Relations of Morbidity Related Groups (MRG), ICD-10 Codes and Age and Gender Structure in Outpatient Treatment

Reinhard Schuster¹, Thomas Ostermann², Marc Heidbreder³ and Timo Emcke⁴

¹Chair of Department of Health Economics, Epidemiology and Medical Informatics, Medical Advisory Board of Statutory Health Insurance in Northern Germany, Katharinenstraße 11, 23554 Lübeck, Germany
²Chair of Research Methodology and Statistics in Psychology, Witten/Herdecke University, Alfred-Herrhausen-Straße 50, 58448 Witten, Germany
³Medical Advisory Board of Statutory Health Insurance in Northern Germany, Katharinenstraße 11, 23554 Lübeck, Germany
⁴Chair of Department of Prescription Analysis, Association of Statutory Health Insurance Physicians, Bismarckallee 1-6, 23812 Bad Segeberg, Germany

Keywords: Morbidity Related Groups (MRG), Outpatient Treatment, ATC Classification System, International Classification of Diseases (ICD-10), Determination of Main Diagnoses, Distance Structure.

Abstract: A patient’s (basic) Morbidity Related Group (MRG) is defined by the drug class (first four characters of the international Anatomic Therapeutic Chemical [ATC] Classification System) with the highest costs per quarter with respect to a physician. The morbidity of a patient is thereby represented by the drug most important economically. We consider the relation of those case groups with diagnoses (ICD-10-GM) on the individual and group level. In analogy to the DRG Systems (Diagnosis-related group) a degree of severity with respect to age, multimorbidity and treatment intensity is defined. We compare multimorbidity and age structures of MRGs and ICD-10 using a distance measure given by the fraction of patients with respect to their MRG and ICD-10. Main diagnoses like in hospital treatment are not given in outpatient care. MRG classification data can be used in order to algorithmically construct an outpatient care equivalent. Individual MRG components as points in a vector space can be used to determine the „biological age“ of groups of individuals with respect to in- or decreased morbidity.

1 INTRODUCTION

In the field of outpatient health services research there is an increasing demand of health policy makers for new tools optimizing patient centered care and for perspective changing information. Such tools enable both health care providers and statutory health insurances to further improve health care services leading to increased benefits for individual patients.

In the past the analysis of prescribing patterns was either done by simply counting for the prescribed remedies (i.e. by means of connecting to practice systems of the physicians cf. (Jeschke, E., Ostermann, T., Vollmar, HC, Tabali, M., Matthes, H., 2012)) or by making use of samples of secondary data. However, both approaches lack of a sustainable modeling approach. because of limitations in the underlying data.

Recently, we described the concept of Morbidity Related Groups (MRG) in order to determine a main drug prescription class for patients with respect to a physician on a quarterly basis cf. (Schuster, R., 2015; Schuster, R., Emcke, T., v. Arnstedt, E., Heidbreder, M., 2016). This concept was constructed in analogy to the Diagnosis Related Groups (DRG) in the hospital setting which are primarily based on diagnoses cf. (InEK, 2016). Prescription analysis therefore utilizes all five resolution levels of the International Anatomic Therapeutic Chemical (ATC) Classification System. Interaction effects, treatment intensities and changes in treatment modalities can be chronologically interconnected by using prescription dates.

While within that mere ATC framework the patient level is of minor importance, the MRG setting takes into account the individual level by looking for the group with the highest drug costs on the third level ATC (four characters) within a quarter for each consulted physician by a certain patient.

322

Schuster, R., Ostermann, T., Heidbreder, M. and Emcke, T.
Relations of Morbidity Related Groups (MRG), ICD-10 Codes and Age and Gender Structure in Outpatient Treatment.
DOI: 10.5220/0006544303220328
In Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2018) - Volume 5: HEALTHINF, pages 322-328
Copyright © 2018 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved
this group is strongly related to the morbidity of the patient hence leading to the term „Morbidity Related Group“. Thereby the costs serve as a proxy for the relative importance of applied drugs. The example given by the Tables 1 and 2 list prescription data of a diabetes patient belonging to the basic group A10A (Insulins and analogues) with total patient cost of 1,154.16 €.

Table 1: Example of the (basic) MRG determination with drug details.

