МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2025
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2025
Proceedings of the Fifty-Fourth Spring Conference
of the Union of Bulgarian Mathematicians
Varna, March 31 – April 4, 2025

FUNCTIONAL MODEL OF PROS DATA COLLECTION SYSTEM

Todor Branzov[1], **Tsvetelin Zaevski**[2]

Institute of Mathematics and Informatics, Bulgarian Academy of Sciences
e-mails: [1]tbranzov@math.bas.bg, [2]t_s_zaevski@math.bas.bg

The article presents an analysis of the documentation and research articles on 8 of the most widely used contemporary methodologies for collecting patient reported outcomes (PRO) data. Based on this analysis, a universalized functional model for collecting PRO data is proposed, which aims to identify the main system roles and present a view independent of a specific methodology, thus facilitating easier software information systems development.

Keywords: functional modeling, digital healthcare, distributed systems, PRO data

ФУНКЦИОНАЛЕН МОДЕЛ НА СИСТЕМА ЗА СЪБИРАНЕ НА ДАННИ ДОКЛАДВАНИ ОТ ПАЦИЕНТИТЕ

Тодор Брънзов[1], **Цветелин Заевски**[2]

Институт по математика и информатика, Българска академия на науките
e-mails: [1]tbranzov@math.bas.bg, [2]t_s_zaevski@math.bas.bg

В статията е представен анализ на документацията и изследвания на 8 от най-използваните съвременни методики за събиране на данни докладвани от пациентите (ДДП). На базата на този анализ е предложен универсализиран функционален модел на система за събиране на ДДП, който цели да бъдат идентифицирани основните системни роли и да предложи изглед независим от конкретна методика.

Ключови думи: функционално моделиране, дигитално здравеопазване, разпределени системи, PRO данни

`https://doi.org/10.55630/mem.2025.54.078-083`

2020 Mathematics Subject Classification: 68M14.

Introduction. Medical institutions have been collecting, storing and analyzing clinical and administrative data for decades. These data are used by medical staff in the treatment of patients. In addition, they are of key importance for the implementation of modern health technologies, the assessment of the quality of provided services and the planning of resources.

The introduction of approaches and concepts of healthcare, oriented towards the value for the patients and the pursuit of optimization of related costs, created a need for modeling and estimating this value. To achieve this goal, it is necessary to supplement the administrative and clinical data collected by medical professionals with data reported directly by the patients (Patient Reported Outcome – PRO). This type of data reflects the patient's own assessment and his subjective perception of the condition and quality of life she leads. The integration of these three pillars of data – administrative, clinical and PRO – provide a basis for a better assessment of the effectiveness of the applied treatment by the medical staff. Managers, funding organizations and state administrative bodies have better understanding of the outcomes in general that can aid the decision-making process regarding the effectiveness of the health technologies and the overall benefit that the healthcare system provides to patients and society. This is especially important for healthcare systems based on the principle of solidarity and operating predominantly with public funds, such as those in the European Union, including the Bulgarian healthcare system.

Patient-reported outcomes data is collected from patients without interpretation by healthcare professionals. These data include patients' subjective assessments of their health status, quality of life, and the impact of treatment on their daily activities. PRO data is used to provide information about health outcomes from the patient's perspective and serve as a complement to objective clinical indicators collected by healthcare professionals. PRO include data of various structures describing physical and emotional state, pain, fatigue, mobility, social activities, and the patient's ability to perform daily tasks. They are collected through specialized tools such as standardized surveys, questionnaires, and using various technical means such as paper, web and mobile applications.

The integration of PRO into health information systems shows significant potential to improve the quality of healthcare by promoting an individualized approach to the patient and supporting decision-making by expanding the availability of relevant and accessible information. This contributes to improving the overall quality of care and increasing patient satisfaction, while also enabling a more accurate assessment of the effectiveness of various health interventions [1]. According to a systematic review of 76 recent studies, the collection of PRO data was found as an effective way to assess the impact of treatment on patients and improve the outcomes in 43 to 71% (according to the specific use of PRO data) of the studies [2].

