Towards a Process Reference Model for Clinical Coding

Ismael Caballero[1(✉)] , Júlio Souza[2,3] , Fernando Lopes[2,3] ,
João Vasco Santos[2,3,4] , and Alberto Freitas[2,3]

[1] Institute of Technology and Information Systems (ITSI), University of Castilla-La Mancha
(UCLM), Ciudad Real, Spain
`Ismael.Caballero@uclm.es`
[2] Center for Health Technology and Services Research (CINTESIS), Faculty of Medicine,
University of Porto, Porto, Portugal
[3] Department of Community Medicine, Information and Health Decision Sciences
(MEDCIDS), Faculty of Medicine, University of Porto, Porto, Portugal
[4] Public Health Unit, ACES Grande Porto VIII–Espinho/Gaia, Espinho, Portugal

Abstract. Coding of medical data is a very important previous step for many
activities in Health Care management, since it is the basis for several activities
ranging from hospital reimbursement to clinical research. Literature identifies
some issues related to coding clinical data, which derives in inadequate levels
of quality leading to some in acceptable situations in health care organizations,
impacting even to their sustainability. To alleviate these undesirable effects, we
posse that the standardization of some best practices around clinical coding can
lead to a better performance of the clinical coding process. One of the most relevant
concerns in the process is the quality of the data used at the various stages of the
data life cycle from its generation by clinicians up to the usage and exploitation
of the data once coded. The main contribution of this work is twofold: on a hand
to identify which are the best practices related to clinical coding, and on the other
hand to investigate how these best practices can be enriched with some other
related to data quality management and data governance. As a result, we produced
CODE.CLINIC, a framework that can be used to support institutions to better code
their medical data. This framework consists of two main components: a Process
Reference Model (PRM) and a Process Assessment Model (PAM). In this paper
we are going to first introduce the CODE.CLINIC PRM, which gather 16 process
grouped in 4 blocks.

Keywords: Coding of medical data process · Data quality management ·
CODE.CLINIC

1 Introduction

Most of the actions happening in health care institutions are based and largely depends
on data. While data regarding patients are typically stored in their corresponding health
records [1], it is also necessary to gather how the medical attendance processes have

been conducted: which the diagnostics tests were driven, which diseases were diagnosed, which treatments were applied and where the patients went once treated. In this sense, relevant diagnoses and treatment procedures should be abstracted and conveniently categorized for ulterior inspections about the performance of the health institution. Examples of these inspection include calculations to reimbursement to health institutions, medical research, epidemiological surveillance, or the calculation of numerous quality indicators, performance monitoring, hospital output, and benchmarking among other purposes [2]. In different countries, specific ways to gather and store all this data once categorized are proposed. For instance, in Portugal, the Hospital Morbidity Database (HMD) maintains a wide range of information on inpatient and outpatient episodes occurring in National Health System (NHS) health institutions, a type of diagnosis-related groups (DRGs) [3].

This comprehensive set of clinical data, the HMD, results from a series of routine processes in hospitals, typically embracing the clinical coding process [4]. These routines begin by documenting all clinical information and services provided to patients through data collection instruments, either in paper or digital format; after patient discharge, this set of information is then accessed by coding physicians, who must translate and classify the diagnoses and procedures according to any of the varieties of the WHO's International Classification of Disease, 10th Revision (ICD-10) codes manually or being supported for automatic classification systems [4, 5]. For instance, in Portugal – the first European country implementing a DRG-based hospital systems in 1989 – the current version used is ICD-10 Clinical Modification (ICD-10-CM/PCS) [6].

Given the importance of this information for the wide range of purposes previously pointed, it is then essential to ensure that the clinical data coded in the different hospitals are of high quality and have accurate, reliable, and fully reported procedure and diagnosis codes to support the reuse of these data at different levels.

In this sense, the literature points to the existence of several quality problems in coded clinical data [7], and many of these stems from barriers that exist in the clinical coding processes themselves in general terms [8], or more specifically in Portugal to some barriers as [9] identified. Differences in coding of diseases and procedures between coding physicians and hospitals, interpretation of guidelines and instruments used, as well as gaps in clinical documentation, delays and lack of health record standards are some of the several problems already mentioned [10].

