3.1 About project Z2 (E)

3.1.1 Title: Central Data and Biosample Management, Biometry and Epidemiology

3.1.2 Principal investigators

Hoffmann, Wolfgang, Prof. Dr., MPH, born 02.02.1963, German
Institut für Community Medicine, Abt. Versorgungsepidemiologie und Community Health, Universitätsmedizin Greifswald, Ellernholzstr. 1-2, 17487 Greifswald, Germany
Phone: +49 3834 867751
Email: wolfgang.hoffmann@uni-greifswald.de

Felix, Stephan B., Prof. Dr., born 03.05.1956, German
Klinik für Innere Medizin B, Universitätsmedizin Greifswald, Ferdinand-Sauerbruch-Str., 17475 Greifswald, Germany
Phone: +49 3834 868 0500
Email: felix@uni-greifswald.de

3.2 Project history

3.2.1 Report

The project Z2 “Central Data and Biosample Management, Biometry and Epidemiology” of the Transregional Collaborative Research Centre 19 (CRC 19) was responsible for the standardized data and biosample collection, the organization and documentation of the data and sample transfers and the standardized collection, archiving and aggregation of data collecting especially in the project Z1.
Therefore a centralized, web-based, federated data and biosample management system was established to allow a quality-assured data collection with standardized instruments while ensuring the privacy and data security of the CRC 19 patients. In close cooperation with the project Z1 and all clinical study centers within the CRC 19 (Greifswald, Berlin, Tübingen), data from the clinical centers are aggregated, documented, quality checked, processed, stored and provided for analyses. A large prospective database with clinical, cardiopulmonary, molecular biological, histological, immunohistochemical, biochemical and imaging parameters and data on therapy and individual course of the disease has been established.

Central Data and Sample Management
The hardware structure of the project Z2 was set up as by two parallel server cluster networks with identical software and hardware features that ensured a high level of data security (Figure 1). Basis of the central data management system was an Oracle Standard Edition 10g database. As an application and web server an Oracle Application Server was used. During the funding period the web application and physical separation of the central data and biosample management of the CRC 19 has been continuously enhanced and optimized regarding performance and ergonomics in strict compliance with the data protection concept to satisfy new requirements (e.g. (1) changing of the primary key to collect patient data, which carry the same case number although the entries refer to different hospital admissions, to ensure an one-unambiguous identification and correct matching of these cases, (2) spatial separation of our production and our back-up system and the configuration of a physically separate backup copy of the data management system, (3) upgrade of the web application software included a significant reduction of loading time and the improvement of the web application performance (4) improving of ergometry by the installation of filtering capabilities on the overview page of the electronic case report forms (5) Expansion of six database tables of the data and biosample management by a total of 352 variables to accommodate simultaneously performed left and right ventricular biopsies (including the management of this a new type of biosample)).