To enable the use of PRO data, standardized methodologies (under the umbrella term patient-reported outcome measurements – PROMs) have been developed and used in practice in the recent decades. The focus of the methodologies is on the structure of the data and their interpretation by specialists and managers. One common feature is that both the management and technical implementation of PROMs collection are described as recommendations and good practices. The design of the collection and delivery process is left to the engineering teams, which indicates a search for maximum flexibility on the part of the creators of the methodologies. As a result, descriptions of functional models and processes are informally considered as part of the reports with examples of implementation cases.

One of the main challenges in the use of PRO data, noted by the research community, is the technical integration of the relevant software solutions into existing health information systems [3]. The studies also propose approaches, such as the use of standardized data exchange interfaces (e.g., HL7 FHIR), as well as the development of integration modules that facilitate interoperability between PRO modules and other components of the health information systems [4, 5].

This paper presents an analysis of 8 of the most widely used modern methodologies for collecting PRO data. Based on this analysis, a functional model of a software system for collecting and storing PRO data is proposed, with aim to identify the main user roles and propose a view that is independent of a specific methodology or questionnaire. The motivation for proposition of the model is to facilitate easier software information system development and system-to-system integration.

Analysis of leading methodologies for collecting PRO data. The standard technique of documentation analysis from information systems requirements engineering was used to examine 8 leading methodologies, created and introduced into practice and healthcare systems in the period from 1990 to 2020. For each methodology, the characteristics described in the main and supporting documentation were analyzed for the presence of a comprehensive data collection process, timeline, key moments and actors in the process. Reports on implementations of PRO collection systems were analyzed, as well as independent studies of implemented systems, processes, applied architectures and approaches to data collection using software information systems.

International Consortium for Health Outcomes Measurement (ICHOM). ICHOM develops standard sets of measures based on clinical and patient research and covering various medical conditions. Each condition includes specific measures such as physical function, pain, fatigue and emotional state, described in “vocabularies” that provide a model of the patient’s health status [6]. The measures are structured in frameworks that cover both general health indicators and condition-specific indicators. The frameworks are accompanied by a detailed timeline that indicates key points at which certain data should be collected. This timeline can be used as a basis for a data collection process, but the frameworks do not contain a formal requirement and description of such a process. Actors are defined implicitly in the descriptions and explicitly in the dictionaries for each of the diagnoses.

EuroQol Five-Dimension Scale (EQ-5D). EQ-5D is a quality-of-life measurement tool that has been introduced and used since the 1990s. It consists of five main dimensions covering mobility, self-care, usual activities, pain and discomfort, and anxiety and depression. The EQ-5D is designed to be easy to use and understand by patients and can be applied to different populations and medical conditions [7]. The EQ-5D does not contain a data collection process description. Actors descriptions can be derived from the so-called “management model”, which describes four possible approaches to data collection.

Patient-Reported Outcomes Measurement Information System (PROMIS) PROMIS was developed by the National Institutes of Health (NIH) and is a dynamic system for collecting PROMs through computer-adapted questionnaires. PROMIS includes a set of measurements that cover physical, mental and social health. The structure of PROMIS is built on three main domains: physical health (including physical function, pain and fatigue), mental health (including emotional state, anxiety, depression) and social health (including social relationships, social role and satisfaction with social support). The questionnaires were designed to be as adaptable as possible by using computer-adaptive questionnaires (CATs), which dynamically select questions depending on the patient's previous responses [8]. The methodology does not contain a description of a data collection process, but accompanying documents describe good practices in its application. Actor descriptions can be extracted from these documents, with the main ones being "patient" and "proxy" to help complete the questionnaires.

