# CIPROS – An Instrument for Evidence-based Evaluation of EDC Systems and the DBFORM Example

Doris Lindoerfer, Thomas H. Müller, Kathrin Halfter, Ulrich Mansmann Ludwig-Maximilians-Universität München Institute for Medical Information Processing, Biometry and Epidemiology (IBE)

Munich, Germany

e-mail: lindoerf@ibe.med.uni-muenchen.de, mueller@ibe.med.uni-muenchen.de, halfter@ibe.med.uni-muenchen.de, mansmann@ibe.med.uni-muenchen.de,

Abstract—The success of medical research projects depends on efficient and powerful information systems. To assess the needed range of functionalities of such systems, we developed the evidence-based Checklist with Items for Patient Registry sOftware Systems (CIPROS), consisting of 72 items, organized within 12 aspects/topics. In this paper, we demonstrate how CIPROS can be used to assess the functionality of an information system. To this end, we evaluated an in-house Electronic Data Capture (EDC) system DataBase FORM generator (DBFORM). The assessment is helpful for project managers and medical scientists to select EDC systems for their own projects. The assessment may also help system developers to assess their systems and inspire them to develop new features. It may also be an efficient tool to evaluate research proposals with respect to the suitability of the selected software.

Keywords-Information Systems; Software; Checklist; Evaluation.

## I. INTRODUCTION

Efficient information systems are essential to successfully perform medical research projects such as clinical trials and medical registries. Various commercial and open-source Electronic Data Capture (EDC) systems support biomedical researchers handle complex data collections. Academic institutions often develop systems in-house in order to accommodate changing requirements. But provide existing information systems sufficient functionalities which are needed in innovative research projects?

An instrument to assess the range of functionalities of register systems is the Checklist with Items for Patient Registry sOftware Systems (CIPROS), consisting of 72 items, organized within twelve aspects/topics [1]. The CIPROS checklist was developed to evaluate existing information systems. While the evaluation has to be done in cooperation with the system specialists the result indicates to the medical scientist how appropriate the system may be for a planned project. A special elaboration paper in which each item is explained and enhanced with examples is prepared which can help to perform the systems assessment [2].

In contrast to common generic Software Requirements Specification (SRS) templates [3]-[5] or standards [6] available for developing software systems for all fields and considering the different steps which must be done in software engineering, CIPROS is a comprehensive assessment tool specifically designed for the evaluation of patient registry software systems and the specification of requirements for patient registry software systems. The evaluation can be done in a few hours or less.

Since requirements engineering is also essential in the medical domain [7], it is of interest, that the evaluation is done by applying a domain specific requirements engineering process.

The purpose of this paper is to demonstrate the use of the CIPROS checklist in assessing a registry system and how the assessment can be used to plan downstream developments of a system. We apply it to assess our in-house EDC system DataBase FORM generator (DBFORM), which has been used to implement medical research projects in recent years.

In Section 2, RELATED WORKS, the in-house developed EDC system DBFORM is presented and some projects in which DBFORM has been used are introduced. In Section III, METHODS, we first introduce the evidencebased CIPROS checklist and then the way the evaluation of the EDC system DBFORM with the CIPROS checklist was done is described. Section IV, RESULTS AND DISCUSSION, first presents the results of this evaluation. Then results are discussed, as well as the implementation of special items in special projects. Section V, CONCLUSION AND FUTURE WORK, gives a conclusion of this work and an outlook to the next planned steps.

## II. RELATED WORKS

In this section we first introduce the in house developed electronic Case Report Form (eCRF) system DBFORM. Then we describe some projects in which the DBFORM system is used to complement the IT infrastructure of the projects.

## A. The EDC System DBFORM

The system was first described in 2002 [8] and its functionality has been extended and adapted during the

following years to the specific requirements of several research projects. In brief, the core system provides a generic implementation of EDC functionality for a variety of data types, including unstructured text, numeric data, coded data, and other more structured types. Major components of DBFORM are a generic configurable web forms generator (dbform.cgi) and a form compiler (ddict.pl) that derives the appropriate configuration data from a tabulated data dictionary. The form generator runs on a platform providing a webserver and DataBase Management System (DBMS) Specifically, Linux, Apache [9], environment. and PostgreSQL [10] are used, but other environments are possible. The major implementation programming language is Perl [11]. A schematic overview of the DBFORM system and how it is used in a project is shown in Fig. 1.

The basic DBFORM project environment consists of development systems and a live (or production) system with nearly the same structure. Each type of system can accommodate multiple DBFORM instances for separate EDC projects. The development systems should be connected to a Common Software Repository (CVS). The deployment process is automated. The system developer develops the core system functionalities, while the project developer develops the project-specific functionalities and communicates with the users.