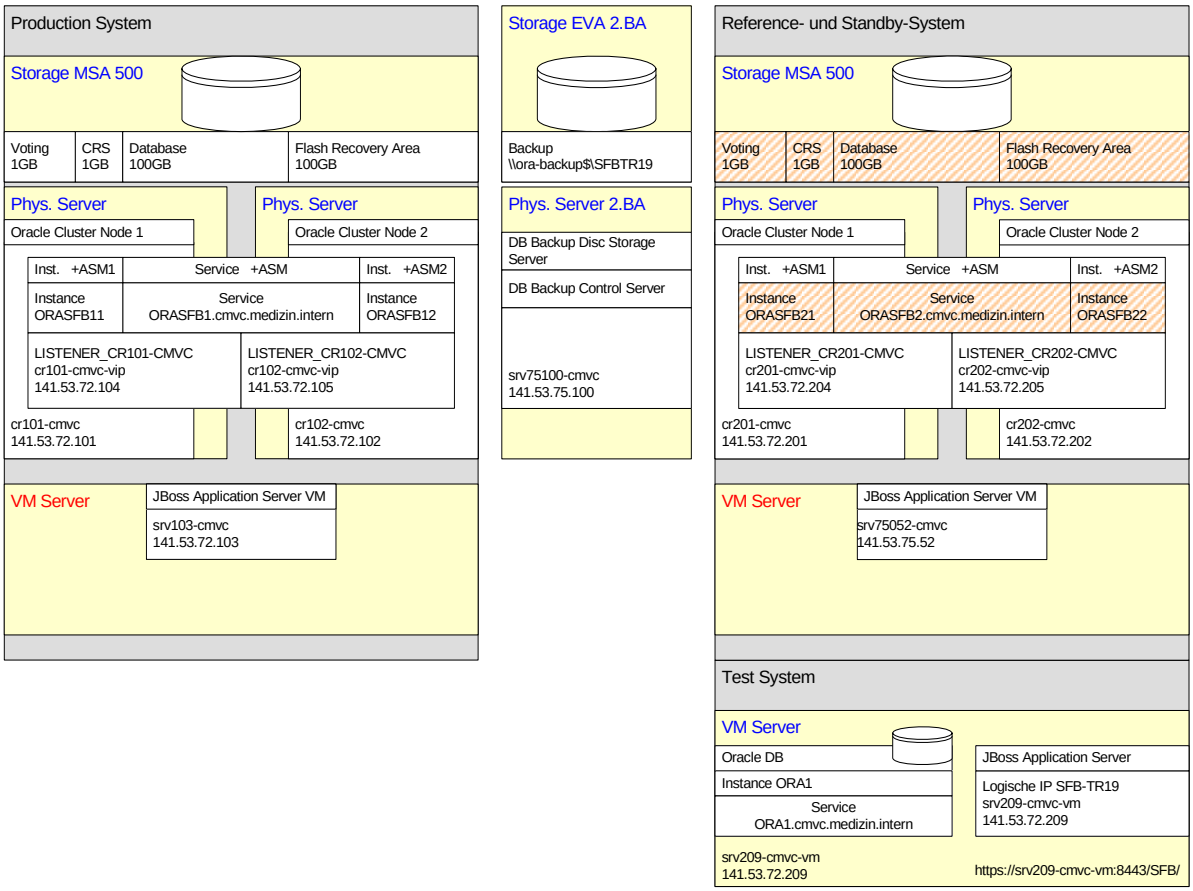


Figure 1: System overview CRC 19, VM Server: virtual machine server. The redundancy of the system with one production and one back-up system allowed refinements and new developments during normal operation of the system.

In Table 1 the solutions for technical requirements, privacy, and data security are presented. Security updates of the software were made according to a strict scheme. Questions regarding the functionalities of the central data and biosample management were handled by employees of Z2 both as on-and offline support. In accordance with the data protection concept of Z2 new staff of the clinical centers, who are involved in the data and biosample collection received individually defined access and permission rights. Corrections and amendments to the system were documented based on scripts according to best practice for comprehensive historisation.

Connection	• Encrypted VPN connection (http/https protocol) • Use of safe intranet connections • Firewall
Application access	• Access only permitted through web-frontend • IP-Verification and user ID / password required • Access limited to a pool of registered and trained users • Automatic user logout in case of inactivity (time range 30 minutes)
Access to identification data	• Additional use of 3DES encryption algorithm; access only for medical staff
High operational availability and advanced system performance	• Central database • Redundant system design and configuration (2 server cluster) • Hard disc array with RAID configuraion • Load balancing • Uninterruptible power supply, airconditioned and protected server room
Quality Management	• Daily Back up • System maintenance by trained personnel, access strictly limited

Table 1 Technical solutions in the Central Data and Sample Management for technical requirements, privacy, and data security

For the standardized data collection the survey instruments Electronic Case Report Form (eCRF) including basic patient related information and patient history, disease related complaints, family history, medications and cardiopulmonary tests as well as two self-reported questionnaires “Minnesota Living with Heart Failure Questionnaire” (german version, self-reported health status) and “Risk and Health Factor Related Questionnaire” (quality of life, health related behavior (smoking, nutrition, physical activity), utilization of medical health care system, women’s health, socioeconomic factors, co-morbidities, lifestyle) were implemented by Z2. The data of the eCRF were collected using a Remote Data Entry and Electronic Data Capture system. Pull down menus were used to simplify data entry. Results were displayed by graphical user interface. Audit trails allow tracking of changes in original data. Both self-reported questionnaires were designed as scanner compatible questionnaires, allowing an efficient documentation and a time saving automated data import (Figure 2).

