THE GAP-DRG MODEL: SIMULATION OF OUTPATIENT CARE FOR COMPARISON OF DIFFERENT REIMBURSEMENT SCHEMES

Patrick Einzinger
Niki Popper
dwh Simulation Services
Neustiftgasse 57–59
A-1070 Vienna, AUSTRIA

Felix Breitenecker
Vienna University of Technology
Wiedner Hauptstraße 8–10
A-1040 Vienna, AUSTRIA

Nina Pfeffer
Reinhard Jung
Gottfried Endel
Main Association of Austrian Social Security Institutions
Kundmanngasse 21
A-1031 Vienna, AUSTRIA

ABSTRACT
In health care the reimbursement of medical providers is an important topic and can influence the overall outcome. We present the agent-based GAP-DRG model, which allows a comparison of reimbursement schemes in outpatient care. It models patients and medical providers as agents. In the simulation patients develop medical problems (i.e., diseases) and a need for medical services. The reimbursement system receives information on the patients’ visits via its generic interface, which facilitates an easy replacement. We describe the assumptions of the model in detail and show how it makes extensive use of available Austrian routine care data for its parameterization. The model design is optimized for utilizing as much of these data as possible. However, many assumptions have to be simplifications. Further work and detailed comparisons with health care data will provide insight on which assumptions are valid descriptions of the real process.

1 INTRODUCTION
The benefits of different reimbursement schemes are a prominent topic in health services research, because how physicians get paid influences their treatment decisions. Separate payment for every single medical service can lead to an expansion of services for patients who would need less (creaming), while physicians with a lump-sum reimbursement might treat more economically but could cut necessary services, especially for costly patients (skimming) (Ellis 1998).

We present an agent-based simulation model that aims at the comparison of different reimbursement schemes for the Austrian outpatient health care sector. It utilizes pseudonymous routine data from practically all insured persons of the Austrian public health insurances. “Pseudonymous” means that, in order to guarantee protection of privacy, the data are linked to individual patients by an artificially generated patient identification number (the pseudonym) and not by the original social security number. The data include the time span of two years (2006 and 2007) and contain medical services in the outpatient sector, drug
Einzinger, Popper, Pfeffer, Jung, Endel, and Breitenecker

data, sick-leaves and hospital stays. The model construction was part of the “General Approach for Patient-oriented Ambulant Diagnosis Related Groups” (GAP-DRG) project (Endel 2011).

One particular objective of the project was the analysis of DRG-like reimbursement schemes. Fetter et al. (1980) developed the original DRG system for the inpatient sector. Also in the 1980s, the first systems for ambulatory care were introduced (Fetter et al. 1984). However, in the outpatient sector it is difficult to assign unique diagnoses to visits, and patients can have long or even chronic episodes of disease (Rogerson 1983). The approach in the project was that episodes last for one quarter of a year, and those diagnoses for which a patient receives medical services during a quarter contribute to his or her classification.

The general idea of the model is that reimbursement is the result of patients consuming medical services from medical providers because of their medical problems (i.e., diseases). It roughly follows the proposed structure of Krol and Reich (1999), which is object-oriented and also contains the classes medical provider and patient. Figure 1 shows the general idea of the model diagrammatically and emphasizes that different reimbursement systems could be the basis for reimbursement. The general idea translates to two types of agents, the patients and the medical providers. The behavior in the model splits into five modules:

1. epidemiology
2. service need
3. provider utilization
4. service provision
5. reimbursement

Figure 1: This diagram shows our viewpoint of the problem. Patients develop medical problems, which— together with the treatment decisions of medical providers—lead to the provision of medical services. These are the basis for the reimbursement from the payer. In Austria’s outpatient care these are the public health insurance providers (Hofmarcher and Rack 2006). Reimbursement can vary based on the reimbursement system.

First, the model has to reproduce the epidemiology of the included medical problems accurately, because most health care services are treatments for diseases (up to now it does not include other services such as medical check-ups). The structure of the model allows each patient to have multiple medical problems at the same time. At the moment it incorporates seven chronic diseases and their parameterization. The selection of chronic diseases is based on their importance (costs and number of cases) in the outpatient sector. A future model version might include acute medical problems, but there are open questions about how to handle them in the framework and they are hard to identify in the routine data.