In this investigation, the researcher team posed that to mitigate or at least to alleviate these problems, it was possible to gather, and group set of best practices (also known as process) that can be used to ensure homogeneous behavior during the clinical coding process and during the use of the resulting data either internally in every health institution or externally when this health institution needs to exchange data with third parties like national governments. This set of best practices should cover aspects of the clinical coding process as well as some other aspects of data quality management and data governance.

In this sense, the research team relied on the idea of the process-based approach set forth in international standards such as ISO/IEC 8000-61 [11] for data quality management or ISO 12207 [12] for the case of software process to develop a framework called CODE.CLINIC to better support the clinical coding. Thus, the idea of CODE.CLINIC is to serve as reference that every health institution can use to customize their own set of

clinical coding practices according to their own restrictions, possibilities, and demands. CODE.CLINIC consists of two main components:

- A **Process Reference Model (PRM)**, containing four groups of processes addressing specific aspects of the clinical coding activities considering the various stages of the clinical coded data lifecycle. This PRM also meets the data governance, data management and data quality management process requirements included in the Alarcos' Model for Data Improvement (MAMDv3.0) [13], an ISO 8000-61-framework compliant. In Sect. 3, explanations about the construction of the PRM has been provided. In addition, this PRM can be also used as a body of knowledge for the various processes of the clinical coding, for instance, to identify relevant stakeholders, specific information systems to better support the processes, or even recommended key process indicators to monitor the institutionalization of the PRM processes. The current version of CODE.CLINIC PRM is the main contribution of this paper.
- A **Process Assessment Model (PAM)**, containing the elements required for organizations to assess and improve their clinical coding activities according to the provided PRM. This PAM will be built by meeting the requirements of ISO/IEC 33003 [14] along with some other parts of ISO/IEC 33000 series. One of the most important components of the PAM is a Maturity Model, in which the processes of the PRM are ordered in an increasing level of difficulty according to the capabilities of the health institution. The CODE.CLINIC Maturity Model has been developed meeting the requirements of ISO/IEC 33004 [15], and it is compliant with MAMD-Maturity Model based on ISO 8000-62 [16]. However, the description of CODE.CLINIC PAM and the corresponding Maturity Model is outside of the scope of this paper.

The paper is structured as follows. Section 2 described the state of the art of the best practices in clinical coding processes. Section 3 summarizes the research method that we have used to produce the framework. Section 4 introduce the Process Reference Model (PRM) of the CODE.CLINIC process, describing the structure of the four groups of processes along the structure of every process. Finally, Sect. 5 introduces discussion, some conclusions, and future works.

2   State of the Art and Related Works

2.1   Existing Works on Clinical Coding

Alonso et al. in [10] identified several problems during the process of the clinical coding in Portugal through the conclusion raised after a the conduction of a focus group like: (1) Variability in clinical coding – (1a) coding process-electronic vs paper, (2) Difficulties in the clinical coding process: (2a) Difficulties in assigning diagnoses code, (2b) Coding process by clinical specialty; (3) Coding delay; (4) Hospital resources made available to coding activities; (5) Clinical coding audits. They also identified two important barriers: (1) limited understanding of medical terminology by coders, and (2) coder experience or problems in health records. All these problems and barriers were grouped in four main categories: (1) the standardization of the documents used for coding an episode,

(2), the adoption of the electronic coding, (3), the development of tools to help coding and audits, and (4) the recognition of the importance of coding by the management.

As of 2014, the CHKS produced the report "*The quality of clinical coding in the UK's NHS*" [17]. In this report, the following areas are identified for a senior management to support with the aim of assuring the adequate levels of quality for the patient care data: (1) source of documentation, (2) discharge summaries, (3) deadline and completeness for the clinical coding, (4) clinical engagement, (5) audit and analysis, (6) staffing, (7) training and guidance, (8) IT systems, (9) assessment units, and (10) broader users.