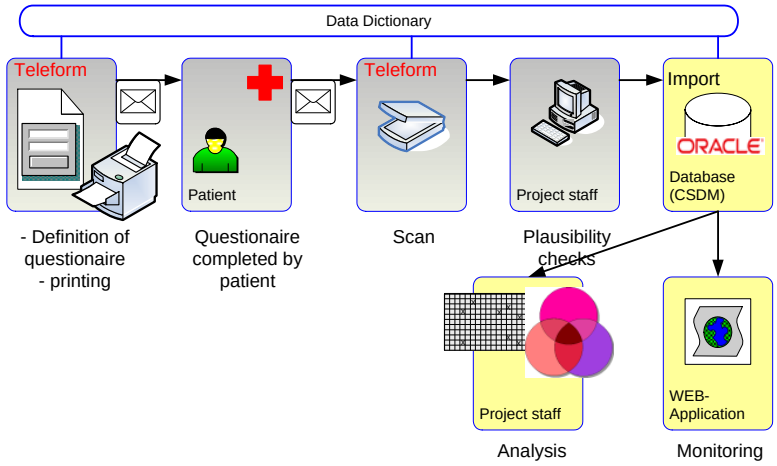
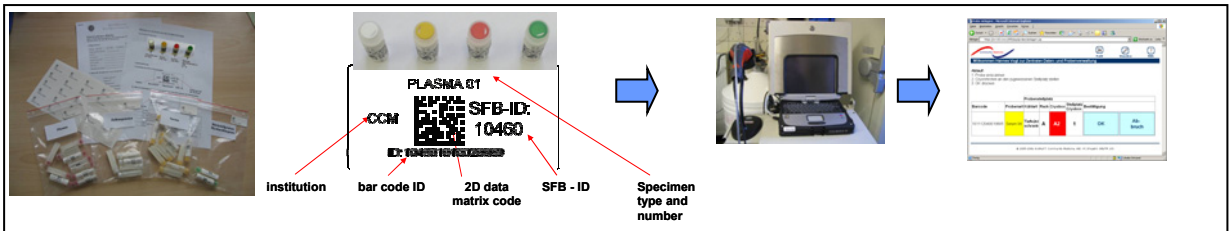


Figure 2: Data flow of scanner compatible questionnaires within the CRC 19

Sample Management

The project Z2 developed a fault-resistant sample documentation system (durable, temperature-insensitive label technique, two-dimensional barcode labeling, online documentation and data transfer via VPN connections) for the logistical administration of biological samples. Each sample vial was colour-coded and coded with a 2D- barcode label (with: SFB ID, barcode ID, sample type, with serial number of the sample, study center, checksum).



- Standard sets of pre-labelled vials
- 2D-bar code labeling and colour coding
- Interactive sample registration / retrieval using 2D barcode Scanner
- Automated calculation of storage position (CSDM system, partitioning algorithm)

- Monitoring of sample registration and retrieval
- Management support in case of cooling capacity failure
- Online and on-site support with the sample archiving software

Figure 3: Sample Management within the CRC 19

All labelled vials were scanned with a 2D-barcode scanner connected with the central database. For each sample a free storage position within in the respective cooling capacity (-80° C freezer, nitrogen storage device) was automatically assigned using a special partitioning algorithm. Data were transmitted via VPN connection to the central data and sample management system allowing the documentation of all sample vials and the storage capacity in real time (Figure 3).

