

IST-2004-027749

HeC

Health-e-Child

Instrument: Integrated Project  
Thematic Priority: IST

## D9.2 Status report on database population and data collection status

Due date of delivery: 30 June 2008  
Actual Submission date: 12 August 2008

Start date of project: 1 January 2006  
Ending: 31 December 2009

Organisation name of lead contractor for this deliverable: I.R.C.C.S. Giannina Gaslini

Revision 4

<b>Project co-funded by the European Commission within the FP6 (2002-2006)</b>		
<b>Dissemination level</b>		
PU	Public	X
PP	Restricted to other programmes participants	
RE	Restricted to a group specified by the Consortium	
CO	Confidential	

### Document Classification

Title	Status report on database population and data collection status
Deliverable	9.2
Reporting Period:	Month 25 - Month 30
Authors	Dr. Anwar Baban, Dr. Gianluca Trocchio, Prof. Giacomo Pongiglione
Work package	9
Version	1
Keywords	Data collection, Database population

### Document History

Name	Remark	Version	Date
Anwar Baban		V 1	12.08.2008
Gianluca Trocchio		V 2	12.08.2008
Giacomo Pongiglione		V 3	12.08.2008

### Health-e-Child Consortium

The partners in this project are:

- 01 Siemens AG (Siemens)
- 02 Lynkeus Srl (Lynkeus)
- 03 I.R.C.C.S. Giannina Gaslini (IGG)
- 04 University College London – Great Ormond Street Children’s Hospital (UCL)
- 05 Assistance Publique Hopitaux de Paris – Necker (APHP)
- 06 Ospedale Pediatrico Bambino Gesù (OPBG)
- 06 European Organisation for Nuclear Research (CERN)
- 09 University of the West of England (UWE)
- 10 University of Athens (UoA)
- 11 Università degli Studi di Genova (DISI)
- 12 The French National Institute for Research on Computer Science and Control (INRIA)
- 13 European Genetics Foundation (EGF)
- 14 Aktsiaselts ASPER BIOTECH (Asper)
- 15 Gerolamo Gaslini Foundation (FGG)
- 16 Maat G Knowledge (MAAT)
17. I.R.C.C.S. Ospedale Pediatrico Bambino Gesù (OPBG)



**List of contributors**

Name	Affiliation	Co-author of:
Giacomo Pongiglione	IGG	
Anwar Baban	IGG	
Gianluca Trocchio	IGG	
Clara Malattia	IGG	
Maria Luisa Garrè	IGG	
Valeria Capra	IGG	
Andrew Taylor	UCL	
Patricia Woo	UCL	
Katherine Owens	UCL	
Younes Boudjemline	APHP	
Julie Blanc	APHP	
Ilona Alova	APHP	
Pierre Quartier	APHP	
Silvina Barcudi	OPBG	
Stephen Sanders	OPBG	
Laura Tanturri	OPBG	
Claudia Bracaglia	OPBG	

**List of reviewers**

Name	Affiliation



## Table of contents

1. Introduction .....	5
2. Data collection status .....	6
2.1. I.R.C.C.S. Giannina Gaslini (IGG) .....	6
2.1.1. Paediatric Heart Diseases .....	6
2.1.2. Inflammatory Diseases .....	6
2.1.3. Brain Tumours .....	6
2.2. Assistance Publique Hopitaux de Paris – Necker (APHP) .....	8
2.2.1. Paediatric Heart Diseases .....	8
2.2.2. Inflammatory Diseases .....	8
2.3. University College London – Great Ormond Street Children’s Hospital (UCL).....	9
2.3.1. Paediatric Heart Diseases .....	9
2.3.2. Inflammatory Diseases .....	9
2.4. Ospedale Pediatrico Bambino Gesù (OPBG).....	10
2.4.1. Paediatric Heart Diseases .....	10
2.4.2. Inflammatory Diseases .....	10
3. Problems encountered during the reporting period.....	11
3.1. IGG .....	11
3.2. APHP .....	11
3.3. UCL .....	11
3.4. OPBG .....	11
4. Plan for the next period for each participating centre.....	12
4.1. IGG .....	12
4.2. APHP .....	12
4.3. UCL .....	12
4.4. OPBG .....	13
5. Status with regard to Self Assessment Plan .....	14
6. References.....	15

## 1. Introduction

This document reports the progress made in work package 9 from month 25 to month 30. The goal of this work package is to collect the necessary clinical, imaging and genetic patient data. This activity supplies core data for the three applications of disease modelling, decision support, and knowledge discovery. Biomedical data are collected for three disease groups, namely Paediatric Heart Diseases, Inflammatory Diseases, Brain Tumours [1].

