethical clearance from all partners' involved in patient recruitment.</td><td>Month 3</td></tr>
<tr><td colspan="2">Estimated % realisation</td><td>Lead</td></tr>
<tr><td>M3</td><td></td><td>BABAN</td></tr>
<tr><td>M6

M9
M12</td><td></td><td>1st draft ready by:</td></tr>
<tr><td rowspan="3">D7.2.1 First report on data collection process
Report on data collection progress, inclusive of analysis of patient data on the basis of inclusion/exclusion criteria and updating of clinical features.</td><td>month 18</td></tr>
<tr><td>Lead:</td></tr>
<tr><td>1st draft ready by:</td></tr>
<tr><td rowspan="3">D7.2.2 Second report on data collection process
Report on data collection progress, inclusive of analysis of patient data on the basis of inclusion/exclusion criteria and updating of clinical features.</td><td>month 36</td></tr>
<tr><td>Lead:</td></tr>
<tr><td>1st draft ready by:</td></tr>
<tr><td colspan="5">Self-Assessment criteria</td></tr>
<tr><td colspan="2">Measurement process and units:</td><td colspan="3">Indicators [Upper and lower limits associated with WP objectives and measurement units]</td></tr>
</table>

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T7.2.1 Cardiology Data from patients with Dilated Cardiomyopathy (DCMP) will be collected from OPBG, UCL and JHU. Genetic testing of disease-genes will be carried out following exclusion of secondary acquired causes of DCMP. Apparently isolated DCMP patients will be clinically evaluated by a trained neurologist, and, when indicated, specifically tested to exclude systemic neuromuscular disorders, as Duchenne and Becker muscular dystrophies and Barth syndrome. After exclusion of secondary causes and multisystemic disorders, the majority of the patients are expected to remain genetically uncharacterized. In fact, only less than the 20% of these patients are affected by familial DCMP and are heterozygous for a mutation in a cardiac sarcomere gene. Next generation sequencing (Genome Wide Analysis – GWA) will be used for searching the underlying genetic background in selected unresolved cases. Clinical assessment will include: family history based on three generations, with specific enquiry about heart failure, sudden death, conduction disorders, stroke, muscular dystrophy and related anomalies, sensorineural deafness, muscle weakness; parental cardiovascular assessment, evaluation of muscle bulk and joint contractures for ruling out multisystemic muscular dystrophies.	BABAN **Estimated % realisation** M3 M6 M9 M12	**D7.3.1 First report on sample storage, DNA extraction and sample analysis processes** First report on DNA extraction and analysis process, inclusive of metagenoma analysis, Operational Taxonomic Unit (OTUs) identification, and phylogenetic analysis of gut microbiota samples.	Month 18 **Lead** BABAN **1st draft ready by:**
		D7.3.2 Second report on sample storage, DNA extraction and sample analysis processes: Second report on DNA extraction and analysis process, inclusive of metagenoma analysis, Operational Taxonomic Unit (OTUs) identification, and phylogenetic analysis of gut microbiota samples.	month 36 **Lead:** BABAN **1st draft ready by:**
		D7.4 Report on integration in the Infostructure Report on the integration of all genetic and meta-genomic input into MD-Paedigree's model-driven infostructure	month 36 **Lead:** BABAN **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP12: Models validation, outcome analysis and clinical workflows			
Tasks	**Lead**	**Deliverables**	**Deadline**
T12.1.1 Clinical assessment of cardiomyopathy models. • As detailed in WP3 and WP8, a series of 180 children with CMD will be analyzed in order to provide clinical and cardiac structural, geometrical and functional data to build the heart model. • The clinical assessment of the heart model, will be an ongoing process which will benefit from the use of the digital repository and the contribution of clinicians and researchers at the point-of-care. • Providers will compare the developed model to the observed patients with regard to cardiac dimensions, parameters of cardiac systolic and diastolic function, hemodynamic variables as well as to clinical and biochemical characteristics. • Data will be acquired and will improve the ability of the model to represent the complete cardiovascular setting of the patient, to predict the progression of the disease and the development of overt cardiac failure and to foresee the effect of personalized treatment strategies. • To further evaluate the predictive power of the multi-physics model developed in WP8 and their clinical use, post-treatment data will be acquired such as after ventricular assist device implant. Model prediction will then be compared with the real outcome. • "In summary, the clinical assessment of the model will result in: a) maximal accuracy of the model, b) identification of strongest markers of outcome prediction c) insights into personalised treatment models.	PONGIGLIONE **Estimated % realisation**	D12.1) Outline of the clinical assessment and validation criteria for all four disease areas: Preliminary analysis of the clinical assessment and validation criteria	Month 18 **Lead:** PONGIGLIONE **1st draft ready by:**
	M3	D12.2.1) First clinical assessment and validation results for all four disease areas: Periodic update at month 24 of clinical assessment and validation outcomes	Month 24 **Lead:**PONGIGLIONE **1st draft ready by:**
	M6		
	M9		
	M12	D12.2.2) Second clinical assessment and validation results for all four disease areas: Periodic update at month 36 of clinical assessment and validation outcomes	Month 36 **Lead:** PONGIGLIONE **1st draft ready by:**
		D12.2.3) Third clinical assessment and validation results for all four disease areas: Periodic update at month 48 of clinical assessment and validation outcomes	Month 48 **Lead:** PONGIGLIONE **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

<table>
<tr><th>Tasks</th><th>Lead</th><th>Deliverables</th><th>Deadline</th></tr>
<tr><td rowspan="6">T12.2.1 Clinical workflows for cardiomyopathy.
The clinical workflow for cardiomyopathy will describe the sequence of operations that start with clinical data acquisition and by using our models ends with a clinically useful diagnostic index and treatment strategy.
The clinical workflow will be subdivided into 4 specific steps: a) acquisition of clinical, structural and functional information, b) integration of all information into a single model, c) similarity search through the digital repository, and d) personalised prediction of disease outcome and optimization of individualized therapy. At the point of care clinical information will be obtained from interview, clinical evaluation and laboratory assessment. Imaging analysis will include either ultrasound or cardiac magnetic resonance imaging or both in order to provide all or part of the needed information on cardiac structure and function. The integration of the gathered information in a cardiac model will provide the researchers and clinicians with patient-specific representation of the heart.
Through the multi-physics modelling, the digital repository will provide the tools for individual and personalized progression of disease prediction, impact of outcome markers and predict the effect of personalised treatment. This will provide optimization of therapy, and thus a complete newly-defined workflow for personalised predictive and clinical medicine.</td><td>PONGIGLIONE</td><td rowspan="6">D12.3) Improved clinical workflows and outcome analysis: Final proposal of innovative clinical workflows based on outcome analysis of all patient cases</td><td>Month 48</td></tr>
<tr><td>Estimated % realisation</td><td>Lead</td></tr>
<tr><td>M3</td><td rowspan="4">PONGIGLIONE

1st draft ready by:</td></tr>
<tr><td>M6</td></tr>
<tr><td>M9</td></tr>
<tr><td>M12</td></tr>
<tr><td colspan="4">Self-Assessment criteria</td></tr>
<tr><td>Measurement process and units:</td><td colspan="3">Indicators [Upper and lower limits associated with WP objectives and measurement units]</td></tr>
<tr><td></td><td>Upper limits (result’s maximum expectation) :</td><td colspan="2">Lower limits (below which result not acceptable):</td></tr>
<tr><td colspan="4">Quality assurance - 1st content check entrusted to:</td></tr>
</table>

WP18: Dissemination & Training

Tasks	Lead		Deliverables	Deadline
18.3 Training Training is considered to be a fundamental task in dissemination. As anecdotal evidence has confirmed via WP4 of the VPH NoE and via feedback from the DISCIPULUS ('Roadmap Towards the Digital Patient') meeting (30/03/2012; Barcelona), training is recognized to be one of the most solid and long-lasting dissemination strategies in place. The training activities within MD Paedigree will consist of 2 'hands-on' workshops to be delivered during years 2 and 4 of the project (at approx. 1 or 1.5 year interval) in order to expose the outcomes achieved both, in disease modelling and in building the infostructure, highlighting the potential for change management and innovation in clinical workflows to the medical/clinical and research community interested in VPH technology. The first workshop will also seek to provide feedback to the research and development activities, so as to refine the outcomes for the final workshop. The workshop participants will fill in a detailed feedback questionnaire that will be passed to the developers. This task will be led by UCL, which has a long-standing commitment with the VPH Community and is involved in several training grants, including the Marie Curie ITN 'MeDDiCA', 'VPH-MIP' and WP4 of the VPH NoE.	DIAZ **Estimated % realization** M3 M6 M9 M12		D18.3) Training event in year 2: Report on the outcomes of the first Training event **D18.6) Training event in year 4:** Report on the outcomes of the second Training event	Month 30 **Lead** DIAZ **1st draft ready by:** Month 42 **Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead		Deliverables	Deadline
T18.4 Seminars, Workshops, Concertation Activities with Other ICT Funded Projects, and Scenario Analysis Sessions The Consortium will identify the most relevant conferences in the area and propose seminars and workshops to be held during these events. It will devote special attention and resources to Concertation Activities with other ICT funded projects and to targeted dissemination actions. Special "Scenario analyses" sessions will be convened, involving the key personnel from both the clinical and the technological partners, with the aim of pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users within MD-Paedigree. The results of the previous workshops will be presented to the Scientific Committee and to the Users' Board in order to assess their relevance and applicability, so as to refine the outcomes for a validation workshop and for a final MD-Paedigree Conference, to be held at the end of the project, targeting both internal and external clinical and research communities as well as patient organisations and the interested media. The participation in any such event will be reported in the periodic reports and the final report.	DIAZ		**D18.4.1) First scenario Analysis Sessions:** First scenario Analyses pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users.	Month 24 **Lead** DIAZ **1st draft ready by:**
	Estimated % realization			
	M3			
	M6			
	M9 M12		**D18.4.2) Second scenario Analysis Sessions:** Second scenario Analyses pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users.	Month 42 **Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T18.7 Engaging Parent and Patient Associations Approaching Parent and Patient associations will become a part of the consortium's dissemination activities. The project will seek to disseminate news of its work, expected results and potential future developments through these channels. It is hoped that the work with Patient associations will help achieve a larger bidirectional knowledge sharing base of clinicians and of patients, and further inform the potential beneficiaries of the ongoing work.	DIAZ **Estimated % realization** M3 M6 M9 M12	**D18.1) Dissemination and training strategy plan and preliminary materials**: Roadmap defining the dissemination and training strategy, indicating the subsequent choice of preliminary materials	Month 12 **Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP19: Exploitation, HTA, and Medical Device Conformity

Tasks	Lead	Deliverables	Deadline
T19.1: Evaluation approach and meaningful indicator development (EMP) • Develop upon and adapt in the VPH and other contexts proven approaches, methods and tools to the specific environment and objectives of this workpackage • Establish a set of meaningful criteria and their measurement process that are robust to demonstrate socio-economic benefit-cost impacts. The focus is • to approach and find measurements for evaluating how virtual collaborations between members of the VPH communities with different expertise are facilitated and • how consequently the uptake and acceleration of model development and integration can find meaningful expression in the overall evaluation framework.	STROETMANN **Estimated % realization** M3 M6 M9 M12	**D19.1 HTA evaluation framework** It reviews proven approaches, methods, and tools which might be relevant to the specific environment and objectives of this workpackage, and establishes a set of meaningful criteria and their measurement process, thereby focusing on evaluating how virtual collaborations between members of the VPH communities with different expertise are facilitated.	Month 12 **Lead** STROETMANN **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

<table>
<tr><th>Tasks</th><th colspan="2">Lead</th><th>Deliverables</th><th>Deadline</th></tr>
<tr><td rowspan="4">T19.3: Benefit-cost scenario for clinical impact assessment (EMP)
In a separate task a high-level, generic benefit-cost scenario for clinical impact assessment will be applied, with the ultimate goal to generate economic and market evidence for true translational medicine.
The benefit-cost scenario will be tested and initially validated with preliminary, exploratory data estimates from the patient-centred workflows that are the basis of the digital repository and Infostructure.
The two main dimensions pertaining to clinical/health impacts focus on the one hand on health service delivery and the health of patients, and on the other on public health/societal outcomes.
To assess such impacts, the scenario development will integrate the following indicators:
• Clinical effectiveness and patient-related outcomes
• Safety (risks associated with applying the technology)
• Organisational and change management aspects
• Human resource implications, knowledge & education needs
• Assessing contributions to the VPH vision of a patient avatar
• Efforts for application (convenience/ease of use; costs for introduction of new technology)
The indicators assessed ultimately prepare for a more targeted and strategically aligned exploitation activities (T19.4) by proving clinical impact of MD-Paedigree with respect to:
• the state-of-the-art in paediatric patient-specific computational modelling,
• improved disease understanding and therapy outcomes that can be applied to both clinical routine and translational clinical research,
• usability by clinicians and clinical researcher,
• transferring technical workflows into clinical workflows,
• the vertical integration of multi-scale patient data and the provision of models, tools, and services readily available to clinicians at the point of care.</td><td>STROETMANN</td><td></td><td rowspan="4">D19.4 Clinical impact assessment scenario
Initial formative evaluation of MD-Paedigree model-driven Infostructure based on a benefit-cost analysis approach, subsequently followed by a generic benefit-cost scenario for clinical impact assessment developed and validated with partners and experts. [month 36]</td><td>Month 36</td></tr>
<tr><td>Estimated % realization</td><td></td><td>Lead</td></tr>
<tr><td>M3
M6
M9
M12</td><td></td><td>STROETMANN

1st draft ready by:</td></tr>
<tr><td colspan="5">Self-Assessment criteria</td></tr>
<tr><td colspan="2">Measurement process and units:</td><td colspan="3">Indicators [Upper and lower limits associated with WP objectives and measurement units]</td></tr>
</table>

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

A.1.1.1 Timing of work packages and their components

The MD-Paedigree project partners have formalized a work plan implementing 4 major phases implying a number of conceptual steps, over 48 months of activity with 4 major milestones. The first milestone is due after 9 months and marks the end of the specification phase; the following milestones are aligned with the reporting periods of the project every 12 months.

Phase 1 (running from month 1 to 9) – Project Set-up, Requirements Elicitation, and Clinical Protocols: During Phase 1 quality assurance guidelines and a self-assessment plan will be prepared, ethical approval will be obtained , and the first dissemination activities will be performed (Step 1) Furthermore, clinical protocols for the selected paediatric applications will be established (Step 2).)Finally, the requirements for models and infostructure implementation will be analysed and documented from an end user standpoint (Step 3).

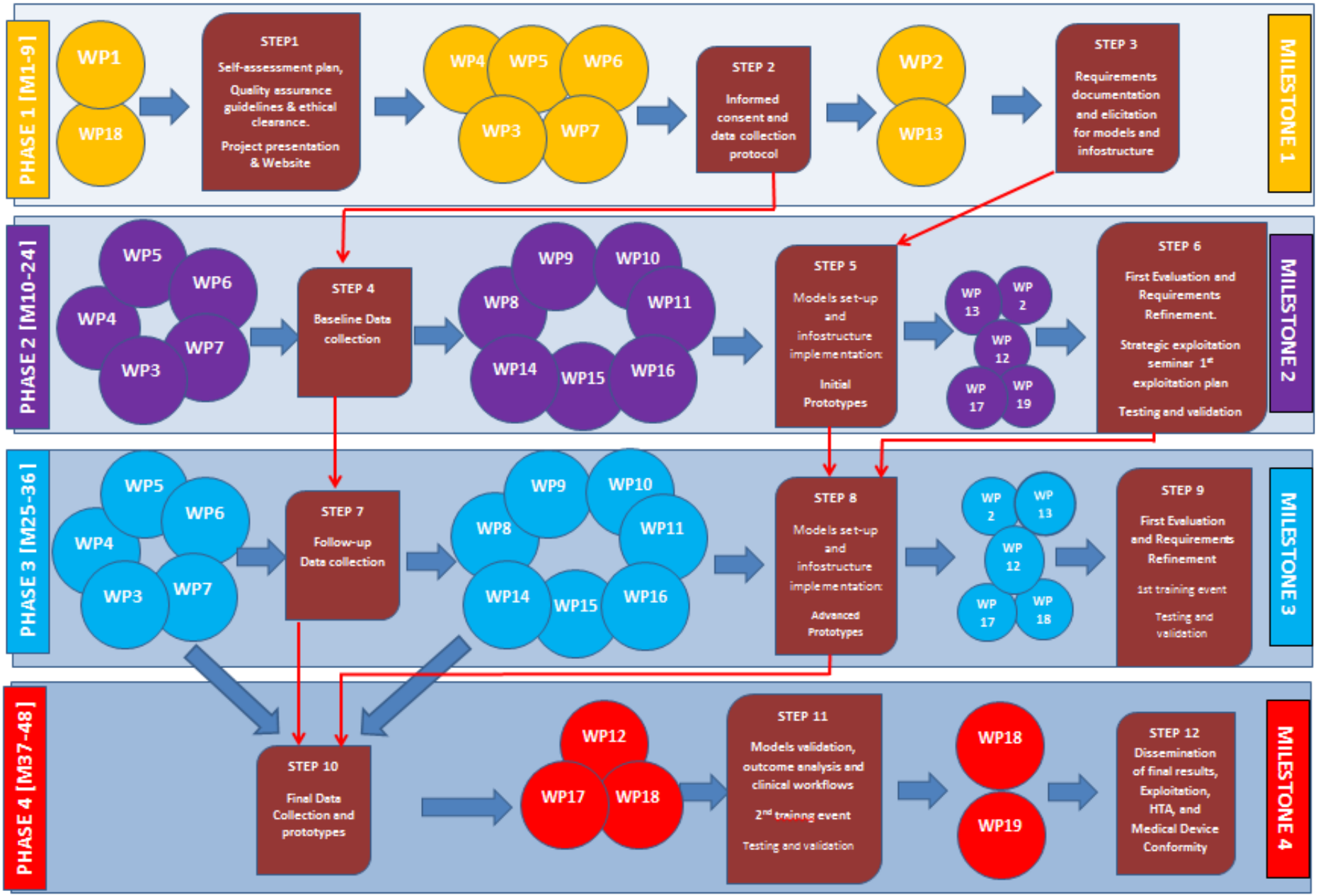
***Phase 2 (running from month 10 to 24) – Baseline Data Collection, Initial Prototypes, First Evaluation and Requirements Refinement*:** Patient enrolment will take place and data acquisition will be started (Step 4). Based on the established requirements, the existing models from Health-e Child and Sim-e-Child projects will be refined and adjusted to the new applications. The open repository for project infrastructure will be introduced and initialized with the current models and data (Step 5). First evaluations will be undertaken and requirements will be refined based on the collected experience; additionally, during this phase, the Strategic Exploitation Seminar will be held and the 1st Exploitation Plan will be drafted (Step 6).

***Phase 3 (running from month 25 to 36) – Follow-up Data Collection, Advanced Prototypes, Evaluation and Requirements Update*:** Follow-up or additional data will be acquired for all clinical applications (Step 7). The respective models will be enhanced to process longitudinal data and refined according to the obtained evaluation results. New functionalities will be integrated into advanced prototypes. The open repository will be improved and updated with content (Step 8). A second set of evaluations will be conducted and requirements will be adjusted for the final system. Furthermore, the 1st Training Event will be held (Step 9).

***Phase 4 (running from month 37 to 48) – Final Data Collection and Prototypes, Clinical Validation, and Deployment*:** In the final year, data collection will be concluded and the clinical validation will take place with the final models and simulation framework (Step 10). Results will be used to propose and disseminate improved clinical workflows. Subsequently, the 2nd Training Event will be held (Step 11). Models for all clinical applications and their respective evaluations will be documented and disseminated, while the implementation plan will be refined and the Health Technology Assessment and the Medical Clearance preparatory activities will be performed (Step 12) .

The timely delivery of all planned deliverables will be the first indicator of the fulfillment of each phase in the expected progress of MD-Paedigree, monitoring what can be demonstrable at each corresponding milestone of the project.

A second and much more detailed means of verification will be provided by the assessment criteria for each milestone and each WP which are to be defined within D1.3 Self-assessment plan on month 3.


PHASE 1 [M1-9]
WP1
WP18
STEP1
Self-assessment plan,
Quality assurance guidelines & ethical clearance.
Project presentation & Website
WP4
WP5
WP6
WP3
WP7
STEP 2
Informed consent and data collection protocol
WP2
WP13
STEP 3
Requirements documentation and elicitation for models and infostructure
MILESTONE 1
PHASE 2 [M10-24]
WP5
WP6
WP4
WP7
WP3
STEP 4
Beseline Data collection
WP9
WP10
WP8
WP11
WP14
WP15
WP16
STEP 5
Models set-up and infostructure implementation:
Initial Prototypes
WP 13
WP 2
WP 12
WP 17
WP 19
STEP 6
First Evaluation and Requirements Refinement.
Strategic exploitation seminar 1st exploitation plan
Testing and validation
MILESTONE 2
PHASE 3 [M25-36]
WP5
WP6
WP4
WP7
WP3
STEP 7
Follow-up Data collection
WP9
WP10
WP8
WP11
WP14
WP15
WP16
STEP 8
Models set-up and infostructure implementation:
Advanced Prototypes
WP 2
WP 13
WP 12
WP 17
WP 18
STEP 9
First Evaluation and Requirements Refinement
1st training event
Testing and validation
MILESTONE 3
PHASE 4 [M37-48]
STEP 10
Final Data Collection and prototypes
WP12
WP17
WP18
STEP 11
Models validation, outcome analysis and clinical workflows
2nd training event
Testing and validation
WP18
WP19
STEP 12
Dissemination of final results, Exploitation, HTA, and Medical Device Conformity
MILESTONE 4

CARDIOMYOPATHIES
1st Year Calendar

March 2013	April 2013	May 2013	June 2013	July 2013	August 2013
	Protocols delivered to Ethical Committee	D3.1 Form of Informed consent and study protocol for DCM: approval by the local Ethical Committees.	Contribution to the Self-Assessment Plan	Interviews to prepare D2.1	First Half-Yealry report
					Self-Assessment Plan
		D7.1 Recruitment protocol with ethical clearance (for genetic Studies)			Check of the enrollment and data collection, analysis and processing.
	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs
	Area Dedicated T&M TC [17th Apr]	Area Dedicated T&M TC [22nd May]	Area Dedicated T&M TC [19th Jun]	Area Dedicated T&M TC [17th Jul]	Area Dedicated T&M TC [21st Aug]

September 2013	October 2013	November 2013	December 2013	January 2014	February 2014
Biannual area meeting	Check of the enrollment and data collection, analysis and processing	First draft of the deliverable D2.1		Internal Review	First Periodic Review
					D2.1 Initial requirements analysis document including priorities for the implementation
Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs
Area Dedicated T&M TC [18th Sep]	Area Dedicated T&M TC [16th Oct]	Area Dedicated T&M TC [20th Nov]	Area Dedicated T&M TC [20th Nov]	Area Dedicated T&M TC [22nd Jan]	Area Dedicated T&M TC [19th Feb]

CARDIOMYOPATHIES	**DELIVERABLES WITHIN MONTH 24**	
	Deliverables	**Month**
	D8.2) Electrophysiological and biomechanical simulation report	M24
	D7.2.1) First report on data collection process	M18
	D12.1) Outline of the clinical assessment and validation criteria for all four disease areas	M18
	D12.2.1) First clinical assessment and validation results for all four disease areas	M24

Cardiovascular Disease Risk in Obese children and adolescents

MD-PAEDIGREE KICK OFF MEETING

CVD RISK WG

Participant's Name	Affiliation

Concept 1 (general)

MD-Paedigree validates and brings to maturity patient-specific computer-based predictive models of various paediatric diseases

- increasing their potential acceptance in the clinical and biomedical research environment
- making them readily available not only in the form of sustainable models and simulations, but also as newly-defined workflows for personalised predictive medicine at the point of care.

These tools can be accessed and used through an innovative model-driven infostructure

- powered by an established digital repository solution
- able to integrate multimodal health data
- entirely focused on paediatrics
- conceived of as a specific implementation of the VPH-Share project, planned to be fully interoperable with it and cooperating, through it, also with p-Medicine.

MD-Paedigree aims at achieving high-level semantic interoperability,

- requiring standards enabling the clinical contents to be interpreted consistently across the different EHR regimes,
- while complete clinical interoperability between systems will require widespread and dependable access to maintained collections of coherent and quality-assured semantic resources,
- including models that provide clinical context,
- mapped to interoperability standards for EHR and PHR and biomedical data,

linked to well specified terminology value sets, derived from high quality ontologies

CONCEPT (SPECIFIC)	**Beyond the state of the art**	**WPs' OBJECTIVES**	**Objectives' Lead**	**Estimated % realisation**
The precise mechanism leading to the development of cardiovascular risk in obesity from childhood to adolescence to adulthood remains largely unsolved. In particular, it is still unclear whether childhood obesity increases CVD risk simply because of the tracking of obesity from childhood to adulthood or via the	**Body mass index, Visceral adipose tissue, and Epicardial adipose tissue** To rate the degree of obesity for clinical diagnostics and studies, the body mass index (BMI) is still the primary measure, also in children. However: • BMI only estimates the general adiposity of a subject, • it does not take into account the distribution of adipose tissue within the body. • Visceral adipose tissue (VAT), the fat between the abdominal organs, has shown to correlate highly	**WP2: Clinical and technical user requirements for disease modelling** • Incorporate into the model the variables that are analysed by the clinicians in their activity. • Ensure that the modeling reflects real clinical needs and is validated against them to assure their robustness and reproducibility. • Provide computational models that can be personalized by adapting the parameters to the		

development of CVD risk factors already present in childhood and adolescence. Many structural and functional changes in the adolescent heart, such as left ventricular (LV) hypertrophy, left atrial (LA) enlargement, and subclinical impairment of LV systolic and diastolic function are believed to be precursors to more overt forms of cardiac dysfunction and heart failure. Cross-sectional studies are able to show correlation between childhood obesity and established surrogate markers for CVD, such as atherosclerosis and cardiac hypertrophy. […] In patients with obesity and/or metabolic syndrome a significantly higher prevalence of left ventricular hypertrophy and left atrial dilation paired with impairment in both systolic and diastolic function is observed. Insulin resistance (IR) is an	with CVD. • Subjects with normal BMI may still have high body fat content, which has proved to be a significant CVD risk factor for adults. • Imaging modalities such as computed tomography (CT) or magnetic resonance imaging (MRI) allow measuring specific adipose tissue types While CT and MRI are the current gold standard for adipose tissue quantification, high costs (and the radiation exposure of CT) restrict these modalities to large-scale studies, Ultrasound (US) is becoming an affordable, non-invasive alternative. Echocardiography allows to measure epicardial adipose tissue (EAT) and has emerged as a novel approach to accurately estimate VAT. Quantification of adipose tissue from image data still mostly performed manually This is a tedious and time-consuming process prone to subjective bias. For the analysis of EAT from MRI, the adipose tissue has to be measured and contoured manually, which leads to noticeable discrepancies between different observers. Measuring the thickness of EAT from US is even more challenging, which is why commonly several manual measurements are performed with electronic callipers and averaged. **Methods for semi- or completely automated image-based quantification of adiposity** The extraction of adipose tissue from MRI has been studied extensively, either for selected body regions or for whole-body scans. Since adipose tissue features high intensities in MRI, many authors use thresholding to separate it from the	integrated data of a patient case • Advance the knowledge about the selected diseases by allowing the simulation of different effects on the evolution of the disease • Predict the effect of therapy. • Ensure that MD-Paedigree models have the highest possible impact at the point of care. • Re-use of models between disease areas to leverage synergies where possible. • Existing standards for modelling and tools will be investigated. • The need for new standards will be evaluated and documented.		
		WP4: Data acquisition and processing for the estimation of CVD risk in obese children • To collect clinical, biochemical and imaging data to estimate cardiovascular risk associated with obesity in adolescents. • To identify significant predictors of increased risk as estimated by changes in arterial stiffness over the time.		
		WP9: Modelling cardiovascular risk in the obese child and adolescent The main objectives of this WP are: • Adaptation of the comprehensive heart model of WP8 to the obese heart; • Model validation; • Automated estimation of the distribution of various adipose		

established determinant in the pathogenesis of CVD; it is constantly observed in patients with hypertension, dyslipidemia and atherosclerosis. Evidence supports firmly that body fat distribution (subcutaneous, visceral, muscle and hepatic fat) modulates IR and cardiovascular risk more than total body adiposity, thus explaining why some individuals who are seemingly equally obese and share common lifestyle and dietary habits tend to have higher IR and CVD risk than others. MD-Paedigree will integrate the variety of known biomarkers for CVD risk assessment into one common framework, enhance body fat distribution biomarker measurement, and analyse interdependencies between the biomarkers. In addition, MD-Paedigree will develop computational models with high predictive power to better understand the	surrounding tissue. Although an automatic selection of thresholds has been proposed, different adipose tissue types (VAT and subcutaneous adipose tissue, SAT) still have to be separated manually. An automatic algorithm for this problem was developed, based on an active contour algorithm proposed to use morphological operations, edge detection, and knowledge-based curvature fitting. In all these approaches, bone marrow is often misclassified as adipose tissue, because it features similar intensities in MRI. Thomas et al. excluded bone marrow by user interaction, while Shen et al. eliminated the paravertebral adiposity tissue automatically. Kullberg et al. used geometrical models of the pelvis and vertebra to exclude these structures and thresholding and morphological operations to automatically separate VAT and SAT. Zhou et al. employed fuzzy c-means clustering and thresholding to quantify VAT and SAT in both water-saturated and non-water saturated MR images. **No automatic algorithms quantifying intraabdominal fat from US** While automated ultrasound segmentation is feasible for a variety of anatomical structures, it has rarely been used on adipose tissue. One of the few approaches was proposed by Ng et al. who used US radiofrequency signals from different locations and beam angles and calculated the spectrum dispersion within the image. Pixels which represent adipose tissue change faster than other areas.	tissue types from MRI and ultrasound data; • Determination of factors contributing to the risk, including metabolic and haemodynamic factors, clinical and family histories, and their interrelation; • Construction of personalised multivariate retrieval-based models for the assessment of cardiovascular risk and therapy selection support, on a selection of surrogate markers, both for cross-sectional and longitudinal studies, including predicting the absolute values and changes in the mitral E/e' ratio, the left-ventricular mass index, and AIx@75 as an indicator of early atherosclerosis; • Interpretation of the models with the purpose of better understanding of the cardiovascular dysfunction • mechanism from childhood to adolescence and adulthood, and quantitative evaluation of predictive performance with cross-validation and sensitivity analysis, and with evaluation on unseen subsequently		