As of March 17th 2014, data on 9,137 patient cases referring to 2,818 patients could be included into the database. A total number of 3,831 myocardial samples and 51,932 other biosamples (serum, plasma, cell suspension) are stored in the federated biobank of the CRC 19 located in Berlin and Greifswald (Table 2).

		Study center		
Data	**and biosamples**	**EMAU**	**CCBF**	**∑**
Patient data	Number of included patients	840	1,978	2,818
	Number of included cases	2,804	6,333	9,137
	Questionnaire Quality of Life (MLHFQ)	988	4,171	5,159
	Health and Risk Factor Related Questionnaire	397	1,252	1,649
Clinical data	Endomyocardial diagnostics	752	1,506	2,258
	Echocardiography	2,641	2,273	4,914
	ECG	2,503	1,935	4,438
	24h ECG	176	651	827
	Spiroergometry	1,955	365	2,320
	Lung function testing	1,904	1,039	2,943
Biosamples	Myocardial biopsies	1,555	2,276	3,831
	Cell suspension	3,125	7,247	10,372
	Serum	6,855	13,764	20,619
	Plasma	7,527	13,414	20,941
	∑Biosamples	19,062	36,701	55,763

Table 2: Data and Biosamples of the CRC (17.03.2014)

For quality assurance and standardized assessment of patient data and biosamples Z2 provided Standard operating procedures (SOPs), standardized study document sets, including data collection instruments, examination scheduling forms, and standard sets with biosample storage material. Comprehensive completeness and plausibility checks add to high data quality: e.g. (1) weekly completeness checks the finalized eCRF (Status: “complete entry”) using a semiautomated-interactive program, (2) multidimensional plausibility checks (comparing data from incoming paper questionnaires and results of digital data with existing database entries). Missing entries of required parameters or differing information from different data sources are requested from the editors, then completed, corrected or noted as permanently missing.
During the last year of the funding period the main aim was the completion and validation of the medical information of the central data and sample management (e.g. haemodynamic parameters, medication, disease duration) by using medical records and the clinical information system.

Digital data collection routines
To improve the data collection process the availability of digital cardio-pulmonary function data of diagnostic devices was identified. Interfaces for the digital catchment of test results were programmed and successfully implemented e.g. for ECG and Holter-ECG examination, pulmonary function testing, spiroergometry testing, and laboratory values from IKDT and clinical laboratory (Figure 4). Data export from interactive devices for cardiopulmonary function testing was supported by short instructions (e.g. cardiopulmonary exercise testing, echocardiography, Holter-ECG) addressing the respective responsible employees. Import procedures were programmed; data were processed and after verification imported into the database of the central data and biosample management. If data extraction of digital resources was not possible due to non-network enabled devices examination results were printed on paper forms and were transferred to Z2 where these sheets were scanned and made available digitally.

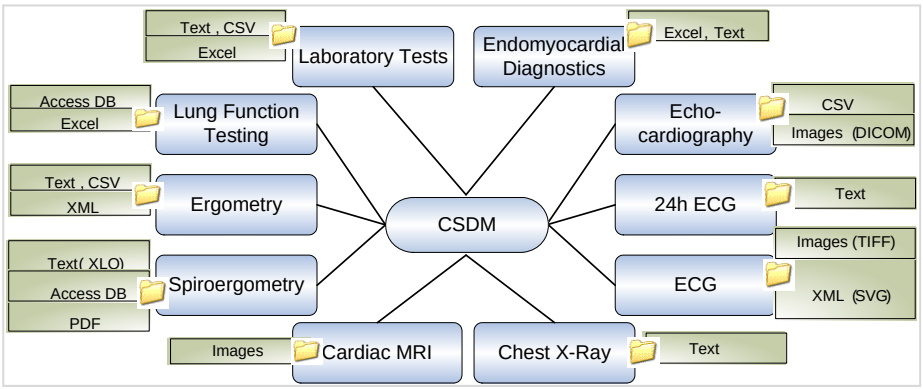


Figure 4: Import of digital data to the CSDM

For survival and time-to-event analyses information on the vital status of the study participants was collected. For patients in Mecklenburg-Western Pomerania, the mortality follow-up was organized in cooperation with the Central Information Registry (ZIR) of this federal state. Therefore an institutional data protection concept was worked out, consented with the data protection officer of Mecklenburg-Western Pomerania and approved by the Ministry of the Interior who is running the ZIR. The vital status of individuals who reside outside of MV or have moved away from MV were determined by the responsible registration offices (Einwohnermeldeamt) in the respective other federal states.