Collecting as comprehensive as possible biomedical data for each patient entering the Health-e-Child system is critical. This step not only serves the purpose of testing the data integration mechanism, but also provides training and testing data for the construction of integrated disease modelling, decision support, and knowledge discovery systems.

Our four hospitals (IGG, Genoa; APHP, Paris; UCL, London; and OPBG, Rome) are collecting patient data for Paediatric Heart Diseases and Inflammatory Diseases. Moreover, Brain Tumours study is ongoing only at IGG.

The following chapters deal with the three disease classes of Health-e-Child, they include: data collection status, problems encountered during the reporting period, plan for the next period for each participating centre and status with regard to the Self Assessment Plan.

## 2. Data collection status

### 2.1. I.R.C.C.S. Giannina Gaslini (IGG)

#### 2.1.1. Paediatric Heart Diseases

One hundred and four patients have already been enrolled. They are divided into 97 RVO which is further subdivided into 50 ASD, 46 TOF, 1 PAPVR, and 7 patients with CMPs being 2 hypertrophic and 5 dilated. For all enrolled patients clinical and imaging data have been gathered (including 32 cardiac MRI). Moreover, exercise testing for 16 patients was undertaken. Three Dimensional echocardiography was been performed in almost all the RVO patients enrolled in the project. For most RVO patients, a series of at least three RV dedicated 3D full-volume acquisitions has been gathered.

Follow-up has been performed for 26 patients (including when appropriate a follow-up MRI). Blood samples for genetic test have also been collected for 88 RVO patients.

Most of the data have been already inputted in the database and released to IT partners. Further data will be shortly inputted in the database and released.

In conclusion regarding Pediatric heart disease recruitment at IGG: the global progress assessment rate was 104.

#### 2.1.2. Inflammatory Diseases

Clinical, radiological (magnetic resonance imaging, ultrasonography and plain radiography) and laboratory data have been collected from 72 JIA patients. 30 patients have 1 year follow-up control and 2 patients have 2 years follow-up control. Imaging evaluations will be assessed to obtain semiquantitative and quantitative scores of pathological findings (synovitis, bone erosions etc.) to be used for vertical integration with other clinical and biological data.

With regard to imaging analysis, all MRI and US performed at IGG have been scored for bone erosions using a purposely devised semiquantitative scale. Semiquantitative assessment of synovitis and bone marrow edema on MRI scan is ongoing. Wrist radiographs were scored according to the modified version of the Sharp/van der Heijde score, recently adapted for use in JIA with the inclusion of new areas. A quantitative assessment of synovitis through the automated computer assisted calculation of synovial membrane volume is ongoing.

Blood samples of all enrolled patients were collected and stored in the biobank to perform immunological and genetic study. Genetic investigations is ongoing on the blood samples of all enrolled patients at IGG. Proteomic analysis is ongoing on synovial fluid and blood samples of 15 patients with clinical indication to perform arthrocentesis. We are entering clinical data on HeC database.

#### 2.1.3. Brain Tumours

The **Brain Tumours** study is ongoing. A total of 49 cases have been enrolled for the initial study and the database has been prepared as well. Clinical data of all 49 patients have been collected. Brain MRI has been performed in all children whereas imaging data have to be completed. In the last period the number of tumour tissue samples was increased to 77 cases. The tumour tissue diagnosis was been performed in all cases as the tissue sampling of all 77 cases. We are currently preparing the RNA from 28 additional cases to be used for expression studies. Hybridization of the chip arrays will be performed in these 28 cases using the Affimetrix Gene-Chip U133 Arrays (IGG). The gene expression data (microarray) will be



then accomplished for the total 77 tumour tissues and the expression data will be analysed by Siemens, DISI and UoA.

From the same cases DNA extractions will be then prepared at IGG from additional tumour samples that will be used for mutation detection in selected genes to be analysed by EGF. Sequence Analysis on PTEN, CDKN2A, PTPN11 and Netrin 1 is currently ongoing.



## **2.2. Assistance Publique Hopitaux de Paris – Necker (APHP)**

### **2.2.1. Paediatric Heart Diseases**

Clinical data is integrated in the HEC database for 102 patients (68 RVO and 33 CMP). All 68 RVO (TOF) patients have had their first visit, 40 have had their second visit, 27 more follow-up visits are scheduled to go on until December 2008. Close to all RVO patients have undergone, exercise testing, ECG monitoring, and MRI. MRI data and exercise testing is currently being integrated in the database. 29 of the 33 cardiomyopathy patients have had their follow up visit, the other follow-up visits are scheduled to go on until October 2008.

### **2.2.2. Inflammatory Diseases**

At the end of June 2008, 72 patients were enrolled. All of them had undergone all imaging tests and for each patient, blood samples had been collected and stored in biobank to perform genetic study. Clinical data were available for all patients. Thirty of these 72 patients have had a one-year follow-up control.