The Australian CCSA produced in 2019 the "*Clinical Coding Practice Framework*" [18], whose main aim is "*to provide guidance in defining and promoting good practices for those involved in the clinical coding process (e.g. clinical coders, clinical documentation improvement specialist, clinical coding auditors, health information managers and managers (at all levels) of the coding process)*". The framework engaged all those involved in the clinical coding process to: (1) gain access to all relevant and pertinent clinical information, (2) ensure the assignment and diagnoses and intervention codes include all the necessary information, (3) apply conveniently the classification conventions stated in the official recommendation of the Australian National Health Services, (4) actively participate in the interdisciplinary meeting for the purpose of clarification of diagnosis and intervention, and (5) improve the clinical understanding of the roles of those involved in the clinical coding process. In addition, the framework establishes that involved in the clinical coding process must not: (1) assign diagnoses or intervention codes without the adequate supporting information, (2) deliberately committing errors in code assignment to minimize financial losses, (3) deliberately manipulating the interdisciplinary engagement inappropriately to maximize financial benefits for the health institution, (4) omit information about an episode of care to prevent the accurately reporting, and (5) submit to pressure from other or to others to misrepresent the patient's episode of care or prevents adherence to the stated classification conventions.

Reid et al. in [19] and in [20] they presented some insights obtained by using four research methods (literature review, workshop, assessment of coding services, and medical record audit) from the state of clinical coding services in the Republic of Ireland as of 2017 studied during one year. The reports raised relevant results and recommended some best practices in several concerns: (1) quality of medical records, (2) coding work allocation and supervision processes, (3) data quality control measures, (4) communication with clinicians, and (5) visibility of clinical coders, their managers, and the coding services. They find that the best managed coding services had the following characteristics: (1) they did larger use of the available checking resources, (2) the medical records were more electronic, and (3) there were much higher levels of direct clinical coders contact with clinicians.

As consequence, it can be said that the coded clinical data often lacks quality enough to successfully be used in the management or research activities.

2.2 Alarcos' Model for Data Improvement (MAMDv3.0)

The Alarcos' Model for Data Improvement (MAMDv3.0) [13] is a framework to assess and improve the maturity of the organization's processes related to data management,

governance and quality. MAMDv3.0, which is publicly accessible[1], is aligned to several international open standards such as ISO/IEC 8000-61 [11], ISO/IEC 8000-62 [16], ISO 38505-1 [21], ISO 38505-2 [22], and it also gather some best practices from standards like COBIT 2019 [23] or DAMA's DMBOK 2 [24].

It consists of two main components:

- A Process Reference Model (PRM) consisting of twenty processes grouped in three categories: Data Management (DM), Data Quality Management (DQM) and Data Governance (DG).
- A Process Assessment Model (PAM), with indications to assess and improve the maturity of an organization regarding the capability of the organization of their DM, DQM and DG processes. One of the most important elements of the PAM is a maturity model which can be used for the both the certification of the maturity model and to outline roadmaps for the improvements.

These components can be used within any type or organization to enrich the design and execution of the business processes with DM, DQM and DG concerns to make them more efficient and effective. In addition, it is important to state that organizations can certify with external thirds their maturity regarding to these disciplines.

3   Research Method

To produce the CODE.CLINIC framework, we are following the Action Research (AR) Method [25] because it is necessary to "*focus on social systems and put knowledge into action as soon as possible*". AR is a collaborative research form that seeks to unite theory and practice between researchers and practitioners through a process that is cyclical in nature, producing new knowledge that is useful in practice. AR has been successfully used in Information Systems [26] and Software Engineering [27], and given the very nature of the medical domain of our research we feel it can be successfully used given the strong component about the "social" fact.

In this case, the research goal is the "*design and testing of clinical coding framework*". To achieve our goal, three AR cycles has been planned. For each AR cycle we identified different critical reference groups, but we maintained as potential beneficiaries, any Portuguese health institution (mainly hospitals) that wants to customize and institutionalize their own clinical coding process.

- An Initial **AR Definition Cycle**, where the main goal is to build a first version of CODE.CLINIC PRM. In this initial cycle, we explored the corresponding literature and existing documentation to identify the common problems in the coding clinical process; once identified the problems, we proposed an initial version of the processes in the PRM grouped in four groups (see Sect. 4) with the idea of tackling the specific problems, along with common ways to address data quality problems by means of the alignment of the alignment of the proposed processes with MAMDv3.0. This

---

[1] MAMDv3.0 can be downloaded for free from https://mamd.dqteam.es.

first version was validated with the Portuguese *Associação dos Médicos Auditores e Codificadores Clínicos* (AMACC)[2], who acted as critical reference group. Through four cycles of meetings with the underlying refinements, we produce the first version of the PRM, which is introduced in this paper.