Data use and access

Z2 has developed, implemented and delivered a standardized, transparent process for the (multi-center) transfer of data and biosamples. To facilitate the application and delivery process Z2 programmed a web interface which enables the digital delivery of all data dictionaries for direct identification and selection of variables required for testing specific hypotheses for internal and external analyses relevant to the scientific goals of the CRC 19 (Figure 5).

Figure 5: Web-interface for data and biosample request

For a transparent, standardized and documented transfer of patient-related data and biosamples between Z2 and the researches Use and Access rules were defined describing in detail the process of applying data and biosamples, the evaluation of the proposal by the directors of the participating clinical institutions and the transfer of data and/or biosamples for all applicants.
In summary four documents were proposed by Z2 and after adoption by the CRC-Board in December 2009 made available online: (1.) Regulations for the use and handling of biosamples and data of CRC 19 "Inflammatory Cardiomyopathy - Molecular Pathogenesis and Therapy", (2.) Request form for data

and/or biosamples from the CRC 19, (3.) Evaluation form for requests for data and/or biosamples from the CRC 19, and (4.) Contract form. The Standard Operating Procedure (SOP) GE04 "Antragsverfahren von Daten und Proben im SFB/TR 19" describes the use of these documents. Up to the end of the funding period the central data and biosample management achieved and handled a total of 30 data-and /or biosamples requests. The Institute for Community Medicine guarantees the further use of all previously collected data and biosamples beyond DFG-funding of the CRC 19 by using available funds of the Institute.

Analysis of data from the central data and sample management, and methodological support (epidemiologic, biometric and statistic advice was solicited for study designs, power estimates, data analyses, statistical support, and graphing) has led to seven co-authorships of publications during the funding period:

Pinkert S, Westermann D, Wang X, Klingel K, Dörner A, Savvatis K, Größl T, Krohn S, Tschöpe C, Zeichhardt H, Kotsch K, Weitmann K, **Hoffmann W**, Schultheiss H-P, Spiller OB, Poller W, Fechner H. Prevention of Cardiac Dysfunction in Acute Coxsackievirus B3 Cardiomyopathy by Inducible Expression of a Soluble Coxsackievirus-Adenovirus-Receptor. **Circulation** 2009, 120(23):2358-2366

Westermann D, Kasner M, Steendijk P, Spillmann F, Riad A, Weitmann K , **Hoffmann W**, Poller W, Pauschinger M, Schultheiss HP, Tschöpe C , et al. Role of left ventricular stiffness in heart failure with normal ejection fraction **CIRCULATION**. 2008; 117(16):2051-2060.

Holzmann M, Nicko A, Kühl U, Noutsias M, Poller W, **Hoffmann W**, Morguet A, Witzenbichler B, Tschöpe C, Schultheiss H-P, Pauschinger M. Complication Rate of Right Ventricular Endomyocardial Biopsy via the Femoral Approach. A Retrospective and Prospective Study Analyzing 3048 Diagnostic Procedures Over an 11-Year Period. **CIRCULATION** 2008, 118(17):1722-1728

Grube M, Ameling S, Noutsias M, Köck K, Triebel I, Bonitz K, Meissner K, Jedlitschky G, Herda L, Reinthaler M, Rohde M, **Hoffmann W**, Kühl U, Schultheiss H-P, Völker U, **Felix S B**, Klingel K, Kandolf R, Kroemer H K. Selective Regulation of Cardiac Organic Cation Transporter Novel Type 2 (OCTN2) in Dilated Cardiomyopathy. **AM J PATHOL** 2011, 178(6): 2547-59

