Exploring Complexity in Health: An Interdisciplinary Systems Approach
A. Hoerbst et al. (Eds.)

doi:10.3233/978-1-61499-678-1-394

Medical Classification and Terminology Systems in a Secondary Use Context: Challenges and Perils

Hauke HUND[a,b,1], Sven GERTH[c], Hugo A KATUS[a] and Christian FEGELER[b]
[a]*Department of Cardiology, University Hospital Heidelberg, Heidelberg, Germany*
[b]*GECKO Research Institute, Heilbronn University, Heilbronn, Germany*
[c]*3M Deutschland GmbH, Berlin, Germany*

Abstract. Since the introduction of diagnosis-related groups in the German healthcare system, classifying patient diagnosis and procedures with controlled vocabularies have become mandatory and thus creating a large dataset for secondary use in biomedical research. In this paper we present the analysis of an ICD dataset with regards to potentially reimbursement motivated classification and the effects on precision and recall when considering the change history of ICD codes.

Keywords. Secondary Use, ICD 10, ICPM, OPS

1. Introduction

Since the introduction of diagnosis-related groups (DRG) in the German healthcare system for the purposes of reimbursement in 2003, classifying patient diagnoses with codes from the German adaptation of the *International Statistical Classification of Diseases and Related Health Problems Version 10* (ICD-10 – German adaptation: ICD-10-GM)[2] and documenting procedures by the use of codes from the German adaptation of the *International Classification of Procedures in Medicine* (ICPM – German adaptation: OPS)[3] have become mandatory. Classifying and documenting inpatients with controlled vocabularies has created a large dataset for secondary use in biomedical research.

Although ICD and sometimes OPS codes are considered a major foundation for identifying patients in a secondary use context [1],[2],[3], the quality of codes documented in the context of reimbursement and the semantics of changing classification systems while querying electronic health records (EHR) deserve a closer look.

Since 2003 sixteen ICD-10-GM and fifteen OPS catalogs have been released by the German governing organization DIMDI[4] with a new release being issued every year since 2004.

[1] Corresponding Author: hauke.hund@med.uni-heidelberg.de
[2] ICD-10-GM: German Modification.
[3] OPS: Operationen- und Prozedurenschlüssel.
[4] DIMDI: Deutsches Institut für Medizinische Dokumentation und Information. http://www.dimdi.de

In this paper we present the analysis of an ICD-10-GM dataset with regards to identifying usage changes of codes due to changes in reimbursement and the effects on precision and recall while querying for patients using ICD codes from only the latest release versus considering the change history of ICD codes. Two hypotheses were used:

- Semantically identical categories with different ICD codes do not affect precision and recall.
- Classification is not affected by changes in reimbursement.

2. Methods

2.1. Open source programming library MEDICATS

To allow convenient access to ICD-10-GM and OPS codes we developed the open source programming library for Java 8 MEDICATS (Medical Classification and Terminology Systems).[5] MEDICATS offers a programming interface to access all major ICD-10-GM, OPS and Alpha-ID[6] releases and provides links between Alpha-ID and ICD-10-GM codes as well as links to previous releases of the same code system.

Our library reads XML, SGML and plain text files provided by DIMDI and stores the supported code systems as graph structures in memory to represent the organization of codes into chapters, blocks and categories as well as links to other code system entries and previous codes. The necessary files can be obtained from DIMDI and are listed on the MEDICATS project website.[5]

Two example use cases for the MEDICATS library are shown in Listing 1: First the lookup of a known ICD code and access of the corresponding previous version code, and second the traversal of the entire OPS code system.

```
package de.gecko.medicats.hec2016;

import de.gecko.medicats.icd10.*;
import de.gecko.medicats.ops.*;

public class Demo {
    public static void main(String[] args) {
        previousCode: {
            IcdService icdService = IcdService.getService();
            IcdNodeFactory nodeFactoryIcdGm2015 = icdService.getNodeFactory("icd10gm2015");
            IcdNodeWalker nodeWalkerIcdGm2015 = nodeFactoryIcdGm2015.createNodeWalker();
            IcdNode icdNodeY34 = nodeWalkerIcdGm2015.getNodeByCode("Y34");
            icdNodeY34.getPrevious().ifPresent(System.out::println);
        }

        traversal: {
            OpsService opsService = OpsService.getService();
            OpsNodeFactory nodeFactoryOps2015 = opsService.getNodeFactory("ops2015");
            OpsNodeWalker nodeWalkerOps2015 = nodeFactoryOps2015.createNodeWalker();
            nodeWalkerOps2015.preOrderStream().forEach(System.out::println);
        }
    }
}
```

Listing 1. Example use cases for the open source programming library MEDICATS.