Each medical problem creates a service need for every quarter of a year in the form of a list of medical services and corresponding frequencies. Patients then try to receive these services in the required quantities from the medical providers (provider utilization). These will perform the services that they have in their portfolios (service provision) and send reimbursement claims to the social insurance. Reimbursement takes
Patients can receive their medical services in one or more visits to providers. From this we distinguish a case: A patient who had at least one visit to a certain medical provider during a quarter of a year counts as one case for this provider and this quarter. The reason for this is that most health insurances in Austria have the quarter of a year as their accounting period (Hofmarcher and Rack 2006), and many reimbursement systems pay a lump sum for patients who visit a provider at least once in a quarter. A patient may represent several cases in one quarter, but only with multiple providers.

2 GENERAL MODEL STRUCTURE

The GAP-DRG model is agent-based. Its implementation is object-oriented (the current implementation uses the simulation software AnyLogic, version 6.9.0, which utilizes the Java programming language). Figure 2 shows the hierarchical structure of objects. All objects on the second level are embedded in the root object Main. MedicalProvider and Patient are agent types, and the simulation contains potentially many of these objects. The reimbursement system (ReimbursementSystem) and the so-called health market (HealthMarket) are single objects. Additionally, patients can develop medical problems (MedicalProblem), which are embedded in the patient and thus represent a third hierarchical level.

![Figure 2](image)

Figure 2: The hierarchical structure of the GAP-DRG model. Everything is embedded in the root object Main. It contains the agents of the two types, medical providers (MedicalProvider) and patients (Patient), as well as one reimbursement system (ReimbursementSystem) and one health market (HealthMarket) object. Patients can develop diseases, which are of the type MedicalProblem.

In the simulation, patients should not consult all medical providers equally likely, because in reality they have a preference for small distances (Finlayson et al. 1999), and therefore the model incorporates a simple quadratic continuous space whose area is approximately equivalent to the area of Austria. This means that its side length equals 290 kilometers times the square root of a scaling factor.

The scaling factor is necessary because memory requirements put an upper bound on the number of patients in the model. So far it was not possible to simulate the whole Austrian population, which would have been about 8.3 million inhabitants on January 1st, 2006 (Statistics Austria 2012), and therefore we performed simulations with a scaling factor of 0.1 (or about 830,000 patients). The number of medical providers scales accordingly. The side length of the space has to scale with the square root of the scaling factor in order to guarantee the right density of medical providers.

The health market handles the patients’ search for providers. Patients can send their service need, caused by their medical problems, to the health market and receive a list of medical providers that fit to their service need. For this task the health market stores the location of every provider. It chooses only
providers that are within a certain range of the patient and that have acceptable waiting times (i.e., not too many patients wait to visit the provider).

ReimbursementSystem is an abstract type. It describes the generic interface that every reimbursement system must have. This includes methods for registering provider visits—medical providers send their own ID, the ID of the patient, the date of the visit and all possible reimbursement claims to the reimbursement system—and for invoicing reimbursement claims. Every implementation of ReimbursementSystem represents a different reimbursement scheme.

One example is ActualReimbursementSystem, which implements a reimbursement system that resembles the current system in Austria. However, it is an artificial system, because the Austrian social insurances use varying reimbursement schemes. There exists a Meta-tariff catalog (Metahonorarordnung), which maps all services that exist in the tariff catalogs of the public health insurances to a common scheme. The implementation of the actual reimbursement system uses this catalog, which is a preferred data source according to guidelines for health economic evaluation in Austria (Walter and Zehetmayr 2006). The reimbursement consists of a lump sum, which is paid on a per-contact base and covers all basic services, and a fee-for-service part (Hofmarcher and Rack 2006). Other reimbursement systems that have capital payment substitute this fee-for-service part with further lump sums.

This architecture facilitates easy substitution of the reimbursement system in the model. Scenarios can use different reimbursement systems without changing any other part of the model structure. This is only possible because the medical providers send all the information that any conceivable reimbursement system could use, even if the reimbursement system in the scenario does not need it. A scheme based on per capita payment, for example, ignores single services for payment, but it receives this information nevertheless.

The parameterization of the model assumes the years 2006 and 2007 as the simulation time interval. Up to now only data from this time span were available to us.