One of DBFORM's key features is its extensibility via a number of interfaces. In several instances data dictionary entries can call on other resources to perform additional processing, for example as SQL or JavaScript phrases, or via a shell-like environment providing access to any other programming language. While DBFORM computes its own form layout for convenience, overrides can be taken from custom HTML/JavaScript templates. These tools allow adaption to project-specific requirements, albeit at different levels of expertise. The basic project development requires managing the data dictionary according to the specifications in the DBFORM documentation. The following paragraphs describe several instances of this adaptability.

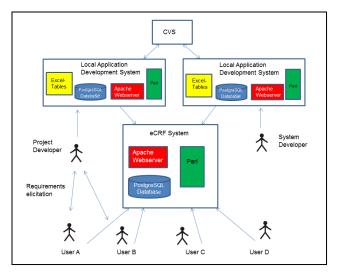


Figure 1.A schematic overview of DBFORM used in a data capture project.

## B. The EUTOS Population-based registry

The EUropean Treatment and Outcome Study (EUTOS) Population-based registry collected baseline, treatment, and outcome data from patients with Chronic Myeloid Leukemia (CML) across Europe [12][13].

Part of the requirements called for a differentiated access authorization scheme for clinical centers in various countries. Role-based authorization is included in DBFORM's core functionality and can be configured in the data dictionary. Furthermore, we complemented DBFORM with an extension to manage offline-generated queries to support this project. The queries were imported into the system ready to be processed online. Implementation made use of the "report"-programming interface, originally intended for reports, but suitable for any additional arbitrary functionality.

The extension of DBFORM with the query module is shown in Fig. 2.

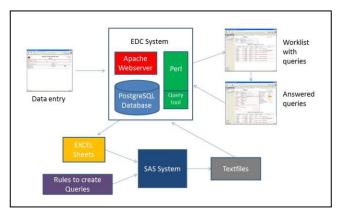


Figure 2.Extension of DBFORM with an automatic Query module.

Initially, the captured data is extracted from the database and verified offline using a SAS software system according specific rules. The query information is stored in a simple tabular format in text files. These are uploaded into the query tool of the EDC patient registry software system. Each query consists of a unique number, a query text with a question about the variable, and the corresponding unit. It is also possible to add new queries during the project. A worklist of queries can be executed by the study groups. The correct replies will be adopted automatically in the database and the query will be removed from the worklist. The described online query module is an efficient tool to communicate a large number of queries at low cost.

Fig. 3 shows a Query screenshot. The wrong value is highlighted. When the Query is answered the correct value will be inserted automatically in the database.

## C. The BreathEase study

The BreathEase study was designed to evaluate palliative care in support of patients with respiratory distress. Participating patients were visited and interviewed by a study nurse in their homes using a variety of questionnaires. The responses were documented in an offline instance of DBFORM, residing on a notebook computer as a mobile system. In this case, the challenge consisted of providing a means to update the mobile "slave" systems and stationary "master" system with the study nurse's assignment details and questionnaire responses.

Datei Bearbeiten Ansicht Gehe	Lesezeichen Extras	Eenster Hife						
🛖 . 🐵 . 🗣 🖸	http://fw.be.	ned.uni-muenchen.	de:900/cgi-bin/re	gistries/eutos/report.pl	-session=T0ux5384	VAAEAAGmittloAAAAD; queries 💦 🔜	Suchen 🚳 -	
Zurück Vor Neuladen Sto	00						Drucken	
Catsete Clesezechen Me						Dne Smoodev.org SeaMonkey deuts	ich 👒 Spitzer_NEngU1992.p.	
v Suchen	EUTOS for CML	Patient selection		EUTOS for CML - Querie	5			
Socie v		Qu	eries			loggi	ed in: gtest	
Section System	Attribute	Filter				Date of laboratory analysis	42020000	
Suchmaschinen verwalten	Country/Region			-			12.5	
SALTER OF RET PERMANENT.				slow-up section		Hemoglobin [g/dl]		
		Baseline section Follow-up section Hematologic data				Hematocrit [%]	37	
		Limit 100 Open gueries only				WBC [10^9/L]	44.9	
						Platelets [10^9/L]	50	
	Country/Region	IT-1/Emilia R IT-1-044	omana			Blasts [%]	1	
	Patient Section					Basophils (%)	4	
	Section Baseline section eCRF part Hematologic data					Eosinophils [%]	1	
	eCRF item(s) Platelets					Was the above analysis carried out before any		
	Court period     C						Ono ⊙yes	
						treatment ?		
	Response date							
		Forgot to ins	to insert.					
	Reason	a second s						
	apply filter save data and send reply close							
	Country/Reg	ion Patient	Section	eCRF part	eCRF	Query date Q	uerv	
	> IT-1/Emilia Romana	IT 1 044	Baseline	Hematologic	item(s) Platelets	26/10/2011 Missing value		
	>IT-1/Emilia Romana	IT 1 050	Baseline section	Hematologic data	Hematocrit	26/10/2011 Please check valu Hematocrit, since	ue for variable it is <10	
> Lesezeichen	T-1/Emilia	IT-1-070	Baseline	Hematologic	Basophils	26/10/2011 Missing value		
> Chronk	Romana		section	data	e-a-o-band	Karranka minabalig todo		
> Adressbuch	>IT-1/Emilia Romana		Baseline	Hernatologic data	Eosinophils	26/10/2011 Missing value		

