

## Design and Implementation of Data Collection Instruments for Neonatology Research

Monica G. HĂȘMĂȘANU<sup>1</sup>, Sorana D. BOLBOACĂ<sup>2,\*</sup>, Lorentz JĂNTSCHI<sup>3</sup>,  
Gabriela C. ZAHARIE<sup>1</sup>, Tudor C. DRUGAN<sup>2</sup>

<sup>1</sup> Iuliu Hațieganu University of Medicine and Pharmacy Cluj-Napoca, Department of Neonatology, 3-5 Clinicilor, 400006 Cluj-Napoca, Romania.

<sup>2</sup> Iuliu Hațieganu University of Medicine and Pharmacy Cluj-Napoca, Department of Medical Informatics and Biostatistics, 6 Louis Pasteur, 400349 Cluj-Napoca, Romania.

<sup>3</sup> Technical University of Cluj-Napoca, 103-105 Muncii Boulevard, 400641 Cluj-Napoca, Romania.  
E-mails: popa.monica@umfcluj.ro; sbolboaca@umfcluj.ro; lorentz.jantschi@gmail.com; gabriela.zaharie@umfcluj.ro; tdrugan@umfcluj.ro

\* Author to whom correspondence should be addressed; Tel.: +4-075-0774506

Received: 28 November 2014/Accepted: 18 December 2014/ Published online: 26 December 2014

### Abstract

*Aim:* The aim of our research was to design and implement data collection instruments to be use in context of an observational prospective clinical study with follow-up conducted on new born with intrauterine growth restriction. *Methods:* The structure of the data collection forms (paper based and electronic based) was first identified and for each variable the best type to accomplish the research aim was established. The code for categorical variables has also been decided as well as the units of measurements for quantitative variables. In respect of good practice, a set of confounding factors (as gender, date of birth, etc.) have also been identified and integrated in data collection instruments. Data-entry validation rules were implemented for each variable to reduce data input errors when the electronic data collection instrument was created. *Results:* Two data collection instruments have been developed and successfully implemented: a paper-based form and an electronic data collection instrument. The developed forms included demographics, neonatal complications (as hypoglycemia, hypocalcemia, etc.), biochemical data at birth and follow-up, immunological data, as well as basal and follow-up echocardiographic data. Data-entry validation criteria have been implemented in electronic data collection instrument to assure validity and precision when paper-based data are translated in electronic form. Furthermore, to assure subject's confidentiality a careful attention was given to HIPPA identifiers when electronic data collection instrument was developed. *Conclusion:* Data collection instruments were successfully developed and implemented as an a priori step in a clinical research for assisting data collection and management in a case of an observational prospective study with follow-up visits.

**Keywords:** paper Case Report Form (pCRF); Electronic Data Collection Instrument (EDCI); Data-Entry Validation; Observational Study; Clinical Research.

### Introduction

Observational longitudinal studies are designs useful in assessment of etiology [1] and prognosis [2,3] and are seen as studies able to produce evidence for clinical decisions [4]. A structured reporting requirements for these studies had been created to assist authors in reporting theirs

research [5] but similar guideline (such as STROBE - STrengthening the Reporting of OBServational studies in Epidemiology [6]) on observational studies reporting results showed no effect on the quality of reports [7].

One important step in medical research is represented by accurate and truthful collection of data to ensure valid results. The validity and reliability of any research project depends to accuracy of collected data [8]. The way in which data are collected in medical research has not been significantly changed since computers become an instrument used in daily practice. Primary data collection are paper-based (pCRF = paper Case Report Form) and later an electronic computer entry take place [9]. In the entry step data could be validated and data errors could be dismissed but the process could also produce entry errors (errors from copying data from pCRF to electronic database) [10]. Even that an alternative to pCRF data collection exists, named electronic data collection (EDC) forms (such as electronic Case Report Forms eCRF), in some health care environments this alternative is not feasible due to the lack of on-site technology, complexity of forms, installation and maintenance, and/or high investment costs [11]. Consequently, despite the fact that EDC have been available they did not become dominant in medical research [12]. Although electronic medical records (EMRs) have already been implemented and become functional in some health care environments, some health-care professionals perceive the implementation of an eCRF in a paper-based working environment as a source of redundancy [13].