3 MODEL ASSUMPTIONS

Simulation models are almost always simplifications of real systems. This manifests itself in implicit or explicit assumptions. The GAP-DRG model describes a large part of the health care system, and it maps service utilization on a single service level. Both patients and medical providers exhibit human behavior, which is difficult to model accurately. All this necessitates simplifying assumptions in some parts of the model. The following subsections list them in detail.

3.1 Demography and Parameters of Patients

The number of patients at simulation start corresponds to the resident population of Austria at January 1st, 2006 (Statistics Austria 2012). There are 40 demographic groups of patients, separated by age and sex, in the model in order to accomplish this: In the interval from 5 to 89, each group spans over five years (5–9, 10–14, . . . , 85–89). Additionally, there are the groups of children under one year and between one and four years old as well as the group 90+. The simulation initializes exactly as many patients in each group at simulation start as were part of the resident population.

Note that the insured persons do not exactly correspond to the resident population. It is compulsory for working and retired people to be insured, and also their relatives (spouses and children) are then co-insured (Hofmarcher and Rack 2006). However, there are some exceptions: For example physicians and lawyers are allowed to opt-out of the public health insurance. Additionally, foreigners who work in Austria are also insured but do not have to be in the resident population. Nevertheless, we decided to use the resident population as it facilitates comparison with other publications and statistics.

The population is dynamic, as it changes through births and deaths. In order to avoid a lot of additional events that are spread out over the entire time span, which would potentially decrease performance, all births and deaths take place only once in a year (on July 1st).
The number of births in the simulation equals the absolute number of births in Austria times the scaling factor. For example there were 38,169 female newborns in 2006, and exactly this number of new female patients adds to the simulation on July 1st, 2006.

Patients have a probability of death that corresponds to their sex and age group and depends on their medical problems. Once in a year the simulation tests for each patient if he or she survives a further year. All other patients die and leave the model.

Additionally to their sex and year of birth all patients have a maximum willingness to wait for a provider visit as a parameter (maxDaysToWait). It defaults to 10 days for all patients.

3.2 Medical Providers

Medical providers have a type, which is a parameter that corresponds to their specialty. 16 different specialties are included. Among them are, for example, general practitioners, internists, ophthalmologists, paediatricians, and laboratories. Hospitals are not included.

We discussed finer classifications of providers within specialties, for example according to treatment style, but this is not yet implemented. However, medical providers of a specialty do differ because they have different service portfolios.

The service portfolio is the set of medical services that the provider can provide. For each service and each specialty there is a probability that a provider of this specialty has the service in his or her portfolio included in the model. Based on this probability the simulation determines the service portfolio of each medical provider at simulation startup.

The model should prevent single providers from treating unrealistically high patient numbers. All providers have patient queues for this purpose. Patients cannot make a visit instantly but have to register to a queue, which providers process in a first in, first out order.

The parameter workingFor of the MedicalProvider type specifies how many cases a provider might treat in one day. It defaults to 20 cases per day on weekdays (except Saturday and Sunday). General practitioners can also treat two cases per day on weekends. This is an average value which accounts for the fact that general practitioners have weekend services on some weekends. Laboratories are an exception: They can process 500 cases on each weekday except Saturday and Sunday.

The numbers above are reasonable but not derived from data. Sensitivity analysis showed that their influence is not negligible, so more accurate estimates are desirable for the future. Until then scenarios should be simulated with different values for these parameters. Note also that all weeks are equal: Providers do not make explicit holidays. The parameter values have to account for that and thus represent averages over the whole year.

3.3 Medical Services

The medical services that are included in the model correspond to the services of the Meta-tariff catalog. Examples are intravenous injection and electrocardiography. The Meta-tariff catalog does not define prices for the services, but instead they derive from the mappings to the service catalogs of the health insurances. The tariff is thus given by

\[
tariff = \frac{cost_{2006} + cost_{2007}}{frequency_{2006} + frequency_{2007}},
\]

where \(cost_{year}\) is the total costs that the health insurances spent for an equivalent service in the corresponding year, and \(frequency_{year}\) is the total quantity of equivalent services that were provided. Therefore, it is a weighted average of the tariffs of those services that are equivalent to the service in the Meta-tariff catalog.