<table>
<thead>
<tr>
<th>cost</th>
<th>nr</th>
<th>ATC</th>
<th>substance</th>
<th>drug</th>
<th>amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>320.74</td>
<td>1</td>
<td>B01A</td>
<td>Rivaroxaban</td>
<td>XARELTO 15 mg</td>
<td>98</td>
</tr>
<tr>
<td>272.61</td>
<td>1</td>
<td>N06A</td>
<td>Duloxetine</td>
<td>CYMVALTA 60 mg</td>
<td>98</td>
</tr>
<tr>
<td>248.02</td>
<td>1</td>
<td>A10A</td>
<td>Insulin Lispro</td>
<td>LIPROLOG Mix 25</td>
<td>10X3</td>
</tr>
<tr>
<td>159.39</td>
<td>1</td>
<td>N02A</td>
<td>Oxycodone</td>
<td>TARGIN 5 mg</td>
<td>100</td>
</tr>
<tr>
<td>124.01</td>
<td>1</td>
<td>A10A</td>
<td>Insulin Lispro</td>
<td>LIPROLOG Mix 50</td>
<td>10X3</td>
</tr>
<tr>
<td>15.41</td>
<td>1</td>
<td>B03A</td>
<td>Levothyroxine Sodium</td>
<td>L-THYROX HXAL 125</td>
<td>100</td>
</tr>
<tr>
<td>13.98</td>
<td>1</td>
<td>C07A</td>
<td>Nebivolol</td>
<td>NEBIVOLOL Glenmark 5 mg</td>
<td>100</td>
</tr>
</tbody>
</table>

The drugs falling in the third level ATC A10A are the most expensive.

Table 2: Example of (basic) MRG determination with ATC4 substance groups.

<table>
<thead>
<tr>
<th>cost</th>
<th>ATC4</th>
<th>substance group</th>
</tr>
</thead>
<tbody>
<tr>
<td>372.03</td>
<td>A10A</td>
<td>insulins and analogues</td>
</tr>
<tr>
<td>320.74</td>
<td>B01A</td>
<td>antithrombotic agents</td>
</tr>
<tr>
<td>272.61</td>
<td>N06A</td>
<td>antidepressants</td>
</tr>
<tr>
<td>159.39</td>
<td>N02A</td>
<td>analgesics</td>
</tr>
<tr>
<td>15.41</td>
<td>H03A</td>
<td>thyroid preparations</td>
</tr>
<tr>
<td>13.98</td>
<td>C07A</td>
<td>beta blocking agents</td>
</tr>
</tbody>
</table>

In the quarter considered the patient is assigned to the MRG A10A.

Out of the various diagnoses in hospital treatment a patient specific main diagnosis is determined by the admitting physician. Today, main diagnoses are unknown in outpatient treatment. Only a rough differentiation between long term and acute diagnoses is documented with respect to quarter and physician. For example patients of above 80 years have on average more than ten diagnoses (three-character level of International Statistical Classification of Diseases and Related Health Problems [ICD], current version: ICD-10-GM 2017).

Interestingly, the gender difference become much more important, if we consider the mean number of physicians per patient and quarter.

In this paper we algorithmically construct a main diagnosis using the MRG. The main diagnosis shall describe the central focus of the treatment for statistical purposes and not as a basis for individual therapeutic decisions. The diagnoses according to ICD-10-GM have only a quarterly resolution differentiating between actual and longterm diagnoses as well as levels of confidence. In the dataset there are no direct links between drug prescriptions and related diagnoses. Nevertheless in some cases there is a strong connection of drug prescriptions and diagnoses, i.e. an prescription of insulin implies a diagnosis of diabetes. If a necessary diagnosis is missing, the quality of documentation has to be improved. In other cases different diagnoses may cause the same prescription behavior. In addition, getting a diagnosis may not always lead to drug prescriptions. Figure 1 illustrates how multimorbidity varies with age and gender con-
sidering the mean number of diseases on the level of the three-digit ICD-10. Additionally, Figure 2 represents the mean number of physicians per patient and quarter. It has to be noted that out of ICD-10 Chapter XIV (Diseases of the genitourinary system), Chapter XV (Pregnancy, childbirth and the puerperium), Chapter XXI (Factors influencing health status and contact with health services) and Chapter XXII (Codes for special purposes) were excluded in the analysis to avoid gender asymmetry and bias caused by administration. Age and gender dependent differences in drug descriptions are considered while matching MRG and ICD-10 (standardization). After determination of a main MRG and ICD group for each patient, we get a n-dimensional vector of age dependent fractions of diseases (n as the number of MRG or ICD groups). The relationship measured by Manhattan distances of such vectors and the age is of interest. The Manhattan distances monotonically increase up to a certain age dependent distance. The reason is that there are characteristic disease profiles for each age. Differences increase if gender is included. If it would be possible to get age information from ICD or MRG vectors, we can determine the „biological age” of population subgroups. One can apply this to existing insurance or social groups.