[5] MEDICATS: https://github.com/hhund/medicats
[6] Alpha ID: A list of stable non classifying alphanumeric diagnoses codes with links to ICD-10-GM.

2.2. ICD-10-GM dataset analysis

Our analysis focuses on a pseudonymized ICD dataset from the Department of Cardiology, University Hospital Heidelberg, Germany containing ICD-10-GM codes for more than 100.000 patients documented from 2003 to 2015.

2.2.1. Effect on precision and recall

The MEDICATS library was used to prepare a code change list starting at codes from the ICD-10-GM 2015 release and going backwards to older releases. From this change list examples of semantically identical categories with varying codes in different ICD-10-GM releases were chosen. Two queries were used to select patients, the first taking the change list into account (baseline) and the second using only ICD codes from the 2015 release (experiment). The results were used to calculate precision, recall and F-measure [4] for the selected examples from the change list.

2.2.2. Classification motivated by reimbursement

During the DRG grouping process Complication and Comorbidity Levels (CCL) (interval [0..7]) of all secondary diagnoses are transformed into a Patient Clinical Complexity Level, which in turn affects the calculated DRG [5]. To identify usage changes of ICD codes the CCL values were used as surrogate criteria to indicate a change in reimbursement.

A documented diagnoses code was thus defined *reimbursement motivated* if the difference in average CCL values between two years (DRG catalog versions) was greater than 1 and the number of documented ICD codes between two years increased by at least 50%, or the difference in average CCL values was less than -1 and the number of documented ICD code decreased by at least 50%.

We selected ICD code pairs from the ICD chapter IX “Diseases of the circulatory system” (codes I00-I99) with equal semantics in different ICD-10-GM releases and only used entries with combined more than 10 documented diagnoses.

3. Results

The ICD-10-GM 2015 catalog contains 15761 categories; of these categories 4808 are based on at least one category with a different previous code when looking at the entire change history; 228 categories were identified having a semantically narrowing change and thus splitting into multiple sub categories.
The majority of the changes (88%) were introduced in the ICD-10 SGB-V 2.0 (2013) to ICD-10-GM 2004 update.

3.1. Effect on precision and recall

Five examples of semantically identical categories with varying codes in different ICD-10-GM releases were used. The selected changes are listed in Table 1.

The baseline experiment used matching codes from the appropriate ICD-10-GM release, while the experiment ignored the matching ICD-10-GM release and only used the new codes for querying. Precision, recall and F-measure results are shown in Table

2. The experiment generally resulted in lower recall then baseline. For the change mappings C_4 and C_5, our experiment also resulted in smaller than 100% Precision.

Table 1. Example ICD-10-GM change mappings.

C_n	Old ICD-10-GM code	New ICD-10-GM code
1	ICD-10-GM 2012: I48.10	ICD-10-GM 2013: I48.0
2	ICD-10-GM 2012: I48.11	ICD-10-GM 2013: I48.1 and I48.2
3	ICD-10-GM 2012: I48.19	ICD-10-GM 2013: I48.9
4	ICD-10-GM 2014: I70.23	ICD-10-GM 2015: I70.24
5	ICD-10-GM 2014: I70.24	ICD-10-GM 2015: I70.25

Table 2. Precision, Recall and F-Measure for ICD release experiments.

C_n	P / R / F-m	Experiment (ignoring ICD release)	Baseline (with ICD release)
1	Precision	100 %	100 %
1	Recall	33 %	100 %
1	F-measure	50 %	100 %
2	Precision	100 %	100 %
2	Recall	30 %	100 %
2	F-measure	46 %	100 %
3	Precision	100 %	100 %
3	Recall	19 %	100 %
3	F-measure	32 %	100 %
4	Precision	12 %	100 %
4	Recall	10 %	100 %
4	F-measure	11 %	100 %
5	Precision	84 %	100 %
5	Recall	7 %	100 %
5	F-measure	13 %	100 %

3.2. Classification motivated by reimbursement

From the pseudonymized ICD dataset 3349 diagnoses pairs with semantically identical ICD-10-GM codes in chapter IX from consecutive years could be identified. 56 pairs matched the inclusion criteria: Combined diagnoses count ≥10, count difference ≥ 50% and average CCL difference ≤ -1 or ≥ 1.