mechanism of CVD development. These models will also allow the simulation of interventions to make personalised predictions for the optimal therapy.	To the best of our knowledge, there are no automatic algorithms quantifying intraabdominal fat from US. **Re-use of proven anatomical organ models to add prior knowledge to image analysis** In MD-Paedigree, we will re-use our proven anatomical organ models developed in Health-e-Child and Sim-e-Child. This will enable us to assess different adipose tissue types automatically from image data and use this information in our further analysis. We will also use established biomarkers such as blood pressure, metabolic and haemodynamic data to estimate the CVD risk. **Multivariate nonlinear models of CVD risk** Currently, most studies that analyse different factors of CVD risk employ univariate or, at best, multivariate but linear models, which represent a major limitation. Univariate models can only identify independent contributors to the risk, while they do not shed much light on the interplay between the factors. Cardiovascular risk can be modelled by multivariate machine learning models with only ten clinical variables (representing commonly acknowledged markers of CVD risk). Kurt et al. successfully modelled the risk of coronary artery disease with a multi-layer perceptron (MLP) and a comparable set of 8 clinical variables. Sumathi and Santhakumaran trained an Artificial Neural Network (ANN) on a set of 15 clinical variables and claimed to use it successfully for early diagnosis of hypertension. **Statistical and machine learning techniques** In MD-Paedigree, we will construct multivariate nonlinear	acquired cases.		
		WP7 Genetic and metagenomic analytics To evaluate the role of genetic (assessed by disease-gene or candidate gene analysis) and metagenome (based on gut microbiota profiling) profiles on the development and progress of diseases and on their outcome.		
		WP12: Models validation, outcome analysis and clinical workflows • To clinically validate derived models • To improve prediction of outcome and risk stratification • To establish integrated clinical workflows and personalised treatment models		
		WP19: Exploitation, HTA, and Medical Device Conformity An early evaluation in the form of health technology assessment (HTA) as well as the development of exploitation strategies is essential for the creation of research related services which can prevail in today's highly competitive markets - be they "academic" and RTD markets, be they health services or		

	models of CVD risk involving state-of-the-art statistical and machine learning techniques. This will not only help to build more accurate models of CVD risk, but also to better understand the mechanism of CVD development via the identification of important risk factors and understanding of their interrelation. Such personalised risk models may become a more reliable alternative or at least a useful complement to the CVD risk prediction charts of WHO, especially since these charts are available for adults only. **Case-based reasoning and discriminative distance learning** A common drawback of the existing works of multivariate modelling is that the underlying techniques like Multi-layer-Perceptron (MLP) or Artificial Neuron Networks (ANN) are basically "black box" models, i.e. the reasons for their results cannot be conveyed to their human users, which leads to low acceptance rates among clinicians. In our modelling, we will focus on case-based reasoning and discriminative distance learning instead. Since these systems base their decisions on concrete patient cases and are able to present the relevant cases (i.e. the ones utilised for decision making) to the user, they provide easy and intuitive decision support and a possibility for personalised therapy planning, based on the clinical history of retrieved similar patients. **CaseReasoner** Our work will be centred on the similarity search based decision support system HeC CaseReasoner developed in the Health-e-Child project. It features recently suggested techniques for discriminative distance learning, including learning from equivalence	commercial markets. The workplan is designed to encourage materializing improved disease understanding and therapy outcomes into both clinical routine and translational research, to deploy early prototypes within the developing VPH Infostructure, and to improve in iterative cycles of specifications, refactoring (i.e. improving the design of existing code), and deployment. **Objectives** • Evaluate the MD-Paedigree's models, workflows, and infostructure based on: o its accessibility, usability and effectiveness for the VPH community o the potential of its contributing to personalised healthcare workflows and integration with EHRs/decision support systems, thereby preparing for the transfer into clinical practice o making models and simulations readily available at the points of care and to researchers • Define effectiveness and usability within the context of sharing		

	constraints and the intrinsic random forest similarity. The basic philosophy behind the design of CaseReasoner is to provide clinicians with a flexible and interactive tool to enable operations such as data filtering and similarity search over a grid of clinical centres, and to facilitate the exploration of the resulting data sets. The major aim is to let clinicians explore and compare the patients' records, regardless of geographical location, and to visualize their place in the distribution of both the whole population of patients, as well as in the distribution of its semantic subsets. The search platform can then be used for several tasks such as case-based retrieval, support for curation and ultimately decision support. HeC CaseReasoner employs a domain-independent technology. With MD-Paedigree, HeC CaseReasoner will be further extended and applied to decision support in the domain of modelling cardiovascular risk in obese children and adolescents. **MD-Paedigree major modelling objectives** In summary, our major objectives with modelling the cardiovascular risk in the obese child and adolescent are (1) automated, objective quantification of different adipose tissue types and their distribution from MRI and ultrasound data, (2) collection of a large number of additional factors contributing to the risk, including metabolic and haemodynamic factors, clinical and family histories, and their interrelation, (3) construction of personalised multivariate retrieval-based models for the assessment of cardiovascular risk using state of-the-art machine learning techniques, both for	"developing ICT tools, services and infrastructure to obtain more elaborate and reusable multi-scale models" (call text) as well as developing an appropriate analytical evaluation framework • Explore the health system and business opportunities o to market concrete project outcomes and results o to prevent diseases and contribute to the safety of care o to identify markets and cost models for the effective diffusion of our models, allowing researchers to exploit, share resources and develop new knowledge • Design business plans that prepare pre-market access and that integrate medical device conformity assessment procedures		

	cross-sectional and longitudinal studies, (4) interpretation of the models with the purpose of better understanding the mechanism of cardiovascular dysfunction from childhood to adolescence and adulthood, and quantitative evaluation of their predictive performance with cross-validation and sensitivity analysis, and with evaluation on unseen subsequently acquired cases.			

Application Scenario

Rose, a 17 years old obese adolescent, suffers from impaired glucose tolerance, high blood pressure, and irregular menstrual periods. Her waist girth is 102 cm. Rose has undergone the fasting measurement of markers of systemic inflammation, the US and MRI evaluation of adiposity (visceral, subcutaneous and epicardic fat estimates including thickness and volume), and the evaluation of cardiac morphology and haemodynamics by echocardiography. Arterial stiffness is estimated by means of the radial applanation tonometry, and a cardiopulmonary exercise test is also performed.

Her waist girth has significantly increased over the last 18 months demonstrating central fat distribution. Her fasting glucose is elevated and her oral glucose tolerance is markedly impaired. The patient is dylipidemic, with increased levels of inflammation markers. Echocardiography demonstrates left ventricular hypertrophy with normal systolic function and impaired cardiac relaxation. The additional evaluation with CMR shows a significant amount of pericardial fat paired with the mild diffuse fibrosis of the cardiac muscle. The cardiopulmonary exercise test on a treadmill highlights reduced tolerance to physical activity, with increased oxygen consumption and an evident pathological blood pressure profile. The applanation tonometry shows reduced arterial compliance, increased wall stress and impaired endothelial function.

MD-Paedigree provides the physician with integrated information on Rose's cardiovascular structure and function, together with the quantitative assessment of fat distribution in the body, and metabolic and genetic data obtained from laboratory tests. The electromechanical model of the obese child's heart allows understanding the mechanism of cardiac muscle and vascular dysfunction by integrating related information on systemic fibrosis, inflammation and cardiovascular mechanics; it also allows prognosis of disease development and predicting the impact of selected therapies and weight loss for the specific cardiovascular function, fat distribution and exercise tolerance. Treatment is personalised and tailored to the cardiovascular and metabolic phenotypes, personal habits and life style, and integrating all available related information on anthropometrics, demographic data, cardiac geometry and function, vascular compliance, genetic and metabolic profiles. Accurate estimation of cardiovascular risk, prognosis of disease development and prediction of the success of a selected therapy, based on the clinical history of previously observed cases in the digital repository, helps a physician in selecting the most effective treatment already at the first evidence of disease.

WP2: Clinical and technical user requirements for disease modelling

Tasks	Lead	Deliverables	Deadline
Task 2.1: Conduct interviews with the clinical and technical partners to obtain a complete list of requirements for the disease modelling that will ensure its usefulness within and beyond the project. All WP Leaders will actively contribute to the requirements documentation while they ensure that the respective WP partners are interviewed. 3. Prioritisation criteria: All requirements will be prioritised ensuring that from the start the most important aspects will be implemented to quickly ensure an operational system. 4. Schedule of requirements updating: The requirements list will be continuously updated on a regular basis such that main requirements and system constraints will be released as deliverables.	CHINALI **Estimated % realisation** M3 M6 M9 M12	**D2.1** Initial requirements analysis document including priorities for the implementation. **Description of the deliverable**: Initial requirements analysis document including priorities for the implementation: Complete interviews with the clinical and technical partners will be collected to obtain a list of variables and requirements for the disease modelling. Requirements will be prioritized ensuring that from the start the most important aspects will be implemented first.	Month 12 **Lead** CHINALI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP4: Data acquisition and processing for the estimation of CVD risk in obese children

Tasks	Lead	Deliverables	Deadline
T4.1 Informed consent & data collection protocol [M 1-4] A 3-year longitudinal study will be performed in three cohorts of obese adolescents (N=180). Parents or responsible guardians will be asked for informed consent. Sixty patients (30 males) for each clinical Centre will be consecutively enrolled. Inclusion criteria will be age between 13 and 18 years; body mass index (BMI) z-score ≥1.645 SDS for age and sex according to the CDC growth charts; no previous treatment for obesity, no systemic and endocrine disease, no previous diagnosis of impaired fasting glucose (IFG, fasting plasma glucose ≤100 mg/dl), impaired glucose tolerance (IGT, 2 h glucose≤140 mg/dl) or diabetes (2 h glucose≤199 mg/dl), and no use of medication. Information will be collected by electronic health recording at baseline (months 1-18) and 18 months later (between months 19 and 36).	TAYLOR **Estimated % realisation** M3 M6 M9 M12	**D4.1) Data collection protocol and ethical clearance:** Study protocol including form of informed consent will be delivered for approval by participating centers' Ethical Committees	Month 4 **Lead** TAYLOR **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T4.2 Clinical data & Routine laboratory test data collection [M 5-40] • T4.2.1. Clinical data and personal history will be collected using standard questionnaires (i.e., history, dietary intake, physical activity and socio-economic status). • **T4.2.2 Anthropometrics: Height will be measured to the nearest 0.5 cm on a standardised height board.** o BMI will be calculated as weight (kilograms) divided by height (meters) squared. o Waist and wrist circumferences will be measured. • **T4.2.3 Routine laboratory tests will include evaluation of:** o fasting glucose, insulin, o c-peptide, lipid profile (total o and HDL cholesterol, triglycerides), o liver function tests (alanine-aminotransferase, aspartate amino transferase, o γ-glutamyl transferase), o white blood cell count; o glucose tolerance by a standard OGTT (1.75 g/kg body weight up o to a maximum of 75 g). o Glucose, insulin and c-peptide will be measured at baseline and 30, 60, 90 and 120 min. • **T4.2.4** Systolic (SBP) and diastolic blood pressure (DBP) will be measured three times while the subjects are seated, and the measurements will be averaged for the analysis.	TAYLOR **Estimated % realisation** M3 M6 M9 M12	**D4.1** Data collection protocol and ethical clearance. **Delivery date**: Month 3. **Description of the deliverable**: Data collection protocol and ethical clearance: Study protocol including form of informed consent will be delivered for approval by participating centers' Ethical Committees. **D4.2** Report on patient recruitment and data collection at baseline study **Description of the deliverable**: Report on patient recruitment and data collection at baseline study: Enrolment of 180 patients and relevant data collection by month 24	Month 4 **Lead** TAYLOR **1st draft ready by:** Month 24 **Lead** TAYLOR **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead	Deliverables	Deadline
• **T4.3 Estimation of adipokines, low-grade inflammation and insulin resistance [M 5-36]** o T4.3.1 Measurements of adipokines and markers of inflammation. Blood samples will be withdrawn to measure fasting plasma adipokines (leptin, adiponectin), circulating markers of inflammation (C-reactive protein, CRP; Tumor-Necrosis Factor-, TNF-; Interleukin 6, IL6) and endothelium dysfunction (e-Selectin, Intercellular Adhesion Molecule 1, ICAM-1). o T4.3.2 Assessment of the renin-angiotensin-aldosterone axis. Dietary sodium intake will be assessed by measuring 24 hour urinary sodium excretion. o T4.3.3 Insulin resistance will be estimated in fasting condition and after OGTT.	TAYLOR **Estimated % realisation** M3 M6 M9 M12	**D4.3) Report on patient follow-up**: Re-evaluation of all patients recruited for D.4.2 based on follow up data collection	Month 36 **Lead** TAYLOR **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

asks	Lead		Deliverables	Deadline
• **T4.4. Image acquisition, clinical annotation and data processing [M 5-40]** o T4.4.1 Image acquisition at the ultrasonography: o Thickness of visceral, subcutaneous and pericardial fat. o B-mode ultrasound of the abdomen will be obtained to measure intra-abdominal and subcutaneous fat. o M-B-mode ultrasound will be obtained to measure epicardial fat by using an echocardiography machine equipped with a 5-MHz transducer. o T4.4.2 Measurement of abdominal and epicardial fat distribution at the MRI [M 5-40]. o A T1-weighted axial 2-dimensional multislice spoiled gradient echo image stack will be centered at the L4/L5 inter-vertebral disk. o Pancreatic (PFF) and hepatic (HFF) fat fractions will be obtained using the Dixon technique o Water and fat image reconstruction from the acquired multi-echo data sets from the diaphragmatic hepatic surface to L5 vertebra will be performed by using ad hoc software (i.e., the Syngo software, Siemens healthcare, Erlangen, Germany) using a three echo two-point Dixon approach enabling voxel-wise correction of T2* decay. o Measurements of hepatic and pancreatic fat fraction will be performed throughout the liver, […] and in the head and tail of the pancreas	TAYLOR **Estimated % realisation** M3 M6 M9 M12		**D4.3) Report on patient follow-up:** Re-evaluation of all patients recruited for D.4.2 based on follow up data collection	Month 36 **Lead** TAYLOR **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead	Deliverables	Deadline
• **T4.5. Systolic and diastolic markers of cardiac dysfunction of US and CMR [M 5-40]** o Echocardiograms will be performed, and reviewed off-line by 2 independent readers using a computerized review station with ad hoc working stations. Left ventricular internal dimension, and septal and posterior wall thickness will be measured […]. o Left atrial volume will also be measured in apical 4- and 2-chamber views and indexed by body height. o Systolic and diastolic markers of cardiac dysfunction and parameters of cardiac morphology will be obtained. o Analysis of LV systolic function will include: ▪ ejection fraction, ▪ endocardial fractional shortening, and ▪ midwall fractional shortening unindexed and indexed by circumferential ▪ end-systolic stress, ▪ stroke volume, ▪ cardiac output and ▪ total peripheral resistance. o Assessment of cardiac geometry will also include relative wall thickness normalised for age (RWTn). o Indices of diastolic function will include the transmitral pulsed Doppler and pulsed Tissue Doppler […] o At the CMR, late-enhancement sequences will be used. o Black-blood fast spin-echo MR images will be used for the morphologic assessment of the heart with high spatial resolution and T2-weighted MR images for the evaluation of the acute myocardial edema. o Flow mapping technique will allow assessing qualitatively and quantitatively flow volumes, velocities, and flow fractions in any oblique cardiac plane of any valvular heart disease and calculation of the stroke volumes from aortic and pulmonary arteries.	TAYLOR **Estimated % realisation** M3 M6 M9 M12	**D4.3) Report on patient follow-up:** Re-evaluation of all patients recruited for D.4.2 based on follow up data collection	Month 36 **Lead** TAYLOR **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	**Lead**		**Deliverables**	**Deadline**
• **T4.6. Measurement of intima media thickness (IMT), arterial stiffness and pulse wave velocity (PWS) [M 5-40]** o **T4.6.1:** The carotid IMT will be measured by ultrasounds using a 14 MHz linear transducer following a standardized protocol. o The measurement is performed at the common carotid artery near the bifurcation at the far wall after a 10 min rest. o The sonographer measures four values on each side and tooks the maximum value for statistical purposes since the strongest association between the different measurements of IMT and coronary risk factors is achieved by using the maximum value of IMT. o **T4.6.2:** Arterial stiffness and PWV will be measured by using the SphygmoCor SCORPVx System (Atcor Medical, Sydney, NSW, Australia). o The average of three measures of the augmentation index (AIx) and PWV will be obtained and used in the analysis. o The device uses a validated generalised transfer function to calculate central (aortic) SBP, DBP, mean arterial pressure (MAP), pulse pressure (PP) and AIx adjusted to a heart rate of 75 bpm. o For PWV, the average of two measures of carotid to sternal notch to femoral artery distance was entered into the software. o Arterial waveforms gated to the R wave on the ECG tracing were recorded from the carotid and then femoral pulse. PWV is the difference in the carotid-to-femoral path length divided by the difference in timing from the R wave on the ECG to the foot of the pressure waveforms.	TAYLOR **Estimated % realisation** M3 M6 M9 M12		**D4.3) Report on patient follow-up:** Re-evaluation of all patients recruited for D.4.2 based on follow up data collection	Month 36 **Lead** TAYLOR **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement**

	units]	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP9: Modelling cardiovascular risk in the obese child and adolescent				
Tasks	**Lead**		**Deliverables**	**Deadline**
T9.1. Heart model adaptation to the obese heart [M 4-36] • The comprehensive multi-physics heart model developed in WP8 will be re-used and adapted to conditions of the obese heart. **Partners involved: SAG, SCR, INRIA, OPBG, JHU, UCL.**	HEIMANN **Estimated % realisation** M3 M6 M9 M12		**D9.1) Report about the adaptation of the heart model: This report will present the results achieved about the design of heart model**	**Month 18** **Lead** HEIMANN **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead	Deliverables	Deadline
T9.2. Automated assessment of body fat distribution from MRI and ultrasound data [M 4-48] In this task, computational models will be developed to automatically assess and quantify the body fat distribution from MRI and ultrasound data acquired in T4. Parameters as LV volume and mass or septal and posterior wall thickness will be extracted by applying the comprehensive heart model adapted in T9.1. In order to measure epicardial fat thickness and volume, this model will be extended by including data about epicardial fat. In addition, existing models of liver and pancreas will be adapted to assess pancreatic and hepatic fat fractions from Dixon MRI sequences. For the automated assessment of visceral and subcutaneous adipose tissue from ultrasound and MRI images of the abdominal region, specific adipose tissue models will be implemented. **Partners involved: SAG, SCR, FhG, OPBG, JHU, UCL.**	HEIMANN **Estimated % realisation** M3 M6 M9 M12	**D9.2) Report about automated assessment of body fat distribution from MRI and ultrasound data:** This report will present the results achieved in the first 14 months of T9.1. T9.1 continues then with model refining and validation on new data until month 36, however this initial period will lay the groundwork for modelling the obese heart.	**Month 24** **Lead** HEIMANN **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead		Deliverables	Deadline
T9.3. Multi-scale data integration and virtual phenotype generation [M 12-48] In this task, all collected clinical data and parameters of related models, such as age, gender, body weight and BMI, clinical and family history, the comprehensive heart model from T9.1, including cardiac geometry, cardiac function, and cardiovascular haemodynamics, fat distribution model from T9.2, endothelial function, metabolic profile, blood pressure profile, cardiopulmonary exercise test, applanation tonometry, etc., are integrated into a single digital repository. A virtual phenotype is built by integration of related heterogeneous data. Integrated patient records are analyzed in order to identify important factors contributing to the CVD risk, most significant associations between clinical variables, and clusters of subjects with similar phenotype and similar CVD risk development. **Partners involved: SAG, INRIA, UoA, OPBG, JHU, UCL, FhG.**	HEIMANN **Estimated % realisation** M3 M6 M9 M12		**D9.3) Report on integrated digital repository, important CVD risk factors and interesting associations:** This report will present the technical advances made in the first 24 months of the project in T9.3, which will constitute the basis for predictive disease and therapy modelling and risk stratification.	**Month 36** **Lead** HEIMANN **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T9.4. Cardiovascular risk stratification and predictive disease and therapy modelling [M 12-48] In this task, predictive models are constructed for cross-sectional and longitudinal data collected. For cross-sectional data, absolute values of known CVD markers are modelled, including insulin sensitivity, left ventricular and left atrial geometry, LV diastolic and systolic function, mitral E/e' ratio and endothelium dysfunction (early atherosclerosis, AIx@75). For longitudinal data, changes in the same CVD markers over the reported period are modelled. Discriminative distance function models are constructed for most important CVD markers. Important risk factors identified in T9.3 are used as input variables for model construction. Similarity-search based DSS CaseReasoner from Health-e-Child is extended and re-used for case-based reasoning, similarity search of patients for decision support with CVD risk prediction and therapy planning, and clustering of more homogenous patient sub-groups. The cross-sectional and longitudinal models are interpreted with the purpose of better understanding of the mechanism of cardiovascular dysfunction via variable importance analysis and finding clinically interesting interrelations. Image registration techniques and the models obtained in T9.2 will be used for longitudinal studies of abdominal adiposity distribution. Then the models are evaluated quantitatively for their predictive performance with cross-validation and sensitivity analysis, and with application to unseen subsequently acquired cases. **Partners involved: SAG, INRIA, UoA, OPBG, JHU, UCL, FhG.**	HEIMANN **Estimated % realisation** M3 M6 M9 M12	**D9.4) Report on predictive risk models and their quantitative evaluation:** This report will present the results achieved in T9.4 regarding cardiovascular risk modelling and will include evaluation of the models on collected clinical data.	**Month 48** **Lead** HEIMANN **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	Indicators [Upper and lower limits associated with WP objectives and measurement units]	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP7 Genetic and metagenomic analytics			
Tasks	**Lead**	**Deliverables**	**Deadline**
T7.1. Informed consent and data collection protocol. Informed consent forms and dedicated testing protocols will be prepared for sample collection, DNA extraction and analysis. **T7.2. Sample collection, storage and DNA extraction.** Samples will be collected from 180 patients for cardiology, 200 for rheumatology, 180 for cardiovascular risk in obesity, and from a control group of 100 unaffected subjects Fecal samples will be stored at -80 °C until further processing and DNA will be extracted according to published methods (Zoetendal, E.G., Heilig, H.G., Klaassens, E.S., Booijink, C.C., Kleerebezem, M., Smidt, H., de Vos, W.M., 2006. Isolation of DNA from bacterial samples of the human gastrointestinal tract. Nat. Protoc. 1, 870–873; Salonen A, Nikilä J, Jalanka-Tuovinen J, Immonen O, Rajilić-Stojanović M, Kekkonen RA, Palva A, de Vos WM., 2010. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. J Microbiol Methods 81(2):127-34), with slight modifications.	OPBG **Estimated % realisation** M3 M6 M9 M12	D7.1) Recruitment protocol with ethical clearance: Completion of the recruitment protocol, consensus and ethical clearance from all partners' involved in patient recruitment.	Month 3 **Lead** OPBG **1st draft ready by:**
		D7.2.1) First report on data collection process: Report on data collection progress, inclusive of analysis of patient data on the basis of inclusion/exclusion criteria and updating of clinical features.	month 18 **Lead:** **1st draft ready by:**
		D7.2.2) Second report on data collection process: Report on data collection progress, inclusive of analysis of patient data on the basis of inclusion/exclusion criteria and updating of clinical features.	month 36 **Lead:** **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and**

	measurement units]	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T7.2.3 Cardiovascular risk in obesity Genetic analysis will be performed on blood samples withdrawn at baseline, in order to build a genetic score of cardiovascular disease (CVD) risk. Metagenome data analysis will be carried out on fecal samples from obese patients collected at baseline and at 18 months, and re-evaluated at the follow-up to investigate the risk to develop CVD associated with specific taxa at the baseline.	MANCO **Estimated % realisation** M3 M6 M9 M12	D7.3.1) First report on sample storage, DNA extraction and sample analysis processes: First report on DNA extraction and analysis process, inclusive of metagenoma analysis, Operational Taxonomic Unit (OTUs) identification, and phylogenetic analysis of gut microbiota samples. [month 18] D7.3.2) Second report on sample storage, DNA extraction and sample analysis processes: Second report on DNA extraction and analysis process, inclusive of metagenoma analysis, Operational Taxonomic Unit (OTUs) identification, and phylogenetic analysis of gut microbiota samples.	Month 18 **Lead** MANCO **1st draft ready by:** Month 36 **Lead** MANCO **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead		Deliverables	Deadline
T7.3 DNA analysis. **T7.3.3 Cardiovascular risk in obesity:** Candidate Single Nucleotide Polymorphisms (SNPs) for estimation of CVD risk in the MD-Paedigree study. DNA analysis. Analysis (DNA extraction and SNPs analysis) of a custom of SNPs (listed in Table 1) in 180 patients plus the statistical analysis in order to build a genetic score of CVD risk. SNPs will be selected among SNPs identified in previous Genome Wide Association (GWAS studies). Selection will be based on either statistical significance threshold of the genetic association with the investigated variable (dyslipidemia, left ventricular hypertrophy, hypertension, type 2 diabetes, increased visceral adiposity and fatty liver) or clinical significance. Two genetic risk scores will be constructed on an a priori basis. Genetic risk scores will be the sum of all cardiovascular risk alleles from all SNPs, both those associated with CVD (increased stiffness/IMT) and those associated with risk factors as done previously (Raynter NP; JAMA 2010; 303: 631-7; Peterson RE, Hum Genet 2011; 129: 221-30). The SNPs affecting more than one phenotype will be included once. T7.3.3.1. Gut microbiota analysis. This study will be performed with pyrosequencing using 454 Junior apparatus as described in T7.3.2. Data will be analysed to seek for any change in the ratio between Firmicutes and Bacteroidetes during the observation period and development of CVD, estimated on the basis of increased arterial stiffness, and for significant relationship among phila representation and cardiovascular risk markers, in particular in respect to increased lipopolysaccharide blood concentration .	MANCO **Estimated % realisation**		D7.4) Report on integration in the Infostructure: Report on the integration of all genetic and meta-genomic input into MD-Paedigree's model-driven infostructure	Month 36 **Lead** MANCO **1st draft ready by:**
	M3			
	M6			
	M9			
	M12			

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP12: Models validation, outcome analysis and clinical workflows

Tasks	Lead	Deliverables	Deadline
T12.1.2 Clinical assessment of obesity models. • As detailed in WP4 and WP9, estimation of the cardiovascular risk associated with obesity in 180 adolescents computational models will be analyzed and used to automatically assess and quantify the body fat distribution, including epicardial fat, from MRI and ultrasound data acquired". • The clinical assessment of the obesity model, will be an ongoing process which will benefit from the use of the digital repository and the contribution of clinicians and researchers at the point-of-care. • Providers will compare the developed model to the observed patients with regard to body habit, clinical and biochemical characteristics (including metabolism and inflammation) as well as cardiac and vascular phenotypes. • Data will be acquired and will improve the ability of the model to predict the complete cardiovascular setting of the patient, to predict the progression of the disease and to foresee the effect of personalised treatment strategies. • In summary, the clinical assessment of the model will result in: a) maximal accuracy of the model, b) identification of strongest markers of outcome prediction and c) insights into personalised treatment models.	PONGIGLIONE **Estimated % realisation**	D12.1) Outline of the clinical assessment and validation criteria for all four disease areas: Preliminary analysis of the clinical assessment and validation criteria	Month 18 **Lead:** PONGIGLIONE **1st draft ready by:**
	M3 M6 M9	D12.2.1) First clinical assessment and validation results for all four disease areas: Periodic update at month 24 of clinical assessment and validation outcomes	Month 24 **Lead:**PONGIGLIONE **1st draft ready by:**
		D12.2.2) Second clinical assessment and validation results for all four disease areas: Periodic update at month 36 of clinical assessment and validation outcomes	Month 36 **Lead:** PONGIGLIONE **1st draft ready by:**
		D12.2.3) Third clinical assessment and validation results for all four disease areas: Periodic update at month 48 of clinical assessment and validation outcomes	Month 48 **Lead:** PONGIGLIONE **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**		**Deliverables**	**Deadline**
T12.2.2 Clinical workflows for CVR in obese children. • The clinical workflow for obese children will describe the sequence of operations that start with clinical data acquisition and by using our models ends with a clinically useful diagnostic index and treatment strategy. • The clinical workflow will be subdivided into 4 specific steps: • a) acquisition of clinical, structural and functional information, • b) integration of all information into a single model, • c) similarity search through the digital repository, and • d) personalised prediction of disease outcome and optimization of individualized therapy. At the point of care clinical information will be obtained from interview, clinical evaluation and laboratory assessment while and imaging analysis include both data on fat distribution and on cardiovascular imaging (including both the heart and the vascular system). • The integration of the gathered information in a model of fatness will provide the researchers and clinicians with comprehensive patient-specific representation of the disease. • Through the similarity search the digital repository will provide the model for individual and personalised progression of disease prediction, impact of outcome markers and predict the effect of personalised treatment. This will provide optimization of therapy, and thus a complete newly-defined workflow for personalised predictive and clinical medicine.	PONGIGLIONE **Estimated % realisation** M3 M6 M9 M12		D12.3) Improved clinical workflows and outcome analysis: Final proposal of innovative clinical workflows based on outcome analysis of all patient cases	Month 48 **Lead** PONGIGLIONE **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement**

	units]	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP18: Dissemination & Training