Short Form-36 (SF-36). The SF-36 is a standardized questionnaire that measures quality of life as a function of the physical and emotional state of patients. It contains 36 questions structured into eight domains: physical health, emotional well-being, social relationships, role in society, pain, fatigue, general health, and daily functioning. Each of these domains provides information about the patient's condition and the impact of various health interventions on his or her quality of life [9]. This instrument does not contain instructions for particular data collection process or actors in it. An online implementation of the questionnaire is available on the developers' website.

Health Utilities Index (HUI). The HUI is a collection of standardized measures used to assess health status and quality of life. It includes measures such as pain perception, mobility, emotional state, and social function. The HUI has been discussed as particularly suitable for long-term studies and evaluation of health interventions [10]. The documentation contains an informal description of a data collection process with at least two actors – patient and "proxy", who replaces the patient in completing the questionnaires.

SF-12 Health Survey. The SF-12 is a shortened version of the SF-36 and is also designed to measure quality of life through physical and psychological components. The SF-12 is popular when quick and less extensive collection of data about the patient's health is needed [11]. Like the SF-36, this instrument does not contain data collection process or actors instructions.

Nottingham Health Profile (NHP). The NHP is a questionnaire used to measure the physical, emotional and social health of patients. It contains six sections: energy, pain, emotional reactions, sleep, social isolation and physical abilities. The NHP is one of the earliest instruments, available since the 1970s and is still popular in Europe [12]. The instrument does not contain data collection process or actors instructions.

Clinical Outcomes in Routine Evaluation (CORE). CORE was developed to assess the mental health of patients in a clinical setting. CORE-OM is one of the most widely used outcome measurement tools in the mental health context [13]. The tool does not contain data collection process or actors instructions.

Functional model for collecting PRO data proposition. The proposed functional model is based on the results of the previously described analysis of the methodologies and is developed by extracting common features and synthesizing them into a model of user roles (derived from the common characteristics of the actors in the analyzed data collection processes) and groups of functionalities. The technique applied for modeling functionalities is a standard UML diagram of user roles and cases ("use case diagram"), (Fig. 1).

The following system roles have been identified:

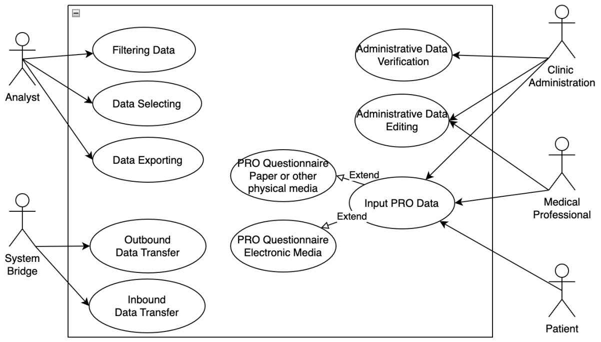

Fig. 1. PRO Data Collection System – Use Case Diagram

- Patient – this role inputs PRO data into the system, on his/her own behalf or on behalf of the patient, if acting as a "proxy".
- Medical Professional – this role inputs PRO data into the system as a patient assistant (applicable to certain diagnoses and methodologies) and edits administrative data if needed.
- Analyst – this role analyzes PRO data for the purposes of management of the healthcare system and determining the effectiveness of care.
- Clinic Administration – this role verifies and edits administrative patient data in the system as needed. The role can input PRO data into the system, either as a patient assistant or data by the patient on a paper (or other physical media) questionnaire.
- System Bridge – this role establishes a connection with other systems or modules and has the ability to download data from the system or upload data to it.

The model allows design of systems in which one user can have more than one role. That enables implementation of more sophisticated or specific scenarios, inherent in some of the methodologies – e.g. a medical professional who is also an analyst.

Discussion and conclusion. The proposed model of the functionalities is verified against the characteristics of each of the described methodologies, through repeated analysis of the documentation and checks for contradictions. The described functionalities allow for the implementation of a complete process for collecting and analyzing PRO data using any of the selected methodologies.