- A Second **AR Acceptance Cycle**, where the main goal is to present the whole framework to several preidentified stakeholders with different responsibilities for the clinical coding process in various Portuguese health institutions (mainly hospitals), to identify its weaknesses and strengths from various points of view. This is to be done by means of surveys and personal interviews. With the obtained feedback, we will produce a more refined PRM, and a refined version of the PAM.
- The Third **AR Validation Cycle**, where the final version of CODE.CLINIC will be applied to a reference hospital following the case study research methodology.

4 The CODE.CLINIC Process Reference Model

In this section the PRM is to be introduced. The main purpose of the PRM is to identify the processes that can be used to describe and characterize the entire lifecycle and formal pathways of coded data in health institution to identify differences and constraints to achieve higher data quality under the perspective of the various stakeholders, ranging from medical coders to clinical coding office managers. This set of processes can be used as body of knowledge to address the required specific practices during the clinical coding. In this sense, every process can be seen as a box where every stakeholder can find the required knowledge (including the common activities and work products, like communication schemas or underlying key process indicators) for the clinical coding activities. In addition, this knowledge can be reviewed over the time to enrich by including new activities and/or work products. This knowledge can be used to outline the clinical coding activities when designed from scratch or to review or enrich existing ones by identifying potential root causes for existing problems.

The structure of the organization of the sixteen processes included in the PRM is adapted from the concept of Primary, Support and Organizational process in ISO/IEC/IEEE 12207:2017 [12]. This enables a better understanding of the purpose of every process, and it can be also used to better determine the contribution of every process to the general goal of the clinical code. As a result, we identified the following groups of processes: (1) Strategic Processes (G Processes), (2) Main Processes (M Processes), (3) Support Processes (S Processes), and (4) Other Processes (O Processes). See Sects. 4.1–4.4 for a larger description of this groups of processes.

As said, the definition of the PRM has been aligned to MAMDv3.0. Due to this reason, and for the sake of interoperability with MAMD, and other process-approaches based on ISO standards, the process description of every process in the CODE.CLINIC-PRM has been done according to ISO/IEC/TR 24774 [28], and it consists of the following elements:

- Title, which is a descriptive heading for the process at task.

---

[2] https://amacc.med.up.pt.

- Purpose, which describes the main goal of the health institution when executing the corresponding process.
- Outcomes, which represent the observable results expected from the successful execution of the process.
- Activities, which is a list of actions (best practices) that can achieve the outcomes.

It is worthy to state that, due to length paper restrictions, unfortunately we will not include the full description (e.g., the full description of the four previous elements for every process). However, we will show the title and purpose of everyone, and we will offer some outcomes and activities for some relevant processes.

In the following subsections we identify the types of process, we listed the title and purpose of every process providing some examples of outcomes and activities.

4.1   The Strategic Process Group

This group of processes (see **Error! Reference source not found.**) address the concerns related to the governance of the clinical coding activities, namely, those related to the creation of internal standards, identification of best practices, norms, guidelines, and policies to rule all details –including data quality concerns from the organizational point of view– regarding the various stages of all data considered for clinical coding. It is also worthy to note the special focus on the organizational structure and human resource parts. It should be addressed not only who should oversee the various activities, but also how the communication should be established. To better support the clinical coding activities, specific competences, and hard and soft skills are required, and health institutions must provide training plan to their workers to achieve such competences and skills for the sake of the sustainability of the organization.

Just an example, we introduce the outcomes we identified for G.01:

- Updated list of best practices in clinical coding.
- Updated list of clinical coding data guidelines.
- Updated list of clinical coding data policies.
- Resources to propagate the best practices, norms, policies, and guidelines.