2 MATERIAL AND METHOD

We utilized prescription and diagnosis data of the most northern federal state of Germany (Schleswig-Holstein) from quarters 3/2015 till 2/2016. The analysis is related to patients, quarters and physicians. That means, that a patient is counted as much as pairs of quarters and physicians appear. With this background there are 8,645 Million patients in the drug prescription data and 11,117 Million patients in the ICD-10 data.

The C-related programming language awk is used for the computations. The visualization was done in Mathematica by Wolfram Reasearch and Microsoft Excel.

As stated in the introduction, the basic MRG is determined by the ATC3 (four characters) with the highest costs with respect to patient, quarter and physician using prescription data. Thus, only patients with drug prescriptions can get a MRG. In analogy to the DRG system in inpatient care the basis MRG is extended by a degree of severity determined by age, multimorbidity (measured by polypharmacy) and prescription intensity.

Hence, relations between MRG and ICD-10 codes with respect to multimorbidity are of interest. In the first step we consider patient with one ATC and one ICD-10 only. The resulting pairs provide ordered lists of ICD-10 per MRG and vice versa. Although the vast majority of drugs is prescribed in the field of multimorbid patients, we can use the obtained lists for additional considerations regarding all patients.

Let \( q_{m}(a, s) \) be the fraction of patients within a certain MRG \( m \) and certain age group \( a \) in 5 year classes and a certain gender value \( s \) and \( q_{s}(a, s) \) the respective fraction within all patients with drug prescriptions. Furthermore let \( p_{m,i}(a, s) \) be the fraction of patients with a certain diagnosis \( i \) within all patients with MRG \( m \), with age and gender values \( a \) and \( s \). Then

\[
p_{1}(m, i) := p_{m,i}(\ast, \ast) = \sum_{a,s} p_{m,i}(a,s)q_{m,i}(a,s)
\]

is the fraction of patients with ICD \( i \) within the group of all patients with MRG \( m \). We compare it with the respective fraction of patients with ICD-10 \( i \) within all patients with drug prescriptions including age and gender standardization:

\[
p_{2}(m, i) := p_{m,i}^{ad}(\ast, \ast) = \sum_{a,s} p_{s,i}(a,s)q_{m,i}(a,s).
\]

Without age standardization we get the fraction of patients with ICD \( i \) as

\[
p_{3}(\ast, i) := p_{s,i}(\ast, \ast) = \sum_{a,s} p_{s,i}(a,s)q_{s,i}(a,s).
\]

The last value may be of interest if there are age related prescription restrictions with certain exceptions. The drug related grouping is done on the physician group level. Looking at medical disciplines or specialists would give different results. The research subject determines which point of view is more relevant. The algorithm is identifying diagnoses leading to a higher probability of acquiring a certain MRG. That means if a certain diagnosis \( i \) is relevant for a given MRG value \( m \), we should demand \( p_{1}(m, i) > p_{2}(m, i) \) or weaker \( p_{1}(m, i) > \min(p_{2}(m, i), p_{3}(\ast, i)) \). This restrictions strongly limit the number of diagnoses positively connected with any given MRG \( m \). The benefit of any of ICD \( i \) with respect to MRG \( m \) is measured absolutely by \( p_{1}(m, i) - p_{3}(\ast, i) \) or relatively by \( p_{1}(m, i)/p_{3}(\ast, i) \). Resulting diagnoses can be ranked by the relevance for every MRG considered. Out of all diagnoses of a patient with a certain MRG we select the highest ranking in the consideration mentioned before. If a matching diagnosis does not exist, we repeat the consideration disregarding physician groups (i.e. general practitioners, surgeons and psychiatrists). If there is no matching at all, it is likely a problem due to documentation, i.e. a prescription of insulin without coding a diagnosis of diabetes.
Let \( r(a) = (r_1(a), r_2(a), \ldots, r_n(a)) \) be a vector with components that are given by the fraction of patients with age \( a \) and MRG \( i \) (\( i = 1, \ldots, n \)) where \( n \) is given by the number of MRGs ordered for instance lexicographically. One can consider this with or without a gender restriction. For the age values \( a \) and \( b \) we consider the Manhattan distance

\[
d(a, b) = \sum_{i=1}^{n} |r_i(a) - r_i(b)|.
\]