Tasks	Lead	Deliverables	Deadline
18.3 Training Training is considered to be a fundamental task in dissemination. As anecdotal evidence has confirmed via WP4 of the VPH NoE and via feedback from the DISCIPULUS ('Roadmap Towards the Digital Patient') meeting (30/03/2012; Barcelona), training is recognized to be one of the most solid and long-lasting dissemination strategies in place. The training activities within MD Paedigree will consist of 2 'hands-on' workshops to be delivered during years 2 and 4 of the project (at approx. 1 or 1.5 year interval) in order to expose the outcomes achieved both, in disease modelling and in building the infostructure, highlighting the potential for change management and innovation in clinical workflows to the medical/clinical and research community interested in VPH technology. The first workshop will also seek to provide feedback to the research and development activities, so as to refine the outcomes for the final workshop. The workshop participants will fill in a detailed feedback questionnaire that will be passed to the developers. This task will be led by UCL, which has a long-standing commitment with the VPH Community and is involved in several training grants, including the Marie Curie ITN 'MeDDiCA', 'VPH-MIP' and WP4 of the VPH NoE.	DIAZ **Estimated % realization** M3 M6 M9 M12	D18.3) Training event in year 2: Report on the outcomes of the first Training event **D18.6) Training event in year 4:** Report on the outcomes of the second Training event	Month 30 **Lead** DIAZ **1st draft ready by:** Month 42 **Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and**

	measurement units]	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T18.4 Seminars, Workshops, Concertation Activities with Other ICT Funded Projects, and Scenario Analysis Sessions **Lead: Vanessa Diaz** The Consortium will identify the most relevant conferences in the area and propose seminars and workshops to be held during these events. It will devote special attention and resources to Concertation Activities with other ICT funded projects and to targeted dissemination actions. Special "Scenario analyses" sessions will be convened, involving the key personnel from both the clinical and the technological partners, with the aim of pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users within MD-Paedigree. The results of the previous workshops will be presented to the Scientific Committee and to the Users' Board in order to assess their relevance and applicability, so as to refine the outcomes for a validation workshop and for a final MD-Paedigree Conference, to be held at the end of the project, targeting both internal and external clinical and research communities as well as patient organisations and the interested media. The participation in any such event will be reported in the periodic reports and the final report.	DIAZ **Estimated % realization** M3 M6 M9 M12	**D18.4.1) First scenario Analysis Sessions:** First scenario Analyses pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users. **D18.4.2) Second scenario Analysis Sessions:** Second scenario Analyses pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users.	Month 24 **Lead** DIAZ **1st draft ready by:** Month 42 **Lead** DIAZ **1st draft ready by:**
Self-Assessment criteria			

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T18.7 Engaging Parent and Patient Associations **Lead: Vanessa Diaz** Approaching Parent and Patient associations will become a part of the consortium's dissemination activities. The project will seek to disseminate news of its work, expected results and potential future developments through these channels. It is hoped that the work with Patient associations will help achieve a larger bidirectional knowledge sharing base of clinicians and of patients, and further inform the potential beneficiaries of the ongoing work.	DIAZ **Estimated % realization** M3 M6 M9 M12	**D18.1) Dissemination and training strategy plan and preliminary materials**: Roadmap defining the dissemination and training strategy, indicating the subsequent choice of preliminary materials	Month 12 **Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria	
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and**

	measurement units]	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP19: Exploitation, HTA, and Medical Device Conformity				
Tasks	**Lead**		**Deliverables**	**Deadline**
T19.1: Evaluation approach and meaningful indicator development (EMP) • Develop upon and adapt in the VPH and other contexts proven approaches, methods and tools to the specific environment and objectives of this workpackage • Establish a set of meaningful criteria and their measurement process that are robust to demonstrate socio-economic benefit-cost impacts. The focus is • to approach and find measurements for evaluating how virtual collaborations between members of the VPH communities with different expertise are facilitated and • how consequently the uptake and acceleration of model development and integration can find meaningful expression in the overall evaluation framework.	STROETMANN **Estimated % realization** M3 M6 M9 M12		**D19.1 HTA evaluation framework** It reviews proven approaches, methods, and tools which might be relevant to the specific environment and objectives of this workpackage, and establishes a set of meaningful criteria and their measurement process, thereby focusing on evaluating how virtual collaborations between members of the VPH communities with different expertise are facilitated.	Month 12 **Lead** STROETMANN **1st draft ready by:**
Self-Assessment criteria				

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**		**Deliverables**	**Deadline**
T19.3: Benefit-cost scenario for clinical impact assessment (EMP) In a separate task a high-level, generic benefit-cost scenario for clinical impact assessment will be applied, with the ultimate goal to generate economic and market evidence for true translational medicine. The benefit-cost scenario will be tested and initially validated with preliminary, exploratory data estimates from the patient-centred workflows that are the basis of the digital repository and Infostructure. The two main dimensions pertaining to clinical/health impacts focus on the one hand on health service delivery and the health of patients, and on the other on public health/societal outcomes. To assess such impacts, the scenario development will integrate the following indicators: • Clinical effectiveness and patient-related outcomes • Safety (risks associated with applying the technology) • Organisational and change management aspects • Human resource implications, knowledge & education needs • Assessing contributions to the VPH vision of a patient avatar • Efforts for application (convenience/ease of use; costs for introduction of new technology) The indicators assessed ultimately prepare for a more targeted and strategically aligned exploitation activities (T19.4) by proving clinical impact of MD-Paedigree with respect to: • the state-of-the-art in paediatric patient-specific computational modelling, • improved disease understanding and therapy outcomes that can be applied to both clinical routine and translational clinical research, • usability by clinicians and clinical researcher, • transferring technical workflows into clinical workflows,	STROETMANN		**D19.4 Clinical impact assessment scenario** Initial formative evaluation of MD-Paedigree model-driven Infostructure based on a benefit-cost analysis approach, subsequently followed by a generic benefit-cost scenario for clinical impact assessment developed and validated with partners and experts. [month 36]	Month 36
	Estimated % realization			**Lead**
	M3			STROETMANN
	M6			**1st draft ready by:**
	M9			
	M12			

• the vertical integration of multi-scale patient data and the provision of models, tools, and services readily available to clinicians at the point of care.				

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

A.1.1.2 Timing of work packages and their components

The MD-Paedigree project partners have formalized a work plan implementing 4 major phases implying a number of conceptual steps, over 48 months of activity with 4 major milestones. The first milestone is due after 9 months and marks the end of the specification phase; the following milestones are aligned with the reporting periods of the project every 12 months.

Phase 1 (running from month 1 to 9) – Project Set-up, Requirements Elicitation, and Clinical Protocols: During Phase 1 quality assurance guidelines and a self-assessment plan will be prepared, ethical approval will be obtained , and the first dissemination activities will be performed (Step 1) Furthermore, clinical protocols for the selected paediatric applications will be established (Step 2).)Finally, the requirements for models and infostructure implementation will be analysed and documented from an end user standpoint (Step 3).

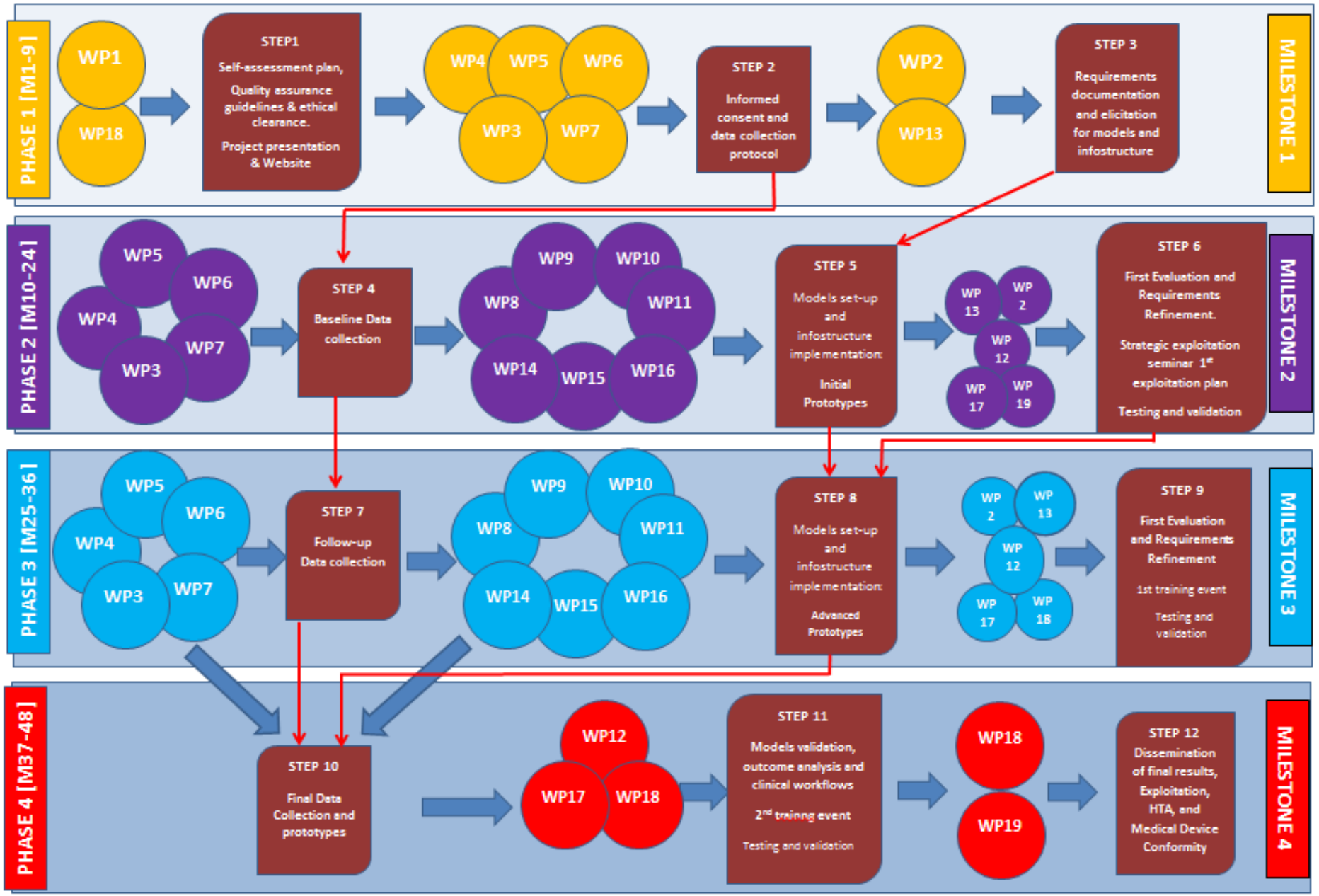
***Phase 2 (running from month 10 to 24) – Baseline Data Collection, Initial Prototypes, First Evaluation and Requirements Refinement*:** Patient enrolment will take place and data acquisition will be started (Step 4). Based on the established requirements, the existing models from Health-e Child and Sim-e-Child projects will be refined and adjusted to the new applications. The open repository for project infrastructure will be introduced and initialized with the current models and data (Step 5). First evaluations will be undertaken and requirements will be refined based on the collected experience; additionally, during this phase, the Strategic Exploitation Seminar will be held and the 1st Exploitation Plan will be drafted (Step 6).

***Phase 3 (running from month 25 to 36) – Follow-up Data Collection, Advanced Prototypes, Evaluation and Requirements Update*:** Follow-up or additional data will be acquired for all clinical applications (Step 7). The respective models will be enhanced to process longitudinal data and refined according to the obtained evaluation results. New functionalities will be integrated into advanced prototypes. The open repository will be improved and updated with content (Step 8). A second set of evaluations will be conducted and requirements will be adjusted for the final system. Furthermore, the 1st Training Event will be held (Step 9).

***Phase 4 (running from month 37 to 48) – Final Data Collection and Prototypes, Clinical Validation, and Deployment*:** In the final year, data collection will be concluded and the clinical validation will take place with the final models and simulation framework (Step 10). Results will be used to propose and disseminate improved clinical workflows. Subsequently, the 2nd Training Event will be held (Step 11). Models for all clinical applications and their respective evaluations will be documented and disseminated, while the implementation plan will be refined and the Health Technology Assessment and the Medical Clearance preparatory activities will be performed (Step 12) .

The timely delivery of all planned deliverables will be the first indicator of the fulfillment of each phase in the expected progress of MD-Paedigree, monitoring what can be demonstrable at each corresponding milestone of the project.

A second and much more detailed means of verification will be provided by the assessment criteria for each milestone and each WP which are to be defined within D1.3 Self-assessment plan on month 3.


PHASE 1 [M1-9]
WP1
WP18
STEP1
Self-assessment plan,
Quality assurance guidelines & ethical clearance.
Project presentation & Website
WP4
WP5
WP6
WP3
WP7
STEP 2
Informed consent and data collection protocol
WP2
WP13
STEP 3
Requirements documentation and elicitation for models and infostructure
MILESTONE 1
PHASE 2 [M10-24]
WP5
WP6
WP4
WP7
WP3
STEP 4
Baseline Data collection
WP9
WP10
WP8
WP11
WP14
WP15
WP16
STEP 5
Models set-up and infostructure implementation:
Initial Prototypes
WP 13
WP 2
WP 12
WP 17
WP 19
STEP 6
First Evaluation and Requirements Refinement.
Strategic exploitation seminar 1st exploitation plan
Testing and validation
MILESTONE 2
PHASE 3 [M25-36]
WP5
WP6
WP4
WP7
WP3
STEP 7
Follow-up Data collection
WP9
WP10
WP8
WP11
WP14
WP15
WP16
STEP 8
Models set-up and infostructure implementation:
Advanced Prototypes
WP 2
WP 13
WP 12
WP 17
WP 18
STEP 9
First Evaluation and Requirements Refinement
1st training event
Testing and validation
MILESTONE 3
PHASE 4 [M37-48]
STEP 10
Final Data Collection and prototypes
WP12
WP17
WP18
STEP 11
Models validation, outcome analysis and clinical workflows
2nd training event
Testing and validation
WP18
WP19
STEP 12
Dissemination of final results, Exploitation, HTA, and Medical Device Conformity
MILESTONE 4

CVD					
March 2013	**April 2013**	**May 2013**	**June 2013**	**July 2013**	**August 2013**
	Protocols delivered to Ethical Committee	D7.1) Recruitment protocol with ethical clearance (for genetic studies)	D4.1 Data collection protocol and ethical clearance	Interviews to prepare D2.1	First Half-Yealry report
			Contribution to the Self-Assessment Plan		Self-Assessment Plan
	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Check of the enrollment and data collection, analysis and processing.
	Area Dedicated T&M TC [24th Apr]	Area Dedicated T&M TC [22nd May]	Area Dedicated T&M TC [26th Jun]	Area Dedicated T&M TC [24th Jul]	Area Dedicated T&M TC [28th Aug]

September 2013	**October 2013**	**November 2013**	**December 2013**	**January 2014**	**February 2014**
Biannual area meeting	Check of the enrollment and data collection, analysis and processing	First draft of the deliverable D2.1			D2.1 Initial requirements analysis document including priorities for the implementation
				Internal Review	First periodic review
Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs
Area Dedicated T&M TC [25th Sep]	Area Dedicated T&M TC 23rd Oct]	Area Dedicated T&M TC [27th Nov]	Area Dedicated T&M TC [18th Dec]	Area Dedicated T&M TC [22nd Jan]	Area Dedicated T&M TC [26th Feb]

CVD	**DELIVERABLES WITHIN MONTH 24**	
	D4.2) Report on patient recruitment and data collection at baseline study	M24
	D9.1) Report about the adaptation of the heart model	M18
	D9.2) Report about automated assessment of body fat distribution from MRI and ultrasound data	M24
	D7.2.1) First report on data collection process	M18
	D7.3.1) First report on sample storage, DNA extraction and sample analysis processes	M18
	D12.1) Outline of the clinical assessment and validation criteria for all four disease areas	M18
	D12.2.1) First clinical assessment and validation results for all four disease areas	M24

Juvenile Idiopatic Arthritis

MD-PAEDIGREE KICK OFF MEETING

JIA WG

Participant's Name	Affiliation

Concept (general)

MD-Paedigree validates and brings to maturity patient-specific computer-based predictive models of various paediatric diseases

- increasing their potential acceptance in the clinical and biomedical research environment
- making them readily available not only in the form of sustainable models and simulations, but also as newly-defined workflows for personalised predictive medicine at the point of care.

These tools can be accessed and used through an innovative model-driven infostructure

- powered by an established digital repository solution
- able to integrate multimodal health data
- entirely focused on paediatrics
- conceived of as a specific implementation of the VPH-Share project, planned to be fully interoperable with it and cooperating, through it, also with p-Medicine.

MD-Paedigree aims at achieving high-level semantic interoperability,

- requiring standards enabling the clinical contents to be interpreted consistently across the different EHR regimes,
- while complete clinical interoperability between systems will require widespread and dependable access to maintained collections of coherent and quality-assured semantic resources,
- including models that provide clinical context,
- mapped to interoperability standards for EHR and PHR and biomedical data,

linked to well specified terminology value sets, derived from high quality ontologies

CONCEPT (SPECIFIC)	**Beyond the state of the art**	**WPs' OBJECTIVES**	**Objectives' Lead**	**Estimated % realization**
The cause and pathogenesis of JIA are still poorly understood, but likely they include both genetic and environmental components. Moreover, disease heterogeneity implies that different factors probably contribute to its pathogenesis and causes [Prakken B et al., 2011][27]. Affected joints develop	**Computerized quantitative measurements of inflammation and destructive changes** In the frame of the Health-e-Child project, a great deal of effort had been spent in order to standardise imaging procedures and devise paediatric-targeted scoring systems for the assessment of disease activity and damage in JIA considering the wrist. The collaboration between clinical and IT partners has enabled the development and validation of computerized quantitative measurements of inflammation and destructive changes that have shown potential value as predictors of future damage. In continuity with Health-e-Child's advanced personalised modelling of disease progression, the goal will be to implement	**WP2: Clinical and technical user requirements for disease modelling** • Incorporate into the model the variables that are analysed by the clinicians in their activity. • Ensure that the modeling reflects real clinical needs and is validated against them to assure their robustness and reproducibility. • Provide computational models that can be personalized by adapting the parameters to the		

synovial proliferation and infiltration by inflammatory cells which may ultimately lead to destructive lesions of joint structures, disability and high disease-related costs. Unfortunately, the present ability to predict the disease course and outcome is limited. Within the FP6 Health-e-Child project, ICT tools for diagnosis and scoring of JIA, based on image data of the wrist, have been developed. This framework is the basis for the developments planned for MD-Paedigree. Comprehensive and accurate computer models derived from patient-specific data across multiple scales covering body, organs, tissues, and molecular levels are developed. This data is gathered and stored in a standardised manner building upon the Health-e-Child software tools developed for wrist	a more robust multi-scale, personalised and predictive computer-based model of JIA – this time focusing on a wider range of joints than the wrist joint. **Pattern discovery in multimodal data** The multi-scale, personalised and predictive computer-based model of JIA will span body, organ, tissue and molecular level with adequate information fusion and in addition information obtained from gait analysis. This allows for pattern discovery in multimodal data through correlations between clinical data, imaging, immunological, metagenomic data (gut microbiota), and a biomechanical gait model. The driving force behind this project stems from the integration of data coming from a new cohort of patients (approximately 200 patients) into the framework developed within the Health-e-child project that will be further extended and adapted to the needs of MD-Paedigree. **Longitudinal design** Initial imaging will be performed at disease onset and followed for 2 years at least, in order to expand predictive multi-scale models in JIA. The longitudinal design of the study will allow a dynamic process of testing multi-scale disease models for each patient at follow-up visits to further personalize treatment strategies. By fusing the information on the anatomy and the physical properties of the tissues provided by the imaging technologies, with the functional information provided by the CGA, it will be possible to personalise a whole body-level model of the musculoskeletal dynamics capable of predicting the forces acting on a given joint during the patient movements. These forces will then be applied to an organ-level finite	integrated data of a patient case • Advance the knowledge about the selected diseases by allowing the simulation of different effects on the evolution of the disease • Predict the effect of therapy. • Ensure that MD-Paedigree models have the highest possible impact at the point of care. • Re-use of models between disease areas to leverage synergies where possible. • Existing standards for modelling and tools will be investigated. • The need for new standards will be evaluated and documented. **WP5: Data acquisition and processing for Juvenile Idiopathic Arthritis** The goal of this work package is to collect clinical, immunological, metagenomic and imaging data for the subsequent integrated analysis. Data collection is set-up as a prospective longitudinal study. The timeframe for patient recruitment spans the first 28 months. The objective is to acquire data from about 200 patients within the first 28 months (baseline acquisitions). For each patient, follow-up data will be collected for monitoring disease course and to identify outcome predictors. The following patient selection criteria will be applied: Inclusion criteria:		

analysis in the context of JIA. These tools are extended for the purpose of integrating model information related to a wider range of joints, covering morphology, gait analysis, bio/genetic data. The tools to be developed will also include the aspect of a multidimensional longitudinal analysis that yields the opportunity to identify potential new outcome measures (imaging or biological biomarkers) for the assessment of treatment efficacy. Furthermore, the prognostic value on an individual level of multidimensional data, including modern imaging modalities, genetic and meta-genetic data will be explored through the development and integration of appropriate data clustering methods. By collecting patient specific multi-scale and	element model of the joint, where the mechanical properties of the tissues will be informed as much as possible from the imaging data. **Cancellous bone anisotropy, cartilage erosion, alteration of the subchondral bone** Among the other things we shall explore the possibility to derive cancellous bone anisotropy from DTI-like MRI imaging, mechanical properties of the cartilage from distribution of the GAG content again obtained by MRI, etc. We shall also correlate the biomechanical predictions with the signatures of the disease that can be quantified, such as the extension and the location of the cartilage erosion, or the alteration of the subchondral bone, to the predictions of stress and strains obtained by the organ-level model. **High-resolution US** High-resolution US will be performed not only at the joints (wrist and ankle) investigated with MRI, but at all the affected joints, in order to better define the extent of the disease. The severity of joint involvement will be judged sonographically by a variety of parameters such as joint effusion, synovial thickening and hyperaemia, cartilage integrity and bone erosions. Quantitative assessments of these parameters will be extracted from the US equipment based on standardised scanning planes by means of 2D imaging. At the same time, using 3D imaging, serial slices will be recorded resulting in a pyramid-shaped volume scan. The acquisition and storage of a number of volume datasets with time would allow better comparison of findings in longitudinal studies and the detection of earlier and subtle predictive signs of damage.	• Patients with JIA according to ILAR criteria and wrist and/or ankle involvement. • Disease duration < 6 months • Parents or legal guardian (and the subject when age is appropriate) must be willing to sign the consent/assent forms. Exclusion criteria • patients requiring general anesthesia or with contraindication to MRI will be excluded from the study. **WP10: Modelling and simulation for JIA** The aims of this work package can be divided into four key areas. Namely: 1. Development of articulated models of the JIA affected joints 2. Automatic extraction of biomarkers 3. Patient-specific biomechanical simulation 4. Multidimensional modelling of the disease course The developments in all of these areas will go beyond the Health-e-Child project. The goal is to gain a better understanding of the inflammation induced anatomical and functional changes in the juvenile joints.		

multi-dimensional information and automating image and data analysis at the point of care, this project has a strong clinical impact on early diagnosis, prediction of disease and of treatment outcome.	**Whole-body Dual X-ray Absorptiometry** In addition to MRI and US, Whole-body Dual X-ray Absorptiometry (DXA) will be performed. Total body DXA provides an accurate measurement of the areal body density over the frontal plane, separating the bone mass, the lean mass (muscles), and the fat mass with good accuracy. This imaging modality will be used to personalise multi-scale models of the musculoskeletal system capable of predicting the forces transmitted at the joints during a given movement. DXA images will be used not only to personalise these generic models anatomically: total body density will inform the inertial properties in the inertial model; lean mass will be used to estimate the muscles cross-section in the musculoskeletal model; bone density will be used to personalise the bone stiffness in the joint models. **Integrative multiscale representation of the patient's musculoskeletal system** All these personalised models will be composed in an integrative multiscale representation of the patient's musculoskeletal system, capable of predicting, for example, the forces being transferred to the joint cartilage during a given movement as captured during the gait analysis. **Multimodal image analysis by means of model-based segmentation of MRI images** The wrist MRI scores, as well as the automated software for the quantitative assessment of disease activity and damage, developed in the frame of Health-e-Child, will be adapted to investigate the ankle. Focusing on the locomotory system, especially the juvenile ankle, enables the physician to study the effects of JIA on the joint motion, which form another scale in the patient-specific	Furthermore this work package wants to create a predictive model, which is able to differentiate given JIA patients based on the extracted biomarkers and the patient-specific model into groups where the disease course is mild or aggressive. For this purpose, the JIA part of the MD-Paedigree project is also defined as a longitudinal study. Additionally to the clinical analysis of the wrist joint in Health-e-Child, this project will also focus on the ankle joint. The developments in the field of model based segmentation, automatic image processing for biomarker extraction, as well as patient-specific biomechanical modelling of the juvenile joint go well beyond the state of the art. During In the course of this project, we are going to develop a modelling technology capable of generating patient-specific multi-scale biomechanical models of the musculoskeletal system. These models will be able to predict the biomechanical conditions to which the articular cartilage and the subchondral bone		

	model. MD-Paedigree aims to automate and extend the multimodal image analysis and therefore standardise the derived biomarkers by means of model-based segmentation of MRI images. **Articulated joint model** An articulated model of the juvenile ankle and wrist will be developed and used. It includes the bones' shape, the spatial relation between the bones and their appearance in MRI images. By simulating the joint articulation, it will allow for the adaption to a specific MRI-scan, resulting in patient-specific models. In order to generate a personalised morphological model for JIA, an articulated joint model –consisting of bones, cartilage and ligaments representing the variation in shape, image appearance and spatial relations trained using machine learning methods – will be developed. It will be built from manual annotations by experts on morphological MRI datasets of patients suffering from JIA. Data from MRI molecular imaging analyses will be also included as well as data from US evaluation. **Integration of image based patient-specific models with gait cycle analysis** Furthermore, the role of the musculoskeletal dynamics and of the mechanical properties of the joint tissues in conditioning disease progression or in response to treatment will be investigated. The integration of image based patient-specific models with gait cycle analysis will allow the generation of highly personalised multiscale models of the musculoskeletal	are exposed during daily life. These techniques enable the evaluation of the associations between anatomo-functional, biomechanical, and clinical indices of the disease progression, in order to elucidate how each determinant contributes to the disease, and if there are complex systemic interactions involved.		
		WP7 Genetic and metagenomic analytics To evaluate the role of genetic (assessed by disease-gene or candidate gene analysis) and metagenome (based on gut microbiota profiling) profiles on the development and progress of diseases and on their outcome.		
		WP12: Models validation, outcome analysis and clinical workflows • To clinically validate derived models • To improve prediction of outcome and risk stratification • To establish integrated clinical workflows and personalised treatment models		

	system capable of elucidating the role of biomechanical properties in onset and/or progression of structural damages. Three-dimensional clinical gait analysis (CGA) is a well-established method enabling, when a strict analysis of causes of errors is carried out and periodical validation procedures are implemented, highly objective and reliable evaluation of gait in both healthy and diseased populations. CGA including kinematics and kinetics, provides more information about gait changes, such as joint angles and moments, which cannot be quantified in a standard clinical setting. The kinematics shows the joint movement, while the kinetics describes the forces involved in movement (e.g. ground reaction forces, joint moments, and joint powers). **Discovering potentially destructive gait deviations** By examining kinetics, the mechanisms of gait deviation can be described and the early use of gait analysis can be instrumental in discovering developments of potentially destructive gait deviations. Patients will be dressed with skin-attached markers that are both visible in MRI imaging, radiopaque (so they appear also in the DXA image) and, successively, reflective markers will be reapplied in the same anatomical positions, so they can be tracked during gait analysis. Whole body imaging and gait analysis will be performed one after the other with the patient dressed with the markers. This will provide a fiducial registration framework between anatomical and functional data. The imaging protocol will be agreed with the modellers, in order to ensure that the highest amount of information is transferred to the predictive models. **Three-dimensional clinical gait analysis, ground force platform, and cutaneous electromyography**	**WP19: Exploitation, HTA, and Medical Device Conformity** An early evaluation in the form of health technology assessment (HTA) as well as the development of exploitation strategies is essential for the creation of research related services which can prevail in today's highly competitive markets - be they "academic" and RTD markets, be they health services or commercial markets. The workplan is designed to encourage materializing improved disease understanding and therapy outcomes into both clinical routine and translational research, to deploy early prototypes within the developing VPH Infostructure, and to improve in iterative cycles of specifications, refactoring (i.e. improving the design of existing code), and deployment. **Objectives** • Evaluate the MD-Paedigree's models, workflows, and infostructure based on: o its accessibility, usability and effectiveness for the VPH community o the potential of its contributing to personalised healthcare workflows and		

	Each patient will be examined using three-dimensional clinical gait analysis (CGA), ground force platform, and cutaneous electromyography (EMG). Depending on the joint of interest, the patient will be asked to repeat a few times a given movement, selected among those most common in daily life (i.e. for lower limb, level walking, stair climbing, sit to stand, etc.), and the relative motions and muscle activation signals are recorded. An expert physiatrist will examine the gait analysis data to exclude specific gait abnormalities. Using the fiducial marker set, the motion data will be fused with the imaging data, and with the internal musculoskeletal and joint models fitted to the imaging data. This will result in a body-organ multi-scale model capable of predicting the forces being transferred to the joint during each of the recorded movements. EMG data will not be used to inform the model, but will be compared with the activation patterns predicted by the models, so as to verify that the model is operating consistently with the patient's neuromuscular activation strategy. **Inverse kinematics** The body model will use inverse kinematics to find the optimal registration framework between the model and the recorded kinematics, so as to reduce as much as possible the so-called skin artefacts. Then, inverse dynamics will be used to compute the joints torque that is required to generate the recorded movement. An optimisation scheme will be used to compute muscle activations and joint forces. This time-varying system of musculo-articular forces will be applied as boundary condition to a finite element model of the joint being investigated. The individualised finite element model will predict the	integration with EHRs/decision support systems, thereby preparing for the transfer into clinical practice o making models and simulations readily available at the points of care and to researchers • Define effectiveness and usability within the context of sharing "developing ICT tools, services and infrastructure to obtain more elaborate and reusable multi-scale models" (call text) as well as developing an appropriate analytical evaluation framework • Explore the health system and business opportunities o to market concrete project outcomes and results o to prevent diseases and contribute to the safety of care o to identify markets and cost models for the effective diffusion of our models, allowing researchers to exploit, share resources and		

	mechanical stresses and strains induced in the various joint tissues by the given movement, and information to be used as an additional "biomarker" in the evaluation of the individual clinical case. **Integration with immunological and metagenomic data** Imaging data will be integrated with immunological and metagenomic data in order to try to identify surrogate parameters for disease activity, disease severity, risk of side effects and treatment outcomes. New particle-based multiplex immunoassay, such as the Luminex technology, allowing the measurement of multiple circulating and/or synovial cytokines, as well as of other immune mediators, will be used to define the individual immunological profile for each patient. **Cytofluorimetric analysis** Furthermore, paired peripheral blood and synovial fluid mononuclear cells subpopulations (naive and effectors T cells, B cells, monocytes, etc.) will be evaluated by cytofluorimetric analysis. We will also look at phenotypic markers, mRNA, epigenetic markers (methylation FOXP3) and functionality (in vitro suppression assays). **Correlation of gut microbiota and immune responses** Analysis of gut microbiota will provide new insight into the environmental factors which regulate innate and adaptive immune homeostasis and affect the development of systemic autoimmune diseases. The gastrointestinal tract is the largest human immune organ and home to a complex community of trillions of bacteria that are engaged in a dynamic interaction with the host immune system. Communication between the microbiota and the host	develop new knowledge • Design business plans that prepare pre-market access and that integrate medical device conformity assessment procedures		

	establishes and maintains immune homeostasis, enabling protective immune responses against pathogens while preventing adverse inflammatory responses to harmless commensal microbes. Correlations have been found between the composition of gut microbiota and some preferential immune responses (i.e. Th17 response). By analysing the gut microbiota of JIA patients collected in specific disease states (at the onset, when patient will achieve clinical remission state, and during flare of the disease) we aim to explore its potential role in conditioning disease susceptibility as well as immune response in the different stages of disease, thus adding a further important dimension to multiscale analysis. Investigating the interaction of gut microbes and the host immune system will improve the understanding of the pathogenesis of this autoimmune disease, and provide innovative foundations for the design of novel immuno- or microbe-based therapies. **Impact of joint mechanical abnormalities on disease progression** The impact of biomechanical property alterations on subsequent progression of structural damage in patients with chronic inflammatory arthritis is not yet characterised. Personalised joint biomechanical modelling allows critical evaluation of the forces within the joint under physiologic and pathological loading conditions. Evaluation of the impact of joint mechanical abnormalities on disease progression is needed for an accurate outcome prediction. The potential of the multi-scale modeling methods proposed, is to make the exploration of complex systemic interactions between the neuromuscular control, the musculoskeletal			

	functional anatomy, and the local biomechanical determinants acting in the joint space at the tissue level, possible. **MRI molecular imaging analysis** The modelling predictions could have significant implications in early diagnosis and therapeutic intervention. In this perspective, early signs of structural damage will be evaluated also using MRI molecular imaging analysis. Molecular imaging allows the detection of microstructural changes in the composition of the cartilage matrix that occurs before morphologic changes can be qualitatively detected by conventional imaging, at stages when damage to the cartilage is potentially still reversible and may be treated. Molecular imaging by providing *in vivo* information beyond morphological changes in articular cartilage, might yield attractive new insights in the biological pathways of cartilage turnover, with the potential to improve our understanding on erosive disease mechanisms and disclose new targets for therapy, thus suggesting a potential role for MRI in the drug development process. **Multidimensionality** Demographic clinical imaging and laboratory data in the form of text, images, annotations, videos, biomarkers and articulated models will be entered in the MD-Paedigree digital repository and will be continuously analysed providing potentially more accurate disease model tools. The combination of different assessment techniques will enable to enhance the value of a multidisciplinary management of JIA. The multidimensionality of the human and microbial phenotypes (and the dynamic, nonlinear interactions) will be explored by means of improved informatics tools, including new approaches for understanding the complexity of the			

	metadata, in order to better understand the implications of gut microbiota variations in human health and disease. The prognostic value on an individual level of multidimensional data, including modern imaging modalities, immunological, metagenomic data, as well as articulated models and biomechanical models will be explored. JIA constitutes an ideal domain for assessing the merits of simulators and predictors based on data generated across different scales. The validity and effectiveness of the proposed solutions will be assessed by using the model to address several open issues in JIA with a strong clinical impact on early diagnosis, prediction of disease and of treatment outcome.			