Beside the usefulness in daily medical practice, EMRs could be seen as a reliable source of data for medical research (retrospective studies or health outcome studies) [14,15] under the limitation that in the EMRs data are not systematically recorded for the research [15,16]. Electronic data collection platforms became lately available as self-platforms (such as Oracle Clinical [17], ClinTrial [18], eClinical [19], REDCap [20], PROMIS [21], etc.) or networks [22]. Despite this fact, researchers from different academic clinical centers prefer to develop and implement customized clinical data management systems due to both specific needs and restrictions of available financial support [12,23,24]. Our objective was to design and implement both paper and electronic versions of CRF for clinical data management in the context of an observational prospective study with follow-up conducted on newborn infants with intrauterine growth restriction.

## Material and Method

The aim of clinical research for which the CRFs were developed and implemented was to evaluate from clinical (anthropometric) and para-clinical (biochemical evaluations and immunologic evaluation by dosing the IGF1 and IGF2 and their associated receptors) point of view the newborn infants with intrauterine growth restriction (IUGR).

The pCRF has been designed to be used in clinical environment to collect and record data related with the research protocol. A corresponding electronic data collection instrument (EDCI) has also been implemented to compile the collected data and to validate them prior to statistical analysis.

The pCRF has been designed to maximize the number of fields per page. To reduce the print costs we aimed to create a one paper CRF for collection of longitudinal data beside a one-page CRF for collection of basic data.

The applied steps were as follows: ① identification of raw variables [25]; ② identification the type of data for each raw variable [26]; ③ codification of categorical variables; ④ identification and definition (in terms of formulas) of needed calculated variables; ⑤ development pCRF; and ⑥ implementation and validation of electronic data collection instrument (EDCI).

The main identified variables along with associated type of data are presented in Figure 1. Whenever specified data will be collected baseline, otherwise data will be collected at baseline, and two follow-up visits (the first one at 1 month and the last one at 6 months).

Three anthropometric calculated variable and three calculated echocardiographic values have been identified to be part of the EDCI along with several classification variables; their names and associated formulas are presented in Table 1.

<p><b>BASELINE: PARENT'S RELATED</b>                  ◇Demographic (for both mother and father): age (years), weight (kg), height (cm), ethnicity                  ◇Baby health related risks: smoking (no. cigarette/day), alcohol consumption (yes/no)                  ◇Personal: history (gesta/para for mother), pathologies  <b>PREGNANCY RELATED DISEASES &amp; CONDITIONS (MOTHER):</b> ◇hypertension (yes/no), ◇dexamethasone (yes/no), ◇IUGR diagnosis week (week), ◇double test (pos/neg), ◇triple test (pos/neg), ◇TORCH related variables.  <b>DELIVERY:</b> ◇ mode of birth (normal / breech presentation / caesarean section), ◇ rupture of membranes (hours), ◇ gestational age at birth (weeks), ◇ Apgar score (1min / 5min / 10min), ◇ resuscitation (yes/no), ◇ respiratory support as type and time (hours).  <b>INFANT COMORBIDITIES (yes/no):</b> ◇trauma, ◇respiratory distress, ◇hyperbilirubinemia, ◇early sepsis, ◇ulceronecrotic enterocolitis, ◇hypoglycemia, ◇hypocalcemia, ◇heart failure, and ves associated procedures</p>	<p><b>LONGITUDINAL DATA COLLECTION: INFANT RELATED</b>  <b>ANTROPOMETRIC:</b> ◇gender, ◇weight (g), ◇height (cm), ◇cranial perimeter (cm), ◇no. days for weight recovery (baseline), and ◇total enteral nutrition (day) (baseline)  <b>TRANSFONTANELLAR ULTRASOUND:</b> ◇haemorrhagia (yes/no) – if yes than grade (1/2/3/4), ◇ischemia (yes/no), ◇IR (baseline), ◇NIRS (baseline)  <b>CARDIAC ULTRASOUND:</b> ◇RVAWT (mm), ◇RVEDD (mm), ◇RVESD (mm), ◇IVSs (mm), ◇IVSd (mm), ◇LVEDD (mm), ◇LVESD (mm), ◇LVPWT (mm), ◇Aorta/LA, ◇E/A, ◇ICT+ IRT+ET, ◇ET (where RVAWT = right ventricular anterior wall thickness; RVEDD = right ventricular end diastolic dimension; RVESD = right ventricular end systolic dimension; IVSs = thickness of interventricular septum at end systole; IVSd = thickness of interventricular septum at end diastole; LVEDD = left ventricular end diastolic dimension; LVESD = left ventricular end systolic dimension; LVPWT = left ventricular posterior wall thickness; LA = left atrium; ICT = Isovolumic Contraction Time; ET = ejection time; IRT = isovolumic relaxation time)  <b>PARACLINICAL TESTS:</b>                  ◇Blood panel: leukocytes (mm<sup>3</sup>), erythrocytes (mm<sup>3</sup>), thrombocytes (mm<sup>3</sup>), hemoglobin (g/L), hematocrit (%).                  ◇Basic metabolic panel: blood urea (mg/dL), blood creatinine (mg/dL), glucose (mg/dL).                  ◇Others: aspartate transaminase (AST, U/L), alanine transaminase (ALT, U/L), creatine kinase (CK, U/L), CK-MB (U/L), lactate dehydrogenase (LDH, U/L).                  ◇Others (baseline): IgM<sub>unspecific</sub>, C-reactive protein (mg/dL), IGF2   IGF1 receptors, IGF2   IGF1.                  ◇ASTRUP (baseline): pH, pCO<sub>2</sub>, pO<sub>2</sub>, base excess, and HCO<sub>3</sub>.  <b>NEUROLOGICAL EXAMINATION:</b> free text</p>
--	---