Figure 3.A selected Query with answer and reason.

The BreathEase implementation was based on a generic synchronization module that can be configured by a special part of the data dictionary.

The additional functionality is found in the web application's patient selection or "home" menu as shown in Fig. 4.

Most clinical research projects can be documented with this repertoire of automatically generated layouts: lists, plain forms, and overview forms. In addition, more specific combinations and options are available.

Fig. 5 shows an example of an ordinary questionnaire or Case Report Form (CRF), documenting the patients Integrated Palliative Outcome Scale (IPOS).

#### D. The HTCR project

Two instances and two major extensions were necessary to fill the requirements of the bio-banking project HTCR [14]. The first is related to identifying and characterizing biomaterial samples, i.e., including it in the data model as a separate entity.

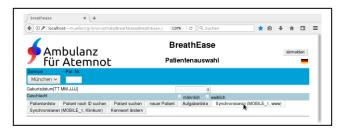


Figure 4.This is the user's root menu where all tasks begin. The mouse pointer marks the button triggering the synchronization process.

1										
🖕 Ambulanz		BreathEa	se					oh	melde	
	Integrated Pa	illiative Outco	e Outcome Scale (IPOS)							
Zentrum - Pat-Nz. Minchen - 12		offen		Viste B:	soine					
Die folgenden Fragen sollen ihre Beschwerden und Sorgen in den letzten 7 Tagen mögl	ichst vollständig erfass							_		
	1. Rente und Finanze	n								
	2. Arbolt									
F2. Unten finden Sie eine Liste mit Symptomen, die Sie unter Umständen haben. Bitte	3.								_	
F2: Unran tinden bie eine Liste mit bymptomen, die bie unter Umstanden haben. Bite kreuzen Sie an, wie sehr diese Symptome Sie in der letzten Woche beeinträchtigt haben.	Schmerzen				<ul> <li>Mailig</li> </ul>					
	Atemnot				<ul> <li>Maßig</li> </ul>					
	Schwäche oder fehle				<ul> <li>MADig</li> </ul>					
	Übelkeit (das Gefühl	erbrechen zu müssen								
	Erbrechen		<ul> <li>Gar nicl</li> </ul>							
	Appetitiosigkeit		<ul> <li>Gar nicl</li> </ul>							
	Verstoplung Mundtrockenheit oder schmerzhalter Mi		Gar nick	ht 🔍 Din meng						
				🔹 Ein wenig						
	Schillingkeit					<ul> <li>Stark</li> </ul>				
	Eingeschränkte Mobi	R.M.	Gar nich	M. 🔍 Din mariy						
Bitte norman Sie weitere Symptome, die Sie in der letzten Woche beeinträchtigt haben und kreuzen Sie an, wie sehr diese Symptome Sie beeinträchtigt haben.	1. Brustschmerzen m	t Engegefühl	Gar nicht	C Ein wenig	O Masig	C Stark	Extrem stark			
und kreuzen sie an, was sehr diese symptome sie beentrachtigt haben.	2.		Car nicht	C Ein wenig	O Millio	C Stark	C Extrem stark			
	3.		C Gar nicht	C Ein wenig	O Masig	C Stark	C Extrem stark			
F3. Waren Sie wegen ihrer Erkrankung oder Behandlung besorgt oder beumuhigt?	• Gar nicht									
F4. Waren ihre Familie oder Freunde ihretwagen besorgt oder beunruhigt?	● Garnicht ○ Se									
F5. Waren Sie traung bedrückt?	O Carriett O Ta	Manchmal	O Manager	<ul> <li>Immer</li> </ul>						
F6. Waren Sie im Frieden mit sich selbst? F7. Konnten Sie ihre Gelühle mit Ihrer Familie oder ihren Freunden teilen, so viel wie Sie	O Immer O Meist	ens O Manchimal		O Gar richt						
wollien?		en Olderchreit		Gar nicht						
F8. Haben Sie so viele Informationen erhalten, wie Sie wollten?				🕽 Gar richt.						
F9. Wurden praktische Probleme angegangen, die Folge Ihrer Erkrankung sind (2.8. Imanzieller oder persönlicher Art)?	Probleme teilweise angegangen 🛛 👻									
F10. Wie haben Sie den Fragebogen ausgelült?	Mit der Hilfe eine	s Mitarbeiters		~						
ändern abschließen weiter mit: Generic HPQL (EQ-5D) zurück zur Visite	zurück zur Übersicht	zurück zur Potierte	insuswahi							

Figure 5.A questionnaire that is part of the "Baseline" documentation.