Herda LR, Trimpert C, Nauke U, Landsberger M, Hummel A, Beug D, Kieback A, Dörr M, Empen K, Knebel F, Ewert R, Angelow A, **Hoffmann W**, **Felix SB**, Staudt A. Effects of immunoadsorption and subsequent immunoglobulin G substitution on cardiopulmonary exercise capacity in patients with dilated cardiomyopathy. **AM HEART J** 2010, 159(5):809-816;

Kasner M, Gaub R, Sinning D, Westermann D, Steendijk P, **Hoffmann W**, Schultheiss H-P, Tschöpe C. Global strain rate imaging for the estimation of diastolic function in HFNEF compared with pressure–volume loop analysis. **Eur J Echocardiogr.** 2010, 11(9):743-751

Kasner M, Westermann D, Steendijk P, Gaub R, Wilkenshoff U, Weitmann K, **Hoffmann W**, Poller W, Schultheiss HP, Pauschinger M, Tschöpe C. Utility of Doppler echocardi-ography and tissue Doppler imaging in the estimation of diastolic function in heart failu-re with normal ejection fraction: a comparative Doppler-conductance catheterization study. **Circulation** 2007 116(6):637-47.

3.2.2 Project-related publications of the investigators

Riad A, Weitmann K, Herda L, Empen K, Groß S, Nauck M, Dörr M, Klingel K, Kandolf R, **Hoffmann W**, **Felix SB**. Initial white blood cell count is an independent risk factor for survival in patients with dilated cardiomyopathy. **Int J Cardiol**. 2013 Sep 30;168(2):1207-13

Riad A, Meyer Zu Schwabedissen H, Weitmann K, Herda L, Dörr M, Empen K, Kieback A, Hummel A, Reinthaler M, Grube M, Klingel K, Nauck M, Kandolf R, **Hoffmann W**, Kroemer H, **Felix SB.** Variants of Toll-like receptor 4 predict cardiac recovery in patients with dilated cardiomyopathy. **J Biol Chem.** 2012 Aug 3;287(32):27236-43

Staudt A, Herda L, Trimpert C, Lubenow L, Landsberger M, Dörr M, Hummel A, Eckerle LG, Beug D, Müller C, **Hoffmann W**, Weitmann K, Klingel K, Kandolf R, Kroemer H, Greinacher A, **Felix SB**. Fcgamma-receptor IIa polymorphism and the role of immunoadsorption in cardiac dysfunction in patients with dilated cardiomyopathy. **Clin Pharmacol Ther.** 2010 Apr;87(4):452-8

Angelow A, Weitmann K, Schmidt M, Schwedler S, Vogt H, Havemann C, Staudt A, **Felix SB**, Stangl K, Klingel K, Kandolf R, Kühl U, Lassner D, von Schlippenbach J, Schultheiss H, **Hoffmann W**. The German Transregional Collaborative Research Centre 'Inflammatory Cardiomyopathy--Molecular Pathogenesis and Therapy'. Methods and baseline results from a 3-Centre Clinical Study. **Cardiology**. 2009;113(3):222-30

Angelow A, Schmidt M, Weitmann K, Schwedler S, Vogt H, Havemann C, **Hoffmann W**. Methods and implementation of a central biosample and data management in a three-centre clinical study. **Comput Methods Programs Biomed**. 2008 Jul;91(1):82-90.

Angelow A, Schmidt M, **Hoffmann W.** Towards risk factor assessment in inflammatory dilated cardiomyopathy: the SFB/TR 19 study. **Eur J Cardiovasc Prev Rehabil**. 2007 Oct;14(5):686-93. Review.

3.3 Funding

Funding of the project within the Collaborative Research Centre started in July 2004. Funding of the project ended in December 2013.