To achieve and manage the previous outcomes for the processes G.01, we identified the following activities:

- AG01.1. To select and define the best coding practices for different medical areas.
- AG01.2. To assign episodes to medical coders according to pre-defined rules.
- AG01.3. To list and define and develop the corresponding controls to the most important norms and regulations regarding data protection/security, quality, and access/use, to comply with the General Data Protection Regulation (EU GDPR).
- AG01.4. To define norms regarding standard reference books, supporting instruments and clinical coding guidelines.
- AG01.5. To define and harmonize patient documentation sources to be considered for clinical coding.

- AG01.6. To ensure continuous awareness, training, and preparation on clinical terminologies and DRG grouper updates.
- AG01.7. To define norms and standards regarding software and hardware resources to be used within the clinical coded data lifecycle (Table 1).

Table 1. The strategic processes.

Process title	Purpose
G.01. Creation or Selection, implementation, and maintenance of standards, best practices, norms, guidelines, and policies	The main goal of this process is to set up the management environment required to execute the coding clinical activities according to the capabilities of the health institutions and following the criteria and recommendations provided by the authorized organizations at a national level
G.02. Development and maintenance of controls to meet policies	The main aim of this process is to develop and maintain the corresponding controls to check if the proposed policies are being met during the execution of the coding clinical activities
G.03. Organizational Structure Management	This process is aimed to establish and maintain a supportive organization for the coding clinical data processes, identifying roles and responsibilities and to also define the competences and skills that are required for every responsibility. In addition, communication means, and protocols are also observed in this process
G.04. Stakeholders' skills and competences management	The process is aimed at maintaining a catalogue of the training required for assuring different stakeholders can achieve the associated competences and skills required to play their roles

4.2 The Main Process Group

This group of seven processes is aimed at covering all the concerns related to the proper clinical coding itself describing the various activities related to the coded data lifecycle, from data acquisition (e.g., health records) to the use and exploitation of the coded data for health institution management or for medical and epidemiologic research. See Table 2 for the purpose of every Main process.

Following the same structure of the previous subsection, the outcomes for two relevant processes are introduced as example. In this sense, the following outcomes for process M.01 are introduced:

- Identification and access to all required patient documentation.
- Reports on the levels of quality of the acquired patient documentation.

    And the corresponding activities for this process M.01 are the following one:

- AM01.1. To list and identify all patient documentation sources to be used for clinical coding.
- AM01.2. To access patient documentation sources for clinical coding.
- AM01.3. To ensure that all medical coders have access to all patient documentation sources for clinical coding.

Table 2. The main processes group

Process title	Purpose
M.01. Data acquisition	This process is aimed at selecting and acquiring the required data from the typical data sources (e.g., health records both paper-based and electronically)
M.02. Data Integration (internal)	In this process, the integration coming from the various data sources should be achieved to create a solid basis for the clinical coding process
M.03. Data Coding	This process is aimed at properly coding the data
M.04. Submission of clinically coded data to the national repository	This process covers the exportation of the results of the codification of the data corresponding to the episodes towards the considered destinations
M.05. Incorporation of Coded Data to APR-DRG (DRG grouper software)	The purpose of this process is to incorporate the coded data into the APR-DRG
M.06. Data exploitation for hospital management, financing (billing), and public health	The objective of this process is to support all the necessary operations for hospital management, billing, and public health
M.07. Data exploitation for clinical and epidemiologic research	The main purpose of this process is to produce research reports on clinical aspects

The main intention of CODE.CLINIC is to be a generic framework valid for any country, for any health institution, and for any technology. CODE.CLINIC should be customized for any context. In this sense, as an example, the activities of process M.04 can be customized for Portugal as follows:

- AM04.1. To submit coded data to SIMH.
- AM04.2. To retrieve coded data from SIMH to perform corrections.

Being SIMH the *Sistema de Informação de para a Morbilidade Hospitalar*[3], whose main goal is to gather, edit and group in Homogeneous Diagnoses Groups the patient episodes.

4.3   The Support Process Group

In this group of four processes, the specifics of quality management of the data used as input (patient documentation) and output (coded data) of the coding clinical is covered. In addition, the concerns related to technological infrastructure management along with the maintenance of the reference data standards are also covered. See Table 3 to see the process title and purpose.

