

Interactive Feedback of Data Quality in Clinical Research. A Case Study from an Infectious Diseases Cohort

Stephan GLÖCKNER^{a,c,1}, Daniela SCHINDLER^{b,c} and André KARCH^{a,c}

^a Helmholtz Centre for Infection Research (HZI), Braunschweig, Germany

^b University Hospital rechts der Isar, Munich, Germany

^c German Center for Infection Research, Braunschweig, Germany

Abstract. To improve data quality, clinical cohort studies require ongoing centralized data monitoring. Results of data quality monitoring are often reported in a general format, without individual feedback for different types of users within the research consortium. Several R packages provide the possibility for an easy to use approach to analyze and communicate information in interactive web-based forms. This article describes a pilot, which is applied in a German infectious disease cohort using interactive feedback and gamification to report data quality, to ultimately increase the quality of research outcomes.

Keywords. Data quality, interactive feedback, dashboards, infectious disease cohort, gamification

1. Introduction

Evaluating data quality (DQ) during a clinical study needs to be an ongoing procedure [1,2]. Unfortunately, DQ assurance is not applied to every clinical study. Clinical trials that apply ongoing data monitoring, often report their results in pdf-Files sent to the principal or local investigators to plan further actions which increase DQ. This approach alone, however, is generally not effective in reducing data inconsistencies.

Evaluating DQ needs to be accompanied by a targeted communication of its results to the different user groups: study nurses, (principal-) investigators and the research community. General forms of communications are text or tables. Figures are an important instrument for research communication, but figures are the results of stories told by its underlying data [3]. All used elements need to have one goal: to motivate the users to increase the value of their data.

In this article, we present an approach to implement targeted individual and interactive web-based feedback using different packages in R.

1.1. Infectious Diseases Cohort Study

The Transplantation Cohort (TxC) of the German Center for Infection Research (DZIF) collects information about infectious complications from patients with different types of organ transplantations (Tx) from five large German university hospitals. The Helmholtz

¹ Corresponding Author: Stephan Glöckner, Helmholtz Centre for Infection Research, Inhoffenstraße 7, 38124 Braunschweig, Germany; E-mail: stephan.gloeckner@helmholtz-hzi.de.

Centre for Infection Research (HZI) performs the epidemiological evaluation of the collected information. For this, data needs to be aggregated to provide an overview of study performance and DQ.

1.2. Evaluation and Visualization of Data Quality

DQ is multidimensional, but still not a standardized term [4]. The Technology, Methods, and Infrastructure for Networked Medical Research (TMF) published a guideline for the evaluation of DQ in clinical cohort studies and disease registries [5]. Here, several DQ indicators (e.g. completeness or value distribution) are defined to support data managers, who need to adapt important indicators to quantify trial performance [6] and to generate feedback for its users.

Web-based feedback about DQ is not a novel approach. It is already applied in private companies to measure metrics that show trends of key performance indicators. However, DQ feedback is rarely implemented in publicly funded clinical cohort studies, and its impact on trial performance is not researched well enough [7–9].

2. Methods

Received data exports consist of multiple csv-files of different records accompanied by metadata (e.g. type of data field, allowed data range) and data dictionaries. Data is twice pseudonymized, still allowing linkage of data between the records.

We developed for our case study a framework for DQ feedback using flexdashboard as visualization approach [10]. For data analysis and visualization, the programming language R was used [11]. For each analysis, a dataframe containing aggregated information and graphical objects were separately stored in an RData-file, allowing the import into the flexdashboard.

2.1. Evaluation of Study Performance and Data Quality with R

For automated DQ assessment we used R tidyverse packages [12]. We used the DQ dimension completeness, which we defined as the ratio of not entered fields divided by all available data fields of a measured variable. For selected continuous variables (e.g. length of ICU stay, surgery time), data distribution was assessed, providing case-based warnings for study nurses, if the data points could be identified as outliers.

Study performance was assessed by descriptive analysis of study population and biosamples. Comparing actual and target count of follow up (FU) visits was conducted to assess FU performance. If patients have an Tx date in the cohort study, regularly FU visits need to be performed, after 3, 6, 9, 12 months and annually thereafter.