Figure 1. Details of raw data collection items

Table 1. Calculated/classification values and associated formulas

Abbreviation (Name)	Formula	Ref
<b>Anthropometric</b>		
BMI <sub>adult</sub> (Body Mass Index - adult)	=weight(kg)/(height(m <sup>3</sup> ))	[27]
cBMI <sub>adult</sub> (classification based on BMI - adult)	0=BMI<18.5 under healthy weight 1=18.5≤BMI≤24.9 healthy weight 2=25≤BMI≤29.9 overweight 3=30≤BMI≤34.9 obesity grade I 4=35≤BMI≤39.9 obesity grade II 5=BMI≥40 obesity grade III	[28]
RPI (Rohrer's ponderal index)	=100*weight(kg)/(Crown-Heel Length(cm <sup>3</sup> ))	[29]
cRPI (classification based on RPI)	0<2 =asymmetric IUGR 1>2= symmetric IUGR	[29]
BBR (Brain:Body weight Ratio)	= 100*[0.037 *(cranial-perimeter) <sup>2.57</sup> ]/(birth-weight(g))	[30]
<b>Echocardiographic</b>		
FS (left ventricular fractional shortening (%))	=(LVEDD-LVESD)/LVEDD*100, where LVEDD = left ventricular end diastolic dimension, LVESD = left ventricular end systolic dimension	[31]
E/A <sub>grade</sub> (grades of diastolic function)	0=Normal diastolic function - 1≤E/A≤2 1=Impaired relaxation - E/A<1 2=Restrictive filling - E/A>2	[31]
Tei (Tei Index of Myocardial Performance)	=(IVCT+IVRT)/ET, where IVCT = iso-lvolumetric contraction time, IVRT = iso-volumetric relaxation time, ET = left ventricular ejection time	[32]

The Excel was used due to its availability and flexibility [33].The EDCI has been implemented by respecting the security rules (45 eCRF Parts 160, 162, and Part 164) established under the Health Insurance Portability and Accountability Act (HIPAA) [34]. In this regards, the EDCI was implemented in such way that did not contain the HIPPA identifiers, such as names, geographical subdivisions, full dates, phone and fax numbers, electronic mail address, social security number, medical record number, biometric identifiers, full face photographic image, etc. Each entry in the EDCI will have assigned a four digits number (for example, the first subject included in the study have assigned the 001 number) to avoid identification of subjects included in the study. Furthermore, data-entry validation rules have been identified and implemented for each variable to

prevent as much as possible data input errors.

The implementation of EDCI has been done using Microsoft Excel platform. The EDCI was created using Microsoft Excel XP for Windows, running under Windows XP on an Intel Pentium Dual CPU @2.20GHz and 2 GB of RAM computer.

## Results

Two data collection instruments have been developed and successfully implemented to allow the management of data for an observational longitudinal study. The first instrument, pCRF has been created to be used by one researcher who will be responsible for collection of observational data related to the subjects included in the study (Figure 2). Two forms have been created as pCRF, one for baseline data collected (Figure 2a) and the other for longitudinal data collection (Figure 2b).

Translation of pCRF to EDCI has been done to prepare clinical data for statistical analysis. To do this translation, each identified variable was documented in a table called a data dictionary (Table 2).