Table 3. The support processes group

Process title	Purpose
S.01. Data quality management of patient documentation	The main aim of this process is to evaluate and improve the level of the quality of the health record documents
S.02. Data quality management of coded data	Once produced the data, the main aim of this process is to evaluate and improve the quality of the resulting clinical coded data
S.03. Reference data management	This process is aimed at maintaining the various reference data involved in the codification of the clinical data (e.g., ICD-10-CM)
S.04. Technological infrastructure management	This process is aimed at establishing the required technology to support the flow of information through all the main processes, as well as to interact with some other agents (e.g., some other hospitals, some other regulatory organizations)

For illustrative purpose, and due to its importance, we proposed the following activities for the S.01 processes:

- AS01.1. Identify the most relevant data quality characteristics/dimensions for health records.
- AS01.2. Define measurement methods to assess the levels of quality of health records.
- AS01.3. Analyze the root causes of inadequate levels of quality.
- AS01.4. To perform internal auditing of health records.
- AS01.5. Analyze the root causes of inadequate levels of quality.
- AS01.6. Improve the levels of quality of health records.
- AS01.7. Generate data quality management reports for health records, with learned lessons.

---

[3] https://www.spms.min-saude.pt/2021/03/simh/.

And for S.02. Data quality management of coded data process, the following activities were proposed:

- AS02.1. To perform internal auditing of coded data according to established norms.
- AS02.2. To retrieve episodes with coding issues from SIMH.
- AS02.3. To correct (recode) and resubmit the episodes to SIMH.
- AS02.4. To define standard auditing controls.

Once again, let us recall that the framework should be customized for the specific reality of the county in which they will be applied. In this sense, the outcomes of the M.06. Data exploitation for hospital management, financing (billing), and public health would be customized for Portugal, by including the following outcome:

- ACSS benchmarking tool[4].

4.4   The Other Process Group

Finally, in the other processes group, consists of the process shown in Table 4.

Table 4. The other group processes group

Process title	Purpose
O1. Health care taking process	This process is aimed at diagnosing diseases and providing the corresponding treatment these diseases. This process should produce and store the corresponding the necessary data to describe the most relevant details

To illustrate this process, the following activities have been proposed:

- AO01.1. To report accurate and complete information in the admission note (e.g., symptoms, comorbidities), following the patient's admission.
- AO01.2. To report accurate and complete information in the discharge notes (e.g., patient's diagnoses, treatment, and disease progression), following the end of the episode.
- AO01.3. To report accurate and complete information in the anesthesia report.
- AO01.4. To report accurate and complete information in the surgical report.
- AO01.5. To report accurate and complete information in the pathology report.
- AO01.6. To report accurate and complete information in the nursing records.

4.5   Customization of the Framework for a Specific Context

We intended the PRM of CODE.CLINIC would be complete and flexible enough to be adapted and suitably customized for various context (e.g., specific countries). In this

---

[4] https://benchmarking-acss.min-saude.pt/.

sense, the outcomes and activities should be selected and reinterpreted for the specific context. This involves, for instance, to identify who are the most relevant actor and stakeholders for the various process groups in every context. In this sense, we have identified specific actors/stakeholders that are relevant for the customization of CODE.CLINIC for Portuguese health institutions, considering the various stages of the coded data lifecycle. We classify them in three large groups:

Consultive Roles and Responsibilities. General policymakers or policy-proposers for health domain, that are typically outside of the organization, mainly in the regional or national government. They provide some hints about general concerns and recommendations for the clinical coding activities. Some of these concerns consists in recommendations for technical support (e.g., like the AMACC), while others are more related to management and interoperability support (e.g., SPMS). This are not typically active roles in the specific case of Portugal we can list:

1. Administração Central do Sistema de Saúde (ACSS).
2. Serviços Partilhados do Ministério da Saúde (SPMS).
3. Ordem dos Médicos.
4. Colégio Competência de Codificação Clínica da Ordem dos Médicos.
5. Associação dos Médicos Auditores e Codificadores Clínicos (AMACC).

   Active Roles and Responsibilities for the Process. Workers that are somehow involved in clinical coding tasks at institutional level for the Strategic and Main and Support Processes.
6. Hospital managers (departments and services).
7. Health care provider.
8. IT (Information Technology) Staff.
9. Hospital manager for clinical coding.
10. Clinical coding office managers (internal auditors).
11. Medical coders.
12. Physicians.