3.3.1 Project staff in the ending funding period

	No.	Name, academic degree, position	Field of research	Department of university or non-university institution	Commitment in hours/week	Category	Funded through:
Available							
Research staff	1	Wolfgang Hoffmann, Prof. Dr., MPH, PI	Epidemiolgy	Dept. of Epidemiology of Health Care and Community Health	5		UG*)
	2	Stephan B. Felix, Prof. Dr., Co-PI	Cardiology	Dept. of Internal Medicine B	3		UG
	3	Marcus, Dörr, Prof. Dr.	Cardiology	Dept. of Internal Medicine B	3		UG
Non-research staff							
Requested							
Research staff	1	Kerstin Weitmann	Epidemiolgy	Dept. of Epidemiology of Health Care and Community Health	39		
Non-research staff	2	Marcus Ernst		Dept. of Internal Medicine B	39	Medical documentalists	
	3	Anna-Juliana Butzek		Dept. of Epidemiology of Health Care and Community Health	19.5		
	4	Stefanie Bohnstädt		Dept. of Internal Medicine B	39		

UG*) University Greifswald

Job description of staff (supported through available funds):

1. Prof. Dr. med. Wolfgang Hoffmann, MPH, Principal Investigator. The applicant Prof. Dr. Hoffmann lead the service project Z2. He coordinated the tasks of the central data and biosample management and was responsible for carrying out the tasks associated with quality assurance, database design

and structuring, coordinating the transfer of patient data and biomaterials and provided advice on design, epidemiologic and biometric issues throughout the partners in the CRC 19. Wolfgang Hoffmann was responsible for coordination of data and/or sample requests to sub-projects. Together with the senior scientists he contributed to a concept to sustain the central data resources of the CRC for long term future use beyond the funding period.

2. Prof. Dr. med. Stephan B. Felix, Co-Principal Investigator
3. Prof. Dr. med. Marcus Dörr, Cardiologist.
Prof. Dr. Felix and Prof. Dr. Dörr both are cardiologists and provided the clinical expertise. They were responsible for the recruiting of CRC patients and the information and solicitation of the informed consent as well as the pre-processing biomaterials. They conducted the initial clinical diagnostic assessment of patients and organized follow-up visits.

Job description of staff (requested):

1 Kerstin Weitmann, Epidemiologist
- Organisation and coordination of the vital status follow-up
- Biometric / statistical support on request
- Coordination and processing of data and/or biosample requests
- Establishment of structures to guarantee the sustainable use of previously collected data and biosamples beyond DFG-funding
- Organisation of the monitoring of hardware (server) and software components (Oracle system, JBoss) (maintenance, up-dates)
- Maintenance of the database within the central data management of the CRC 19
- Quality assurance of routine data collection instruments and digital export routines
- to provide technical assistance to the medical documentalist
- Instructions for the data entry, on-site and online support for users of the CRC
- Monitoring and maintenance of the Remote Data Entry System (import of up-dates, user management, setting up new users, etc.)

2-4 Marcus Ernst (01/2013-05/2013), **Anna-Juliana-Butzek** (06/2013-07/2013), **Stefanie Bohnstädt** (07/2013-12/2013), Medical documentalist
- Documentation of all data-related operations (incoming data exports, biosample logistics, etc.)
- Providing and processing of machine-readable questionnaires (printing, distribution, reading of completed questionnaires, completeness and accuracy control of the machine -read data, importing of data into the database of Z2)
- Providing of standardized study document sets (questionnaires, examination applications, standard sets with biosample storage material, etc.)
- Monitoring of the hardware components of the biobank and regular checks of biosample storage and biosample position in the cooling capacities
- Coordination of data exports from digital resources and import of data into the database and storage and documentation of imaging information
- Participation in the database monitoring and database maintenance
- Implementation of plausibility and completeness checks
- Completion, validation, and clinical quality control of the central database base for the technical clinical examination, imaging, and pharmacotherapy
- Validation and completion of the medical information (e.g. haemodynamic parameters) from medical records and the clinical information system