**MODELINFIUGR**

Multidimensional Modeling in Evaluation of Infants with Intrauterine Growth Restriction

„Iuliu Hațieganu” University of Medicine and Pharmacy Cluj-Napoca

**ID**

--	--	--	--

**DATA COLLECTION FORM**

**PARENTS DATA**

	Mother	Father
Age (years)	<input type="text"/>	<input type="text"/>
Weight	<input type="text"/>	<input type="text"/>
Birth (g)	<input type="text"/>	<input type="text"/>
Beginning of pregnancy (kg)	<input type="text"/>	<input type="text"/>
Height (cm)	<input type="text"/>	<input type="text"/>
Smoking (no. cigarette per day)	<input type="text"/>	<input type="text"/>
Alcohol consumption (yes/no)	<input type="text"/>	<input type="text"/>
Ethnicity	<input type="text"/>	<input type="text"/>
Personal History (gesta/para)	<input type="text"/>	<input type="text"/>
Personal Pathologic History	<input type="text"/>	<input type="text"/>
Phone:	<input type="text"/>	
Address:	<input type="text"/>	
	<input type="checkbox"/>	Rural/Urban

**PREGNANCY RELATED DISEASES & CONDITIONS**

Hypertension: Yes  No  If yes – week of diagnosis:  Dexamethasone: Yes  No

IUGR diagnosis week:  Double test: Positive  Negative

Triple test: Positive  Negative

TORCH Toxo: Positive  Negative  IgM (U/L):  IgG (U/L):

Rubeola: Positive  Negative  IgM (U/L):  IgG (U/L):

Citomegalo: Positive  Negative  IgM (U/L):  IgG (U/L):

Herpes: Positive  Negative  IgM (U/L):  IgG (U/L):

**DELIVERY**

Mode of birth: normal vaginal  breech presentation  caesarean section

Rupture of membranes (hours):  Gestational age at birth (weeks):

Apgar score: 1 min  5 min  10 min

Baby resuscitation: Yes  No  If yes: bag and mask  drugs given  IPPV

Respiratory support (hours): SIMV  nCPAP  O<sub>2</sub>  FIO<sub>2</sub>

**INFANT COMORBIDITIES**

Trauma: Yes  No  Type:

Respiratory distress: Yes  No  Curosurf: Yes  No

Hyperbilirubinemia: Yes  No  Phototherapy (hours):

Early sepsis: Yes  No  Late sepsis: Yes  No

Ulceronecrotic enterocolitis: Yes  No

Hypoglycemia: Yes  No  Max glucose (mg/kg/min):

Hypocalcemia: Yes  No

Heart failure: Yes  No  If Yes: Day of Dopamina/Dobutamina  Maximum dose:

**a)**

Figure 2. ModelInFIUGR pCRF forms: a) Initial raw data collection form.

The image shows a complex data collection form titled 'ModelInflUGR' with a subtitle 'Multidimensional Modeling in Evaluation of Infants with Intrauterine Growth Restriction'. It includes an ID field and a 'DATA COLLECTION FORM' section. The form is organized into several key areas:
 

- INFANT DEMOGRAPHICS:** Fields for First name, Family name, Gender (M/F), Date of birth (dd/mm/yy), and Date of discharge of baby (dd/mm/yy).
- VISIT DATA:** A table with columns for Baseline, Visit 1, and Visit 2, and rows for Date (DD/MM/YY).
- ANTHROPOMETRIC DATA:** Fields for Weight (g), Height (cm), Cranial perimeter (cm), No days for weight recovery, and Total enteral nutrition (day).
- TRANSFONTANELAR ULTRASOUND:** A grid for recording 'Yes' or 'No' for various conditions like Haemorrhagia, Ischemia, and IR, with sub-columns for grades 1, 2, 3, and 4.
- CARDIAC ULTRASOUND:** A grid for recording measurements such as RVAWT, RVEDD, RIVSD, RAS, RASd, LVEDD, LVESD, LVPWT, Aorta/IA, I/A, ICT+IRT+ET, and ET.
- PARACLINICAL TESTS:** A table with columns for Baseline, Visit 1, and Visit 2, and rows for various tests including Blood panel (Leukocytes, Erythrocytes, Thrombocytes, Hemoglobin, Hematocrit), Basic metabolic panel (Blood urea, Creatinine, Glucose), and others (Alp, Amin, Creat, CK-MB, LDH, IgM, CRP, IGF2, IGF1).
- ASTRUP:** Fields for pH, pCO<sub>2</sub>, pO<sub>2</sub>, Base Excess, and HCO<sub>3</sub>.
- NEUROLOGICAL EXAMINATION:** A large empty box for notes, with a date field (DD/MM/YY).