   Benefited Roles, played by workers that will use the coded data for the various purposes.
13. Public Health authorities and health care administrators (health management users).
14. Researchers (health research users).

Customization of CODE.CLINIC for Health Institutions. The institutionalization of CODE.CLINIC involves the identification and customization of the various outcomes and activities for the reality of the health institutions, by considering the specifics for the regular flow of data (episodes and coded data) following the main processes. During the presentation of the processes in Sects. 4.1–4.4 we introduced some examples of customization. As part of the customization of the activities, the specific positions in the health institutions corresponding to the previously introduced roles must be identified and their responsibilities conveniently assigned, and the communication protocol and means adequately established. Regarding outcomes and work products, every health institution should raise consensus on how to name, store and establish guidelines by means of policies and procedures to exploit them.

Finally, it is important to note that the customization will be better supported by the CODE.CLINIC PAM along with the underlying maturity model, as it will enable the creation of roadmaps.

5 Discussion, Conclusions and Future Work

This research has been motivated by the hypothesis that having and implementing a framework for clinical coding will not only contribute to a greater homogenization of clinical coding processes not only internally in health institutions, but also to enable better interoperability between other organizations, enabling even benchmarking. In addition, the institutionalization of the framework will make the clinical coding much more efficient, interoperable, and error-free, by facing the list of problems listed in Sect. 2.1 in a unified and controlled way from an organizational point of view.

The main expected impact of this PRM is not only to identify, map and structure the various processes and activities related to clinical coding in day-to-day life at health institutions (mainly hospitals), but also to provide a working tool to promote data reuse, good clinical coding practices and to organize processes, promoting data quality improvement, and body of knowledge.

As part of our near-term future work, we will conduct the AR acceptance cycle, in which we aim to gain buy-in from the medical community involved in the clinical coding process. To this end, we will conduct a series of questionnaires and interviews with representatives of the various stakeholders described in Sect. 4.5. The feedback obtained will be used to refine both the process reference model and the evaluation model. Once the framework will be refined, we will disseminate it in different forums.

References

1. Gesulga, J.M., Berjame, A., Moquiala, K.S., Galido, A.: Barriers to electronic health record system implementation and information systems resources: a structured review. Procedia Comput. Sci. **124**, 544–551 (2017)
2. Alonso, V., et al.: Health records as the basis of clinical coding: is the quality adequate? A qualitative study of medical coders' perceptions. Health Inf. Manag. J. **49**(1), 28–37 (2020)
3. Fetter, R.B.: Diagnosis related groups: understanding hospital performance. Interfaces **21**(1), 6–26 (1991)

4. Stanfill, M.H., Williams, M., Fenton, S.H., Jenders, R.A., Hersh, W.R.: A systematic literature review of automated clinical coding and classification systems. J. Am. Med. Inform. Assoc. **17**(6), 646–651 (2010)

5. Hazelwood, A.C.: ICD-9 CM to ICD-10 CM: implementation issues and challenges. In: ICD-9 CM ICD-10 CM: Implementation Issues and Challenges/AHIMA, American Health Information Management Association (2003). http://library.ahima.org/doc?oid=59978

6. CMS: ICD-10-CM Official Guidelines for Coding and Reporting. Centers for Medicare and Medicaid Services (2021). https://www.cms.gov/files/document/2021-coding-guidelines-upd ated-12162020.pdf

7. Carvalho, R., et al.: Analysis of root causes of problems affecting the quality of hospital administrative data: a systematic review and Ishikawa diagram. Int. J. Med. Inf. **156**, 104584 (2021). https://doi.org/10.1016/j.ijmedinf.2021.104584

8. de Lusignan, S.: The barriers to clinical coding in general practice: a literature review. Med. Inform. Internet Med. **30**(2), 89–97 (2005). https://doi.org/10.1080/14639230500298651

9. Alonso, V.: A Codificação Clínica e os problemas associados à qualidade dos dados: perspetiva dos codificadores. Maestrado em Informática Médica. Faculty of Medicine. University of Porto, Porto (2018). https://repositorio-aberto.up.pt/bitstream/10216/118231/2/306324.pdf