Application Scenario

Chiara and Simona are affected by JIA. Age at disease onset was 4 years old for both girls. Antinuclear antibodies were positive in both cases. Both of them had an asymmetric involvement of knee and ankle at disease onset. Within 1 year from disease onset both patients showed wrist involvement and started a treatment with second line agent (methotrexate). Chiara experienced a severe and irreversible structural damage progression, as revealed by a 4-years follow-up plain radiography. Conversely, Simona developed a milder course of the disease and the 4-years follow-up plain radiography showed no significant signs of structural damage. Notably, at disease onset demographic, routine clinical and laboratory data did not significantly differ between these two patients. It was not possible, therefore, to distinguish, at disease onset, which of the patients would develop a more aggressive disease.

By integrating in a multi-scale integrated model clinical, MRI and US evaluations, immunological and meta-genetic data (microbiote), as well as the results of biomechanical analysis, we aim to identify outcome predictors and discriminate, early after disease onset, patients who will develop a more severe course of the disease and will require an earlier and more aggressive treatment.

A comprehensive model of JIA-related changes in two joints – wrist and ankle – will be available. Different imaging modalities (MRI, DXA) provide information to classify the degree of bone erosion and synovitis in both regions. For this, automated image analysis tools will be developed to reduce the time necessary for performing that task but also becoming independent from the individual observer who does the exam. Enhanced biomechanical models are generated by

adapting highly sophisticated standard models to the individual case and thus predicting locomotive changes caused by JIA.

These are tested against results of a personalised gait analysis that further enriches this model by providing more details about the locomotive constraints for the considered joints. In addition, a selected sub-group is examined for a second time in order to record the data basis necessary for modelling also the progression of the disease.

The related database is built on the outcomes of the Health-e-Child project and contains information about the morphological changes visible in the image data but also about clinical, immunological and genetic and metagenetic (microbiote) data.. Having access to a large repository of such classified individual JIA cases that are described in great detail – also longitudinal aspects – makes it possible to better predict the progression of the disease and to provide the best adapted medication for these patients.

WP2: Clinical and technical user requirements for disease modelling

Tasks	Lead	Deliverables	Deadline
Task 2.1: Conduct interviews with the clinical and technical partners to obtain a complete list of requirements for the disease modelling that will ensure its usefulness within and beyond the project. All WP Leaders will actively contribute to the requirements documentation while they ensure that the respective WP partners are interviewed. Prioritisation criteria: i. All requirements will be prioritised ensuring that from the start the most important aspects will be implemented to quickly ensure an operational system. Schedule of requirements updating: ii. The requirements list will be continuously updated on a regular basis such that main requirements and system constraints will be released as deliverables.	CHINALI **Estimated % realization** M3 M6 M9 M12	**D2.1 Initial requirements analysis document including priorities for the implementation** Complete interviews with the clinical and technical partners will be collected to obtain a list of variables and requirements for the disease modelling. Requirements will be prioritized ensuring that from the start the most important aspects will be implemented first.	Month 12 **Lead** CHINALI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP5: Data acquisition and processing for Juvenile Idiopathic Arthritis

Tasks	Lead	Deliverables	Deadline
T5.1 Data collection protocols and informed consent [M 1-3] The first three months will be dedicated to the preparation of the data collection protocols to be submitted to the Local Ethics Committee, together with the parents and patients (when applicable for age) informed consents in which the design, study purposes, and the privacy issues related to the data management will be duly explained. The project will be carried out in accordance with the applicable EU and national data privacy protection laws and regulations.	MARTINI **Estimated % realization** M3 M6 M9 M12	**D5.1) Report on data collection protocols and parents and patients informed consents** Submission of protocol and informed consent to the institutional review boards.	Month 4 **Lead** MARTINI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T5.2 Clinical data collection [Month 4-40] The following clinical data will be acquired at 6 months follow up intervals for the first two years from patient enrollment: • demographic data such as gender, age at disease onset, JIA subtype according to ILAR classification, etc. • clinical variables including standardised and validated measures of disease activity and disease damage (e.g. number and site of inflamed joints, presence of systemic feature, functional ability, the Juvenile Arthritis Disease Activity Score, the Juvenile Arthritis Damage Index etc)) will be collected at enrolment and every 6 months. • Information concerning previous and ongoing treatment will be recorded.	MARTINI **Estimated % realization** M3 M6 M9 M12	**D5.2) Report on baseline data collection status** Collection of data (clinical, imaging, laboratory samples) in the first 80 patients. **D5.3) Report on baseline and intermediate follow-up data collection status** Completion of data collection	Month 16 **Lead** MARTINI **1st draft ready by:** Month 28 **Lead** CHINALI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T5.3 Routine laboratory tests [M4-40] • Routine laboratory tests to extract markers of inflammation such as ESR, CRP, antinuclear antibodies, and rheumatoid factor will be performed at enrolment and every 6 months.	MARTINI **Estimated % realization** M3 M6 M9 M12	**D5.4) Report on longitudinal data collection status** Follow-up data at one year in the large majority of patients	Month 40 **Lead** MARTINI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T5.4 Synovial and blood Cytokine and inflammatory mediators profile[M 4-40] Biological samples (blood, and synovial fluid from patients with clinical indication to perform local steroid injection) will be collected at disease onset, when patient will achieve clinical remission state (according to Wallace criteria for remission in JIA) and during flare of the disease. For biomarkers we will use a high throughput methodology, namely the multiplex immuno assay or Luminex®. This is a bead-based assay that allows the detection of more than 100 soluble mediators in a single sample of 50 microliter of body fluid, such as plasma or synovial fluid. Partner UMCU is an international expertise centre for this technology and has developed a "home-brew" assay for the determination of over 100 soluble factors, mostly cytokines and all directly related to inflammation, and thus potential co-determining risk factors for inflammation. We will measure a set of markers related to inflammation and/or cardiovascular risk, mostly adipokines and cytokines. We will perform pilot experiments in small proof-of principle cohorts (max 20 patients) to determine the panel that will be measured in a large validation cohort. These markers will be measured in peripheral blood plasma, and, if available in synovial fluid. In a smaller subpopulation of patients , based on the results from the previous studies, we will perform T cell characterization in paired peripheral blood and synovial fluid derived mononuclear cells focusing on regulatory T cells (natural and induced regulatory T cells expressing FOXP3, Tr1 cells) and effector T cells (Th17, Th1 cells). We will both look at phenotypic markers, mRNA, epigenetic markers (methylation FOXP3) and functionality (in vitro suppression assays).	MARTINI **Estimated % realization** M3 M6 M9 M12	**D5.4) Report on longitudinal data collection status** Follow-up data at one year in the large majority of patients	Month 40 **Lead** MARTINI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T5.5 Metagenome data analysis (gut microbiota) [M 4-40] Microbioma analysis will provide an opportunity to understand how the gut microbiota regulates innate and adaptive immune homeostasis and affects the development of systemic autoimmune diseases. Dysregulation of host responses as a consequence of dysbiosis in the gut lumen could affect distant anatomical sites through the activation of host immune responses. Stool samples which will be collected at disease onset, when patient will achieve clinical remission state(according to Wallace criteria for remission in JIA) and during flare of the disease. (See WP7 for protocols and analyses.) The results of gut microbiota analysis will be integrated with clinical immunological and imaging data to assess how it does affect human health, and in particular to explore the prognostic value of the presence of major clustering patterns at the gastrointestinal tract in conditioning disease susceptibility as well as the immune response in the various phases of the disease.	MARTINI **Estimated % realization** M3 M6 M9 M12	**D5.4) Report on longitudinal data collection status** Follow-up data at one year in the large majority of patients	Month 40 **Lead** MARTINI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T5.6 Image acquisition and clinical annotation [M 4-40] The following imaging techniques will be performed: • Digital plain radiography: wrist and/or ankle plain radiography will be performed at the enrolment and after 1 and 2 years to assess the presence and the degree of local growth disturbances and abnormal joint alignment (i.e joint subluxation, dislocation and flexion/extension defects). • Magnetic Resonance (MR): wrist and/or ankle MRI will be performed on a 1.5 Tesla MR scanner at the time of patient enrollment and after 1 and 2 year from baseline evaluation. • The following image sequences will be used in the study protocol: Morphological study: TSE T1 3D; TSE T2 fat sat; GRE 3D fat sat. MRI detectable pathological findings will be extracted using both a semi-quantitative and a quantitative approach (using scoring systems and automated tools that have been developed in the frame of Health-e-Child project). • Cartilage damage will be investigated also through the analysis of its ultra structural composition (molecular imaging). • For the ultrastructural study, the following protocols will be used: TSE Multi IR (TSE Multi IR, TR 2400, TE80;TI:50,80,180,350,700,1400,2200) sequence for the T1 mapping of cartilage (dGEMRIC technique);TSE MultiTE (SEQUENCE TSE, TR 1000, TE16,28,40,55,70,85) sequence for the T2 mapping. MR images will be centralized and annotated at IGG. • Ultrasound imaging: high-resolution ultrasound (U/S) evaluation of joints will be performed using a machine, equipped with broadband linear-array transducers (frequency band, 12-5MHz and 17-5MHz). At the patient enrollment all the joints evaluated by the rheumatologic clinical examination will be also investigated with U/S for a more accurate assessment of disease extension. • U/S follow-up data will then be acquired at 6 months follow-up intervals for the first two years from patients enrollment. The U/S scanning protocols will be based on the standardised technical guidelines issued by the European Society of Musculoskeletal Radiology and the OMERACT US group. The severity of joint involvement may be judged sonographically by a variety of gray-scale parameters, including the amount of joint effusion, the presence of synovial thickening, the degree and duration of synovial hyperemia, the occurrence of cartilage abnormalities and bone erosions. • In our protocols, quantitative assessments of these parameters will be extracted from the U/S equipment based on standardised scanning planes by means of 2D imaging. At the same time serial slices will be recorded resulting in a pyramid-shaped volume scan also using 3D imaging. • The acquisition and storage of a number of volume datasets with time will allow better comparison of findings in longitudinal studies. • This retrospective 3D review could enable to discover very early and subtle predictive signs of damage. Correlation between the site of gray-scale damage and the site of hyperemia will be performed in order to identify patterns of hyperemia that may be predictive of disease progression. the site of hyperemia will be performed in order to identify patterns of hyperemia that may be predictive of disease progression.	MARTINI **Estimated % realisation** M3 M6 M9 M12	**D4.3) Report on patient follow-up** Re-evaluation of all patients recruited for D.4.2 based on follow up data collection	Month 36 **Lead** MARTINI **1st draft ready by:**

• Dual X-ray Absorptiometry (DXA) : will be performed at the time of patient enrolment and after 1 year. Total body DXA provides an accurate measurement of the areal body density over the frontal plane, separating with good accuracy the bone mass, the lean mass (muscles), and the fat mass. Being a radiological image, it provides a fairly accurate spatial location of the joint centres, and of other skeletal landmarks.				

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead	Deliverables	Deadline
T5.7 Gait cycle analysis [M 4-40] • Quantitative gait assessment will be carried out with CGAs installed at the labs and the reflective markers will be attached bilaterally on the participant's skin at the head, shoulders, trunk, arms, pelvis, legs and feet according to the common biomechanical gait model. • Children with JIA will be evaluated the second time after treatment. The same examiner will perform the clinical measurement and marker placement in children with JIA. • To evaluate kinetic and kinematic variables in all three anatomical planes we plan to calculate from five gait cycles beginning with the left foot strike and five gait cycles beginning with the right foot strike. Differences will be evaluated in children with JIA between pre- and post-treatment gait analyses using a Repeated Measures Analysis of Variance (ANOVA). • Non-parametric statistical (Mann–Whitney) tests will be used to determine differences between children with JIA before treatment and controls, and between children with JIA after treatment and controls. • Moreover the gait analysis will be integrated with pressure matrix sensor outputs to quantify static and dynamic pressures exerted by foot during static posture or gait. • The values obtained via the matrix pressure will permit, together with CGA outputs, the development of disease-staged targeted treatment .	MARTINI **Estimated % realization** M3 M6 M9 M12	**D4.3) Report on patient follow-up:** Re-evaluation of all patients recruited for D.4.2 based on follow up data collection	Month 36 **Lead** MARTINI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	Indicators [Upper and lower limits associated with WP objectives and measurement units]	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP10: Modelling and simulation for JIA

Tasks	Lead	Deliverables	Deadline
T10.1. Patient-specific anatomical modelling based on image data [M 4-48] • An articulated model of the wrist and ankle region – containing bones, cartilage and ligaments – will be developed based on morphological MRI. • For each structure, a representation of the variation in shape and image appearance is obtained by the means of machine learning methods based on manually annotated training images. • In addition, the variability of the spatial relations between the single structures is described by the model. • The model will improve the detection and quantification of JIA related pathologies like bone erosion and synovitis by facilitating and automating the segmentation of the relevant anatomical structures. • The anatomical information will be fused with functional MRI/molecular imaging that highlights inflamed regions. In addition to the static MRI imaging of the joint, dynamic ultrasound images will be acquired showing the relationship between cartilage and bone during articulation. • The automated segmentation of these structures from the dynamic imaging as well as the derived biomarkers from automated analysis of the ultrasound and functional MRI images will provide important input for the patient specific bio-mechanical model of the joints. • Finally, based on DXA images the T-score provided by the machine will be used to calibrate the bone density using a generic tissue density distribution. • Multimodal image registration methods for 3D MRI, 2D functional MRI and 2D dynamic ultrasound will be adapted to the particular needs for building the patient-specific wrist and ankle models. • There, the articulated model will provide a set of landmarks together with their probable intra-joint shifts that will help to better model expected deformations between the different image acquisitions – for initial imaging as well as longitudinal studies. **Partners involved: FhG, USFD, IGG**	VICECONTI **Estimated % realization** M3 M6 M9 M12	**D10.1) Report about initial modelling results** At PM24 T10.1, T10.2 and T10.3 have been running for at least 12 months. Thus, it will be possible to report some preliminary results relative to image-based modelling, extraction of biomarkers and biomechanical simulation. **D10.2) Report about image based patient-specific modelling** This deliverable will summarise the activity of WP10 by reporting on the final workflow developed to transform imaging and functional data of each patients into predictive models and quantifications capable to modelling the disease.	**Month 24** **Lead** VICECONTI **1st draft ready by:** **Month 48** **Lead** FRANGI **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T10.2. Automatic biomarker extraction [M 7-42] [FhG, IGG] Existing tools for biomarker extraction related to JIA (i.e. automated assessment of inflamed synovial volume, quantitative assessment of inflammation based on the analysis of the time course of signal changes following gadolinium injection) will be integrated and enhanced. Based on the knowledge about differences between the anatomical and functional model of a healthy joint and the adaptation of said model to the JIA patient, the extraction process of interesting biomarkers from MRI and ultrasound images can be automated. One goal is the classification of the functional MRI and ultrasound images into regions of interest and an automation of the analysis process leading to the extraction of important biomarkers for the staging of the JIA progression. This will help in standardization of the extraction process for ultrasound imaging and therefore overcome inter-observer variability which is highly pronounced due to the steep learning curve of ultrasound imaging. Furthermore the registration of three dimensional anatomical MRI with two dimensional functional MRI and dynamic ultrasound imaging allows the correlation of biomarkers from different modalities describing the same or complement characteristics of the JIA progression.	VICECONTI **Estimated % realization** M3 M6 M9 M12	**D10.3 Report on biomarker extraction** Although T10.2 will run until PM42, at PM36 we expect to have a sufficient understanding to report on biomarkers extraction.	**Month 36** **Lead** VICECONTI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T10.3. Biomechanical simulation based on image based modelling and gait analysis [M 13-42] • The attention will be focused on the ankle joint of the lower limb, for which the loading regime is more predictable, based on gait analysis observations. • Patients will be dressed with skin-attached markers that are radiopaque, for imaging analysis, and reflective markers, for CGA analysis. • The patients are examined with whole body DXA imaging to obtain the whole body mass composition and the general anatomy of the skeleton. • A generic musculoskeletal model of the lower body will be morphed on the DXA image: the mass distribution information will be used to inform the inertial properties of the model; the lean mass information will be used to scale the muscle tetanic forces, and the bone mineral density distribution to define the heterogeneous module of elasticity of bones. • Additionally to DXA, a full gait analysis session will be conducted on these patients. Also superficial EMG signals will be recorded. The skin markers radiopaque and reflective will be used as fiducial points to fuse the imaging and gait data into a body level model of the musculoskeletal dynamics that will predict the muscle and articular forces transmitted to the hip, knee, and ankle joint which is to be examined. • These forces will be imposed to organ level finite element models of the ankle joint, individualised using MRI and ultrasound data. The resulting multi-scale model will be able to predict for each patient the stresses and strains induced by normal physical activities; such as level walking in the articular cartilage and in the subchondral bone. • This information will be combined with the rest of the clinical data to explore the role of biomechanical determinants in the development and the severity of the disease. E • MG data will be used to verify that the optimisation function used to estimate the muscle forces, which is correct for healthy subjects, is also reasonable for JIA patients; in case this is not confirmed alternative cost functions will be explored to find the one appropriate for these patients. **Partners involved: USFD, MOTEK, FhG, IGG, URLS**	VICECONTI **Estimated % realization** M3 M6 M9 M12	**D10.4) Report about biomechanical simulation based on image based modelling and gait analysis**: T10.3 also end at PM42, but at PM36 we should be able to see which will be the achievements in term of biomechanical modelling and report on them.	**Month 36** **Lead** VICECONTI **1st draft ready by:**

Self-Assessment criteria	
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T10.4. Multidimensional modelling of disease course [M 24-42] [USFD, FhG, IGG] The patient-specific modelling will be extended by a longitudinal component that allows detecting and describing bone erosion and synovitis caused by the progression of JIA. For this, homogeneous sub-groups of patients are identified employing the output of the initial patient-specific modelling and biomechanical simulation (T10.1, T10.2). A patient-specific finite element model of the joint of interest, built on stress and strain values derived from all imaging modalities, will be created and loaded with predicted musculo-articular forces. To predict how the joint forces are distributed across the cartilage, as a first approximation the generalised joint forces are taken from inverse dynamics assuming the bone to be infinitely rigid and only the non-linear contact deformation of the articular cartilages is modelled. Here, data from MRI molecular imaging about the mapping of the GAG content (T10.1) will help to refine the model and to personalise it. In particular we want to explore the role that the biomechanical determinants have on the disease severity. Quantification of erosion extension and location will be correlated with the stresses and strain distributions; the longitudinal changes in the subchondral bone will be correlated with the bone strains induced by normal physical activities, to see if a causal relationship can be proposed. The same model will be employed to explore the role of synovial inflammation (evaluated in terms ofsynovial volume, volume of synovial fluid etc.) in conditioning structural damage progression.	VICECONTI **Estimated % realization** M3 M6 M9 M12	**D10.5) Report on multidimensional modelling of disease course:** This deliverable will provide the final results of T10.4, relative to the longitudinal analysis of bone erosion and synovitis caused by JIA, and their association with biomechanical factors as predicted by joint finite element models	**Month 42** **Lead** VICECONTI **1st draft ready by:**

Self-Assessment criteria	
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

<table>
<tr><th colspan="4">WP7 Genetic and metagenomic analytics</th></tr>
<tr><th>Tasks</th><th>Lead</th><th>Deliverables</th><th>Deadline</th></tr>
<tr><td rowspan="4">T7.1. Informed consent and data collection protocol.
Informed consent forms and dedicated testing protocols will be prepared for sample collection, DNA extraction and analysis.
T7.2. Sample collection, storage and DNA extraction.
Samples will be collected from 180 patients for cardiology, 200 for rheumatology, 180 for cardiovascular risk in obesity, and from a control group of 100 unaffected subjects
Fecal samples will be stored at -80 °C until further processing and DNA will be extracted according to published methods (Zoetendal, E.G., Heilig, H.G., Klaassens, E.S., Booijink, C.C., Kleerebezem, M., Smidt, H., de Vos, W.M., 2006. Isolation of DNA from bacterial samples of the human gastrointestinal tract. Nat. Protoc. 1, 870–873; Salonen A, Nikilä J, Jalanka-Tuovinen J, Immonen O, Rajilić-Stojanović M, Kekkonen RA, Palva A, de Vos WM., 2010. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. J Microbiol Methods 81(2):127-34), with slight modifications.</td><td>OPBG
Estimated % realization
M3
M6
M9
M12</td><td>D7.1) Recruitment protocol with ethical clearance: Completion of the recruitment protocol, consensus and ethical clearance from all partners’ involved in patient recruitment.</td><td>Month 3
Lead
OPBG
1st draft ready by:</td></tr>
<tr><td>D7.2.1) First report on data collection process: Report on data collection progress, inclusive of analysis of patient data on the basis of inclusion/exclusion criteria and updating of clinical features.</td><td>month 18
Lead:
1st draft ready by:</td></tr>
<tr><td>D7.2.2) Second report on data collection process: Report on data collection progress, inclusive of analysis of patient data on the basis of inclusion/exclusion criteria and updating of clinical features.</td><td>month 36
Lead:
1st draft ready by:</td></tr>
<tr><td colspan="4">Self-Assessment criteria</td></tr>
<tr><td colspan="2">Measurement process and units:</td><td colspan="2">Indicators [Upper and lower limits associated with WP objectives and measurement units]</td></tr>
</table>

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T7.2.2 Rheumatology In order to analyze the taxonomic gut content of JIA patients, a targeted approach based on sequencing of the variable regions V1 and V3 of 16S rRNA locus will be used. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. Fecal samples will be collected and analysed at onset of disease, at time of clinical remission, and during disease flares.	PUTIGNANI **Estimated % realization** M3 M6 M9 M12	**D7.3.1) First report on sample storage, DNA extraction and sample analysis processes:** First report on DNA extraction and analysis process, inclusive of metagenoma analysis, Operational Taxonomic Unit (OTUs) identification, and phylogenetic analysis of gut microbiota samples. [month 18]	Month 18 **Lead** PUTIGNANI **1st draft ready by:**
		D7.3.2) Second report on sample storage, DNA extraction and sample analysis processes: Second report on DNA extraction and analysis process, inclusive of metagenoma analysis, Operational Taxonomic Unit (OTUs) identification, and phylogenetic analysis of gut microbiota samples.	Month 36 **Lead** CHINALI **1st draft ready by:**

Self-Assessment criteria	
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**		**Deliverables**	**Deadline**
T7.3 DNA analysis. **7.3.2 Rheumatology** • Analysis of microbiome of fecal samples will be carried out following DNA extraction (see T7.2), with pyrosequencing using a 454 Junior apparatus and sequence analysis; comparison will be performed with therecently developed MEGAN 4 software (available at http://www-ab.informatik.unituebingen.de/software/megan)in order to identify the microbiota operational taxonomic units (OTUs).	PUTIGNANI		**D7.4) Report on integration in the Infostructure:** Report on the integration of all genetic and meta-genomic input into MD-Paedigree's model-driven infostructure	Month 36
	Estimated % realization			**Lead**
	M3			PUTIGNANI **1st draft ready by:**
	M6			Month 36
	M9			**Lead**
	M12			CHINALI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP12: Models validation, outcome analysis and clinical workflows				
Tasks	**Lead**		**Deliverables**	**Deadline**
T12.1.3 Clinical assessment of musculoskeletal biomechanics models (JIA and NND). • In two of the target clinical applications, JIA and NDD, we shall explore the use of complex multiscale biomechanical models of the musculoskeletal system personalised for each patient using as much as possible of the information available from medical imaging, molecular imaging, and gait analysis. • We shall also establish appropriate reference framework to make possible correlative explorations between clinical signatures of the disease that can be quantified using clinical, imaging, or instrumental assessment, and the prediction of the biomechanical models, as a support for the ethiopatological speculation (JIA) and a more effective scoring of the disease severity and for treatment planning (NND). • But before we can use the predictions of these models, we need to conduct an extensive clinical validation on the various elements that form them. • Medical imaging protocols will be tightly controlled, and period quality assessment conducted on all systems in use for the project, with particular reference to spatial calibration, and densitometry calibration for x-ray imaging. • To validate the fusion of imaging and gait analysis data superficial skeletal landmarks such as knee epicondyles will be located both by palpation in the gait lab and on the MRI images, and used to verify the accuracy of fiducial registration with the skin markers. • All image processing and image modelling methods will be tested using an alternative source of information, typically CT scans to validate bone reconstruction, etc. In particular DTI processing for cancellous bone will be validated on a small cohort patients recruited at USFD, that are undergoing wrist or ankle HRpQCT, which provides a very detailed information of the bone tissue spatial organization. Some of these patients will be examined also with the MRI at the same site, and the tissue orientation computed from DTI-like processing of the MRI images, to be verified against the HRpQCT data assumed as true value.	PONGIGLIONE **Estimated % realization**		D12.1) Outline of the clinical assessment and validation criteria for all four disease areas: Preliminary analysis of the clinical assessment and validation criteria	Month 18 **Lead:** PONGIGLIONE **1st draft ready by:**
	M3		D12.2.1) First clinical assessment and validation results for all four disease areas: Periodic update at month 24 of clinical assessment and validation outcomes	Month 24 **Lead:**PONGIGLIONE **1st draft ready by:**
	M6			
	M9			
	M12		D12.2.2) Second clinical assessment and validation results for all four disease areas: Periodic update at month 36 of clinical assessment and validation outcomes	Month 36 **Lead:** PONGIGLIONE **1st draft ready by:**
			D12.2.3) Third clinical assessment and validation results for all four disease areas: Periodic update at month 48 of clinical assessment and validation outcomes.	Month 48 **Lead:** PONGIGLIONE **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T12.2.3 Clinical workflows in JIA. • Clinical workflow for JIA will describe the sequence of operations that start with clinical data acquisition and by using our models ends with a clinically useful outcome predictors which are crucial to personalise treatment strategy. • The clinical workflow will be subdivided into 4 specific steps: a) acquisition of clinical, structural and functional information, b) integration of all information into a model, c) similarity search through the digital repository, and d) personalised prediction of disease outcome and optimization of individualized therapy. Data will be acquired through the use of validated standardised clinical measures of disease activity and damage. • Immunological and genetic and metagenetic data will be collected at the disease onset, during remission and at disease flare. • Imaging analysis will include plain radiography, MRI, US and Dxa. Imaging information will be integrated with the results of gait analysis and clinical evaluation in order to build an articulated joint model and a biomechanical model able to predict the mechanical stresses and strains induced in the various joint tissues. • The prognostic value on an individual level of multidimensional data including modern imaging modalities, immunological, meta-genetic data, as well as articulated models and biomechanical models will be explored. • The model will work across scales from molecular function through to joint function, scaled-up to whole body, and incorporated into changing lives. The risk-benefit ratio will be measurable and incorporated into the model and the clinical decision-making process. The model will allow individual planning of interventions with subsequent consequences for functioning and quality of life of the affected children. • JIA constitutes an ideal domain for assessing the merits of simulators and predictors based on data generated across different scales. The validity and effectiveness of the proposed solutions will be assessed by using the model to address several open issues in JIA with a strong clinical impact on early diagnosis, prediction of disease and of treatment outcome.	PONGIGLIONE **Estimated % realization** M3 M6 M9 M12	**D12.3) Improved clinical workflows and outcome analysis:** Final proposal of innovative clinical workflows based on outcome analysis of all patient cases	Month 48 **Lead** PONGIGLIONE **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP18: Dissemination & Training