Figure 2. ModelInflUGR pCRF forms: b) Longitudinal raw data collection form.

Table 2. Data dictionary by example

Column	Variable name	Label (units)	Format	Codes/Ranges	Missing value
A	ID	Subject ID number	Text (4)	0001-9999	Not allowed
B	Group	Group under evaluation	Numeric	1=case 2=control	Not allowed
C	Gender	Subject gender	Text (1)	F = female M = male	X
D	PD_AgeM	Age of mother (years)	Numeric	Range 10 - 50	
F	PD_WeightBPM	Weight of mother at beginning of pregnancy (kg)	Numeric		
H	PD_BMI_M	Mother body mass index (kg/m <sup>3</sup> )	Numeric	Range 0-70	
I	PD_IF-BMI_M	BMI classification	Text (1)	0=BMI<18.5 1=18.5≤BMI≤24.9 2=25≤BMI≤29.9 3=30≤BMI≤34. 4=35≤BMI≤39. 5=BMI≥40	

To respect confidentiality and to protect medical data from disclosure, no identification variables were included in the implemented EDCI (called ModelInflUGR = Multidimensional Modeling in Evaluation of Infants with Intrauterine Growth Restriction); each subject included in the study has assigned a four digits identification number (Figure 3a). To avoid data input errors, lists (Figure 3b) or data between imposed values (Figure 3c) have been used when EDCI was implemented. Furthermore, for each variable an input message was defined to guide the user in the

process of data collection (Figure 3b) while an error message has been defined and to appear when the input data would not correspond with the imposed type and range (Figure 3c).

The implemented calculated values (such as 'PD BMI M') or classifications (such as 'PD IF-BMI M') appears automatically in the electronic data collection form once the raw data based on which are computed are introduced (Figure 4).

The implemented EDCI works properly on Microsoft Excel XP, 2003, 2007 and 2010. The EDCI was checked to assure that all rules and limitations were implemented properly and all errors were corrected. Furthermore, the accuracy (defined as 'true'/'correct' output) of the calculated values or of classifications was also carefully verified to assure a valid output of these EDCI variables.

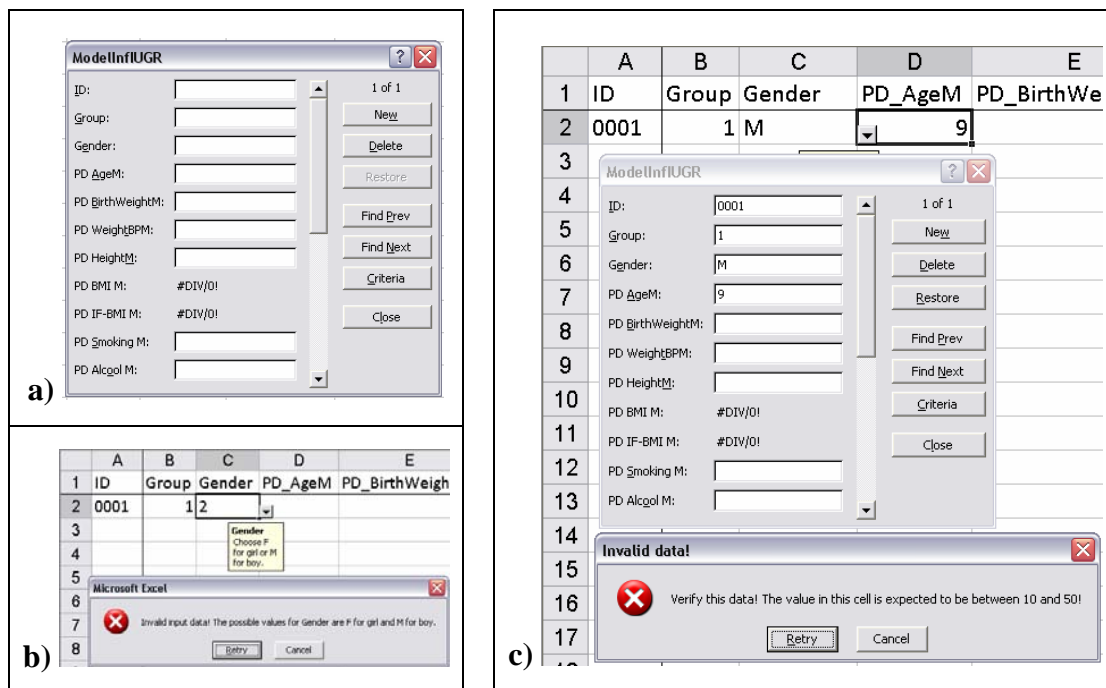


Figure 3. ModelInfiUGR EDCI: a) Electronic data collection form; b) Input data validation with list; c) Input data validation by restricting the range of possible values.