## **2.3. University College London – Great Ormond Street Children’s Hospital (UCL)**

### **2.3.1. Paediatric Heart Diseases**

The total baseline data obtained is 120 patients. 15 MRIs were carried out between April and June 2008 and Xrays and ECG for almost all these patients as well. Data of 17 patients were inputted on the database. The following work is still ongoing: performing and interpretation of cardiopulmonary exercise testing, physical examination + patients’ history; performing analysis and interpretation of echocardiograms; MR scanning in patients with TOF (assessment of biventricular function; identification, morphological description and quantification of severity of right ventricular outflow tract dysfunction). Work with Siemens researcher on the validation of an automated model for right ventricular volumes has continued.

### **2.3.2. Inflammatory Diseases**

Recruitment has been productive over the last 3 months At the end of June 2008, 66 patients were enrolled. There is baseline data for 58 of these patients. Since March there have been: 18 MRI scans (1 of these without contrast due to allergic reaction), 19 US scans and 19 X-ray carried out, unfortunately since then 7 patients dropped-out of the study leaving 51 patients for follow-up (parents felt that their children’s condition had improved markedly and they did not want to visit the hospital more than was necessary for the management of the child). 16 patients have had a six month follow-up



## 2.4. Ospedale Pediatrico Bambino Gesù (OPBG)

### 2.4.1. Paediatric Heart Diseases

Because of technical and organizing issues **OPBG** will begin the enrolment in July.

(this is what was stated in the QR 9 p. 8: I think we should be consistent in our declarations so either we declare the patients – without saying they were enrolled in July – or we repeat the sentence of the QR.)

Ten patients have been enrolled. The aim of our study was to enrol patients with corrected Tetralogy of Fallot (TOF) only. For all enrolled patients clinical data, including patient's history, medications and physical examination, were entered. Imaging data have been gathered including: EKG, Holter 24h, exercise testing, spirometry and color-Doppler echocardiography. Only one of the enrolled patients underwent cardiac MRI in this period. Blood samples have been collected and stored in a biobank to perform genetic studies including the three candidate genes (TXB5, GATA4, NKX 2.5). Only one patient has required other specialist consultation with an hematologist because of a C677T (MTHFR) homozygous mutation. Therefore 9 out of 10 patients did not require any specialist consultation.

### 2.4.2. Inflammatory Diseases

From the beginning of month 28, 13 JIA patients aged from 5 to 17 years, were enrolled. Clinical and laboratory data were collected from all patients. All of them had undergone radiological (X-Rays, US and MRI) studies. We studied wrists in 10 patients and hips in 8 patients with the following parameters: In patients with wrist involvement, in order to evaluate radiographic damage, we measured carpal length using the Poznanski score. About US, we used the same score proposed by IGG's colleagues to assess: presence or absence and grade of synovitis and tenosynovitis (including Doppler scoring), joint effusion and bone erosions. Regarding MRI, we evaluated synovial enhancement, tenosynovitis of extensor and flexor tendons, joint effusion, subchondral cysts, bone edema, cartilage damage and bone erosions; we have not yet validated a quantitative score for the assessment of synovial membrane volume. We are working in this direction and we think that this will represent a main goal of our study.

### 3. Problems encountered during the reporting period

#### 3.1. IGG

No significant problem was encountered for **Inflammatory Diseases** and **Brain tumours studies**.

Regarding **Paediatric Heart Diseases**, we still have not obtained a computer readable family pedigree.

#### 3.2. APHP

For **Paediatric Heart Diseases**, specific patient problem: one of the TOF patients has moved abroad and follow up will not be carried out at APHP.

Regarding the **Inflammatory Diseases** study, no major problems presented

#### 3.3. UCL

There was a minor delay adding patients to the database. This was because we needed to verify that the HeC numbers allocated to the MRs corresponded correctly with the data added to the database and also the consent forms. This has now been completed and entering data on to the database will continue at a faster pace.

**Inflammatory Diseases:** No significant problems encountered. Whilst a structure is in place to ensure that the patient numbers are met, it remains an issue and will continue to be, as parents do voice these concerns about participating in research that requires extra time at hospital when they are already losing school/personal time with standard hospital procedures. MR time has now been secured for an additional day a week in order that follow-ups scans can continue at a suitable pace to meet task objectives.

#### 3.4. OPBG

For **Paediatric Heart Diseases** data input was slow due to recurrent difficulties in connecting with HEC database. This is now resolved. Also, we believe it would be important to have a function in the system which allows to save the content in our own data backup or to print it. The other problems we have had have been related to being new in the project and we are solving them with the experience. At the moment we do not perform Cardio-pulmonary exercise testing with O<sub>2</sub> consumption evaluation at our centre. In order to gather extra information, we add the spirometry test to the exercise test.