Tasks	Lead	Deliverables	Deadline
18.3 Training Training is considered to be a fundamental task in dissemination. As anecdotal evidence has confirmed via WP4 of the VPH NoE and via feedback from the DISCIPULUS ('Roadmap Towards the Digital Patient') meeting (30/03/2012; Barcelona), training is recognized to be one of the most solid and long-lasting dissemination strategies in place. The training activities within MD Paedigree will consist of 2 'hands-on' workshops to be delivered during years 2 and 4 of the project (at approx. 1 or 1.5 year interval) in order to expose the outcomes achieved both, in disease modelling and in building the infostructure, highlighting the potential for change management and innovation in clinical workflows to the medical/clinical and research community interested in VPH technology. The first workshop will also seek to provide feedback to the research and development activities, so as to refine the outcomes for the final workshop. The workshop participants will fill in a detailed feedback questionnaire that will be passed to the developers. This task will be led by UCL, which has a long-standing commitment with the VPH Community and is involved in several training grants, including the Marie Curie ITN 'MeDDiCA', 'VPH-MIP' and WP4 of the VPH NoE.	DIAZ **Estimated % realization** M3 M6 M9 M12	**D18.3) Training event in year 2:** Report on the outcomes of the first Training event **D18.6) Training event in year 4:** Report on the outcomes of the second Training event	Month 30 **Lead** DIAZ **1st draft ready by:** Month 42 **Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T18.4 Seminars, Workshops, Concertation Activities with Other ICT Funded Projects, and Scenario Analysis Sessions **Lead: Vanessa Diaz** The Consortium will identify the most relevant conferences in the area and propose seminars and workshops to be held during these events. It will devote special attention and resources to Concertation Activities with other ICT funded projects and to targeted dissemination actions. Special "Scenario analyses" sessions will be convened, involving the key personnel from both the clinical and the technological partners, with the aim of pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users within MD-Paedigree. The results of the previous workshops will be presented to the Scientific Committee and to the Users' Board in order to assess their relevance and applicability, so as to refine the outcomes for a validation workshop and for a final MD-Paedigree Conference, to be held at the end of the project, targeting both internal and external clinical and research communities as well as patient organisations and the interested media. The participation in any such event will be reported in the periodic reports and the final report.	DIAZ **Estimated % realization** M3 M6 M9 M12	**D18.4.1) First scenario Analysis Sessions:** First scenario Analyses pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users. **D18.4.2) Second scenario Analysis Sessions:** Second scenario Analyses pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users.	Month 24 **Lead** DIAZ **1st draft ready by:** Month 42 **Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead	Deliverables	Deadline
T18.7 Engaging Parent and Patient Associations **Lead: Vanessa Diaz** Approaching Parent and Patient associations will become a part of the consortium's dissemination activities. The project will seek to disseminate news of its work, expected results and potential future developments through these channels. It is hoped that the work with Patient associations will help achieve a larger bidirectional knowledge sharing base of clinicians and of patients, and further inform the potential beneficiaries of the ongoing work.	DIAZ **Estimated % realization** M3 M6 M9 M12	**D18.1) Dissemination and training strategy plan and preliminary materials**: Roadmap defining the dissemination and training strategy, indicating the subsequent choice of preliminary materials	Month 12 **Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP19: Exploitation, HTA, and Medical Device Conformity				
Tasks	**Lead**		**Deliverables**	**Deadline**
T19.1: Evaluation approach and meaningful indicator development (EMP) • Develop upon and adapt in the VPH and other contexts proven approaches, methods and tools to the specific environment and objectives of this workpackage • Establish a set of meaningful criteria and their measurement process that are robust to demonstrate socio-economic benefit-cost impacts. The focus is • to approach and find measurements for evaluating how virtual collaborations between members of the VPH communities with different expertise are facilitated and • how consequently the uptake and acceleration of model development and integration can find meaningful expression in the overall evaluation framework.	STROETMANN **Estimated % realization** M3 M6 M9		**D19.1 HTA evaluation framework** It reviews proven approaches, methods, and tools which might be relevant to the specific environment and objectives of this workpackage, and establishes a set of meaningful criteria and their measurement process, thereby focusing on evaluating how virtual collaborations between members of the VPH communities with different expertise are facilitated.	Month 12 **Lead** STROETMANN **1st draft ready by:**

	M12			

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**		**Deliverables**	**Deadline**
T19.3: Benefit-cost scenario for clinical impact assessment (EMP) In a separate task a high-level, generic benefit-cost scenario for clinical impact assessment will be applied, with the ultimate goal to generate economic and market evidence for true translational medicine. The benefit-cost scenario will be tested and initially validated with preliminary, exploratory data estimates from the patient-centred workflows that are the basis of the digital repository and Infostructure. The two main dimensions pertaining to clinical/health impacts focus on the one hand on health service delivery and the health of patients, and on the other on public health/societal outcomes. To assess such impacts, the scenario development will integrate the following indicators: • Clinical effectiveness and patient-related outcomes • Safety (risks associated with applying the technology) • Organisational and change management aspects • Human resource implications, knowledge & education needs • Assessing contributions to the VPH vision of a patient avatar	STROETMANN **Estimated % realization** M3 M6 M9		**D19.4 Clinical impact assessment scenario** Initial formative evaluation of MD-Paedigree model-driven Infostructure based on a benefit-cost analysis approach, subsequently followed by a generic benefit-cost scenario for clinical impact assessment developed and validated with partners and experts. [month 36]	Month 36 **Lead** STROETMANN **1st draft ready by:**

• Efforts for application (convenience/ease of use; costs for introduction of new technology) The indicators assessed ultimately prepare for a more targeted and strategically aligned exploitation activities (T19.4) by proving clinical impact of MD-Paedigree with respect to: • the state-of-the-art in paediatric patient-specific computational modelling, • improved disease understanding and therapy outcomes that can be applied to both clinical routine and translational clinical research, • usability by clinicians and clinical researcher, • transferring technical workflows into clinical workflows, • the vertical integration of multi-scale patient data and the provision of models, tools, and services readily available to clinicians at the point of care.	M12			

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

A.1.1.3 Timing of work packages and their components

The MD-Paedigree project partners have formalized a work plan implementing 4 major phases implying a number of conceptual steps, over 48 months of activity with 4 major milestones. The first milestone is due after 9 months and marks the end of the specification phase; the following milestones are aligned with the reporting periods of the project every 12 months.

Phase 1 (running from month 1 to 9) – Project Set-up, Requirements Elicitation, and Clinical Protocols: During Phase 1 quality assurance guidelines and a self-assessment plan will be prepared, ethical approval will be obtained , and the first dissemination activities will be performed (Step 1) Furthermore, clinical protocols for the selected paediatric applications will be established (Step 2).)Finally, the requirements for models and infostructure implementation will be analysed and documented from an end user standpoint (Step 3).

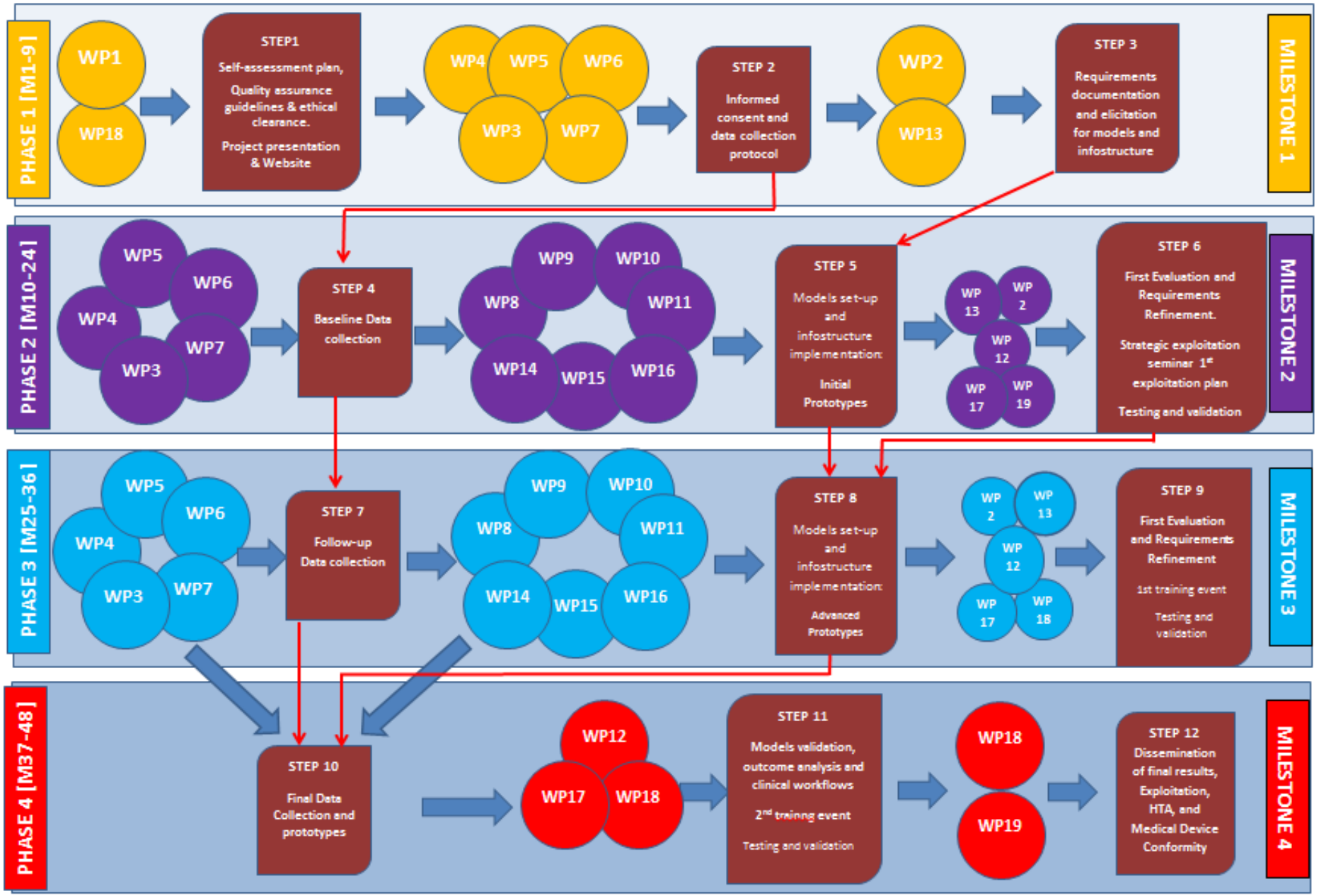
***Phase 2 (running from month 10 to 24) – Baseline Data Collection, Initial Prototypes, First Evaluation and Requirements Refinement*:** Patient enrolment will take place and data acquisition will be started (Step 4). Based on the established requirements, the existing models from Health-e Child and Sim-e-Child projects will be refined and adjusted to the new applications. The open repository for project infrastructure will be introduced and initialized with the current models and data (Step 5). First evaluations will be undertaken and requirements will be refined based on the collected experience; additionally, during this phase, the Strategic Exploitation Seminar will be held and the 1st Exploitation Plan will be drafted (Step 6).

***Phase 3 (running from month 25 to 36) – Follow-up Data Collection, Advanced Prototypes, Evaluation and Requirements Update*:** Follow-up or additional data will be acquired for all clinical applications (Step 7). The respective models will be enhanced to process longitudinal data and refined according to the obtained evaluation results. New functionalities will be integrated into advanced prototypes. The open repository will be improved and updated with content (Step 8). A second set of evaluations will be conducted and requirements will be adjusted for the final system. Furthermore, the 1st Training Event will be held (Step 9).

***Phase 4 (running from month 37 to 48) – Final Data Collection and Prototypes, Clinical Validation, and Deployment*:** In the final year, data collection will be concluded and the clinical validation will take place with the final models and simulation framework (Step 10). Results will be used to propose and disseminate improved clinical workflows. Subsequently, the 2nd Training Event will be held (Step 11). Models for all clinical applications and their respective evaluations will be documented and disseminated, while the implementation plan will be refined and the Health Technology Assessment and the Medical Clearance preparatory activities will be performed (Step 12) .

The timely delivery of all planned deliverables will be the first indicator of the fulfillment of each phase in the expected progress of MD-Paedigree, monitoring what can be demonstrable at each corresponding milestone of the project.

A second and much more detailed means of verification will be provided by the assessment criteria for each milestone and each WP which are to be defined within D1.3 Self-assessment plan on month 3.


PHASE 1 [M1-9]
WP1
WP18
STEP1
Self-assessment plan,
Quality assurance guidelines & ethical clearance.
Project presentation & Website
WP4
WP5
WP6
WP3
WP7
STEP 2
Informed consent and data collection protocol
WP2
WP13
STEP 3
Requirements documentation and elicitation for models and infostructure
MILESTONE 1
PHASE 2 [M10-24]
WP5
WP6
WP4
WP7
WP3
STEP 4
Baseline Data collection
WP9
WP10
WP8
WP11
WP14
WP15
WP16
STEP 5
Models set-up and infostructure implementation:
Initial Prototypes
WP 13
WP 2
WP 12
WP 17
WP 19
STEP 6
First Evaluation and Requirements Refinement.
Strategic exploitation seminar 1st exploitation plan
Testing and validation
MILESTONE 2
PHASE 3 [M25-36]
WP5
WP6
WP4
WP7
WP3
STEP 7
Follow-up Data collection
WP9
WP10
WP8
WP11
WP14
WP15
WP16
STEP 8
Models set-up and infostructure implementation:
Advanced Prototypes
WP 2
WP 13
WP 12
WP 17
WP 18
STEP 9
First Evaluation and Requirements Refinement
1st training event
Testing and validation
MILESTONE 3
PHASE 4 [M37-48]
STEP 10
Final Data Collection and prototypes
WP12
WP17
WP18
STEP 11
Models validation, outcome analysis and clinical workflows
2nd training event
Testing and validation
WP18
WP19
STEP 12
Dissemination of final results, Exploitation, HTA, and Medical Device Conformity
MILESTONE 4

JIA					
March 2013	**April 2013**	**May 2013**	**June 2013**	**July 2013**	**August 2013**
	Protocols delivered to Ethical Committee	D7.1 Recruitment protocol with ethical clearance: Completion of the recruitment protocol, consensus and ethical clearance from all partners' involved in patient recruitment	D5.1 Report on data collection protocols and parents and patients informed consents	Interviews to prepare D2.1	First Half-Yealry report. Delivery date
					Self-Assessment Plan
					Check of the enrollment and data collection, analysis and processing
			Contribution to the Self-Assessment Plan		
	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs
		Area Dedicated T&M TC [1st May]	Area Dedicated T&M TC [5th Jun]	Area Dedicated T&M TC [3rd Jul]	Area Dedicated T&M TC [7th Aug]

September 2013	**October 2013**	**November 2013**	**December 2013**	**January 2014**	**February 2014**
Biannual area meeting	Check of the enrollment and data collection, analysis and processing.	First draft of the deliverable **D2.1**		Internal Review	D2.1 Initial requirements analysis document including priorities for the implementation
					First periodic review
Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs
Area Dedicated T&M TC [4th Sep]	Area Dedicated T&M TC [2nd Oct]	Area Dedicated T&M TC [6th Nov]	Area Dedicated T&M TC [4th Dec]	Area Dedicated T&M TC [8th Jan]	Area Dedicated T&M TC [5th Feb]

JIA	**DELIVERABLES WITHIN MONTH 24**	
	D5.2) Report on baseline data collection status	M16
	D10.1) Report about initial modelling results	M24
	D7.2.1) First report on data collection process	M18
	D7.3.1) First report on sample storage, DNA extraction and sample analysis processes	M18
	D12.1) Outline of the clinical assessment and validation criteria for all four disease areas	M18
	D12.2.1) First clinical assessment and validation results for all four disease areas	M24

Neurological and Neuro muscular Diseases (NND)

MD-PAEDIGREE KICK OFF MEETING

NND WG

Participant's Name	Affiliation

Concept (general)

MD-Paedigree validates and brings to maturity patient-specific computer-based predictive models of various paediatric diseases

- increasing their potential acceptance in the clinical and biomedical research environment
- making them readily available not only in the form of sustainable models and simulations, but also as newly-defined workflows for personalised predictive medicine at the point of care.

These tools can be accessed and used through an innovative model-driven infostructure

- powered by an established digital repository solution
- able to integrate multimodal health data
- entirely focused on paediatrics
- conceived of as a specific implementation of the VPH-Share project, planned to be fully interoperable with it and cooperating, through it, also with p-Medicine.

MD-Paedigree aims at achieving high-level semantic interoperability,

- requiring standards enabling the clinical contents to be interpreted consistently across the different EHR regimes,
- while complete clinical interoperability between systems will require widespread and dependable access to maintained collections of coherent and quality-assured semantic resources,
- including models that provide clinical context,
- mapped to interoperability standards for EHR and PHR and biomedical data, linked to well specified terminology value sets, derived from high quality ontologies

CONCEPT (SPECIFIC)	**Beyond the state of the art**	**WPs' OBJECTIVES**	**Lead**	**Estimated % realisation**
Conventional clinical gait analysis (CGA) is already an important tool in the treatment of children with CP that aims to improve or sustain walking performance, but its potential is under-utilised and recent developments need full exploration. For ambulant SMA patients, new methods for functional	**Protocol definitions for clinical gait analysis** The potential of gait analysis to serve clinical decision making in NDD is generally under-used for several reasons. These will be taken up within the MD-Paedigree project. Established and clinically authorised protocols (technical, markers and procedures) of CGA will be an important step forward for the NND paediatric care in the EU, along with the establishment of a reliable MD-Paedigree database for typically developing children. Three levels of protocol definitions are needed to assure multicentre reliable data for the repository: **Technical Quality assurance for CGA laboratories**	**WP2: Clinical and technical user requirements for disease modelling** • Incorporate into the model the variables that are analysed by the clinicians in their activity. • Ensure that the modeling reflects real clinical needs and is validated against them		

motor evaluation based on gait modelling would allow to increase sensitivity to change in assessing weakness and fatigability. In the last few years, following a rapidly increasing number of potentially effective therapeutic approaches for DMD, the request for validated and sensitive outcome measures to be used in clinical trials has increased. Walking implies a complex involvement of inputs from several senses (visual, vestibular, proprioceptive, somatosensory), partly automated by the so called spinal central pattern generator (CPG). These inputs are known to interact with each other, but the way in which this is performed is not fully exploited at present. Nevertheless, the current insights are certainly at an advanced state that allows for meaningful application towards pathological walking, where decision	It is important to realise that for accurate data from the experimental systems a strict analysis of causes of errors and periodical validation procedures needs to be implemented in the gait labs. If the adopted experimental procedure permits the gathering of valid data, the first important prerequisite for reliable and accurate results from a particular subject is fulfilled. Within MD-Paedigree these quality assurance (QA) procedures will therefore be formalized between laboratories for clinical gait analysis. MD-Paedigree will constitute a European standard for technical QA and have this approved by the important European bodies on clinical gait analysis, i.e. the ESMAC. A consensus meeting will be part of this. **Standardisations of gait analysis protocols: Marker placements** One of the main non-technical sources of error in CGA using OptoElectronic Movement Analysis systems is caused by marker artefacts, resulting from skin movement relative to the bone. In the case of well-trained staff, errors due to marker placements errors and skin movement artefacts will stay within a few degrees of error of the joint kinematics graphs. This error level is considered to be just clinically acceptable. This means that all gait labs should fulfil the requirements to be qualified for MDPaedigree graded gait analysis. In analogy with the Technical Quality Assurance (TQA), MD-Paedigree will strongly promote interoperability and constitute a protocol for standardised marker placement, as well as standard procedures to evaluate this within and between laboratories. In parallel, we shall explore the possibility to use imaging/gait analysis protocols, where patients are dressed with radiopaque/MRI opaque and reflective markers attached to the skin as used in gait analysis protocols, while the imaging protocol is conducted. These data will make possible to use sophisticated inverse kinematics modelling methods to minimise the skin artefacts, and to obtain accurate estimations of the skeletal kinematics.	to assure their robustness and reproducibility. • Provide computational models that can be personalized by adapting the parameters to the integrated data of a patient case • Advance the knowledge about the selected diseases by allowing the simulation of different effects on the evolution of the disease • Predict the effect of therapy. • Ensure that MD-Paedigree models have the highest possible impact at the point of care. • Re-use of models between disease areas to leverage synergies where possible. • Existing standards for modelling and tools will be investigated. • The need for new standards will be evaluated and documented. **WP6: Data acquisition and processing for NND** • To clinically authorize		

support is needed. In the clinical practice of specialised centres, CGA is used to evaluate the joint and muscle functions in their functional context, i.e. during gait. Common CGA measures 3D kinematics (by 3D optoelectronic registration of skin mounted markers). Each relevant degree of freedom (DOF) is expressed as a function of the gait cycle. Moreover, using a mass distribution model and measuring ground reaction forces, the net moments for each DOF are calculated using inverse dynamics analysis. Muscle activation patterns, for all relevant muscles, are measured using electromyography (EMG) for each targeted muscle. Finally, the energy cost of walking can be evaluated using metabolic measurements. Unfortunately, the output of CGA is not yet in a format that permits clear,	**Standardisations of gait analysis protocols: operational protocols** The results of kinematics and kinetics of CGA are also dependent on the use of standard protocols for instruction on walking targets. In particular, the enforcement of a precise walking speed is of major influence on the output. As such, instructions should be carefully standardised and protocols developed that use multiple walking speeds. It has been suggested and shown by previous studies, that these protocols are necessary to detect important pathological features of the NMSS of the subject, especially in patients with CP. EMG recordings and oxygen consumption will be part of the overall assessment procedures. Moreover, in order to feed the development of probabilistic models a standardised description of therapies will be completed. This description will be used to longitudinally describe the applied clinical workflows that are currently used to improve gait performance in children with NND. **Neuro-Musculo-Skeletal models and clinical gait analysis** For clinical gait analysis the use of Neuro-Musculo-Skeletal (NMS) models is an important step forward in the interpretation of its results, aiming to inform the clinical decision-making. Because of the modelling based interpretation, the physician no longer needs to interpret the results of clinical gait analysis, within his own informal frame of interpretation. Using NMS models the results of CGA are quantitatively "translated" into the function and performance of the underlying structures, i.e. muscle activation, muscle forces, and joint loads that make possible to unravel the aetiology of the pathological gait pattern of the subject under study. The EU project “Personalised models of the Neuro-Musculo-Skeletal Physiome” (NMS Physiome) is moving towards the development of PPI (Predictive, Personalised and Integrative) musculoskeletal medicine. A key result of this project is the integration of an advanced software application	(technical-, marker- and procedures-) protocols used for gait analysis, data collection and quality assurance. • Establish reliability levels for gait analysis protocols, that provide data for modelling. • To acquire sets of data (gait analysis and images) for the repository, probabilistic modelling and biophysical modelling.		
		WP11: Modelling and simulation for NND To construct an accurate personalised biophysical model for the paediatric population , driven from the needs in clinical practice to estimate muscle forces and joint loads in the gait of NND populations. Specifically, this involves: • Extraction of subject-specific bone and muscle anatomy from DXA and MRI images • Development of novel, accurate scaling methods for musculo-skeletal modelling		

unambiguous interpretation, because of the redundancy of the Neuro-Musculo-Skeletal System (NMSS) which obstructs distinguishing cause from compensation. Even though recent developments in modelling the NMS Physiome as a part of EU funded Virtual Physiological Human efforts are at an advanced state, their results have not yet been implemented in clinical practice, and the full potential of CGA still needs to be reaped. A combination of standard protocols of gait analysis, biophysical modelling and large scale statistical analysis can therefore be expected to provide a powerful framework for meaningful interpretation.	for the pre-processing of imaging and gait analysis data into a full musculoskeletal model (NMS Builder) and the OpenSIM musculoskeletal modelling environment developed by Stanford University. NMS Builder is already available in prototypical form to all partners of the MD-Paedigree consortium. **Applying to paediatrics the HBM model and the NMS Physiome** Although NMS computational models are thus well known in the biomechanical research community, as yet only one company, MOTEK, has incorporated gait analysis and model based interpretation of gait for market delivery. Their model (the HBM model) is computationally very efficient: even without high performance computers it can run in real time. More complex modelling activities can be conducted using the NMS Physiome tools. The actual problem of accuracy of NMS models is that all models currently used in paediatric gait analysis are based on data scaled from a single cadaver in a simple way. Sensitivity studies have shown that such a gross simplification in applying generic models is too inaccurate, and, especially in the case of children, dedicated and validated models, fused with medical imaging data, should be developed in order to yield reasonable accuracy for clinical application in this population. **Mass distribution model of body segments** The first level of MS models in CGA is the mass distribution model of body segments. Mass distribution means that the masses, centre of mass and inertial properties of each segment need to be known for accurate calculation of inverse dynamics resulting in valid joint kinetics. What is needed is a method for scaling that allows application, in clinical workflows, to enable personalised medicine. MDPaedigree will develop and evaluate a scaling method for the NMSS of children, to be applied in existing NMS models that are used in CGA.	• Adaption of existing musculoskeletal model to subject-specific and pathology specific data • Design of models driven by the dynamics of gait perturbations **WP7 Genetic and metagenomic analytics** To evaluate the role of genetic (assessed by disease-gene or candidate gene analysis) and metagenome (based on gut microbiota profiling) profiles on the development and progress of diseases and on their outcome. **WP12: Models validation, outcome analysis and clinical workflows** • To clinically validate derived models • To improve prediction of outcome and risk stratification • To establish integrated clinical workflows and personalised		

	Validation will be based on MRI measures. **Whole-body Dual X-ray Absorptiometry** Next to anthropometrics scaling is the alternative to use a 2D image, generated by a whole body DXA image, morphed to a generic 3D skin model of a child. The advantage is that DXA provides accurate measurement of the areal density of the bone, fat, and lean tissues the inertial properties of each segment. **Subject specific bony deformities** The second level of personalised MS models in CGA are to account for the subject specific bony deformities. The bony deformities that should be accounted for can be limited to the clinically well known deformities in CP. These deformities have significant influence on the output of NMS model calculations (i.e. femoral anteversion and tibial torsion). These effects could primary be modelled by morphing the generalised bony structures towards the actual morphology of the bone. The most important effects of bony deformities should be parameterized by the effects on axis alignment: i. introducing a skewness of the principal axes of rotation of the joints in the kinematic chain of linked segments, ii. the altered lever arms of muscles with respect to these principal axes of rotation of the joint. Again antropometric measures and DXA will be explored. **Pathology specific muscle parameters** The third level of personalised modelling is to account for pathology specific muscle parameters. These models should focus on the parameters that are known to be of large influence on the second step in inverse dynamics, i.e. the estimation of muscle forces based on optimisation criteria on how to explain the net joints moments from CGA.	treatment models		

	This means that especially muscle contractures, altered muscle structure and hypertonia (in CP) as well as muscle weakening (in DMD and SMA) must be targeted. US measures of the muscle belly, along with fibre directions will enable estimates of the muscle Physiological Cross sectional Area (PSCA), while dynamometric evaluations will yield measures of muscle belly length and optimal fibre length. Supporting probabilistic models, despite the strong potential of biophysical models of the NMSS, will only hold a certain amount of predictive value, i.e. as far as their assumed accuracy will allow. **Generating decision rules from dataset** However, in clinical practice, even if the pathology cannot be fully explained by biophysical modelling, the use of probabilistic models is still extremely powerful in supporting clinical decision making. Until now only two gait laboratories in the world (Gillette Children's, Minneapolis, US and Pellenberg, Leuven, Belgium) have explored the possibilities of generating decision rules from their dataset. These laboratories are the only ones that have created a large enough set of reliable data to make such an effort worthwhile. In MD-Paedigree the clinical partners will collect data, according to the dataset and quality protocols defined on the basis of standardised formats, for feeding into the repository.			

Application Scenario

François is a 10 year old boy with cerebral palsy (CP). The medical records of the milestones of his motor development show a consistent significant delay with respect to his peers with normal growth. His current level of the so-called Gross Motor Function Classification (GMFCS) is 2, meaning that he can walk unsupported but has difficulties while walking outside. Actually his complaints are frequently falling on the playground, and very limited walking distance, due to early fatigue. Moreover, his physiotherapist, who has been specifically trained and has ample experience with children with CP, is concerned whether François' walking pattern will deteriorate in the coming future, resulting eventually in wheelchair dependency.

François is referred to a specialist paediatric centre of rehabilitation medicine in Nantes, where his walking pattern is analysed in the gait laboratory. A complete recording of his gait pattern, using 3D kinematics (i.e. in fact 4D), joint kinetics and muscle activation patterns, is performed along with metabolic measurements of the energy cost of walking (ECW).