The second enhancement is related to data protection requirements, and consists of separating identification data and medical data into two separate databases. These are linked using a common arbitrary key and an encrypted version as a logical link between corresponding records. This link can only be accessed by authorized users of both datasets, thus providing additional privacy protection. Again, logical links were implemented as an extension subject to configuration in the respective data dictionaries.

## E. The PASTURE project

The PASTURE project [15] investigated the conditions for allergy development during childhood in four different countries and required three different languages (Finnish, French and German), as well as English as the common language. While translated forms are part of the DBFORM core functionality, site-specific questionnaire dependencies and fixups at the template level were needed to correctly implement some of the country-specific variants (e.g., slightly different questions asked, different units of measure, etc.) of the questionnaires used.

## F. The RESIST study

The RESIST study is a translational research project using xenograft mouse models to find an individualized therapy for secondary resistance for colorectal carcinoma patients. The RESIST study consists of two parts. The first part is a registry for an avatar model. Patients are registered and have to agree to give a tumor tissue sample to be used in the avatar model. The second part is the clinical study. Special challenges in this project are the combination of the two studies with clinical and biomedical data, the collaboration of different user-groups, and a sophisticated data protection model. We extended DBFORM to provide various pseudonyms.

## III. METHODS

In this section we first introduce the evidence-based CIPROS checklist. Then we describe how the CIPROS checklist is used to evaluate the DBFORM system.

## A. The evidence-based CIPROS checklist

CIPROS is an evidence-based checklist, to assess patient registry software systems [1]. It consists of 72 items, organized within twelve aspects/topics. The CIPROS checklist was developed after an initial Systematic Literature Review (SLR) in PubMed. The papers were analyzed using Qualitative Content Analysis (QCA) methods published for social research [16] and adapted to the field of patient registry software systems. In an iterative process the items were assorted in the CIPROS checklist.

In this paper we describe the use of CIPROS to evaluate the in house developed EDC system DBFORM.

#### B. Evaluation of the EDC system DBFORM with CIPROS

The evaluation was performed by two raters (DL and TM). First, each rater evaluated DBFORM with the CIPROS items independently. Each item of the CIPROS checklist was considered and whether or not the item was implemented in DBFORM or not. The evaluation showed that a simple yes/no answer is not for all items possible. Some items are implemented fully in DBFORM and some are not. Some items are implemented for individual projects, while a few items are only partly implemented. Other items can be configured for special projects, if necessary. Both raters discussed their results and reached an agreement in the answers for all items. For some items it was easy to find an answer, while other items needed to be discussed to find an appropriate answer. In the latter case it was very helpful to refer to the elaboration paper [2] and review the item to clarify the meaning of the item.

In order to avoid misunderstandings, and for unambiguous assignment, we recommend using the full CIPROS checklist [1] in the evaluation where each item is described and not just the Aspect/Topic list, as shown in Table I. If there are still ambiguities in special items, we recommend using the elaboration paper [2], in which each item includes examples from the literature and explanations by the authors, for the evaluation.

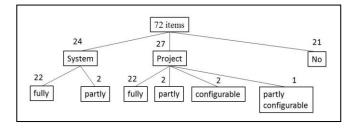


Figure 6.Evaluation result of the EDC system DBFORM with CIPROS.

## IV. RESULTS AND DISCUSSION

In this section we first present the cumulative evaluation results and the detailed results for each item. Then we give further insights in the implementation of the items and discuss them.

## A. Overall evaluation results

The evaluation of DBFORM shows substantial conformity with CIPROS. Compliance is achieved in 44 out of 72 items: 22 are fully implemented in the system and another 22 are implemented at the project level.

In four items DBFORM agrees partly with CIPROS, two are implemented partly in the system and two are implemented partly in projects. Two items can be configured in projects and one item can be configured partly at project level. Only 21 items of CIPROS are not available in DBFORM. The evaluation result is shown in Fig. 6.

The overall result of the evaluation for each CIPROS item is shown below in Table I. Because of the limited space only the Aspect/Topic and the Item-No., of each item and the corresponding result if it is implemented in DBFORM or not is listed. For a description of each item we refer to the full CIPROS checklist [1].