10. Alonso, V., et al.: Problems and barriers during the process of clinical coding: a focus group study of coders' perceptions. J. Med. Syst. **44**(3), 1–8 (2020). https://doi.org/10.1007/s10916-020-1532-x

11. ISO: ISO/IEC 8000-61:2016: Data quality – Part 61: Data quality management: Process reference model. ISO (2016). https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/30/63086.html. Accessed 4 Aug 2021

12. ISO: ISO/IEC/IEEE 12207:2017 – Systems and software engineering – Software life cycle processes. ISO/IEC/IEEE 12207:2017 (2017). https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/37/63712.html. Accessed 11 Apr 2022

13. DQTeam: MAMD: Modelo Alarcos Mejora Datos (2020). https://mamd.dqteam.es. Accessed 11 Apr 2022

14. ISO: ISO/IEC 33003:2015: Information technology – Process assessment – Requirements for process measurement frameworks. ISO (2015). https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/05/41/54177.html. Accessed 11 Apr 2022

15. ISO: ISO/IEC 33004:2015: Information technology – Process assessment – Requirements for process reference, process assessment and maturity models. ISO (2015). https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/05/41/54178.html. Accessed 11 Apr 2022

16. ISO: ISO 8000-62:2018: Information technology – Process assessment – Requirements for process reference, process assessment and maturity models. ISO (2018). https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/53/65340.html. Accessed 11 Apr 2022

17. Capita: The quality of clinical coding in the NHS: payment by results data assurance frameworks. Capita Health and Wellbeing Limited (2014). https://www.chks.co.uk/userfiles/files/The_quality_of_clinical_coding_in_the_NHS.pdf

18. CCSA: Clinical Coding Practice Framework. Clinical Coders' Society of Australia (2019). https://www.ccsofa.org.au/wp-content/uploads/2021/05/HIMAA-CCSA-CCPF-FIN AL5-Sep2019.pdf

19. Reid, B.A., Ridoutt, L., O'Connor, P., Murphy, D.: Best practice in the management of clinical coding services: insights from a project in the Republic of Ireland, Part 1. Health Inf. Manag. J. **46**(2), 69–77 (2017)

20. Reid, B.A., Ridoutt, L., O'Connor, P., Murphy, D.: Best practice in the management of clinical coding services: insights from a project in the Republic of Ireland, Part 2. Health Inf. Manag. J. **46**(3), 105–112 (2017)

21. ISO: ISO/IEC 38505-1:2017 Information technology – Governance of IT – Governance of data – Part 1: Application of ISO/IEC 38500 to the governance of data. ISO/IEC 38505-1:2017 Information technology – Governance of IT – Governance of data – Part 1: Application of ISO/IEC 38500 to the governance of data (2017). https://www.iso.org/standard/56639.html. Accessed 9 May 2021

22. ISO: ISO/IEC TR 38505-2:2018 Information technology – Governance of IT – Governance of data – Part 2: Implications of ISO/IEC 38505-1 for data management. ISO/IEC TR 38505-2:2018 Information technology – Governance of IT – Governance of data – Part 2: Implications of ISO/IEC 38505-1 for data management (2018). https://www.iso.org/standard/70911.html. Accessed 23 May 2021

23. ISACA: COBIT 2019 Framework. Introduction and methodology. Schaumburg, IL. EE.UU (2018)

24. DAMA: DAMA-DMBOK: Data Management Body of Knowledge. Technics Publications, LLC (2017)

25. Wohlin, C., Runeson, P.: Guiding the selection of research methodology in industry–academia collaboration in software engineering. Inf. Softw. Technol. **140**, 106678 (2021). https://doi.org/10.1016/j.infsof.2021.106678

26. Avison, D.E., Davison, R.M., Malaurent, J.: Information systems action research: debunking myths and overcoming barriers. Inf. Manage. **55**(2), 177–187 (2018). https://doi.org/10.1016/j.im.2017.05.004

27. Staron, M.: Action Research in Software Engineering. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-32610-4

28. ISO: ISO/IEC/IEEE 24774:2021 Systems and software engineering – Life cycle management – Specification for process description. ISO (2021). https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/07/89/78981.html. Accessed 11 Apr 2022