The physiatrist in charge of interpreting the results of the gait analysis of François concludes that hyperactivation of the calf muscles (m. gastrocnemius) is present, while at the same time this young boy is walking with slightly flexed knees during stance. This positioning causes compensatory hyperactivation of muscles at other levels, resulting in an overall increased ECW. The analysis is clear, but now the physiatrist should decide on the therapy. Unfortunately he can only rely on scattered information derived from some single cases, he learned to know at some courses, rather than explicit design rules. Current knowledge tells him that chemo denervation of the calf muscles (using local Botox injections) should normalize its hyperactivation, end hence compensatory activation and decrease the enhanced ECW. However, it would also contribute to a higher knee flexion moment, that would drive the knee in further flexion during stance, resulting in a further worsening of his walking, an outlook the therapist is afraid of. This latter phenomenon alone would call for another therapy, i.e. stiff carbon ankle-foot orthoses, that is known to be effective to counteract knee flexion in stance. So now the physiatrist is faced with the dilemma of what to do.

It is exactly this kind of clinical decision making problem that profits from being informed by multiscale reusable models to be built upon the results of the Health-e-Child and the Sim-e-Child projects. This means (a) the development and use of a paediatric musculoskeletal model applied to gait, and (b) building a repository of many clinical cases, based on standardised gait analysis protocols, that generates decision rules derived by probabilistic modelling.

Fortunately, the gait laboratory of Nantes has committed itself to apply the EU standards of clinical gait analysis and became a registered user of the Model-Guided European Paediatric Digital Repository since 2016. An additional DXA scan is made, and all information regarding the case of François is uploaded into the system. In return, the disease modelling analysis shows that chemo denervation of the calf muscles is likely to solve the problem. However, the effect on knee flexion remains undecided, within acceptable model accuracy, for this particular case. Running the probabilistic model supports the treatment choice, and points towards two matched cases from KU Leuven. In those cases, chemo denervation of the calf muscles with additional intensive physiotherapy to prevent enhanced knee flexion proved to be successful. The physiatrist is now confident that François will profit from the treatment as indicated on the short and long term.

WP2: Clinical and technical user requirements for disease modelling

Tasks	Lead	Deliverables	Deadline
Task 2.1: Conduct interviews with the clinical and technical partners to obtain a complete list of requirements for the disease modelling that will ensure its usefulness within and beyond the project. All WP Leaders will actively contribute to the requirements documentation while they ensure that the respective WP partners are interviewed. 5. Prioritisation criteria: All requirements will be prioritised ensuring that from the start the most important aspects will be implemented to quickly ensure an operational system. 6. Schedule of requirements updating: The requirements list will be continuously updated on a regular basis such that main requirements and system constraints will be released as deliverables.	CHINALI **Estimated % realization** M3 M6 M9 M12	**D2.1 Initial requirements analysis document including priorities for the implementation.** Initial requirements analysis document including priorities for the implementation: Complete interviews with the clinical and technical partners will be collected to obtain a list of variables and requirements for the disease modelling. Requirements will be prioritized ensuring that from the start the most important aspects will be implemented first.	Month 12 **Lead** CHINALI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP6: Data acquisition and processing for NND

Tasks	Lead	Deliverables	Deadline
T6.1 QA on data collection and clinical protocols [M 1-18] • Task T6.1 will start with a complete description of the protocols used in the clinical institutes, which is the base for a common descriptive format and its default values. Three levels will be used: o Technical Quality assurance (TQA) protocols in Gait analysis laboratories o Marker placement protocols (MPP) in 3D Optoelectronic CGA o Operational protocols and workflow (OPWF) used in clinical practice • Then a survey will be setup by the partners. This survey will be taken from CGA laboratories in EU, based on the network, provided by ESMAC (European Society of Movement Analysis in Adults and Children). • The results from the survey will constitute a complete EU inventory on the protocols (TQA, MPP, OPWF) used in Clinical Gait Analysis CGA. • A Consensus Proposal for EU CMA gait labs for all three levels will be drawn up • For the TQA and MPP, the clinical partners will perform reliability measures of the protocols, to assure quantitative levels of reliability. • In parallel with the inventory, a dataformat (syntaxis & semantics) will be defined **Partners involved: VUA, OPBG, KU Leuven, URLS**	HARLAAR **Estimated % realization** M3 M6 M9 M12	**D6.1) CGA standard protocol** A standard protocol of clinical gait analysis is described based on a representative inventory along the EU , and describes both clinical and technical procedures.	Month 18 **Lead** HARLAAR **1st draft ready by:**
		D6.2) A standard protocol of clinical gait analysis is described based on a representative inventory along Standard minimal dataset of clinical gait analysis outcome measures and associated context parameters needed for data exchange and modelling.	Month 24 **Lead** HARLAAR **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	Indicators [Upper and lower limits associated with WP objectives and measurement units]	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T6.2 Gait analysis collection for CP [M 1-36] • Gait analysis data will be provided to the work packages that are involved in biophysical and probabilistic modelling. • A complete dataset related to clinical gait analysis consists of: o A standardised anamnesis o Standard clinical testing: Physical Examinations and Tests; Questionnaires o Xray s if applicable o From gait analysis: • Kinematic data; • Kinetic data; • EMG Data; • O2 Data. o Contextual data, like treatments received • KULeuven will provide 400 sets of data (O2 data for 100 patients only) from its current database, followed up by another 200. • OPBG will provide 200 sets of data (kinematics, kinetics, and electromyography) from its current data base. • Data Quality checks will be performed for each subject. • Complete data sets of gait analysis in Cerebral Palsy (CP) will be acquired, to serve as an input for the biophysical and probabilistic modelling. • Criteria for selection are based on children with CP that are routinely measured in the gait lab: classified as GMFCS 1-3 ; diplegic or hemiplegic; sufficient cognitive skills; without relevant visual deficit; and older than 6 years. 1. Complete data sets of 10 CP patients for each clinical center (VUA, OPBG, KULeuven) will be provided for biophysical modelling. 2. For the probabilistic modelling, as many as the clinical load would allow, can be included, the aim is 50 patients per center (VUA, OPBG, KULeuven) before month 36.	HARLAAR **Estimated % realization** M3 M6 M9 M12	**D6.3) Report on the collection of 130 CP patients clinical gait dataset** A clinical gait dataset according to defined standards of 130 CP patients reprocessed form existing databases (100) and new measurements (30)	Month 36 **Lead** HARLAAR **1st draft ready by:**

Self-Assessment criteria	
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T6.3 Gait analysis collection for DMD and SMA [M 12-48] the protocols developed in T6.1 apply for Duchenne Muscular Dystrophy (DMD) and Spinal Muscular Atrophy (SMA) • For SMA clinical data will be collected by o OPBG, o KU Leuven o and VUA • from 20 ambulant patients (severity grade type 3); • 10 patients will be selected among the 3a subgroup (symptoms of weakness appearing before age 3 years), • 10 patients will belong to the 3b group (weakness appearing after the age of 3 years). • Besides considering type of severity in the selection of patients for data analysis, we will include children having an age range of 5 to 10 years of age. • Particularly, we will recruit children of 5-6 years with the diagnosis of SMA type 3a and children with age range of 5-10 years with SMA type 3b. • All patients will receive a longitudinal full control evaluation at baseline (0), after 12-18 month (1) and 2-3 years(2). **Measurements:** • Functional motor scales: Expanded Hammersmith functional motor scale to measure strength 6 minutes walk test to measure strength and fatigue, hand held myometer (CITEC) to measure strength (knee flexors and extensors) • Gait analysis according to protocols T6.1	HARLAAR **Estimated % realization** M3 M6 M9 M12	**D6.3) Report on the collection of 130 CP patients clinical gait dataset** A clinical gait dataset according to defined standards of 130 CP patients reprocessed form existing databases (100) and new measurements (30)	Month 36 **Lead** HARLAAR **1st draft ready by:**

Compared to SMA, DMD is a rather a homogeneous disorder with well defined natural history endpoints, although the standardised use of steroid treatment and progress in standards of care has changed the natural history of the disease prolonging walking by 2 to 5 years, in relation to natural history data known before systematic steroid treatment, when patients generally lost walking ability between ages of 7-12 years.

- Clinical data will be collected by OPBG, KU Leuven and VUA from 20 ambulant genetically confirmed DMD patients treated with the same steroid regimen of daily deflazacort 0.75mg/kg/day and with the most common mutations in the dystrophin gene.
- Age range of patients will be between 5 and 7 years. In particular, we will recruit 10 patients with age between 5 and 6 years, and additional 10 patients with age between 7 and 8 years. In this second DMD group we will observe longitudinally the progression of the disease in the time span of 4 years, because it is known from current natural history data that DMD patients start a downhill progression of function after age of 7-8 years.
- All patients (10 from OPBG and 10 from KU Leuven) will receive a longitudinal full control evaluation at baseline (0), after 12-18 month (1), and 2-3 years(2).

Measurements:

1. Functional motor scales:
   - the North Star Ambulatory Assessment (NSAA)
   - 6 minutes walk test (6MWT) to measure strength and fatigue,
   - hand held myometer (CITEC) to measure strength (knee flexors and extensors)
2. Gait analysis according to protocols identified in T6.1
3. In addition OPBG, KU Leuven and VUA will acquire electrocardiographic and echocardiographic data from all the 20 DMD patients

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead	Deliverables	Deadline
T6.4 Image acquisition [M 3-36] • In WP 11 some advanced modelling is developed, that the fusion of multimodal sources of data (MRI, DXA and CGA). As an input to this WP, each clinical center (VUA, OPBG, KU Leuven) will acquire at least 10 subjects with both MRI and DXA, including the markers that are needed for gait analysis. • Volume of interest includes pelvis, femur, tibia, foot. The first three subjects should be acquired within the first year of the project. Images will have to be anonymiszed before making them available for the technical partners.	HARLAAR **Estimated % realization** M3 M6 M9 M12	**D6.4) A clinical gait dataset according to defined standards of 130 CP patients reprocessed form existing** A comprehensive clinical dataset of gait analysis data for CP, DMD and SMA and MRI and DXA data sets of 30 CP patients.	Month 44 **Lead** HARLAAR **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP11: Modelling and simulation for NND

Tasks	Lead		Deliverables	Deadline
T11.1 Construction of a scalable mass distribution model suitable for the paediatric population [M 1-48] • The aim here is to develop methods for scaling that may be based on simpler/cheaper methods than complete MRI investigation with semi-automatics interpretation. • The simplification consists of two levels: - scaling based on common anthropometric measures, when possible - scaling based on 2D DXA, if needed • MRI will be used as reference method to validate the two simplified approaches. • So the scaling will be based on full MRI, when absolutely needed. • Firstly we plan to obtain more accurate mass distribution models for musculoskeletal modelling of children and adolescents. In order to automatically analyse the involved image data, statistical shape models of the required bone and muscle anatomy will be built jointly by SAG and USFD, using readily available paediatric MRI scans of the lower limb. • SAG will use these models to automatically extract these structures from MRI data. • For this effort 100 complete MRI images, from retrospective analysis will be available from OPBG, VUA and KU Leuven. • SAG will use the resulting models to automatically extract these structures from MRI data. • An alternative approach to building models completely based on MRI images, is to use a generic 3D skin model of a child based on a database of 3D MR data, and morph it to the child-specific 2D image generated by a whole body DXA image (using multiple images acquired from different directions if required).	STEENBRINK **Estimated % realization** M3 M6 M9 M12		**D11.1) Automatic extraction method of mass distribution and muscle volumes** Automatic extraction of mass distribution and muscle volumes from DXA and MRI images and prepare it to be suitable as input for musculoskeletal models of children and adolescents. **D11.2) Development of novel scaling method** Development of novel, accurate scaling methods based on regression equations between the actual data (Deliverable 4.1), and simple anthropometric measures	Month 18 **Lead** STEENBRINK Month 36 **Lead** STEENBRINK

• The advantage is that the DXA provides accurate measurement of the areal density of the bone, fat and lean tissues. We will develop share regression models enabling to estimate muscle insertion points based on the information available from DXA based on DXA/MR regression learned from our training database. • When this information is integrated over the morphed 3D shape, we shall be able to estimate the inertial properties of each segment with an higher level of accuracy. • For this effort we need to collect MRI and DXA measures from the same subject, which will be acquired in WP6 (T6.4). USFD and SAG will employ the extracted geometry to generate subject specific mass distribution models for lower limbs. • MOTEK and DUT will develop a scaling method for mass distribution and muscle volumes based on regression equations between the actual data as extracted from MRI (T11.1), and simple anthropometric measures available without images (e.g. total weight, segment length, joint width and circumference). • T11.1.2 means that three different mass distribution models will be available per subject , based on the data from WP6 (T6.4). One aspect to be analyzed in this task is in how far and for which subjects the added effort of DXA or MRI imaging is required to build an accurate model. • Since the data from WP6 (T6.4) also includes the marker positions (used for CGA) within the images, a sensitivity analysis for these parameters can be performed by running a musculoskeletal model based on the different data sets. **Partners involved: MOTEK, OPBG, KU Leuven, VUA, USFD, SAG, DUT.**			available without imaging techniques	**1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T11.2 Development of a personalized disease specific skeletal model [M 12-36] [M 30-48] • The aim of this task is to construct personalized models for musculoskeletal simulation from MRI and DXA images. The statistical shape models constructed in T11.1 will be refined to differentiate between different muscles and to include muscle insertion points. • Disease-specific statistical shape models that explicitly account for variation, such as the femur anteversion in CP patients, will be explored. • Again, SAG will apply these models on MRI and USFD on DXA data. Interactive methods will be used to correct muscle anatomy or insertion points in case the automatic detection fails. In order to use the extracted muscle geometry and insertion points in the HBM model, MOTEK will adapt the HBM model to the new parameters. • Firstly we plan to construct a personalized models for musculoskeletal simulation from MRI images. The statistical shape models constructed will be refined to differentiate between different muscles and to include muscle insertion points. • Disease-specific statistical shape models that explicitly account for variation, such as the femur anteversion in CP patients, will be explored. Interactive methods will be used to correct muscle anatomy or insertion points in case the automatic detection fails. • Then we plan to construct personalized models for musculoskeletal simulation from DXA images. The statistical shape models constructed will be refined to differentiate between different muscles and to include muscle insertion points. Alternatively, the statistical shape models constructed by SAG in T11.1 could be used to be morphed to the DXA data. • Disease-specific statistical shape models that explicitly account for variation, such as the femoral anteversion in CP patients, will be explored. Interactive methods will be used to correct muscle anatomy or insertion points in case the automatic detection fails. **Partners involved: MOTEK, OPBG, KU Leuven, VUA, USFD, SAG, DUT.**	STEENBRINK **Estimated % realization** M3 M6 M9 M12	**D11.3) Adaption of existing musculoskeletal model: Adaption of existing musculoskeletal model to subject and disease specific data. Pathology related parameters clinically measured are to be included in the model's optimalization routines [month 36]**	Month 12 **Lead** STEENBRINK **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T11.3 Construction of a disease specific muscle model [1-36] • Since most pathologies in NDD typically affect muscle parameters, it is necessary to develop pathology specific muscle models. • These models should focus on the parameters that are known to be of large influence on the estimation of muscle forces, based on optimisation criteria on how to explain the net joints moments from CGA. • This means that especially muscle contractures, altered muscle structure and stiffness (in CP) as well as muscle weakening (in DMD and SMA) must be targeted. • The data from DXA imaging provide an accurate measurement of the areal density of the lean tissue (muscle) in the frontal plane. This information will be used to refine and further personalisation the muscle model. • The use of muscle ultrasound (MUS) enables to measure the fibre direction, and consequently the PCSA, as well as muscle belly length. • Structural information from images (i.e. MRI, DXA or MUS), can only partly reveal the behavior of the muscle. • The parameterisation of the neuromuscular complex can be estimated using mechanical perturbations of the joint, which will identify the system behavior. Parameters like contractures, optimal muscle length, viscoelastic stiffness are the result of these measurements, using a robot manipulator, or joint dynamometer, based on existing models developed by URLS and DUT. • There will be a focus on the ankle, because of the clinical importance of the calf muscle in walking. • The personalized muscle parameters must be used to adapt the muscle parameters of the HBM models. The aim is to construct and evaluate inverse dynamics muscle force estimation, using these disease muscle models, that can be personalized for the most important muscle parameters. **Partners involved: DUT, URLS**	VEEGER **Estimated % realization** M3 M6 M9 M12	**D11.3) Adaption of existing musculoskeletal model:** Adaption of existing musculoskeletal model to subject and disease specific data. Pathology related parameters clinically measured are to be included in the model's optimalization routines	Month 36 **Lead** STEENBRINK **1st draft ready by:**

Self-Assessment criteria	
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T11.4 Design of models driven by the dynamics of gait perturbations [M12-36] • Only a system disturbance will make it possible to identify system responses, beyond the phenomenology of the apparent suboptimal solution in the pathological case. The aim is to construct a model in which the various walking speeds are combined to separate speed dependent contributions (determinants of adaptive motion control) from speed-independent variables (determinants of pathological gait). This also includes protocols for stability analysis after single perturbations. • The first model adaptation that needs to be developed should handle the adaptations to various walking speeds. • This model should identify the speed regulation in terms of: (i) neuromuscular reflex modulation; (ii) adapted central motion control (synergy/selectivity); and (iii) compensatory movement strategies. EMG data will be used to model function of the neurological system (especially aberrant control: contraction synergies, co-contraction and spasticity). • The second model adaptation that needs to be developed will handle the adaptations to small mechanical perturbations. This model should identify the responses to perturbations (small changes of treadmill speed in terms of: (i) adaptive, coping movement responses; (ii) neuromuscular reflexes; (iii) adapted central motion control (synergy/selectivity); and (iv) (lack of) adapted central motion control. • EMG data will be used to model function of the neurological system (especially aberrant control: contraction synergies, co-contraction and spasticity).	STEENBRINK **Estimated % realization** M3 M6 M9 M12	**D11.4) Disease-specific muscle model** Construction of a disease specific muscle model designed of models driven by the dynamics of mechanical and visual gait perturbations [month 48]	Month 4 **Lead** STEENBRINK **1st draft ready by:**
Self-Assessment criteria			
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**		

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP12: Models validation, outcome analysis and clinical workflows				
Tasks	**Lead**		**Deliverables**	**Deadline**
T12.1.3 Clinical assessment of musculoskeletal biomechanics models (JIA and NND). • In two of the target clinical applications, JIA and NDD, we shall explore the use of complex multiscale biomechanical models of the musculoskeletal system personalised for each patient using as much as possible of the information available from medical imaging, molecular imaging, and gait analysis. • We shall also establish appropriate reference framework to make possible correlative explorations between clinical signatures of the disease that can be quantified using clinical, imaging, or instrumental assessment, and the prediction of the biomechanical models, as a support for the ethiopatological speculation (JIA) and a more effective scoring of the disease severity and for treatment planning (NND). • But before we can use the predictions of these models, we need to conduct an extensive clinical validation on the various elements that form them. • Medical imaging protocols will be tightly controlled, and period quality assessment conducted on all systems in use for the project, with particular reference to spatial calibration, and densitometry calibration for x-ray imaging. • To validate the fusion of imaging and gait analysis data superficial skeletal landmarks such as knee epicondyles will be located both by palpation in the gait lab and on the MRI images, and used to verify the accuracy of fiducial registration with the skin markers. • All image processing and image modelling methods will be tested using an alternative source of information, typically CT scans to validate bone reconstruction, etc. In particular DTI processing for cancellous bone will be validated on a small cohort patients recruited at USFD, that are undergoing wrist or ankle HRpQCT, which provides a very detailed	PONGIGLIONE		D12.1) Outline of the clinical assessment and validation criteria for all four disease areas: Preliminary analysis of the clinical assessment and validation criteria	Month 18 **Lead:** PONGIGLIONE **1st draft ready by:**
	Estimated % realization			
	M3		D12.2.1) First clinical assessment and validation results for all four disease areas: Periodic update at month 24 of clinical assessment and validation outcomes	Month 24 **Lead:**PONGIGLIONE **1st draft ready by:**
	M6			
	M9			
	M12		D12.2.2) Second clinical assessment and validation results for all four disease areas: Periodic update at month 36 of clinical assessment and validation outcomes	Month 36 **Lead:** PONGIGLIONE **1st draft ready by:**
			D12.2.3) Third clinical	Month 48

information of the bone tissue spatial organization. Some of these patients will be examined also with the MRI at the same site, and the tissue orientation computed from DTI-like processing of the MRI images, to be verified against the HRpQCT data assumed as true value.			assessment and validation results for all four disease areas: Periodic update at month 48 of clinical assessment and validation outcomes.	**Lead:** PONGIGLIONE **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	**Lead**		**Deliverables**	**Deadline**
T12.2.4 Clinical workflows for NND. • Clinical workflow for NND will describe the sequence of operations that start with clinical data acquisition and by using our models ends with a clinically useful outcome predictors which are crucial to personalise treatment strategy. • The clinical workflow will be subdivided into 4 specific steps: a) acquisition of clinical, structural and functional information, b) integration of all information into a model, c) similarity search through the digital repository, and d) personalised prediction of disease outcome and optimization of individualized therapy. Data will be acquired through the use of validated standardised clinical measures of disease activity (WP6). Physical • Examination data, Imaging data and gait analysis data (Kinematics, kinetics, EMG and O2 data). • For CP data will be collected at common critical point during growth (between 8 and 14 years) , when clinical decision on therapies is urgently needed to maintain walking function. • For DMD and SMA longitudinal data at three points during the critical period of functional recline are taken. Imaging information will be integrated with the results of gait analysis and clinical evaluation in order to build an articulated joint model and a biomechanical model able to predict the muscle forces. • The prognostic value on an individual level of multidimensional data including modern imaging modalities, as well as muscle models and biomechanical skeletal models will be explored. • The model will work across scales from muscle function, bony deformities scaled-up to whole body walking • performance, and incorporated into changing lives. The risk-benefit ratio will be measurable and incorporated	PONGIGLIONE **Estimated % realization** M3 M6 M9 M12		**D12.3) Improved clinical workflows and outcome analysis:** Final proposal of innovative clinical workflows based on outcome analysis of all patient cases	Month 48 **Lead** PONGIGLIONE **1st draft ready by:**

into the model and the clinical decision-making process. The model will allow individual planning of interventions with subsequent consequences for functioning and quality of life of the affected children This will provide optimization of therapy, and thus a complete newly-defined workflow for personalised predictive and clinical medicine.				

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP18: Dissemination & Training

Tasks	Lead		Deliverables	Deadline
18.3 Training Training is considered to be a fundamental task in dissemination. As anecdotal evidence has confirmed via WP4 of the VPH NoE and via feedback from the DISCIPULUS ('Roadmap Towards the Digital Patient') meeting (30/03/2012; Barcelona), training is recognized to be one of the most solid and long-lasting dissemination strategies in place. The training activities within MD Paedigree will consist of 2 'hands-on' workshops to be delivered during years 2 and 4 of the project (at approx. 1 or 1.5 year interval) in order to expose the outcomes achieved both, in disease modelling and in building the infostructure, highlighting the potential for change management and innovation in clinical workflows to the medical/clinical and research community interested in VPH technology. The first workshop will also seek to provide feedback to the research and development activities, so as to refine the outcomes for the final workshop. The workshop participants will fill in a detailed feedback questionnaire that will be passed to the developers. This task will be led by UCL, which has a long-standing commitment with the VPH Community and is	DIAZ **Estimated % realization** M3 M6 M9 M12		D18.3) Training event in year 2: Report on the outcomes of the first Training event **D18.6) Training event in year 4:** Report on the outcomes of the second Training event	Month 30 **Lead** DIAZ **1st draft ready by:** Month 42 **Lead**

involved in several training grants, including the Marie Curie ITN 'MeDDiCA', 'VPH-MIP' and WP4 of the VPH NoE.				DIAZ **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	**Lead**		**Deliverables**	**Deadline**
T18.4 Seminars, Workshops, Concertation Activities with Other ICT Funded Projects, and Scenario Analysis Sessions **Lead: Vanessa Diaz** The Consortium will identify the most relevant conferences in the area and propose seminars and workshops to be held during these events. It will devote special attention and resources to Concertation Activities with other ICT funded projects and to targeted dissemination actions. Special "Scenario analyses" sessions will be convened, involving the key personnel from both the clinical and the technological partners, with the aim of pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users within MD-Paedigree. The results of the previous workshops will be presented to the Scientific Committee and to the Users' Board in order to assess their relevance and applicability, so as to refine the outcomes for a validation workshop and for a final MD-Paedigree Conference, to be held at the end of the project, targeting both	DIAZ **Estimated % realization** M3 M6 M9		**D18.4.1) First scenario Analysis Sessions:** First scenario Analyses pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users. **D18.4.2) Second scenario Analysis Sessions:** Second scenario Analyses pre-	Month 24 **Lead** DIAZ **1st draft ready by:** Month 42

internal and external clinical and research communities as well as patient organisations and the interested media. The participation in any such event will be reported in the periodic reports and the final report.	M12		empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users.	**Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T18.7 Engaging Parent and Patient Associations **Lead: Vanessa Diaz** Approaching Parent and Patient associations will become a part of the consortium's dissemination activities. The project will seek to disseminate news of its work, expected results and potential future developments through these channels. It is hoped that the work with Patient associations will help achieve a larger bidirectional knowledge sharing base of clinicians and of patients, and further inform the potential beneficiaries of the ongoing work.	DIAZ **Estimated % realization** M3 M6 M9	**D18.1) Dissemination and training strategy plan and preliminary materials**: Roadmap defining the dissemination and training strategy, indicating the subsequent choice of preliminary materials	Month 12 **Lead** DIAZ **1st draft ready by:**

	M12			

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP19: Exploitation, HTA, and Medical Device Conformity

Tasks	Lead	Deliverables	Deadline
T19.1: Evaluation approach and meaningful indicator development (EMP) • Develop upon and adapt in the VPH and other contexts proven approaches, methods and tools to the specific environment and objectives of this workpackage • Establish a set of meaningful criteria and their measurement process that are robust to demonstrate socio-economic benefit-cost impacts. The focus is • to approach and find measurements for evaluating how virtual collaborations between members of the VPH communities with different expertise are facilitated and • how consequently the uptake and acceleration of model development and integration can find meaningful expression in the overall evaluation framework.	STROETMANN **Estimated % realization** M3 M6 M9 M12	**D19.1 HTA evaluation framework** It reviews proven approaches, methods, and tools which might be relevant to the specific environment and objectives of this workpackage, and establishes a set of meaningful criteria and their measurement process, thereby focusing on evaluating how virtual collaborations between members of the VPH communities with different expertise are facilitated.	Month 12 **Lead** STROETMANN **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T19.3: Benefit-cost scenario for clinical impact assessment (EMP) In a separate task a high-level, generic benefit-cost scenario for clinical impact assessment will be applied, with the ultimate goal to generate economic and market evidence for true translational medicine. The benefit-cost scenario will be tested and initially validated with preliminary, exploratory data estimates from the patient-centred workflows that are the basis of the digital repository and Infostructure. The two main dimensions pertaining to clinical/health impacts focus on the one hand on health service delivery and the health of patients, and on the other on public health/societal outcomes. To assess such impacts, the scenario development will integrate the following indicators: • Clinical effectiveness and patient-related outcomes • Safety (risks associated with applying the technology) • Organisational and change management aspects • Human resource implications, knowledge & education needs • Assessing contributions to the VPH vision of a patient avatar • Efforts for application (convenience/ease of use; costs for introduction of new technology) The indicators assessed ultimately prepare for a more targeted and strategically aligned exploitation activities (T19.4) by proving clinical impact of MD-Paedigree with respect to: • the state-of-the-art in paediatric patient-specific computational modelling, • improved disease understanding and therapy outcomes that can be applied to both clinical routine and translational clinical research, • usability by clinicians and clinical researcher, • transferring technical workflows into clinical workflows, • the vertical integration of multi-scale patient data and the provision of models, tools, and services readily available to clinicians at the point of care.	STROETMANN **Estimated % realization** M3 M6 M9 M12	**D19.4 Clinical impact assessment scenario** Initial formative evaluation of MD-Paedigree model-driven Infostructure based on a benefit-cost analysis approach, subsequently followed by a generic benefit-cost scenario for clinical impact assessment developed and validated with partners and experts. [month 36]	Month 36 **Lead** STROETMANN **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

A.1.1.4 Timing of work packages and their components

The MD-Paedigree project partners have formalized a work plan implementing 4 major phases implying a number of conceptual steps, over 48 months of activity with 4 major milestones. The first milestone is due after 9 months and marks the end of the specification phase; the following milestones are aligned with the reporting periods of the project every 12 months.

Phase 1 (running from month 1 to 9) – Project Set-up, Requirements Elicitation, and Clinical Protocols: During Phase 1 quality assurance guidelines and a self-assessment plan will be prepared, ethical approval will be obtained , and the first dissemination activities will be performed (Step 1) Furthermore, clinical protocols for the selected paediatric applications will be established (Step 2).)Finally, the requirements for models and infostructure implementation will be analysed and documented from an end user standpoint (Step 3).