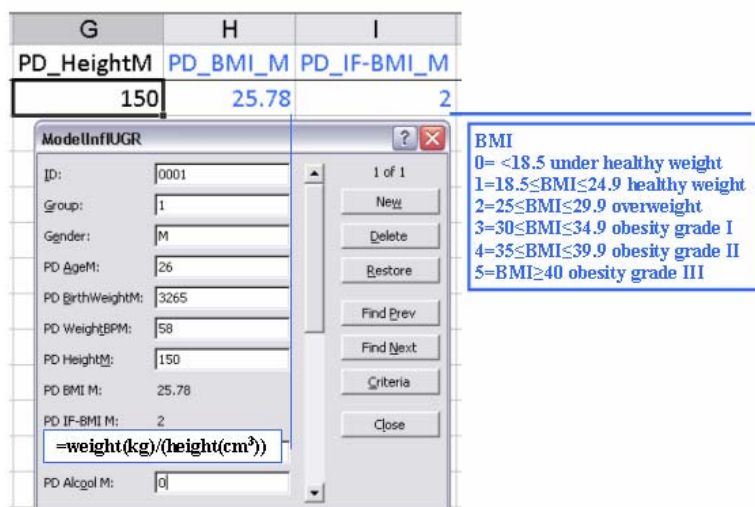


Figure 4. EDCI: implementation by example of BMI calculated value and BMI classification

## **Discussion**

The objective of our study was successfully achieved and two instruments (paper-based and electronic) has been designed and implemented to assist the researcher in data management on the context of an observational prospective study two follow-up visits conducted on newborn infants with intrauterine growth restriction.

Data management is very important in clinical research since is responsible by proper and high quality statistical sound data with direct link with the outcome of any study. Planning for collection of research data is an ethics thing that need clear identification of all data which will be collected before seeing the first research subject. The design of our pCRF was a team effort which implied specialists in different fields that allowed identifying the proper data to be collected for this research. The pCRFs (Figure 2) were developed to be used by a single researcher during a clinical observational study which will be responsible by data collection and to assure the confidentiality of medical data. Data included in the pCRF are in compliance with regulatory standard in regards of subject's data collection [19].

Development of the EDCI as an intermediary step for data analysis was also achieved for the ModelInfiUGR. The EDCI was developed to assure accuracy, reliability, and analyzability. The Excel has been used for this task since it is a convenient and accessible tool. Basic computer skills are needed to properly use the EDCI while the input errors were limited using validations rules. Data validation, the process of testing the validity of data in accordance with the protocol specifications [19], is not a new feature in Excel and could find its usefulness in creation of rules that help users to input accurate data by using limits and dropdowns lists and thus to improve the accuracy of the input raw data. In our EDCI, if a cell is not filled with proper data type a message appears (Figure 3) and the introduction of data is stopped. Furthermore, to reduce errors in derived variables from raw data, formulas were implemented for all calculated and classification variables (see Table 1). EDCI was encrypted and could be accessed with proper authentication to assure confidentiality and security of research data.

The main disadvantages of the implemented EDCI is related with the easiness to corrupt data or delete data but this could be prevent by creation of backup copies on each utilization. Even if the implementation of data collection in Excel has certain disadvantages it has its usefulness in management of clinical data. Several articles presented the use of Excel in collection of research data, such as in blood pressure management in practice and research [35], creation and implementation of questionnaires [36], etc. However, different application of Excel could be seen in analysis of research data (such as regression analysis [37], as decision support systems [38,39], simulations [40], etc.).

It is not doubt that the existent technology allow a wide range of possibilities in regards of clinical data collection systems, such as PDAs, Tablet PCs, d-pen (digital pen), DTP hybrid (hybrid handheld/pen-on-paper), each with its advantages and disadvantages [41]. One important step in research management is to identify a proper data collection form to well fit to your own study design [42] and when the identified systems did not accomplish with the study requirements a creation of a new one is mandatory.