The package lubridate [13] was used to evaluate the FU. Lubridate uses interval class objects. The analysis was performed by measuring, whether the expected FU is within the time interval (+/- 28 days) of the required FU. Fig. 1 provides an example, where the first two FU visits match, but the third visit was performed after the allowed time interval, and was subsequently reported as a warning.

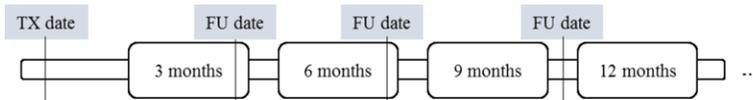


Figure 1. Follow up (FU) evaluation in accordance to entry date.

2.2. Interactive Visualization of Information with R

We used knitr and datatables (DT) packages [14] for rendering tables. For the DT package, we used an extension, providing the users an option to directly export the table into standard data formats (CSV/XLS/PDF). Graphs were rendered using ggplot2 and plotly [15]. For colorblind purposes, we applied the color palette viridis [16] on graphical elements.

3. Results

The general frame of the dashboard contained the overview of study population, biosamples, FU visits and DQ. Fig 2 shows an example of the dashboard DQ frame.

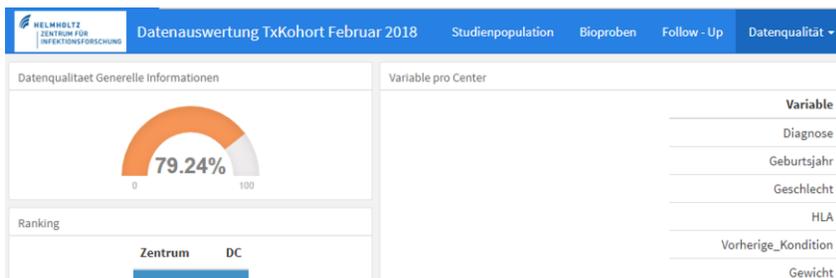


Figure 2. Example of the upper DQ screen.

3.1. Demographic and Population Information

First page of visualized information is the overview of study performances. Here, the number of overall enrolled participants is shown as a bar graph, distinguished into every type of organ, each bar graph stratified by center. The graph is scripted using the syntax of plotly. The interactive element grouped by center, which provides the possibility to show a bar graph of all types of organs by a single or multiple selected centers.

Second element is a DT element, providing an additional overview that shows total numbers of patients by center and by type of organ.

Another important element is the recruitment chart, which highlights positive or negative trends of patient enrollment per center. We used a plotly line graph, to provide the possibility of an interactive graph to show lines of selected centers.

Fourth element is a set of important values of all other forms, including total number of biosamples, FU performance values and overall DQ rates. This is helpful to show the user all important information on the landing page without working through forms.

3.2. Feedback on Biosamples

The head of the page provides three value boxes, summarizing total number of blood samples, total number of other samples and the total number of unique patients that provide these samples.

Below the value boxes two tables are provided to the user, showing different types of blood samples (EDTA, PBMC, RNA, or Serum blood samples) per center. The second table has the same structure, showing types of other samples (e.g. stool or urine samples), by center. Both tables have tab stops that subset the data by type of organ. When moving to the biosample page, first tab stop shown, is the total number of samples, which is the sum of all subsets (types of organs).

During data modelling it was discovered, that some patients provide samples for two organs. These patients are provided as a case-based list, allowing study nurses to check their plausibility. It is often the case that kidney transplanted patients could also have a pancreas transplantation. However, all combinations should be checked, therefore a list of warnings is exported.

3.3. Feedback on Follow-Up

Two main elements are reported; graphs, which show the actual and target number of visits of all patients with entry date, and tables with total numbers and percentages by center. Graphs and tables have tab-stops, which allow showing results per type of organ.

The graph is a line graph with colored area to $y=0$. We are using green to visualize the correctly performed FU and red for the not correctly performed FU (not performed or not matching the required interval). The larger the red area, the lower is the FU performance. Each center has an individual graph.

3.4. Feedback on Data Quality

Elements used for feedback about DQ are gauges, tables with center rankings (gamification element), line graphs with data quality trends and tables showing item completeness for each data item per center (see Fig. 2).