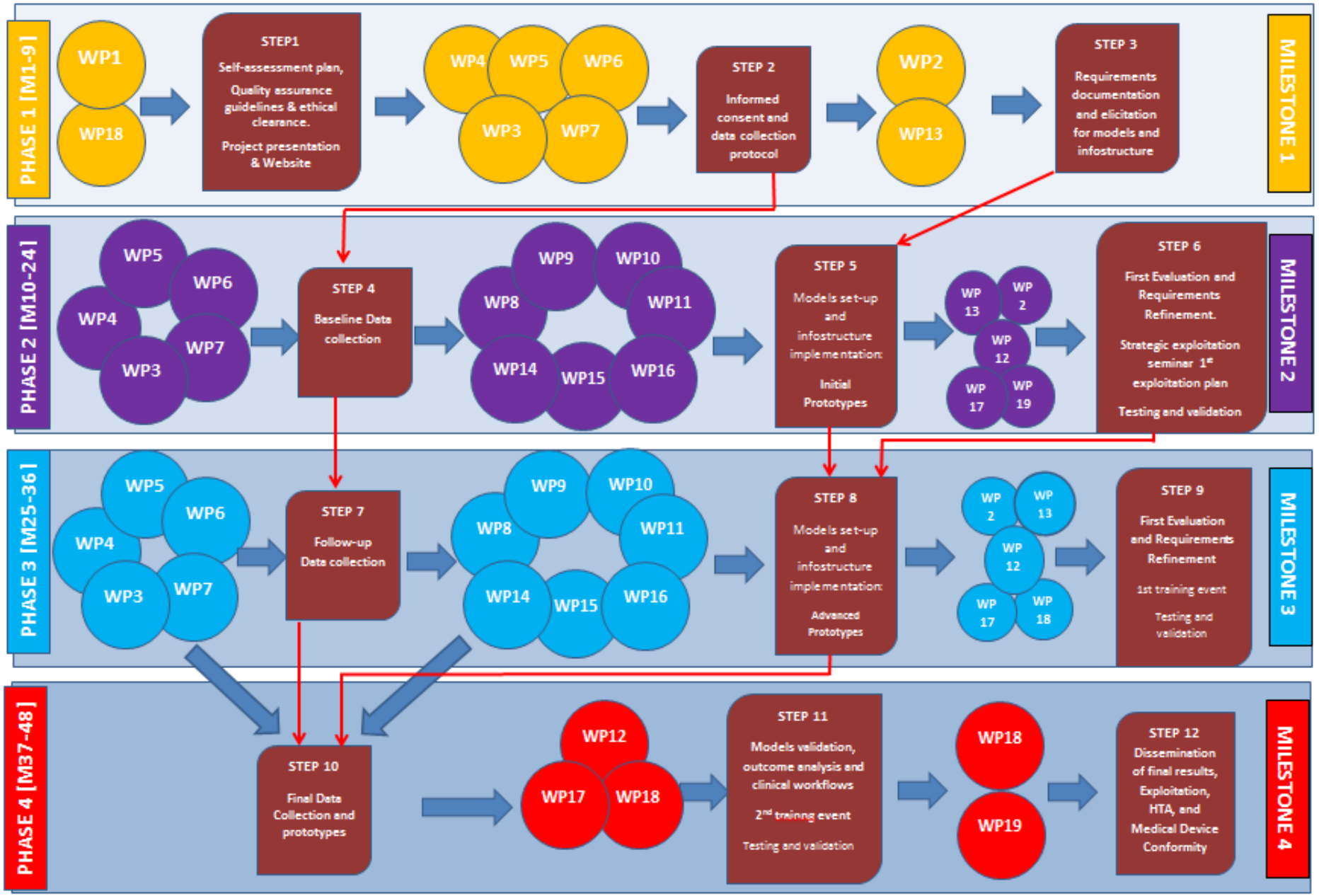
***Phase 2 (running from month 10 to 24) – Baseline Data Collection, Initial Prototypes, First Evaluation and Requirements Refinement*:** Patient enrolment will take place and data acquisition will be started (Step 4). Based on the established requirements, the existing models from Health-e Child and Sim-e-Child projects will be refined and adjusted to the new applications. The open repository for project infrastructure will be introduced and initialized with the current models and data (Step 5). First evaluations will be undertaken and requirements will be refined based on the collected experience; additionally, during this phase, the Strategic Exploitation Seminar will be held and the 1st Exploitation Plan will be drafted (Step 6).

***Phase 3 (running from month 25 to 36) – Follow-up Data Collection, Advanced Prototypes, Evaluation and Requirements Update*:** Follow-up or additional data will be acquired for all clinical applications (Step 7). The respective models will be enhanced to process longitudinal data and refined according to the obtained evaluation results. New functionalities will be integrated into advanced prototypes. The open repository will be improved and updated with content (Step 8). A second set of evaluations will be conducted and requirements will be adjusted for the final system. Furthermore, the 1st Training Event will be held (Step 9).

***Phase 4 (running from month 37 to 48) – Final Data Collection and Prototypes, Clinical Validation, and Deployment*:** In the final year, data collection will be concluded and the clinical validation will take place with the final models and simulation framework (Step 10). Results will be used to propose and disseminate improved clinical workflows. Subsequently, the 2nd Training Event will be held (Step 11). Models for all clinical applications and their respective evaluations will be documented and disseminated, while the implementation plan will be refined and the Health Technology Assessment and the Medical Clearance preparatory activities will be performed (Step 12) .

The timely delivery of all planned deliverables will be the first indicator of the fulfillment of each phase in the expected progress of MD-Paedigree, monitoring what can be demonstrable at each corresponding milestone of the project.

A second and much more detailed means of verification will be provided by the assessment criteria for each milestone and each WP which are to be defined within D1.3 Self-assessment plan on month 3.


PHASE 1 [M1-9]
WP1
WP18
STEP1
Self-assessment plan,
Quality assurance guidelines & ethical clearance.
Project presentation & Website
WP4
WP5
WP6
WP3
WP7
STEP 2
Informed consent and data collection protocol
WP2
WP13
STEP 3
Requirements documentation and elicitation for models and infostructure
MILESTONE 1
PHASE 2 [M10-24]
WP5
WP6
WP4
WP7
WP3
STEP 4
Baseline Data collection
WP9
WP10
WP8
WP11
WP14
WP15
WP16
STEP 5
Models set-up and infostructure implementation:
Initial Prototypes
WP 13
WP 2
WP 12
WP 17
WP 19
STEP 6
First Evaluation and Requirements Refinement.
Strategic exploitation seminar 1st exploitation plan
Testing and validation
MILESTONE 2
PHASE 3 [M25-36]
WP5
WP6
WP4
WP7
WP3
STEP 7
Follow-up Data collection
WP9
WP10
WP8
WP11
WP14
WP15
WP16
STEP 8
Models set-up and infostructure implementation:
Advanced Prototypes
WP 2
WP 13
WP 12
WP 17
WP 18
STEP 9
First Evaluation and Requirements Refinement
1st training event
Testing and validation
MILESTONE 3
PHASE 4 [M37-48]
STEP 10
Final Data Collection and prototypes
WP12
WP17
WP18
STEP 11
Models validation, outcome analysis and clinical workflows
2nd training event
Testing and validation
WP18
WP19
STEP 12
Dissemination of final results, Exploitation, HTA, and Medical Device Conformity
MILESTONE 4

NND					
March 2013	**April 2013**	**May 2013**	**June 2013**	**July 2013**	**August 2013**
	Protocols delivered to Ethical Committee	D7.1 Recruitment protocol with ethical clearance (for genetic Studies)	D11.4 Disease-specific muscle model	Interviews to prepare D2.1	First Half-Yealry report.
					Self-Assessment Plan
			Contribution to the Self-Assessment Plan		Check of the enrollment and data collection, analysis and processing.
	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs
		Area Dedicated T&M TC [8th May]]	Area Dedicated T&M TC [12th Jun]	Area Dedicated T&M TC [10th Jul]	Area Dedicated T&M TC [14th Aug]

September 2013	**October 2013**	**November 2013**	**December 2013**	**January 2014**	**February 2014**
Biannual area meeting	Check of the enrollment and data collection, analysis and processing.	First draft of the deliverable **D2.1**			D2.1 Initial requirements analysis document including priorities for the implementation
					D11.3 Adaption of existing musculoskeletal model
Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCsq	Individual WPs' TCs	Individual WPs' TCs
Area Dedicated T&M TC [11th Sep]	Area Dedicated T&M TC [9th Oct]	Area Dedicated T&M TC [13th Nov]	Area Dedicated T&M TC [11th Dec]	Area Dedicated T&M TC [8th Jan]	Area Dedicated T&M TC [12th Feb]

NND	**DELIVERABLES WITHIN MONTH 24**	
	D6.1) CGA standard protocol	M18
	D6.2) A standard protocol of clinical gait analysis is described based on a representative inventory along	M24
	D11.1) Automatic extraction method of mass distribution and muscle volumes	M18
	D12.1) Outline of the clinical assessment and validation criteria for all four disease areas	M18
	D12.2.1) First clinical assessment and validation results for all four disease areas	M24

Infostructure

MD-PAEDIGREE KICK OFF MEETING

INFOSTRUCTURE WG

Participant's Name	Affiliation

Concept (general)

MD-Paedigree validates and brings to maturity patient-specific computer-based predictive models of various paediatric diseases

- increasing their potential acceptance in the clinical and biomedical research environment
- making them readily available not only in the form of sustainable models and simulations, but also as newly-defined workflows for personalised predictive medicine at the point of care.

These tools can be accessed and used through an innovative model-driven infostructure

- powered by an established digital repository solution
- able to integrate multimodal health data
- entirely focused on paediatrics
- conceived of as a specific implementation of the VPH-Share project, planned to be fully interoperable with it and cooperating, through it, also with p-Medicine.

MD-Paedigree aims at achieving high-level semantic interoperability,

- requiring standards enabling the clinical contents to be interpreted consistently across the different EHR regimes,
- while complete clinical interoperability between systems will require widespread and dependable access to maintained collections of coherent and quality-assured semantic resources,
- including models that provide clinical context,
- mapped to interoperability standards for EHR and PHR and biomedical data,

linked to well specified terminology value sets, derived from high quality ontologies

CONCEPT (SPECIFIC)	**WPs' OBJECTIVES**	**Objectives' Lead**	**Estimated % realisation**
B.1.2.3 The MD-Paedigree VPH Infostructure and Digital Repository Building on the Sim-e-Child/PCDR digital repository, MD-Paedigree will implement the Service-Oriented Knowledge Utility (SOKU) vision, to facilitate the design and development of innovative new predictive models as reusable and adaptable workflows of data mining applications, and turning the latter into clinically validated decision support tools, made available at the point of care to physicians.	**WP13 Requirements and Compliance for the MD-Paedigree Infostructure** • Manage throughout the project lifetime the requirements for the infostructure,in collaboration with the clinical WPs as well as external stakeholders, to ensure that MD-Paedigree becomes an integral part of the European data infrastructures ecosystem dealing with accessibility, interoperability and exchange.		

B.1.2.3.1 Service-Oriented Knowledge Utility (SOKU) • MD-Paedigree translates the domain-specific applications and data into services and associated knowledge that can be further published, discovered and semi-automatically orchestrated in the grid/cloud, by physicians and medical data integration experts. This way leads to the development of new workflows and enables their personalisation to real patient cases. • The MD-Paedigree system consists of standard services to hybrid services, and to finally more complex high-level entities, which produce knowledge. • MD-Paedigree exploits recent ground-breaking European research on semantic modelling, ontology-based data access and scalable query execution to develop an extensible platform based on open standards and protocols to deliver a complete and generic solution able to tackle the targeted paediatric disease areas. **B.1.2.3.2 Data Access and Query Formulation** • One important goal of the MD-Paedigree infostructure is to provide the necessary tools and applications to assist users in accessing and foraging the wealth of heterogeneous data available in the digital repository in an easy, intuitive and seamless way across the care continuum via enhanced connectivity with other hospital information systems and the patient's electronic health records. • Technologies and research related to Ontology-Based Data Access (OBDA) are applied, such as the new forms of query by navigation based on ontologies[125] and the extensible declarative query language supporting linked data (e.g. SPARQL endpoint). • Interactive search based on relevance feedback will be applied to	• Special attention will thus be given to assure users and VPH community acceptance by delivering appropriate tools and services addressing usability, accessibility and maintenance of the system thanks to, among others, open environments and open-source software. • Separate tasks will assure compliance with two major initiatives in the field, i.e. VPH Share and OpenAIRE. • Requirements will be published as deliverables and will be made available to all stakeholders. **WP14 Grid-Cloud Services Provision and GPU Services Integration** • Building on the experience and the work done in Health-e-Child, Sim-e-Child and PCDR, this WP aims at • deploying, maintaining the MD-Paedigree physical infrastructure at participating centers and topping it up with a technological glue of horizontal services under the form of a Service Oriented Knowledge Utility (SOKU), • harmonize the use of and access to all infrastructure resources from computational power, to data, • information, knowledge and applications. • this work package will evaluate, identify and when necessary design, develop and test a semantically enriched framework of predictive models and other utilities as software services. • It will therefore extend the Health-e-Child/Sim-e-Child Science Gateway to integrate the semantic • framework of data mining and composition tools from WP15 and WP16, while closely following the prioritized requirements formalized in WP13. • The SOKU infrastructure will be the environment directly perceived and manipulated by users and client applications, the latter being developed in		

improve data recall in the infostructure. **B.1.2.3.3 Distributed Processing and GPU support** MD-Paedigree extends the distributed processing capabilities of the Sim-e-Child platform in two major axes: • On the one hand, it develops compatibility with GPU processing and makes it possible to execute validated models onto real patient data, thus providing real-time support to physicians at the point of care in the 7 participating centres. • Indeed, the introduction of non graphics application programming interfaces (APIs) for GPUs brought a new perspective on GPUs, transforming them into generalpurpose units. • On the other hand, MD-Paedigree will experiment with the operation of a sustainable translational service for healthcare professionals and other external centres, by integrating an open Cloud API (i.e. the OCCI) in its abstraction layer, thereby allowing the infrastructure to elastically adapt according to faced requests from end-users. • The Athena Distributed Processing (ADP) Engine is considered to more easily integrate and adapt algorithms distribution, through the newly integrated abstraction APIs. **B.1.2.3.4 Intelligent Mining, Modelling, Reasoning and Simulation Framework** • MD-Paedigree integrates AITION, an evolutionary information processing and knowledge discovery framework developed by the University of Athens (UoA) for biomedical research, which is able to provide highly accurate predictive and statistical simulation models combining a bottom-up datadriven process to analyse heterogeneous demographic, phenotypic, clinical, molecular, and genomic biomedical data, images and streams; and a top-down model-driven process to	the A2 Modelling and Simulation activity. It will be made of a set of dynamic and meaningful utilities, which can be used as standard atomic and/or composite services. **WP15 Semantic Data Representation and Information access** • Define and implement the data catalogue of the project using standard ontological/terminological resources [T15.1, T15.2]; • Develop the interfaces and services (query reformulation, search, feed-back...) needed to answer user information requests (clinical research, drug development, patient safety...) [T14.3, T14.4]. **WP16 Biomedical Knowledge Discovery and Simulation for Model-guided Personalised Medicine** • This WP will integrate and further extend data analysis tools and personalisation techniques stemming from • former EC-funded research (especially the FP6 IP Health-e-Child and the FP7 Sim-e-Child projects), in order to provide a comprehensive information processing, knowledge discovery and simulation framework delivering multi-scale statistical simulation models that capture the whole disease information. • In addition, it will integrate them with specialized VPH models developed in other WPs and explore different adaptation and combination schemes based on specific patient profiles, complementing WP12 efforts for model-driven clinical workflows targeting a holistic framework for model-guided personalised medicine. • In more detail, the objectives of this WP are: • to provide general tools and techniques for		

incorporate external knowledge coming from domain experts, literature, or model-guided processes and relational/semantic models. • AITION integrates Probabilistic Graphical Models (PGMs) as a unifying patient/disease modelling approach providing an integrated framework for multi-scale vertical integration, feature selection, simulation, knowledge discovery and decision support. • AITION is based on state-of-the-art techniques for Bayesian Network Learning, Markov Blanket induction and real-time inference. • Moreover, ontologies and a priori knowledge will also be incorporated automating causal discovery and feature selection, providing semantic modelling under uncertainty. • In MD-Paedigree, hierarchical architectures, as well as, Granular Computing (GrC) and Statistical Relational Learning (SRL) techniques will be extended. SRL is an emerging research area which aims at combining statistical learning and probabilistic reasoning (such as PGMs). • Moreover, the Hierarchical Layered Architecture incorporating hidden (latent) layers/variables and GrC techniques allows to build an efficient multi-resolution computational model targeting complex applications consuming large amounts of data, information and knowledge. **B.1.2.3.5 Holistic Model-Guided Personalised Medicine** • Ultimately, MD-Paedigree will provide an evolvable framework for holistic model-driven medicine and personalised treatment combining knowledge constructs from observational data analysis, statistical and specialized VPH patient- or disease-specific simulation models, domain knowledge representations, as well as patient/disease-specific profiles. • The goal will be to find efficient ways to optimize and combine multiple	intelligent querying, data analysis and knowledge discovery on vertically integrated data (i.e. analysing all dimensional scales from genetic and molecular levels to clinical &behavioural) across disease areas; • to deliver highly accurate and reusable predictive – patient or disease specific – statistical simulation models combining bottom-up data driven analysis with top-down modelling and domain knowledge inclusion, capturing vertical integration and temporal evolution that will be validated and utilized on WP12; • to integrate statistical models with VPH models and explore different adaptation and combination schemes based on specific patient profiles; • to complement and evaluate WP12's model-driven clinical workflows; • to provide advanced scaling capabilities utilizing the underlying infrastructure; • to provide clinical trial support [T16.4]. **WP19: Exploitation, HTA, and Medical Device Conformity** An early evaluation in the form of health technology assessment (HTA) as well as the development of exploitation strategies is essential for the creation of research related services which can prevail in today's highly competitive markets - be they "academic" and RTD markets, be they health services or commercial markets. The workplan is designed to encourage materializing improved disease understanding and therapy outcomes into both clinical routine and translational research, to deploy early prototypes within the		

statistical and/or specialized VPH simulation models in prediction tasks supporting the creation and validation of model-driven clinical workflows. • Utilizing the PAROS personalization platform, clinicians and domain experts will create ontology-based patient and disease-specific profiles capturing high-level concepts and common characteristics. • Similarity search techniques will then be developed mapping specific medical cases to pertinent patient/disease profiles. These profiles will be used to adapt and optimise individual simulation models by transformations, as well as to explore their combinations and re-use in different disease areas. • Finally, a holistic scheme for model-driven personalised medicine will be developed that will allow analysing and testing scientific hypotheses, predicting disease evolution and treatment responses (e.g. early diagnosis of poor outcome that needs aggressive treatment) and elaborating individualized treatment plans. **B.1.2.3.6 Compliance with Guidelines for Model Based-Drug Development (MBDD)** • special attention to having functional databases available to assist drug developers. • MD-Paedigree can support the drug discovery process. In particular, it can help in identifying biomarkers likely to characterise a particular pathology or dysfunction. • Second, it can help to design clinical trial protocols (i.e. exclusion/ inclusion criteria, statistical power, and cohort identification) by providing a feasibility testbed to conduct clinical research studies, as currently explored by IMI projects such as EHR4CR.	developing VPH Infostructure, and to improve in iterative cycles of specifications, refactoring (i.e. improving the design of existing code), and deployment. **Objectives** • Evaluate the MD-Paedigree's models, workflows, and infostructure based on: o its accessibility, usability and effectiveness for the VPH community o the potential of its contributing to personalised healthcare workflows and integration with EHRs/decision support systems, thereby preparing for the transfer into clinical practice o making models and simulations readily available at the points of care and to researchers • Define effectiveness and usability within the context of sharing "developing ICT tools, services and infrastructure to obtain more elaborate and reusable multi-scale models" (call text) as well as developing an appropriate analytical evaluation framework • Explore the health system and business opportunities o to market concrete project outcomes and results o to prevent diseases and contribute to the safety of care o to identify markets and cost models for the effective diffusion of our models, allowing researchers to exploit, share resources and develop new knowledge		

• Last but not least, the longitudinal follow up of MD-Paedigree populations can help to monitor longer term effects of therapeutic treatments, including -drug response, phenotype evolution (e.g. neoplastic processes), as well as rare adverse effects. The resulting views can ultimately help to cluster populations according to specific genotypic variations (pharmacogenomics).	• Design business plans that prepare pre-market access and that integrate medical device conformity assessment procedures		

WP2: Clinical and technical user requirements for disease modelling

Tasks	Lead	Deliverables	Deadline
Task 2.1: Conduct interviews with the clinical and technical partners to obtain a complete list of requirements for the disease modelling that will ensure its usefulness within and beyond the project. All WP Leaders will actively contribute to the requirements documentation while they ensure that the respective WP partners are interviewed. Prioritisation criteria: i. All requirements will be prioritised ensuring that from the start the most important aspects will be implemented to quickly ensure an operational system. Schedule of requirements updating: ii. The requirements list will be continuously updated on a regular basis such that main requirements and system constraints will be released as deliverables.	CHINALI **Estimated % realisation** M3 M6 M9 M12	**D2.1 Initial requirements analysis document including priorities for the implementation** Complete interviews with the clinical and technical partners will be collected to obtain a list of variables and requirements for the disease modelling. Requirements will be prioritized ensuring that from the start the most important aspects will be implemented first.	Month 12 **Lead** CHINALI **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP13: Requirements and Compliance for the MD-Paedigree Infostructure

Tasks	Lead	Deliverables	Deadline
T13.1 Requirement elicitation and documentation (HES-SO, URLS, Maat, SAG, Lynkeus) [M1-18] Task 13.1 will conduct interviews with the clinical and research partners to obtain a complete list of requirements for the infostructure that will ensure its usefulness within and beyond the project.	RUCH **Estimated % realisation** M3 M6 M9 M12	**D13.1Initial list of main requirements after stakeholder interviews including priority domains**	Month 9 **Lead** RUCH **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T13.2 Requirement revision and management (HES-SO, SAG, HES-SO, Lynkeus) [M18-36] In Task 13.2, the initial requirements analysis will be managed for the remainder of the project and contact with selected stakeholders will ensure the relevance of the developed infostructure, that is not only based on past research projects but particularly aims at taking into account future developments and make the obtained research data available to allow for efficient and effective data mining, modelling on the data and interoperability of the various existing initiatives. Task 13.2 will also assure that all requirements regarding VPH Share and Open AIRE are taken into account in the design, and a close connection between the requirements management and the compliance tasks exists. Task 13.2 will also control the implementation of all requirements in the course of project developments.	RUCH **Estimated % realisation** M3 M6 M9 M12	**D13.2 Compliance outcomes for VPH-Share and OpenAIRE influencing the infostructure**	Month 12 **Lead** RUCH **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead		Deliverables	Deadline
T13.3 Compliance with VPH-Share (UoS, MAAT, UoA, LYN) [M1-24] Task 13.3 will assure that the infostructure is in full compliance with VPH-Share. The goal is that all data acquired in the project and from the partners in the various disease areas can be used and made useful in the long term and for a large number of partners inside and outside the project. Compliance with VPH Share will guaranty the interoperability of the infostructure with the entire VPH community. Whilst ensuring the data formatting and storage mechanisms are consistent is important, the key objective will be interoperability at the semantic layer to support effective indexing and search of the data collections. This interoperability will cover the full spectrum of outputs from both projects, i.e. data, tools and models so any of these objects can be accessed and used from within their respective environments.	HOSE **Estimated % realisation** M3 M6 M9 M12		**D13.3 Complete list of functionalities for compliance and the system functionality**	Month 24 **Lead** RUCH **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T13.4 Data policy definition and implementation (UoA, URLS, SAG, MAAT, HES-SO, LYN) [M1-36] T13.4 will investigate the governing policies for the infostructure resulting data and publications, ensuring its compliance with the European efforts and policies currently evolving in the scientific research area. The re-use, share and correct citation/crediting of data in an Open Science environment require the establishment of appropriate intellectual property rights policies. MD-Paedigree will investigate such policies in the entire data flow and will consider some for adoption, to properly characterize the quality of its data products and to promote interoperability with other data infrastructures. T13.4 will assure compliance of the infostructure with the two EU FP7 Open Access OpenAIRE and OpenAIREplus projects, which deal with the implementation of appropriate scientific results policies. MD-Paedigree will thus take advantage of the cross-linking discovery tools (i.e. publications-data-funding schemes) offered through OpenAIRE that are based on diverse forms of data-mining (i.e. textual, usage, etc.) and access best practices in similar domains (e.g. EBI and UKPMC). Compliance with horizontal infrastructures such as OpenAIRE will increase MD-Paedigree's visibility and on-line exploration by scientists and non-scientists. More specifically, the compliance will be pursued by i. implementing schemes for making the data identifiable and accessible, ii. following metadata guidelines improving data exploration by 3rd parties iii. organising a workshop and/or other networking/training initiatives with clinicians to discuss the benefits of open access and open science, and to investigate win-win solutions attractive to both bio-medical and biomechanical researchers within paediatric institutions and application providers, in line with the EC's expected impact in terms of: • improved interoperability of biomedical & biomechanical information and knowledge, • increased acceptance and use of realistic and validated models that allow researchers from different disciplines to exploit, share resources and develop new knowledge, • large-scale benefits of having both the data and models readily available.	RUCH **Estimated % realisation** M3 M6 M9 M12	**D13.4 Update on the requirements and compliance requirements including priorities for the implementation**	Month 36 **Lead** RUCH **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP14: Grid-Cloud Services Provision and GPU Services Integration

Tasks	Lead	Deliverables	Deadline
T14.1 Adaptation and Extension of Sim-e-Child Platform (MAAT, UoA, LYN) [M1-48] Starting from the platform available in the Health-e-Child/Sim-e-Child/PCDR projects, MAAT and UoA will extend it to address the MD-Paedigree requirements, as specified by WP13. It will in particular facilitate integration with new developments from WP15 and WP16. Partner MAAT will extend the Sim-e-Child network by deploying new Science Gateways in the Cloud/Grid, as needed. A load capacity analysis will thus be carried out to determine the number of Science Gateways needed versus the number of users and average load of integrated applications.	MANSET **Estimated % realisation** M3 M6 M9 M12	**D14.1 MD-Paedigree, Ground Truth Infrastructure Setup Report** Report on MD-Paedigree ground truth infrastructure sites deployment. The list of new Science Gateways will be published. (Phase 1)	Month 9 **Lead** MANSET **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead		Deliverables	Deadline
T14.2 Open Cloud-API and GPU Integration (MAAT, UoA, HES-SO, TBV, SAG) [M1-24] MD-Paedigree aims to make the latest validated models available at the point of care. There are needs in at least 2 directions: fast responsive equipment for emergency or complex diagnostics, respective simulation and prediction of clinical procedures or treatment effects. To respond to such time critical requirements, MD-Paedigree will extend the Health-e-Child/Sim-e-Child Science Gateway on two major axes. On the one hand, it will develop compatibility with GPU processing and make it possible for MD-Paedigree model applications to be locally and very quickly run onto live patient data, thus providing time critical support to physicians at the point of care. On the other hand, MD-Paedigree will integrate a Cloud API to its abstraction layer, thus allowing elastic adaption of the infrastructure to evolving users needs, and to work on a longer term business plan and sustainable model. MAAT, in cooperation with UoA and HES-SO, will be in charge of integrating the selected Cloud/Grid APIs, while TBV and SAG will contribute to the GPU processing layer.	MANSET **Estimated % realisation**		**D14.2 MD-Paedigree, Alfa version Infrastructure Deployment Report** Report on MD-Paedigree Alfa infrastructure deployment. The first release of the platform will be accompanied with information on software packages versions, software repository location and associated pertinent data.	Month 24 **Lead** MANSET **1st draft ready by:**
	M3			
	M6			
	M9			
	M12			

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T14.3 Athena Distributed Processing (ADP) Engine Integration (UoA, MAAT, TBV) [M1-48] Current algorithms made available through Health-e-Child and Sim-e-Child are not developed to take advantage of the new opportunities delivered by GPU or to be properly parallelized and distributed through cloud/grid facilities. MAAT, UoA and TBV will thus investigate to greatly improve their efficiency and distribution capacity, while defining a gold standard for MD-Paedigree's new applications. More particularly, this task will integrate and further enhance UoA's open-source Athena Distributed Processing (ADP) Engine in order to provide distributed querying over federated heterogeneous sources and big data management on the grid/cloud, distributed processing and parallelization of resource/time-consuming algorithms related to knowledge discovery, simulation and data mining. ADP is a system for complex dataflow processing that acts as a mediating middleware placed between the infrastructure and other components of the system, simplifying their "view" of the underlying infrastructure, supporting distributed querying and versatile execution of distributed algorithms on ad-hoc clusters, clouds, or grids. ADP provides several services on top that will be immediately useful to MD-Paedigree, including: an SQL engine (AdpDB), a MapReduce engine (AdpMR) and a data mining library (AdpDM). In more detail, this task will support distributed execution of complex queries over the heterogeneous federated data sources of MD-Paedigree. The system will be tightly coupled with the query translator and related ontology-based data access components, as well as, the PAROS personalisation platform (WP16). In addition, it will support both classic (request/response) and long running queries. The ADP engine will integrate the Health-e-Child/Sim-e-Child/PCDR underlying query abstraction API, provided by partner MAAT. Finally, besides querying, this task will also utilize ADP for distributed processing and parallelizationof knowledge discovery, simulation and data pre-processing & mining algorithms, such as the ones used by AITION (in WP16) and DCV/madIS (in WP15), providing advanced scaling based on massively parallel execution over elastic, cloud-based platforms. The ADP engine will integrate the Health-e-Child/Sim-e-Child/PCDR underlying distributed computing abstraction API, provided by partner MAAT, thus ensuring proper access to the repository and associated computational resources.	DIMITROPOULOS **Estimated % realisation** M3 M6 M9 M12	**D14.3 MD-Paedigree, Beta version Infrastructure Deployment Report** Report on MD-Paedigree Beta infrastructure deployment. This second release of the platform will be accompanied with information on software packages versions, software repository location and associated pertinent data.	Month 36 **Lead** MANSET **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T14. 4 SOKU Implementation (MAAT, HES-SO, UoS, LYN) [M13-48] Define a service as basic and generic as possible, to be the nutshell pattern for reuse in all servitisation processes. Develop semi-automated mechanisms and accompanying interfaces to enable on-demand servitisation of new resources. This task will also take care of the semantic enrichment of the servitised utilities. T14.4.1 SOA design for VPH-Share and open source technologies integration (UoS, MAAT) This sub-task will take care of ensuring interoperability with and conformance to the VPH-Share project recommendations. T14.4.2 SOA Governance Layer (SGL) development (HES-SO, MAAT) This sub-task will take care of implementing security and privacy requirements within the MD-Paedigree platform governance layer.	MANSET **Estimated % realisation** M3 M6 M9 M12	**D14.4 MD-Paedigree, Final Release Report** Report on MD-Paedigree final infrastructure deployment. This last release of the platform will be accompanied with information on software packages versions, software repository location and associated pertinent data	Month 48 **Lead** MANSET **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T14.5 Privacy and Security Issues (SAG, ALL) [M1-48] In this task a systematic review will be undertaken on data protection and other applicable ethical rules and regulations. Privacy guidance and recommendations may be extracted from: • World Medical Association Declaration of Helsinki, adopted by the 18th World Medical Assembly, Helsinki, Finland June 1964. Revised 1975, 1983, 1989, 1996 and on October 6, 2000 in Edinburgh, Scotland (www.wma.net). • ICH-GCP Guidelines; Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95), Sept. 1997 • (www.emea.eu.int). • International Ethical Guidelines for Biomedical Research involving Human Subjects, Council for International Organizations of Medical Sciences (CIOMS), Geneva 2002. (www.cioms.ch). • International Guidelines for Ethical Review of Epidemiological Studies, Council for International Organizations of Medical Sciences (CIOMS), Geneva 1991 (www.cioms.ch). • WHO: Operating Guidelines for Ethics Committee that Review Biomedical Research, Geneva, 2000, www.who.int/tdr/publications. • Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (www.europa.eu.int/comm/internal_market/privacy). • Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 on privacy and electronic communications (www.europa.eu.int/comm/internal_market/privacy). • EU_Article29_Data Protection Working Party Feb_2007 http://ec.europa.eu/justice_home/fsj/privacy/index_en.htm) • Bernice Elger, Jimison Iavindrasana, Luigi Lo Iacono, Henning Müller, Nicolas Roduit, Paul Summers, Jessica • Wright, Health Data Depersonalisation for Prospective research in the life sciences, Computer Methods and Programs in Biomedicine, volume 99, number 3, pages 230-251, 2010 • NHIN Slipstream Use Case for Medical Product Safety Surveillance using EHR • Other relevant regulations, codes of conduct on data protection in Europe incl. local rules and regulations in participating countries.	COSTA **Estimated % realisation** M3 M6 M9 M12	**D14.4 MD-Paedigree, Final Release Report**	Month 48 **Lead** MANSET **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP15: Semantic Data Representation and Information access