From the included diagnoses pairs, 35 could be identified as potentially motivated by changes in reimbursement. The identified 35 pairs make up 1% of all semantically identical pairs in chapter IX included in the analyzed data set.

4. Discussion

In this paper we describe two potential sources of errors while using ICD-10-GM codes from a DRG dataset in a secondary use context: Changing classification systems and under or over estimation of patient counts due to reimbursement motivated classification of patients.

An impact on precision and recall of changes in the ICD-10-GM classification could be observed for the examples described in this paper. Querying the dataset with ICD codes from matching ICD-10-GM releases could increase precision and recall. The chosen examples from the ICD-10-GM 2014 to 2015 update represent a worst case scenario for retrieval: Identical ICD codes with different semantics. With the open source programming library MEDICATS corresponding ICD codes can easily be retrieved. We are currently working on a web service implementation of MEDICATS, which should be available in the summer of 2016.

To measure possible impacts of reimbursement changes on classifications we used Complication and Comorbidity Levels (CCL) as surrogate criteria. Only 1% of all semantically identical ICD diagnoses pairs from different releases included in the analyses could be observed with a significant change in CCL and corresponding significant change in classification frequency.

A disadvantage of using the CCL values as surrogates is the lack of huge changes in the CCL matrix of diagnosis in the last years, making it harder to prove a correlation between increasing or decreasing CCL and the coding rate of specific ICDs.

Our analysis focuses on ICD codes from the chapter "Diseases of the circulatory system" and as such might not be transferable to other ICD code chapters. Using only data from a single department introduces another bias.

A different approach might use diagnosis related additional reimbursement fees. For example, the ICD D68.4 is one of the diagnoses required for getting additional fees. Based on the calculation data, published by the InEK Institute annually, an increase of coding rates of this ICD from 22,38% in 2012 to 34,74% 2013 could be shown for the DRG F05Z [6][7].

In conclusion the first hypothesis *semantically identical chapters with different ICD codes do not affect precision and recall* could be disproven, while the second hypothesis *classification is not affected by changes in reimbursement* could not be ruled out.

To further validate our results other ICD-10-GM chapters should be analyzed using the described methodology in a hospital wide data set. The analysis of OPS data from the department of cardiology is work in progress.

References

[1] A. Boeck Jensen, P.L. Moseley, T.I. Oprea, S. Gade Ellesøe, R. Eriksson, H. Schmock, P. Bjødstrup Jensen, L Juhl Jensen, S Brunaka, Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients, *Nature communications* **5** (2014).

[2] S. Reaa, J. Pathakb, G. Savovac, T.A. Onikid, L. Westberge, C.E. Beebeb, C. Taob, C.G. Parkera, P.J. Hauga, S.M. Huffd, C.G. Chuteb, Building a robust, scalable and standards-driven infrastructure for secondary use of EHR data: The SHARPn project, *Journal of Biomedical Informatics* **4** (2012), 763–771.

[3] M. Dugas, M. Lange, C. Müller-Tidow, P. Kirchhof, HU. Prokosch, Routine data from hospital information systems can support patient recruitment for clinical studies, *Clinical Trials* **7** (2010), 183-189.

[4] C.D. Manning, P. Raghavan, H. Schütze, Introduction to Information Retrieval, Cambridge University Press, (2008), 154ff.

[5] InEK, German Diagnosis Related Groups Version 2016, Definitionshandbuch Kompaktversion, Band 3, Anhang C, 1078-1083. German.

[6] InEK, Report-Browser 2013/2015, G-drg.de (2016), InEK GmbH, Online: http://www.g-drg.de/cms/G-DRG-System_2015/Abschlussbericht_zur_Weiterentwicklung_des_G-DRG-Systems_und_Report_Browser/Report-Browser_2013_2015 last accessed on (last visited 2016-03-12). German.

[7] InEK, Report-Browser 2012/2014, G-drg.de (2016), InEK GmbH, Online: http://www.g-drg.de/cms/Archiv/DRG_Systemjahr_2014_Datenjahr_2012#sm7 (last visited 2016-03-12). German.