We can consider an inversion problem. If there is given a vector of disease fractions \( s = (s_1, s_2, \ldots, s_n) \) we want to determine the respective age by

\[a = \text{Min}_d d(r(a), s)!
\]

A vector of a certain subgroup of patients with a certain social status or insurance type with given age may optimally match a vector of a different age group. This can be interpreted as a higher or lower biological age. It has been already remarked, that polypharmacy is on factor for the determination of severity levels. A alternative model can be built applying polypharmacy instead of MRGs. We consider an age dependent polypharmacy vector \( v(a) = (v_1(a), v_2(a), \ldots, v_{20}(a)) \) where the component \( v_i(a) \) describes the fraction of patients with \( i \) different drug groups (ATC3). More than 20 groups are included in the \( v_{20}(a) \) value. Again, a Manhattan distance can be defined as

\[
d_2(a, b) = \sum_{i=1}^{20} |v_i(a) - v_i(b)|.
\]

with a related inversion problem

\[a = \text{Min}_d d_2(v(a), s)!
\]

Raising the question if the MRG-based or the polypharmacy-based model is more suitable for determination of the biological age of any chosen subgroup.

3 RESULTS

For three example MRGs those diagnoses having a higher conditional probability then in the unconditional case are listed. Within the basis MRG M01A (Antiinflammatory and antirheumatic products) 33.0% of the patients (\( p_1 \)) have „dorsalgia“ (M54). In an age and gender adjusted patient group without the condition M54 only 17.8% have a M54 diagnosis (\( p_2 \)). Without age and gender adjustment we have 18.2% (\( p_3 \)).

### Table 3: Antiinflammatory and antirheumatic drugs (M01A).

<table>
<thead>
<tr>
<th>ICD</th>
<th>( p_1 )</th>
<th>( p_2 )</th>
<th>( p_3 )</th>
<th>ICD label</th>
</tr>
</thead>
<tbody>
<tr>
<td>M54</td>
<td>33.0%</td>
<td>17.8%</td>
<td>18.2%</td>
<td>Dorsalgia</td>
</tr>
<tr>
<td>M17</td>
<td>12.6%</td>
<td>10.8%</td>
<td>8.6%</td>
<td>Gonarthrosis [arthrosis of knee]</td>
</tr>
<tr>
<td>M51</td>
<td>10.9%</td>
<td>7.6%</td>
<td>7.5%</td>
<td>Other intervertebral disc disorders</td>
</tr>
</tbody>
</table>

In patients with MRG M01A (antiinflammatory and antirheumatic drugs) only three diagnoses resulted in an increase in the conditional probability. For betablocking agents the same hold true for six ICD-10. The top ranking diagnosis is I10 („essential primary hypertension“).

### Table 4: Beta blocking agents (C07A).

<table>
<thead>
<tr>
<th>ICD</th>
<th>( p_1 )</th>
<th>( p_2 )</th>
<th>( p_3 )</th>
<th>ICD label</th>
</tr>
</thead>
<tbody>
<tr>
<td>I10</td>
<td>80.3%</td>
<td>67.3%</td>
<td>61.4%</td>
<td>Essential (primary) hyperten-sion</td>
</tr>
<tr>
<td>E74</td>
<td>38.4%</td>
<td>35.7%</td>
<td>31.9%</td>
<td>Disorders of lipoprotein me-tabolism and other lipidemi-as</td>
</tr>
<tr>
<td>I25</td>
<td>28.3%</td>
<td>23.5%</td>
<td>20.7%</td>
<td>Chronic ischemic heart dis-ease</td>
</tr>
<tr>
<td>I48</td>
<td>14.9%</td>
<td>11.4%</td>
<td>10.1%</td>
<td>Atrial fibrillation and flutter</td>
</tr>
<tr>
<td>J50</td>
<td>10.7%</td>
<td>10.5%</td>
<td>9.6%</td>
<td>Heart failure</td>
</tr>
<tr>
<td>E79</td>
<td>10.6%</td>
<td>10.5%</td>
<td>9.5%</td>
<td>Disorders of purine and pyrimidine metabolism</td>
</tr>
</tbody>
</table>

The most significant diagnosis for patients with antidepressants is a F32 („depressive episode“).