3.4 Medical Problems

At the moment the model incorporates seven chronic medical problems. A medical problem corresponds to certain codes in the International Classification of Diseases (ICD) (DIMDI 2012). Table 1 shows the
included medical problems with their ICD-9 codes. Note that other, more refined classifications than ICD-9 (and of course the more up-to-date ICD-10 classification) might be beneficial, but we had to use ICD-9 since some of the diagnosis data from hospitals and sick leaves are encoded in it, and thus Weisser et al. (2010) also used ICD-9 for their results. These results are an input for our analyses (see below). A unique translation from ICD-10 to ICD-9 is possible, but not vice versa.

### Table 1: Medical Problems that are included in the model and their corresponding ICD-9 codes that were used for data analyses. Some analyses used the ICD group level.

<table>
<thead>
<tr>
<th>Medical Problem</th>
<th>ICD-9 codes</th>
<th>ICD-9 group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes</td>
<td>250</td>
<td>250–259</td>
</tr>
<tr>
<td>Hypertension</td>
<td>401–405</td>
<td>401–405</td>
</tr>
<tr>
<td>Coronary Heart Disease</td>
<td>410–414</td>
<td>410–414</td>
</tr>
<tr>
<td>Asthma and COPD</td>
<td>491–493, 495, 496</td>
<td>490–496</td>
</tr>
<tr>
<td>Degenerative Joint Disease</td>
<td>710–719</td>
<td>710–719</td>
</tr>
<tr>
<td>Chronic Back Pain</td>
<td>720–724</td>
<td>720–724</td>
</tr>
<tr>
<td>Incontinence</td>
<td>618, 625, 629</td>
<td>617–629</td>
</tr>
</tbody>
</table>

Medical problems generate a need for medical services in the patient. Need factors such as disease severity are one of the main reasons why patients consult physicians (de Boer, Wijker, and de Haes 1997). Ideally a model would describe the course of a disease and the treatment pathway the patient takes, i.e., the sequence of provider visits. However, this would require expert knowledge and register or trial data on all incorporated diseases.

The approach taken in our model is thus a different one: Based on the results of a previous project it is possible to assign a probability for having a certain medical problem to patients in routine care data (Weisser et al. 2010). (In Austria, data on hospital stays and sick leaves include the diagnosis, but data in outpatient care, apart from sick leaves, do not.) This assignment is based on the drug prescriptions of the patient. Patients who got insulin would for example have a high probability of having diabetes. In a further step, a list of common medical services (a service bundle) was generated for each included medical problem with the following scheme:

1. Select all patients who have a positive probability for a diagnose of the medical problem, according to their medication and the assignment probabilities (described above).
2. Count the frequency of every medical service that the patients received in a time span of 92 days before and after a diagnose.
3. Include the most frequent services in the service bundle. In order to do this, determine a cutoff, which is based on a statistical procedure that tries to find out significant jumps in the frequency distribution of the services. Thus, the frequencies of excluded services are significantly lower than the frequencies of the services in the bundle.
4. For each service in the bundle, calculate the corresponding frequency distribution for the period of a quarter of a year.

In the model each medical problem of a patient samples its service need for each quarter of a year from these frequency distributions.

### 3.5 Prevalence and Incidence

Two measures determine how many patients have a certain medical problem: The prevalence at the start of simulation and the incidence, i.e., the onset of new diseases. The former is a proportion and the latter is given as incidence rates (Miller and Homan 1994).
In the model the (initial) prevalence is the probability for a patient of having the disease at simulation start. Therefore, the simulation initially performs a probability experiment for each patient and each medical problem.

The incidence rates determine how long it takes—on average—for a patient to develop a medical problem that he or she does not have already. The time to development of the disease is exponentially distributed and the incidence rate is the parameter of the distribution. All incidence rates are independent from the medical problems that a patient already has, which means that morbid patients do not have a higher risk of developing further diseases than healthy patients. This is probably not realistic for some of the diseases. Hypertension, for example, is a known risk factor for coronary heart diseases. However, it is a difficult task to evaluate these dependencies statistically, and the main objective of the model is the comparison of different reimbursement systems and not the exact reproduction of the epidemiology.