If the item is fully implemented in the DBFORM system, the answer is "Yes, System" if it is implemented for special projects, the answer is "Yes, Project", if the item is partly implemented the answer is "Partly System" or "Partly, Project", respectively. If it is not implemented, the answer is "No", if the item is configurable for Projects it is stated with "Conf. Project".

A summary of the items by the possible answers is given in Table II. The first column contains the possible evaluation answers for the items. The second column contains the item numbers and the third column gives the summary number of the items for this answer.

#### B. Features generally implemented in DBFORM

Here we describe shortly the items which are generally implemented in the system DBFORM.

DBFORM provides a multi-tier system architecture and a framework for the development of new projects. DBFORM also provides a table-based questionnaire builder. The framework supports early field tests. The system provides a web-interface which is compatible with the major web-browsers. In addition E-mail alerts are possible. The system provides a programming interface with third-party access and an Application Programming Interface (API) for inserting and retrieving data. Extensibility is possible. All data types are supported in the system. The system has an interface for a manual data check. Datasets can be downloaded, complete or selected cohorts only. The system has a role-based authorization module and supports encrypted data transfer. All changes are documented in an audit trail. The server is behind a firewall and located in a server room. The costs are controlled, because we have to pay no license costs for the system.

CIPROS ITEMS.						
Aspect / Topic	Item- No.	In DBFORM implemented				
System Architecture	1.1	Yes, System				
Platform independence	1.2	No				
Open Source	1.3	Partly, System				
Design model	2.1	No				
Framework-based design	2.2	Yes, System				
Questionnaire builder	2.3	Yes, System (table-based)				
Usability testing	2.4	Yes, System				
Performance testing	2.5	Partly, Project (estimated)				
Web Interface	3.1	Yes, System				
Compatibility	3.2	Yes, System				
Email-alert	3.3	Yes, System				
Messaging interface	3.4	Yes, Project				
Online discussion forum	3.5	No				
Mobile interface	3.6	Partly, Project (offl. No app)				
Patient interface	3.7	No				
Third party access	3.8	Yes, System				
API for inserting data	3.9	Yes, System				
API for retrieving data	3.10	Yes, System				
Data update mechanism	3.11	Yes, Project				
Interface to HIS / CIS	3.12	No				
Integration of biological data	3.13	Yes, Project				
Extensibility is possible	3.14	Yes, System				
CRFs	4.1	No				
Data	4.2	No				
Metadata	4.3	No				
Vocabularies	4.4	No				
XML Schema	4.5	No				
Multilingualism	5.1	Yes, Project				
Pseudonymous patient identifier	6.1	Yes, Project				
CRF is divided in parts	6.2	Yes, Project				
Customizable CRF parts	6.3	Yes, Project				
Minimal and extended dataset	6.4	Yes, Project				
All data types are supported	6.5	Yes, System				
Special data types are possible	6.6	Yes, Project				
Multiple choice is used	6.7	Yes, Project				
No predefined selection	6.8	Yes, Project				

Aspect / Topic	Item- No.	InDBFORM implemented
Data validation components	6.9	Yes, Project
Data query tool	6.10	Yes, Project
Interface for manual data check	6.11	Yes, System
Manual data queries	6.12	No
Data Query Flags	6.13	No
Plausibility Flags	6.14	No
Insertion of unplanned visits	6.15	Yes, Project
Software ergonomics	6.16	Partly, System
Query builder for researchers	7.1	No
Report generation	7.2	No
Download of datasets	7.3	Yes, System
Graphical Presentation of results	7.4	No
Risk Analysis	7.5	No
Authorized users	8.1	Yes, System
Role-based access	8.2	Yes, System
Encrypted data transfer	8.3	Yes, System
Encrypted data storage	8.4	No
Audit trail	8.5	Yes, System
Master-Slave replication	8.6	Conf.,Project
Backup management	8.7	Yes, System
Firewall	8.8	Yes, System
Server room	8.9	Yes, System
Data Protection concept	9.1	Yes, Project
Double pseudonymization	9.2	Conf., Project
Costs	10.1	Yes, System
	1011	(no licence costs)
Multi-client capability	10.2	No
Update mechanism	10.3	No
Source documentation in pdf	10.4	No
Compliance with regulations	11.1	Yes, Project
Informed Consent	11.2	Yes, Project
Rights on the data	11.3	Yes, Project
Data protection guidelines All	11.4	Yes, Project
User manuals	12.1	Yes, Project
User training	12.2	Yes, Project
User feedback	12.3	Yes, Project
Online help	12.4	Partly conf., Project