Even with all existent technology, using an electronic data entry system will just be able to minimize data entry errors resulting from an extra data entry step but the researcher will need to know properly how to use these technologies. However, by implemented data validation, our EDCI withdrawn this disadvantage while using a well-known and available program. Excel provided a useful, flexible and easy to use interface for implementation of our EDCI and should be considered by research students as a legitimate alternative for clinical research data management.

## Conclusion

A paper-based collection report form has been created for clinical data collection for a study called multidimensional modeling in evaluation of infants with intrauterine growth restriction a priori to beginning of the study. The implemented data collection instrument done using Microsoft Excel proved reliable alternative as research data management tool.

## Conflict of Interest

The authors declare that they have no conflict of interest.

## Acknowledgements

This paper was published under the frame of European Social Fund, Human Resources Development Operational Programme 2007-2013, project no. POSDRU/159/1.5/S/138776.

## References

1. Whyte J, Gosseries O, Chervoneva I, DiPasquale MC, Giacino J, Kalmar K, et al. Predictors of short-term outcome in brain-injured patients with disorders of consciousness. *Prog Brain Res* 2009;177(C):63-72.
2. Ioannidis J, Haidich A, Lau J. Any casualties in the clash of randomised and observational evidence? *BMJ* 2001;322:879-80.
3. Gladman DD, Chandran V. Observational cohort studies: Lessons learnt from the University of Toronto Psoriatic arthritis program. *Rheumatology* 2011;50(1):25-31.
4. Carpenter L, Nikiphorou E, Young A. Importance of registries in informing clinical practice for arthritis. *Clin Pract* 2013;10(6):723-36.
5. Tooth L, Ware R, Bain C, Purdie DM, Dobson A. Quality of Reporting of Observational Longitudinal Research. *Am J Epidemiol* 2005;161(3):280-8.
6. von Elm E1, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61(4):344-9.
7. Bastuji-Garin S, Sbidian E, Gaudy-Marqueste C, Ferrat E, Roujeau J-C, Richard M-A, et al. Impact of STROBE Statement Publication on Quality of Observational Study Reporting: Interrupted Time Series versus Before-After Analysis. *PLoS ONE* 2013;8(8):e64733.
8. Weinstein J. Spine Patient Outcomes Research Trial. Hanover, NH: Trustees of Dartmouth College, 1999.
9. Cole E1, Pisano ED, Clary GJ, Zeng D, Koomen M, Kuzmiak CM, et al. A comparative study of mobile electronic data entry systems for clinical trials data collection. *Int J Med Inform* 2006;75:722-9.
10. Day S, Fayers P, Harvey D. Double data entry: what value, what price? *Control Clin Trials* 1998;19:15-24.
11. Welker JA. Implementation of electronic data capture systems: barriers and solutions. *Contemp Clin Trials* 2007;28:329-36.
12. Kuchinke W1, Ohmann C, Yang Q, Salas N, Lauritsen J, Gueyffier F, et al. Heterogeneity prevails: the state of clinical trial data management in Europe - results of a survey of ECRIN centres. *Trials* 2010;11:79.
13. Lium J-T, Tjora A, Faxvaag A. No paper, but the same routines: a qualitative exploration of experiences in two Norwegian hospitals deprived of the paper based medical record. *BMC Med*