Thirteen different records are analyzed (e.g. general information, immunosuppressive drugs, bacterial infections, laboratory results, etc.). Every form has the same structure to provide the user a simplified overview of its DQ. Primary assessment is the evaluation of data completeness, which is accompanied by single plausibility checks (e.g. scheme of immunosuppressive drug). The mean of data completeness and data plausibility is defined as DQ.

4. Discussion

This regularly created report is provided via encrypted email as an html file to the research consortium. Other options should be discussed with regards to methods (e.g. providing a server with login to the user) and data security.

Flexdashboards have also limited flexibility (e.g. the usage of tab stops in only one element instead of all), other more code intensive feedback options like shiny

applications or dashboards options using JavaScript solutions (e.g. angular.js in combination with d3.js) need to be considered.

One mandatory feature that needs to be added is that case-based lists need to be provided as check lists, simplifying work flows of study nurses.

5. Conclusion

Further research needs to assess the impact of interactive (web-based) communication on DQ. For the TxC, a quasi-experimental study design in accordance to [17] needs to be developed, e.g. by using a stepped wedge design and providing interactive feedback for a period of time to one center, after that adding other centers and evaluating the change of DQ.

6. Conflict of Interest

The authors declare no conflict of interest.

References

- [1] FDA, Guidance for industry: oversight of clinical investigations—a risk-based approach to monitoring, *Silver Spring MD FDA*. (2013).
- [2] D.G. Arts, N.F. De Keizer, G.-J. Scheffer, Defining and improving data quality in medical registries: a literature review, case study, and generic framework, *J. Am. Med. Inform. Assoc.* **9** (2002), 600–611.
- [3] C.N. Knaflic, *Storytelling with data: A data visualization guide for business professionals*, John Wiley & Sons, 2015.
- [4] T. Dasu, T. Johnson, *Exploratory data mining and data cleaning*, John Wiley & Sons, 2003.
- [5] M. Nonnemacher, D. Nasseh, J. Stausberg, *Datenqualität in der medizinischen Forschung*, Medizinisch Wissenschaftliche Verlagsgesellschaft, 2014.
- [6] C.O. Schmidt, C.E. Krabbe, J. Schössow, K. Berger, C. Enzenbach, P. Kamtsiuris, G. Schöne, R. Houben, C. Meisinger, F. Bamberg, and others, Qualitätsstandards für epidemiologische Kohortenstudien, *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz* **61** (2018) 65–77.
- [7] S. De Lusignan, P.N. Stephens, N. Adal, A. Majeed, Does feedback improve the quality of computerized medical records in primary care?, *J. Am. Med. Inform. Assoc.* **9** (2002), 395–401.
- [8] H. Brouwer, P. Bindels, H. Weert, Data quality improvement in general practice, *Fam. Pract.* **23** (2006), 529–536.
- [9] S. Glöckner, S.E. Toikkanen, J. Rui, W. Hu, R. Sinnott, A Case Study of the International Niemann-Pick Disease Registry (INPDR), in: *E-Health Netw. Appl. Serv. Heal. 2017 IEEE 19th Int. Conf. On, IEEE, 2017*: pp. 1–6.
- [10] B. Borges, J.J. Allaire, *flexdashboard: R Markdown Format for Flexible Dashboards*, 2017.
- [11] W.N. Venables, D.M. Smith, The R development core team, *Introd. R R Found. Stat. Comput. Vienna Austria* (2006).
- [12] H. Wickham, tidyverse: Easily Install and Load the “Tidyverse,” 2017.
- [13] G. Grolemund, H. Wickham, Dates and Times Made Easy with lubridate, *J. Stat. Softw.* **40** (2011), 1–25.
- [14] Y. Xie, DT: A Wrapper of the JavaScript Library “DataTables,” 2018.
- [15] C. Sievert, C. Parmer, T. Hocking, S. Chamberlain, K. Ram, M. Corvellec, P. Despouy, plotly: Create Interactive Web Graphics via “plotly.js,” 2017.
- [16] S. Garnier, viridis: Default Color Maps from “matplotlib,” 2018.
- [17] A.D. Harris, J.C. McGregor, E.N. Perencevich, J.P. Furuno, J. Zhu, D.E. Peterson, J. Finkelstein, The use and interpretation of quasi-experimental studies in medical informatics, *J. Am. Med. Inform. Assoc.* **13** (2006), 16–23.