### TABLE I. EVALUATION OF THE EDC SYSTEM DBFORM WITH THE CIPROS ITEMS.

Is the Item implemented in DBFORM?	Item numbers	Σ
Yes, System	1.1, 2.2, 2.3, 2.4, 3.1, 3.2,         3.3, 3.8, 3.9, 3.10, 3.14, 6.5,         6.11, 7.3, 8.1, 8.2, 8.3, 8.5,         8.7, 8.8, 8.9, 10.1	22
Yes, Project	3.4, 3.11, 3.13, 5.1, 6.1, 6.2, 6.3, 6.4, 6.6, 6.7, 6.8, 6.9, 6.10, 6.15, 9.1, 11.1, 11.2, 11.3, 11.4, 12.1, 12.2, 12.3	22
Partly, System	1.3, 6.16	2
Partly, Project	2.5, 3.6	2
Configurable, Project	8.6, 9.2	2
Partly configurable, Project	12.4	1
No	1.2, 2.1, 3.5, 3.7, 2.12, 4.1,         4.2, 4.3, 4.4, 4.5, 6.12, 6.13,         6.14, 7.1, 7.2, 7.4, 7.5, 8.4,         10.2, 10.3, 10.4	21
		72

TABLE II. SUMMARY OF THE EVALUATION OF DBFORM WITH CIPROS.

## C. Features partly fulfilled in DBFORM

Open-source components are used to create DBFORM but it is not yet made open-source, so this item is only partly fulfilled. Software ergonomics, defined in ISO 9241-110 are only partly fulfilled in DBFORM, because they are mostly dependent on the implementation of web-forms for the special project.

## D. Features implemented in DBFORM in special Projects

In this sub-section, we describe the items which are implemented in special projects. This generally indicates that while the feature is not strictly part of DBFORM, there is an appropriate configuration or at least a workaround to obtain a satisfactory result.

Many of the Project items are fulfilled in several projects. For example item 6.1, pseudonymous patient identifier, is implemented in several projects, for example in EUTOS and in RESIST. Also item 6.2, CRF is divided in parts, item 6.5, all data types are supported, item 6.7, multiple-choice is used, and item 6.8, no predefined selection, are used in EUTOS and in RESIST. Also the regulatory items 11.1, compliance with regulations, 11.2, informed consent and 11.4, data protection guidelines, are fulfilled in EUTOS and in RESIST.

## 1) The EUTOS Population-based registry

In the EUTOS population-based registry we implemented an automatic query tool to perform queries as explained in Fig. 2. This enhancement is implemented in the system at the project level. The questions can be answered by the users and the given results were automatically inserted into the database, so Items 3.11 and 6.10 of CIPROS are implemented in EUTOS with this feature. We have also a messaging interface in EUTOS to send messages to the users, this fulfilled item 3.4. Item 6.15, insertion of unplanned visits is also implemented in EUTOS, since it was necessary to collect the results of all performed cytogenetic and molecular samples.

Since EUTOS was a pan-European project with many participants we held a user training session at the study start. We also provided a user manual and collected user feedback. Hence the items 12.1, 12.2 and 12.3 of CIPROS are implemented in EUTOS.

All participating study-groups had the right to access their own data at any time during the project phase from the central database by placing a request via the project manager. With this statement item 11.3 was fulfilled in EUTOS. In the EUTOS population-based registry we had many studygroups and users, we determined the performance of the system before we started, so item 2.5 performance testing was estimated in this project.

## 2) The BreathEase study

The mobile system used in the BreathEase study to collect the answers from the patients needed to be synchronized with the central database. The synchronization mechanism was implemented as a DBFORM extension subject to configuration in the data dictionary [17]. So with this study item 3.6, mobile interface, from the CIPROS list was partly fulfilled. This means that data can be collected offline and integrated in the central database at a later time, however, there is no special app for mobile phones.

3) The HTCR project

Due to the implementation of an identification module of biomedical data and a special privacy model item 3.13, integration of biological data and item 6.1, pseudonymous patient identifier are fulfilled within this project [18]. Identification data and medical data are stored in two separate databases. Corresponding records are linked using a common arbitrary key and an encrypted version as a logical link, which can only be accessed by authorized users.

## 4) The PASTURE project

DBFORM supports multilingualism, which means that the complete eCRF can be displayed in different languages. Since this study was performed in Germany, France, and Finland, all items were translated and inserted in different languages in addition to English. The user can select the language in which the eCRF will be displayed. With this feature item 5.1, multilingualism, of CIPROS is fulfilled.