- Inform Decis Mak 2008;8:2.
14. Lin J, Jiao T, Biskupiak JE, McAdam-Marx C. Application of electronic medical record data for health outcomes research: A review of recent literature. *Expert Rev Pharmacoecon Outcomes Res* 2013;13(2):191-200.
  15. Klumpp TR. Electronic medical records and quality of cancer care. *Curr Oncol Rep* 2013;15(6):588-94.
  16. Jansen AC, van Aalst-Cohen ES, Hutten BA, Büller HR, Kastelein JJ, Prins MH. Guidelines were developed for data collection from medical records for use in retrospective analyses. *J Clin Epidemiol* 2005;58(3):269-74.
  17. Kunithala VK, Kabeer A, Vemula SK. Advance Oracle Clinical Software tool used in clinical research from patient data collection to data base lock. *Clin Res Regul Aff* 2013;30(1-2):1-13.
  18. Tai BC, Seldrup J. A review of software for data management, design and analysis of clinical trials. *Ann Acad Med Singapore* 2000;29(5):576-81.
  19. Krishnankutty B, Bellary S, Kumar NBR, Moodahadu LS. Data management in clinical research: An overview. *Indian J Pharmacol* 2012;44(2):168-72.
  20. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42(2):377-81.
  21. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care* 2007;45:S3-11.
  22. Demotes-Mainard J, Ohmann C. European Clinical Research Infrastructures Network: promoting harmonisation and quality in European clinical research. *Lancet* 2005;365(9454):107-8.
  23. Danciu I, Cowan JD, Basford M, Wang X, Saip A, Osgood S, et al. Secondary Use of Clinical Data: the Vanderbilt Approach. *J Biomed Inform* 2014;52:28-35.
  24. El Emam K, Jonker E, Sampson M, Krleza-Jeric K, Neisa A. The use of electronic data capture tools in clinical trials: web-survey of 259 Canadian trials. *J Med Internet Res* 2009;11:e8.
  25. Zaharie G. Neonatologie. Editura Didactică și Pedagogică 2007.
  26. Drugan T, Achimaș A, Țigan S. Aplicații medicale ale statisticii. Editura Medicală Univ. Iuliu Hațieganu 2008.
  27. Quetelet GEA (1796–1874)-the average man and indices of obesity. *Nephrol Dial Transplant* 2007;23(1):47-51.
  28. BMI classification. [online] © Copyright 2006 World Health Organization [Last update: 18/02/2014]. Available from: [http://apps.who.int/bmi/index.jsp?introPage=intro\\_3.html](http://apps.who.int/bmi/index.jsp?introPage=intro_3.html) [cited 3 November 2013]
  29. Tamim H, Beydoun H, Itani M, Khogali M, Chokr I, Yunis KA. Predicting neonatal outcomes: birthweight, body mass index or ponderal index? *J Perinatal Med* 2005;32(6):509-13.
  30. McLennan JE, Gilles FH, Neff RK. A model of growth of the human fetal brain. In: Gilles FH, Leviton A, Dooling EC. *The developing human brain: growth and epidemiologic neuropathy*. Wright PG 1983.
  31. de Simone G, Devereux RB, Roman MJ, Ganau A, Saba PS, Alderman MH, et al. Assessment of left ventricular function by the midwall fractional shortening/end-systolic stress relation in human hypertension. *J Am Coll Cardiol* 1994;23(6):1444-51.
  32. Tei C. New non-invasive index of combined systolic and diastolic ventricular function. *J Cardiol* 1995;26:135-6.
  33. Kaparathi S, Power DJ. Spreadsheets. In: Bidgoli H. (Ed.) *Encyclopedia of Information Systems USA*, Academic Press, 2003, pp. 171-186.
  34. HIPAA Administrative Simplification [Internet] U.S. Department of Health and Human Services Office for Civil Rights [Cited 3 November 2013]. Available from <http://www.hhs.gov/ocr/privacy/hipaa/administrative/combined/hipaa-simplification-201303.pdf>.
  35. Peleska J. Home blood pressure Excel logbook in practice and research. *Eur J Intern Med* 2013;24:e34.

36. Black IR, Efron A, Ioannou C, Rose JM. Designing and Implementing Internet Questionnaires Using Microsoft Excel. *Australasian Marketing Journal* 2005;13(2):61-72.
37. Wang JP, Huang D, Chang S-C, Brant L. On-site earthquake early warning with multiple regression analysis: Featuring two user-friendly applications for Excel. *Comput Geosci* 2013;58:1-7.
38. Palocsay SW, Markham IS, Markham SE. Utilizing and teaching data tools in Excel for exploratory analysis. *J Bus Res* 2010;63(2):191-206.
39. Hyde KM, Maier HR. Distance-based and stochastic uncertainty analysis for multi-criteria decision analysis in Excel using Visual Basic for Applications. *Environ Modell Softw* 2006;21(12):1695-710.
40. Lambert RJW, Mytilinaios I, Maitland L, Brown AM. Monte Carlo simulation of parameter confidence intervals for non-linear regression analysis of biological data using Microsoft Excel. *Comput Methods Programs Biomed* 2012;107(2):155-63.
41. Cole E, Pisano ED, Clary GJ, Zeng D, Koomen M, Kuzmiak CM, et al. A comparative study of mobile electronic data entry systems for clinical trials data collection. *Int J Med Inform* 2006;75(10-11):722-9.
42. Pace WD, Staton EW. Electronic data collection options for practice-based research networks. *Ann Fam Med* 2005;3(Suppl 1):S21-9.