Tasks	Lead	Deliverables	Deadline
T15.1 Data curation and validation tool (UoA, MAAT, URLS) [M6-24] Data curation/validation aims at ensuring that the submitted data is relevant, syntactically and semantically well-formed, and properly linked into the system. A specific Data Curator and Validator (DCV) tool, developed by UoA during the European FP6 Health-e-Child project, will be used as the foundation to build an advanced (semi)-automatic data curation and validation system, which is able to handle the heterogeneous MD-Paedigree data. In MD-Paedigree, DCV will be enhanced with intelligent data curation mechanisms, able to learn/adapt to the available user feedback and the a-priori information available from the semantic data representation/ontologies of the project. some of the functionalities provided by UoA's madIS complex data analysis and processing system will also be incorporated, since it is able to easily handle millions of rows on a single desktop/laptop computer. DCV will integrate the query engine API, as provided by partner MAAT in WP14.	DIMITROPOULOS **Estimated % realisation** M3 M6 M9 M12	**D15.1 A prototype for the case- and ontology-based retrieval service** The first prototype makes services available for the case- and ontology–based retrieval, so these can be integrated in the infostructure by month 24. Report on delivered services functions.	Month 18 **Lead** RUCH **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T15.2 Semantic data representation and interoperability (SAG, UoA, HES-SO, LYN) [M1-48] • The task ensures that the project uses wherever possible standard terminological resources (such as ICD-10, LOINC, WHO-ATC, SNOMED, FMA, or RadLex) and appropriate representation languages (e.g. OWL, RDF) for encoding all data items. • Task 15.2 will take into account current developments and trends in semantic interoperability of medical data. • In addition to the ontology development, the system will thus maintain mappings that specify: • the relationship between the initial bio-medical data sources and the ontology; the relationship between the MD-Paedigree ontology and other related bio-medical ontologies and taxonomies (e.g. UMLS, Gene Ontology, etc.); • the relationship between the ontology and related MD-Paedigree infrastructure components, in particular the Sim-e-Child/PCDR repository services and models and their parameters. • This task will also provide methods and tools for developing and maintaining the ontology including the mappings and supporting their evolution. • Initial ontology development and mapping specification will be based on existing domain-specific data models capturing bio-medical data sources, following an appropriate mapping scheme (Local As View, Global As View, or Global Local As View). • In addition, semi-automated reasoning techniques will be used for ontology analysis as well as mapping with external bio-medical ontologies and VPH or statistical models, evaluating covering of the data sources, equivalence with another set of mapping or ontology fragment, checking redundancy, etc. • Finally, special attention will be given to automate the update and maintenance procedures of the ontologymappings, addressing both ontology evolution and data source modifications.	COSTA **Estimated % realisation** M3 M6 M9 M12	**D15.2 DCV curation tools and services to automatically and manually acquire high-quality curated data** The deliverable makes available services that allow creating curated data from available data sources. Report on delivered DCV tools and interfaces	Month 24 **Lead** RUCH **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T15.3 Ontology-based querying (UoA, HES-SO, SAG, LYN) [M1-48] • This task will provide a flexible querying front-end for the MD-Paedigree platform, addressing Ontology-based Query Formulation that will give the ability to the users to formulate queries using familiar vocabularies and conceptualisations. • In more detail, the goal is to develop a flexible module that will support query formulation by different types of users (ranging from clinicians and researchers to IT experts and computer scientists), as well as, to provide a querying interface to other MD-Paedigree subsystems like KDD tools and the Sim-e-Child/PCDR repository. • The proposed module will support both ontology-based querying by navigation that will give the ability to the user to pose a query while exploring the ontology in a GUI, as well as direct query formulation based on an extensible declarative query language. • Such language could also be used by external sub-systems. In addition, a querying translation module will transform posed queries based on mappings of T15.2 in order to generate data source oriented queries that can be planned and executed through the Athena Distributed Processing and Querying Engine (T14.3). • Finally, the proposed system will incorporate the PAROS personalisation platform (T16.2), in order to generate user profiles and to personalise query formulation, query transformation as well as the query results.	DIMITROPOULOS **Estimated % realisation** M3 M6 M9 M12	**D15.2 DCV curation tools and services to automatically and manually acquire high-quality curated data** The deliverable makes available services that allow creating curated data from available data sources. Report on delivered DCV tools and interfaces	Month 24 **Lead** RUCH **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T15.4 Multimodal case-based retrieval and query reformulation (HES-SO, SAG, UoA, URLS, MAAT) [M1-48] • Task 15.4 will develop services to allow for searching the infostructure for a large variety of information needs that will also be specified in the requirements analysis. The query possibilities include visual (content-based),textual and semantic information needs, as well as multi-modal combinations. • Retrieval of similar cases will be facilitated and will include search for similar cases based on all available data including images and incomplete data, as well as various models for data fusion. Search for images with visually similar regions of interest will also be possible based on open source tools developed in the European FP7 project Khresmoi. • As the queries to the MD-Paedigree infostructure are likely to combine several modalities, e.g. a list of symptoms in free-text and a set of structured data (e.g. age, gender, RadLex descriptors, etc), the basic approach will consist of retrieving all information available in the Sim-e-Child/PCDR data repository corresponding to features found in the original query. • Algorithms such as Rocchio relevance feedback and latent semantic indexing will mainly be used to perform the query reformulation either automatically (blind feedback) or interactively in collaboration with the user.	RUCH **Estimated % realisation** M3 M6 M9 M12	**D15.3) A multimodal case- and ontology-based retrieval service, powered with relevance feedback** The application will receive as input a user information request. The service will output similar cases. It will use these similar cases to suggest refinements in order to reformulate the input query. Final release of the case- and ontology-based retrieval service for integration and test in the infostructure by month 48. Report on delivered application and interfaces.	Month 42 **Lead** RUCH **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

WP16: Biomedical Knowledge Discovery and Simulation for Model-guided Personalised Medicine

Tasks	Lead	Deliverables	Deadline
T16.1 General data analysis and knowledge discovery tools (UoA, SAG, URLS, TBV, LYN) [M25-48] • This task will mainly focus on general – non-domain specific – reusable knowledge discovery techniques that will be used for the analysis of vertically integrated data (i.e. analysing all dimensional scales from genetic and molecular levels to clinical and behavioural) across disease areas. • This approach will include unsupervised techniques for high dimensionality reduction, similarity analysis and clustering, targeting problems such as grouping similar patients or detecting genes in microarray data that are expressed together. • In more detail, techniques such as hierarchical affinity propagation, based on message passing between data points and rough set based analysis, will be used to identify representative cases ("exemplars"), most informative features and detect patterns in data, grouping similar cases in clusters based on well-defined similarity measures. • In addition, this task will incorporate and extend the CaseReasoner application developed in Health-e-Child, in order to provide clinicians with a flexible and interactive tool to enable operations such as data filtering and similarity search over the repository to facilitate the exploration of the resulting data sets.	DIMITROPOULOS **Estimated % realisation** M3 M6 M9 M12	D16.1) First report on Biomedical knowledge discovery and simulation for model-guided personalized medicine: Overview of the tools/platforms integrated in this WP, describing both user/business requirements as well as incorporated algorithms and techniques.	Month 30 **Lead** DIMITROPOULOS **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**

Quality assurance - 1st content check entrusted to:

Tasks	Lead	Deliverables	Deadline
T16.2 PAROS Personalisation Platform (UoA) [M13-48] • PAROS (Profiling Adaptation Recommendation Service) is a system under development at UoA, whose goal is to offer personalisation, recommendation, and other adaptation services to information providing systems. PAROS' goal will be twofold in MD-Paedigree: **T16.2.1: Personalised querying:** • The goal of this task is to provide user or application specific results/context, capturing different needs, preferences or demands, depending on user type (e.g. researcher, clinicians, etc.) or discipline, as well as, querying application. This sub-task will complement Ontology-Based Data Access (T15.3). **T16.2.2: Patient profiles modelling:** • The goal of this task is to generate specific patient profiles that can be used to further adapt the provided simulation models. In this task, partners will work closely with clinicians and domain experts to study and understand how to create specific patient and disease specific profiles capturing high-level concepts and common characteristics, using ontologies as a basis. As a next step, they will develop specific similarity search techniques, mapping a specific medical case to a particular patient/disease profile.	DIMITROPOULOS **Estimated % realisation** M3 M6 M9 M12	D16.2) Beta Prototype of KDD & Simulation platform: Set of GUIs, as well as - whenever applicable - a related Application Programming Interface (API) for: • General data analysis and KDD tools, • PAROS: User and patient profile modelling, and personalization platform, • AITION: KDD and simulation platform for Model-Guided Medicine, • Data-driven drug and trial design tools to generate drug-disease associations and assess the feasibility of a given trial protocol	Month 36 **Lead** DIMITROPOULOS **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead	Deliverables	Deadline
T16.3. AITION Knowledge Discovery & Simulation Framework (UoA) [M18-48] • UoA's AITION is an advanced multi-scale KDD, DSS and simulation framework, developed as part of the Health-e-Child project, that incorporates Probabilistic Graphical Models (PGMs) as a unifying patient/diseasemodelling approach. • Its goal is to deliver highly accurate and reusable predictive – patient or disease specific – statistical simulation models combining a: • bottom-up data-driven process, i.e. analysing heterogeneous, vertically-integrated demographic, phenotypic, clinical, molecular, and genomic bio-medical data, images and streams; and a • top-down model/concept-driven process, i.e. incorporating external knowledge coming either from domain experts, literature, or model-driven processes and relational/semantic models. • This task will further extend current techniques related to PGMs, incorporating more advanced hierarchical architectures, as well as, Granular Computing (GrC) and Statistical Relational Learning (SRL) techniques, in order to provide multi-scale, multi-entity probabilistic modelling and simulation. • The ultimate goal of this task is to provide personalised disease prediction models, as well as, adapted and individualized therapy plans.	DIMITROPOULOS **Estimated % realisation** M3 M6 M9 M12	D16.2) Beta Prototype of KDD & Simulation platform: Set of GUIs, as well as - whenever applicable - a related Application Programming Interface (API) for: • General data analysis and KDD tools, • PAROS: User and patient profile modelling, and personalization platform, • AITION: KDD and simulation platform for Model-Guided Medicine, • Data-driven drug and trial design tools to generate drug-disease associations and assess the feasibility of a given trial protocol	Month 36 **Lead** DIMITROPOULOS **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead	Deliverables	Deadline
T16.4 Data-driven drug and trial design (HES-SO, LYN) [M18-48] • In T16.4, the plan is to mainly address two specific subtasks of a standard drug discovery pipeline: **T16.4.1 Hypothesis generation.** • For each disease model generated by MD-Paedigree, this task should help identifying a set of specific biomarkers likely to be targeted to positively influence the clinical synopsis of the pathology. • This tasks aims at delivering a holistic association model, directly extracted from the project's clinical data (clinical WPs and WP16) and from legacy databases (e.g. Comparative Toxicogenomic Database, PharmaGKB…). • The discovery system will include the following types of data: demographic (age, sex, etc.), phenotypic (vital signs, symptoms, laboratory results, diagnosis…), genotypic (e.g. SNPs), and therapeutic descriptors (drugs, surgery interventions…). • It will make it possible to navigate from any combination of data type (e.g. age + diagnosis +drugs) to the most strongly associated neighboring types (e.g. co-morbidities, SNPs). Each predicted association will be ranked according to a similarity estimate to express the confidence of the association. **T16.4.2 Design of clinical trials.** • A selection of trials, such as those available in the EHR4CR project (sources: ClinicalTrials.gov or TREC Medical track) will be used to design and assess the appropriateness of the project's data repository for protocol design; the main estimate being the population likely to be recruited on a given clinical site. The final interface should provide basic interaction to refine/relax the original exclusion and inclusion criteria.	RUCH **Estimated % realisation** M3 M6 M9 M12	**D16.3 Final Release of KDD & Simulation platform** Final release of related tools and API, as well as final report on Biomedical knowledge discovery and simulation for model-guided personalized medicine	Month 48 **Lead** DIMITROPOULOS **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP18: Dissemination & Training

Tasks	Lead	Deliverables	Deadline
18.3 Training Training is considered to be a fundamental task in dissemination. As anecdotal evidence has confirmed via WP4 of the VPH NoE and via feedback from the DISCIPULUS ('Roadmap Towards the Digital Patient') meeting (30/03/2012; Barcelona), training is recognized to be one of the most solid and long-lasting dissemination strategies in place. The training activities within MD Paedigree will consist of 2 'hands-on' workshops to be delivered during years 2 and 4 of the project (at approx. 1 or 1.5 year interval) in order to expose the outcomes achieved both, in disease modelling and in building the infostructure, highlighting the potential for change management and innovation in clinical workflows to the medical/clinical and research community interested in VPH technology. The first workshop will also seek to provide feedback to the research and development activities, so as to refine the outcomes for the final workshop. The workshop participants will fill in a detailed feedback questionnaire that will be passed to the developers. This task will be led by UCL, which has a long-standing commitment with the VPH Community and is involved in several training grants, including the Marie Curie ITN 'MeDDiCA', 'VPH-MIP' and WP4 of the VPH NoE.	DIAZ **Estimated % realization** M3 M6 M9 M12	D18.3) Training event in year 2: Report on the outcomes of the first Training event **D18.6) Training event in year 4:** Report on the outcomes of the second Training event	Month 30 **Lead** DIAZ **1st draft ready by:** Month 42 **Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**

	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T18.4 Seminars, Workshops, Concertation Activities with Other ICT Funded Projects, and Scenario Analysis Sessions **Lead: Vanessa Diaz** The Consortium will identify the most relevant conferences in the area and propose seminars and workshops to be held during these events. It will devote special attention and resources to Concertation Activities with other ICT funded projects and to targeted dissemination actions. Special "Scenario analyses" sessions will be convened, involving the key personnel from both the clinical and the technological partners, with the aim of pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users within MD-Paedigree. The results of the previous workshops will be presented to the Scientific Committee and to the Users' Board in order to assess their relevance and applicability, so as to refine the outcomes for a validation workshop and for a final MD-Paedigree Conference, to be held at the end of the project, targeting both internal and external clinical and research communities as well as patient organisations and the interested media. The participation in any such event will be reported in the periodic reports and the final report.	DIAZ **Estimated % realization** M3 M6 M9 M12	**D18.4.1) First scenario Analysis Sessions:** First scenario Analyses pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users. **D18.4.2) Second scenario Analysis Sessions:** Second scenario Analyses pre-empting unforeseen technical uptake problems and establishing a smooth and proactive dialogue between technology developers and end-users.	Month 24 **Lead** DIAZ **1st draft ready by:** Month 42 **Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria

Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	Lead	Deliverables	Deadline
T18.7 Engaging Parent and Patient Associations **Lead: Vanessa Diaz** Approaching Parent and Patient associations will become a part of the consortium's dissemination activities. The project will seek to disseminate news of its work, expected results and potential future developments through these channels. It is hoped that the work with Patient associations will help achieve a larger bidirectional knowledge sharing base of clinicians and of patients, and further inform the potential beneficiaries of the ongoing work.	DIAZ **Estimated % realization** M3 M6 M9 M12	**D18.1) Dissemination and training strategy plan and preliminary materials**: Roadmap defining the dissemination and training strategy, indicating the subsequent choice of preliminary materials	Month 12 **Lead** DIAZ **1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

WP19: Exploitation, HTA, and Medical Device Conformity			
Tasks	**Lead**	**Deliverables**	**Deadline**
T19.1: Evaluation approach and meaningful indicator development (EMP) • Develop upon and adapt in the VPH and other contexts proven approaches, methods and tools to the specific environment and objectives of this workpackage • Establish a set of meaningful criteria and their measurement process that are robust to demonstrate socio-economic benefit-cost impacts. The focus is • to approach and find measurements for evaluating how virtual collaborations between members of the VPH communities with different expertise are facilitated and • how consequently the uptake and acceleration of model development and integration can find meaningful expression in the overall evaluation framework.	STROETMANN **Estimated % realization** M3 M6 M9	**D19.1 HTA evaluation framework** It reviews proven approaches, methods, and tools which might be relevant to the specific environment and objectives of this workpackage, and establishes a set of meaningful criteria and their measurement process, thereby focusing on evaluating how virtual collaborations between members of the VPH communities with different expertise are facilitated.	Month 12 **Lead** STROETMANN **1st draft ready by:**

	M12			

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T19.2: Benefit-cost evaluation of MD-Paedigree infostructure (EMP) The benefit-cost evaluation will analyse and perform initially formative and, towards the end of the project, a summative evaluation of VPH Infostructure developments and outputs along dimensions like: • potential to develop newly-defined workflows for personalised predictive medicine • accessibility and usability for simulation and modelling efforts; • accessibility and ability to interface with other infrastructures (resources, tools and methods); • potential interfaces to and integration with EHR systems; • perceived and experienced benefits by type of user. Our HTA analysis tools should help us qualify, quantify, and structure findings through which we could gauge the performance of the VPH Infostructure at the project level, ultimately proving "the large scale benefits of	STROETMANN **Estimated % realization** M3	**D19.4 Clinical impact assessment scenario:** Initial formative evaluation of MD-Paedigree model-driven Infostructure based on a benefit-cost analysis approach, subsequently followed by a generic benefit-cost scenario for clinical impact assessment developed and validated with partners and experts. [month 36]	Month 36 **Lead** STROETMANN **1st draft ready by:**

having both the data and models readily available". A specific focus will be put on measuring initial ("time zero") benchmarks. Gathering feedback and monitoring performance will be based on indicators of usage, effectiveness, user satisfaction, etc. Toward the end of the project, a more in-depth survey and analysis of power users will be undertaken as well. Close collaboration with the exploitation strategies, the workflow design engines, and the HTA approaches of the VPH-Share project will be ensured.	M6 M9 M12			

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

Tasks	**Lead**	**Deliverables**	**Deadline**
T19.5 Preparing market access and medical device conformity assessment procedures (EMP) • This task, closely linked with T19.4, will prepare for the services, tools, and models of MD-Paedigree to accelerate the respective markets entrance, thereby directly supporting the notion to impact on the leadership of European industry, advancing innovative medical care beyond the end of the project, and proving the large scale benefits of having both the data and models readily available. • The task will lay the groundwork for market access by assisting partners in meeting all of the regulatory obligations required to market medical products in Europe.	STROETMANN **Estimated % realization**	**D19.6 Socio-economic impact and HTA report:** Summative evaluation of VPH Infostructure developments based on a benefit-cost analysis approach. Furthermore, it explores the relevance and processes of medical device conformity assessment procedures	Month 48 **Lead** STROETMANN

• By interpreting, in particular, EU Medical device Directive 2007/47/EC and other directives (currently under revision) and, of particular relevance to the VPH community, the consequences arising from the new definition of software as medical device. • For devices which incorporate software or for standalone software that are devices in themselves, the software must be validated according to the state of the art taking into account the principles of development lifecycle, risk management, verification and validation. • A device may only be made available on the market when it complies with the requirements laid down in this regulation and having undergone the respective medical device conformity assessment procedures -- those will be analysed, translated and made available to all partners in supporting individual and project-wide business plans such as: • safety and performance (usefulness/efficacy) • technical, semantic and legal interoperability • ethico-legal aspects of data usage. Prior to placing MD-Paedigree technologies on the market, we will undertake an assessment of the conformity of the device in accordance with the provisions of the European and national device regulations and the conformity assessment procedures set out.	M3 M6 M9 M12		for the outputs of this project, in support of partners to attempt to facilitate and accelerate market access.	**1st draft ready by:**

Self-Assessment criteria		
Measurement process and units:	**Indicators [Upper and lower limits associated with WP objectives and measurement units]**	
	Upper limits (result's maximum expectation) :	**Lower limits (below which result not acceptable):**
Quality assurance - 1st content check entrusted to:		

A.1.1.1 Timing of work packages and their components

The MD-Paedigree project partners have formalized a work plan implementing 4 major phases implying a number of conceptual steps, over 48 months of activity with 4 major milestones. The first milestone is due after 9 months and marks the end of the specification phase; the following milestones are aligned with the reporting periods of the project every 12 months.

Phase 1 (running from month 1 to 9) – Project Set-up, Requirements Elicitation, and Clinical Protocols: During Phase 1 quality assurance guidelines and a self-assessment plan will be prepared, ethical approval will be obtained , and the first dissemination activities will be performed (Step 1) Furthermore, clinical protocols for the selected paediatric applications will be established (Step 2).)Finally, the requirements for models and infostructure implementation will be analysed and documented from an end user standpoint (Step 3).

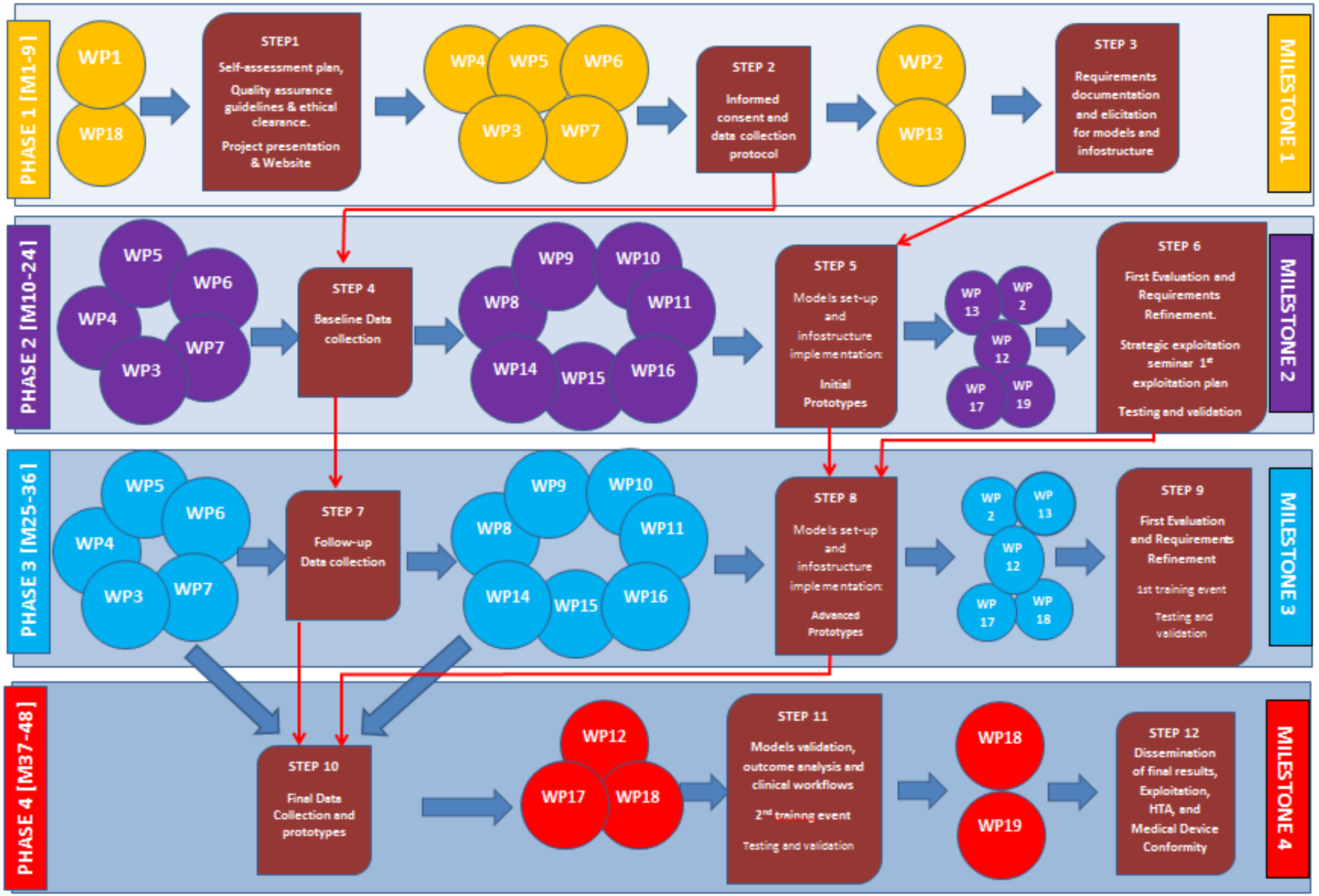
***Phase 2 (running from month 10 to 24) – Baseline Data Collection, Initial Prototypes, First Evaluation and Requirements Refinement*:** Patient enrolment will take place and data acquisition will be started (Step 4). Based on the established requirements, the existing models from Health-e Child and Sim-e-Child projects will be refined and adjusted to the new applications. The open repository for project infrastructure will be introduced and initialized with the current models and data (Step 5). First evaluations will be undertaken and requirements will be refined based on the collected experience; additionally, during this phase, the Strategic Exploitation Seminar will be held and the 1st Exploitation Plan will be drafted (Step 6).

***Phase 3 (running from month 25 to 36) – Follow-up Data Collection, Advanced Prototypes, Evaluation and Requirements Update*:** Follow-up or additional data will be acquired for all clinical applications (Step 7). The respective models will be enhanced to process longitudinal data and refined according to the obtained evaluation results. New functionalities will be integrated into advanced prototypes. The open repository will be improved and updated with content (Step 8). A second set of evaluations will be conducted and requirements will be adjusted for the final system. Furthermore, the 1st Training Event will be held (Step 9).

***Phase 4 (running from month 37 to 48) – Final Data Collection and Prototypes, Clinical Validation, and Deployment*:** In the final year, data collection will be concluded and the clinical validation will take place with the final models and simulation framework (Step 10). Results will be used to propose and disseminate improved clinical workflows. Subsequently, the 2nd Training Event will be held (Step 11). Models for all clinical applications and their respective evaluations will be documented and disseminated, while the implementation plan will be refined and the Health Technology Assessment and the Medical Clearance preparatory activities will be performed (Step 12) .

The timely delivery of all planned deliverables will be the first indicator of the fulfillment of each phase in the expected progress of MD-Paedigree, monitoring what can be demonstrable at each corresponding milestone of the project.

A second and much more detailed means of verification will be provided by the assessment criteria for each milestone and each WP which are to be defined within D1.3 Self-assessment plan on month 3.


PHASE 1 [M1-9]
WP1
WP18
STEP1
Self-assessment plan,
Quality assurance guidelines & ethical clearance.
Project presentation & Website
WP4
WP5
WP6
WP3
WP7
STEP 2
Informed consent and data collection protocol
WP2
WP13
STEP 3
Requirements documentation and elicitation for models and infostructure
MILESTONE 1
PHASE 2 [M10-24]
WP5
WP6
WP4
WP7
WP3
STEP 4
Baseline Data collection
WP9
WP10
WP8
WP11
WP14
WP15
WP16
STEP 5
Models set-up and infostructure implementation:
Initial Prototypes
WP 13
WP 2
WP 12
WP 17
WP 19
STEP 6
First Evaluation and Requirements Refinement.
Strategic exploitation seminar 1st exploitation plan
Testing and validation
MILESTONE 2
PHASE 3 [M25-36]
WP5
WP6
WP4
WP7
WP3
STEP 7
Follow-up Data collection
WP9
WP10
WP8
WP11
WP14
WP15
WP16
STEP 8
Models set-up and infostructure implementation:
Advanced Prototypes
WP 2
WP 13
WP 12
WP 17
WP 18
STEP 9
First Evaluation and Requirements Refinement
1st training event
Testing and validation
MILESTONE 3
PHASE 4 [M37-48]
STEP 10
Final Data Collection and prototypes
WP12
WP17
WP18
STEP 11
Models validation, outcome analysis and clinical workflows
2nd training event
Testing and validation
WP18
WP19
STEP 12
Dissemination of final results, Exploitation, HTA, and Medical Device Conformity
MILESTONE 4

INFOSTRUCTURE					
March 2013	**April 2013**	**May 2013**	**June 2013**	**July 2013**	**August 2013**
			Contribution to the Self-Assessment Plan		First Half-Yealry report. Delivery date: Month 6.
					Self-Assessment Plan
	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs
	Area Dedicated T&M TC	Area Dedicated T&M TC	Area Dedicated T&M TC	Area Dedicated T&M TC	Area Dedicated T&M TC

September 2013	**October 2013**	**November 2013**	**December 2013**	**January 2014**	**February 2014**
		D13.1Initial list of main requirements after stakeholder interviews including priority domains		Internal review	First periodic review
					D2.1 Initial requirements analysis document including priorities for the implementation
		D14.1 MD-Paedigree, Ground Truth Infrastructure Setup Report			D13.2 Compliance outcomes for VPH-Share and OpenAIRE influencing the infostructure
Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs	Individual WPs' TCs
Area Dedicated T&M TC	Area Dedicated T&M TC	Area Dedicated T&M TC	Area Dedicated T&M TC	Area Dedicated T&M TC	Area Dedicated T&M TC