5) The RESIST study

In the RESIST study a sophisticated data protection concept was established, this included double pseudonymization. DBFORM was enhanced to support this data protection concept [19]. A schematic overview how the role-based access and the double-pseudonymization with different identifiers for different users were implemented is shown in Fig. 7. The users in the clinics have access to the medical data with a Patient IDentifier (PID). They retrieve the tumor\_no and send it with the tumor to the pathology. The users in the pathology have access to the molecular tumor analysis data with the tumor\_no. They are able to retrieve the av\_no and send a tumor sample with this av\_no to the laboratory with the AVATARMODEL.

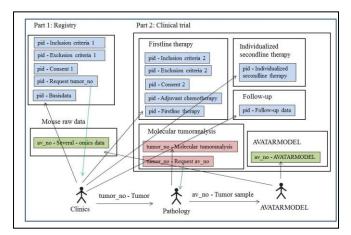


Figure 7.Simplified data model of the RESIST study.

The users in the AVATARMODEL have access to the data in the AVATARMODEL with the av\_no. Using this mechanism the data is strictly separated through different identifiers for different users and role-based user access. So in this project, beside other items mentioned before, items 9.1 and 9.2 of CIPROS are fulfilled.

#### 6) Influence of the project features for CIPROS

Since some projects were performed before the CIPROS checklist was developed, the project features also influenced the development of the CIPROS checklist by bringing own experiences into the development of the CIPROS items. For example item 6.10 Data query tool, was implemented in the EUTOS Population-based registry and then introduced in the CIPROS checklist as own experiences.

The mobile interface which was implemented in the BreathEase study is also part of the CIPROS checklist, but this item is not solely based on own experiences, since it was also found by the SLR.

#### E. Configurable Features

Item 8.6, Master-slave replication is not yet implemented but it is configurable, the same applies to item 9.2, doublepseudonymization. We have no complete online-help implemented, but some features, for example mouse-overfield help can be configured, so we considered item 12.4 online-help as partly configurable.

#### F. Features not implemented in "DBFORM"

Our system runs on Linux, therefore we decided to reply "no" for item 1.2 platform independence. Since all major components (DBMS, webserver, and the main programming language) are available on multiple platforms, DBFORM should be portable to other environments with reasonable effort, but this has not yet been attempted. The system was not developed following a design model, item 2.1, but we have our own design model for new projects. The idea is shown in Fig. 1, which is an adaption of agile software development. We have no implemented online discussion forum, item 3.5, since we have mainly relied on email correspondence with the users until now. We have no patient interface, item 3.7 implemented, because there has not been any project until now for which it was necessary. We have no interface to Hospital Information Systems / Clinical Information Systems (HIS/CIS), item 3.12. We replied "no" standardization of CRFs, data, metadata, and for vocabularies (items 4.1 - 4.4) because we have no implemented thesaurus. However, we use standardized answers with multiple choice menus. We have no eXtensible Markup Language (XML)-procedure for data exchange (item 4.5.), manual data queries (item 6.12), are not possible also data query flags (item 6.13), and plausibility flags (item 6.14), are not implemented. We have no query builder for researchers (item 7.1), and report generation (item 7.2), graphical presentation of results (item 7.3), and risk analysis (item 7.5), are not possible. Encrypted data storage at the single-item level (item 8.4) is not possible, except for user passwords. Multi-client capability, (item 10.2), is not supported, there is no update mechanism, (item 10.3), and source documentation of CRFs in pdf format (item 10.4) is also not possible.

## G. Planned improvements of DBFORM

The evaluation of DBFORM with the CIPROS checklist showed some shortfalls of DBFORM. In the near future we want to implement some of the features proposed in the CIPROS checklist. For example, source documentation in pdfs (item 10.4), was regarded as very helpful and will be implemented soon. Also item 3.7, patient interface, is on the list of features to be implemented. It will most likely rely on the more general tool of form templates. Also the standardization of CRFs, data, metadata and vocabularies is regarded as very important and is considered for implementation. If a project includes monitoring visits in the study centers it is also very helpful to document the results of these visits in the system near the captured data. Therefore the system should provide data query flags (item 6.13).

### V. CONCLUSION AND FUTURE WORK

The evaluation detailed above has shown that the CIPROS checklist is a practical solution for the assessment of patient registry software systems. It also showed that a simple yes/no answer is not possible for many of the items. So a more differentiated assessment was applied (implemented at System or Project level, partly implemented, etc.) see Table II. It is also highly recommended to use the full CIPROS checklist [1] for the assessment, and not just the aspect/topic list as shown in Table I of this paper. If the two raters chose different answers when the intention of the item was not clear the elaboration paper [2] helped to clarify the meaning of the items.

The evaluation also showed that it would be very helpful to have a quantification mechanism for the answers to generate a rating scale. So if different patient registry software systems would be compared the users would have an objective instrument to choose the appropriate system for their projects. This rating scale may be developed in the future.