**Inflammatory Diseases:** clinical, laboratory and radiological data collection of Inflammatory Diseases presented no significant problem. Unfortunately, as also reported by APHP colleagues, we can benefit of only 2 MRI slots per week. Moreover, in younger patients, especially in those with simultaneous wrist and hip involvement, quality of the images was partially limited by patients movement artifacts due to the length of the study.

## 4. Plan for the next period for each participating centre

### 4.1. IGG

**Paediatric Heart Diseases:** the target number of enrolled patients has been achieved. Clinical, imaging and genetic data collection will continue in the next months for those with missing and / or follow-up data. Simultaneously, the data collected will be progressively inserted in the database and delivered to IT partners.

**Inflammatory Diseases:** to date it is not possible to share imaging data among the Centres. This issue is crucial in order to allow IT partner to proceed with the quantitative analysis of JIA pathological findings and in order to go on with the scoring system validation process (i.e. to assess the multicentre reliability of the MRI scoring system etc). We asked IT partners their support in order to share MRI and US images among the three Centres.

**Brain Tumours,** in the coming 3 months the aim will be to complete the statistical and biological analysis of the expression data in the initial 49 cases and the integration with the additional 28 cases to arrive to have the expression data on 77 tumour samples totally. All the biological information will be correlated with the clinical information. The list of identified genes according to the selected clinical variables will be then validated using quantitative PCR. Furthermore, the number of cases studied for mutation of selected genes will be implemented by EGF researcher Daniela Turchetti through the sequencing of (CDKN2A, PTPN11, NETRIN – 1, PTEN), LOH analysis (NF1), Re-sequencing chip array (P53) (Asper). Collection of further MRI data is still ongoing.

### 4.2. APHP

**Paediatric Heart Diseases:** data collection for follow up visits will continue in the coming months according to the protocol, and will be progressively inserted in the database and delivered to IT partners for analysis. No particular problem is anticipated.

**Inflammatory Diseases:** clinical, imaging and genetic data collection will continue in the next months; contemporary, the data collected will be progressively inserted in the database

### 4.3. UCL

**Paediatric Heart Diseases:** no major problems anticipated for next period. Minor setbacks may occur with getting data on to the database as some medical records required are prioritised for clinical use (as opposed to research) so can not always retrieve the records on demand but this should not have a significant impact on getting the data on the database in a timely manner.

**Inflammatory Diseases:** given that the procedures carried out for HeC at UCL are not standard NHS procedures, parents can be less keen to go out of there way in order to participate. In order to accommodate the patient's wishes, we coordinate participation with patients' clinical visits as much as possible. However this is not always feasible due to patients' clinical time. Recruitment in the first month of the next phase may be more challenging due to the impending school holidays in the UK, where families traditionally go away on vacation. Nevertheless, we have been able to recruit some for July already and will



continue to make this a priority, working with the patients in order to recruit as many as possible in the next phase.

#### 4.4. OPBG

**Paediatric Heart Diseases:** clinical, imaging and genetic data collection will continue during the study period. At this point it is obvious we are not going to include follow-up. We do not have any other major problems.

**Inflammatory Diseases:** patients enrolment is in progress. Clinical, laboratory and imaging data collection will continue in the next months. For each patient, we prepared a written document with radiographic, US and MRI findings. Moreover, at the moment we filled a form only for US. We are going to insert data collected also in the HeC database. We also think that, particularly for MRI, it's is crucial to validate a quantitative analysis of JIA pathological findings and to use scoring systems shared by all the Centers involved in the Project.

## 5. Status with regard to Self Assessment Plan

According to the self assessment plan, data collection at the four hospitals at month 30 (4a data collection at the four hospitals at month 30 in Paediatric Heart Diseases; 4b data collection at the three hospitals at month 30 in Inflammatory Diseases; 4c data collection at IGG at month 30 in Brain Tumours) is to be evaluated [2].

Data collection for Inflammatory Diseases started on month 6 at IGG, on month 10 at APHP, on month 19 at UCL and on month 28 at OPBG. Data collection for Paediatric Heart Diseases started on month 10 both at IGG and APHP, on month 17 at UCL and on month 30 at OPBG. Data collection for Brain Tumours started on month 11 at IGG.

The number of cases expected to be enrolled by the first three hospitals by month 30 was supposed to be completed in each subgroup while the deadline for terminating patient's enrolment at OPBG is month 42. This was respected for Paediatric Heart Diseases (120%) and Inflammatory Diseases (105%) subgroups. In the Brain Tumours subgroup 99% of the cases were collected by month 30.



## 6. References

- [1] Health-e-Child "Project Proposal, Annex I: Description of Work, Project Phase II"
- [2] Health-e-Child D.1.5.a Self Assessment Plan