An important feature of the "Patient" role, which should be reflected in the implementation of the system, is the possibility of filling out the questionnaire not directly by the patient, but by a "proxy". According to some of the methodologies, the "proxy" is an extension of the patient that may have some unique characteristics. For example, in HUI methodology there is a difference in the questionnaires, if they are filled out by the patient or by a proxy. There is no discrete role of proxy modeled, however, since the functionalities are the same as those of the patient.

The "System Bridge" role is not described in any of the methodologies, but was added to allow integration with other systems. The argumentation is based on the research of the implementations case studies and the proposed reference models of systems of systems, especially the keystone "Advancing the Collection and Use of Patient-Reported

Outcomes through Health Information Technology" report [4].

One natural path of development of the results is to broaden the model with UML sequence and data flow diagrams to both make the model more sophisticated and to further the claim for universality and measurement methodology independence. Another path to add even more value is to supplement with descriptions and proposition of several of the key business processes involved in PRO data collection. An empirical approbation of the models is also planned to be exercised through prototyping a software system based on these designs and implementing it in the practice of one medical organization.

References

[1] M. E. Porter, E. O. Teisberg. Redefining Health Care: Creating Value-based Competition on Results. Harvard Business Press, 2006.

[2] J. M. Bonsel, A. J. Itiola, A. S. Huberts, G. J. Bonsel, H. Penton. The use of patient-reported outcome measures to improve patient-related outcomes – a systematic review. *Health Qual. Life Outcomes*, **22** (Nov. 2024), p. 101, `https://doi.org/10.1186/s12955-024-02312-4`.

[3] C. A. Harle et al. Overcoming barriers to implementing patient-reported outcomes in an electronic health record: a case report. *J. Am. Med. Inform. Assoc. JAMIA*, **23**, no. 1 (Jan. 2016), 74–79, `https://doi.org/10.1093/jamia/ocv085`.

[4] Advancing the Collection and Use of Patient-Reported Outcomes through Health Information Technology, 2021. Accessed: Dec. 15, 2024. [Online]. Available: `https://aspe.hhs.gov/sites/default/files/private/pdf/259016/ONCPROFinalReportFinal.pdf`.

[5] E. Basch, C. Snyder. Overcoming barriers to integrating patient-reported outcomes in clinical practice and electronic health records. *Ann. Oncol.*, **28**, no. 10 (Oct. 2017), 2332–2333, `https://doi.org/10.1093/annonc/mdx506`.

[6] M. E. Porter, S. Larsson, T. H. Lee. Standardizing Patient Outcomes Measurement. *N. Engl. J. Med.* **374**, no. 6 (Feb. 2016), 504–506, `https://doi.org/10.1056/NEJMp1511701`.

[7] M. Herdman et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil.*, **20**, no. 10, (Dec. 2011), 1727–1736, `https://doi.org/10.1007/s11136-011-9903-x`.

[8] D. Cella et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J. Clin. Epidemiol.*, **63**, no. 11 (Nov. 2010), 1179–1194, `https://doi.org/10.1016/j.jclinepi.2010.04.011`.

[9] J. E. Ware, C. D. Sherbourne. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med. Care*, **30**, no. 6 (Jun. 1992), 473–483.

[10] D. Feeny et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Med. Care*, **40**, no. 2 (Feb. 2002), 113–128, `https://doi.org/10.1097/00005650-200202000-00006`.

[11] J. Ware, M. Kosinski, S. D. Keller. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med. Care*, **34**, no. 3 (Mar. 1996), 220–233, `https://doi.org/10.1097/00005650-199603000-00003`.

[12] S. M. Hunt, J. McEwen, S. P. McKenna. Measuring health status: a new tool for clinicians and epidemiologists. *J. R. Coll. Gen. Pract.*, **35**, no. 273 (Apr. 1985), 185–188.

[13] M. Barkham, N. Gilbert, J. Connell, C. Marshall, E. Twigg. Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings. *Br. J. Psychiatry*, **186**, no. 3 (Mar. 2005), 239–246, `https://doi.org/10.1192/bjp.186.3.239`.