It is also planned to update the CIPROS checklist with a new SLR and with input from other persons, for example

performing a workshop or establishing a focus group to get input.

The evaluation was successful for both, the system DBFORM because it showed the advantages and the deficiencies, and the CIPROS checklist because it showed that there are possibilities for improvements, such as a scale to rate the replies.

This was the first evaluation of an EDC system using the CIPROS checklist.

The presented evaluation of DBFORM with CIPROS can be a template for other researchers to evaluate their systems using the CIPROS checklist. It can also be an inspiration for scientists and system developers to develop new features of their own systems.

#### REFERENCES

- D. Lindoerfer, U. Mansmann, "A comprehensive assessment tool for patient registry software systems: the CIPROS checklist," Methods Inf Med. vol. 54(5), pp. 447-54, 2015, doi: 10.3414/ME14-02-0026.
- [2] D. Lindoerfer, U. Mansmann, "Data for the elaboration of the CIPROS checklist with items for a patient registry software system: Examples and explanations," Data in Brief, 2017 Aug 3; 14:494-497. doi: 10.1016/j.dib.2017.07.075. eCollection 2017 Oct.
- [3] K. Wiegers, J. Beatty, Software Requirements, Third Edition. Redmont, Washington. Microsoft Press. 2013
- [4] AD. McKinnon, "Software requirements specification template," [Online] Available from: http://users.tricity.wsu.edu/~mckinnon/cpts322/cpts322-srsv1.doc. 2017.08.18
- [5] Volere Requirements Specification Template. [Online]. Available from: http://www.volere.co.uk/template.htm. 2017.05.07
- [6] IEEE STANDARD. 29148-2011 Systems and software engineering -- Life cycle processes --Requirements engineering. [Online] Available from: https://standards.ieee.org/findstds/standard/29148-2011.html. 2017.08.18
- [7] Kossmann M, The significance of requirements engineering for the medical domain, J Med Eng Technol. 38 (2014), 238-43. doi: 10.3109/03091902.2014.918199.

- [8] TH. Müller, K. Adelhard, "A web-based central diagnostic data repository," Stud Health Technol Inform. vol. 90 pp. 246-250, 2002.
- [9] Apache HTTP SERVER PROJECT. [Online]. Available from: https://httpd.apache.org/ 2017.08.18
- [10] PostgreSQL. [Online] Available from: https://www.postgresql.org/ 2017.08.18
- [11] The perl programming language. [Online]. Available from: https://www.perl.org/ 2017.08.18
- [12] D. Lindörfer, TH. Müller, "Integration und analyse von patienten mit chronischer myeloischer leukämie (CML) in 25 europäischen ländern mit webbasierter elektronischer datenerfassung," in: P. Schmücker, K-H. Elsässer, S. Hayna, Eds.: Tagungsband der 55. GMDS Jahrestagung 2010, Mannheim, pp. 494-5, 2010.
- [13] VS. Hoffmann, et al., "The EUTOS population-based registry: incidence and clinical characteristics of 2904 CML patients in 20 european countries," Leukemia. vol. 29(6) pp. 1336-43, Jun 2015, doi: 10.1038/leu.2015.73.
- [14] WE. Thasler, TS. Weiss, K. Schillhorn, PT. Stoll, B. Irrgang, KW. Jauch, "Charitable state-controlled foundation human tissue and cell research: ethic and legal aspects in the supply of surgically removed human tissue for research in the academic and commercial sector in germany," Cell Tissue Bank. vol. 4(1) pp. 49-56 2003.
- [15] U. Gehring, et. al., "Endotoxin levels in cow's milk samples from farming and non-farming families - the PASTURE study," Environ Int. vol. 34(8) pp. 1132-6 Nov 2008, doi: 10.1016/j.envint.2008.04.003.
- P. Mayring, "Qualitative content analysis. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research," Vol. 1, No. 2, Art. 20, June 2000. ISSN 1438-5627. [Online]. Available from: http://www.qualitativeresearch.net/index.php/fqs/article/view/1089/2386 2017.08.18
- [17] TH. Müller, "Synchronising mobile electronic case report forms with a central database," Stud Health Technol Inform. vol. 228 pp. 307-11, 2016.
- [18] TH. Müller, R. Thasler, "Separation of Personal Data in a Biobank Information System," Stud Health Technol Inform. vol. 205, pp. 388-92, 2014.
- [19] D. Lindoerfer, U. Mansmann, "IT infrastructure of an oncological trial where xenografts inform individual second line treatment decision," Stud Health Technol Inform. vol. 235 pp. 226-30, 2017.