### Table 5: Antidepressants (N06A).

<table>
<thead>
<tr>
<th>ICD</th>
<th>( p_1 )</th>
<th>( p_2 )</th>
<th>( p_3 )</th>
<th>ICD label</th>
</tr>
</thead>
<tbody>
<tr>
<td>F32</td>
<td>45.4%</td>
<td>14.8%</td>
<td>13.6%</td>
<td>Depressive episode</td>
</tr>
<tr>
<td>M54</td>
<td>19.6%</td>
<td>18.8%</td>
<td>18.2%</td>
<td>Dorsalgia</td>
</tr>
<tr>
<td>F41</td>
<td>14.8%</td>
<td>5.1%</td>
<td>4.4%</td>
<td>Other anxiety disorders</td>
</tr>
<tr>
<td>F45</td>
<td>13.0%</td>
<td>6.8%</td>
<td>6.2%</td>
<td>Somatoform disorders</td>
</tr>
<tr>
<td>G47</td>
<td>12.5%</td>
<td>7.3%</td>
<td>7.9%</td>
<td>Sleep disorders</td>
</tr>
<tr>
<td>R52</td>
<td>12.1%</td>
<td>8.4%</td>
<td>8.4%</td>
<td>Pain, not elsewhere classified</td>
</tr>
<tr>
<td>N40</td>
<td>11.4%</td>
<td>10.9%</td>
<td>5.0%</td>
<td>Hyperplasia of prostate</td>
</tr>
<tr>
<td>F33</td>
<td>10.9%</td>
<td>3.4%</td>
<td>3.0%</td>
<td>Recurrent depressive disorder</td>
</tr>
</tbody>
</table>

For each example MRG the model has determined corresponding a top level diagnosis, for MRG M01A (Antiinflammatory and antirheumatic producs) the diagnosis M54 („dorsalgia“), for MRG C07A (Beta blocking agents) the diagnosis I10 („essential primary hypertension“) and for MRG N06A (Antidepressants) the diagnosis F32 („depressive episode“). Figures 3 and 4 show the different age distributions on both the MRG and the ICD-10 level.
Next we compare the age distribution of the considered related MRG and ICD values (Figures 5-7).

Generally the age dependent vectors $r(a)$ determine the subgroup with corresponding age $a$. Age distances at least locally but also for the age between 25 and 95 years the Manhattan distance increases monotonically (independence of $b$ while fixing $a$ or vise versa respectively).

Results again can be applied to analyze if the bio-
logical age of a subgroup of patients (i.e., with a certain social status) is better adapted than calendaric age (Figure 9). We observe special local distance maxima for childhood and adolescence.

Fixing one side of distance measurement (setting parameter $b$ to 40, 60 and 80 years), we observe monotonic behaviour of $d(a, b)$ for $a \in [20, b]$ and $a \in [b, 95]$. For $b \in [25, 95]$ with a local maximum for $a = 19$.

**Figure 9:** Manhatten distance of MRG-fractions for patients of age 40, 60 and 80.

A different curve is observed for $b = 14$ and $b = 19$ years. Regarding patients with $b = 14$ there is only a clear distance information to other patients with $a \in [2, 19]$.

**Figure 10:** Manhatten distance of MRG-fractions for patients of age 14, 19.

At least for the ages between 30 and 90 we observe similar results for age dependent vectors $v(a)$ regarding polypharmacy or MRGs (Figure 11).

In large areas we detect age sharply. This allows for comparison of models with and without gender components.

**Figure 11:** Manhatten distance of multimorbidity with respect to age.

### 4 CONCLUSIONS

The MRG determines an unique type of patient based on drug prescription data labeled by a drug group. Furthermore we can construct another unique type using the number of prescribed drug groups (polypharmacy). If we want to analyze diagnosis structures in relation to age, gender, geographical regions or social status, a well defined patient type might be useful. Reversely, starting with a special diagnosis and asking for the probability of getting a special drug or drug group again with age and gender standardization might be of interest. The presented modeling approach can be applied in both directions. This flexibility offers a wide range of applications especially when patient orientation is necessary for the development of new forms of care. The need of an individualized medicine in certain patient subgroup can also be met and synchronized with the present risk adjustment scheme in the German statutory health insurance. This risk adjustment scheme might also profit by redefinition of patient groups and underlying parameters.

### REFERENCES