3.6 Service Need, Provider Search and Service Consumption

Once in a quarter of a year all medical problems of a patient determine their service need (see Subsection 3.4). Then the patient sends a request to the health market and attaches every service from any service need of his or her medical problems. The health market returns at most one medical provider per specialty, based on the following criteria:

- The distance between all returned providers and the patient is not more than
  - 5 kilometers for general practitioners
  - 25 kilometers for all other specialties except laboratories, which have an unrestricted distance.
- The expected waiting time is acceptable for the patient (not more than maximumDaysToWait).
- For specialties where the patient already visits a provider and this provider fulfills the criteria above only this “preferred” provider is available.
- The returned providers cover the service need in an optimal way.

Here “in an optimal way” means that the service portfolios of the providers cover as many services of the service need as possible, and that the number of providers is minimal. Mathematically this is equivalent to a minimum set cover problem, where the set of services should be covered with a minimum number of service portfolios. The simulation approximates the best solution with a greedy algorithm (Vazirani 2003), so the provider whose service portfolio covers the most remaining services of the service need is picked iteratively.

Patients register in the queues of every provider they get returned from the health market. When it is their turn they perform an average number of visits. At every visit the provider performs each service that is in his or her portfolio and in the service need of the patient exactly one time, and each provided service subtracts from the remaining service need. At the end of the visit the provider sends the visit information to the reimbursement system.

As mentioned, patients always perform an average number of visits, which is derived from data separately for each specialty. (In general the average number of visits for the specialty is not an integer. It is then statistically decided if the patient performs a rounded up or rounded down number of visits.) Probably a visit frequency that is dependent on the service need and sampled from a distribution would be more realistic, but also more difficult to determine.

4 PARAMETERS

The model has several types of parameters that were parameterized from data:

- Number of patients in each group (age and sex)
- Number of medical providers in each specialty
- Prevalence and incidence for each medical problem in each group (age and sex)
• Death rates in each group for every disease combination
• Average provider contact rates
• Lump sums for basic services in the actual reimbursement system
• Service portfolio probabilities for each specialty and medical service
• Frequency distributions for each service in the service bundle of a medical problem

The demographic information was derived from national statistics (Statistics Austria 2012) and the number of medical providers from data of the Main Association of the Austrian Social Security Institutions. Statistical analyses on the main database of the project yielded the other parameters. Especially the calculations for prevalence, incidence and mortality rates ground on the statistical relationship between drug prescriptions and diagnoses (Weisser et al. 2010). The detailed methodology is described in (Einzinger, Jung, and Pfeffer 2012).

5 CONCLUSION

The model facilitates comparisons of different reimbursement systems in outpatient care. These are easily exchangeable in the proposed framework. Moreover, implemented reimbursement systems are reusable because often only slight changes are necessary.

The basis for the model are extensive routine data sources. It integrates demographic data, epidemiological parameters, service utilization and tariff data. Its structure is thus also strongly oriented towards structures of the main database that was used. In this database the data are linked to patient pseudonyms and to medical providers. These are also the most important objects (or agents) in the model.

The model is not only useful for the comparison of reimbursement systems but also for testing assumptions. Future work will include validation of individual aspects, as for example the provider utilization that arises from the provider search algorithm described in Subsection 3.6. This should lead to new insights on how the health care system works and which models can describe it accurately. It will also show which assumptions might be oversimplifications.

There are several limitations of the model. So far it includes only a limited number of chronic diseases and not the full spectrum of diseases. Treatments do not influence the course of diseases. Additionally, providers do not adapt or optimize their behavior based on reimbursement, utilization, etc.

Lessons learned from the project include the following: The model building process can reveal in which areas there is a lack of knowledge. It is therefore in itself valuable as an input for further empirical research. What also became clear is that a lot of simplifying assumptions have to be made. Furthermore, concerning output of the model there is a conflict between level of detail and ease of interpretation. The model is capable of producing output that is nearly as complex as the routine care data. Thus, it is not always straightforward to interpret results. One last aspect is that the model can also help with formulations of new reimbursement systems, because some effects that are logically inherent but not immediately obvious show quickly in simulations.

Future work will include further tests of the model, especially its correspondence with available data. So far we concentrated on verification and on validation of the epidemiology. It showed that the parameterization with routine care data leads to results that are close to the overall prevalence of the included diseases in a national health survey (Statistics Austria 2007).

Previous scenarios only considered changes in the reimbursement system. Future scenarios might also include organizational changes, which could for example require patients to take alternative treatment pathways.

Finally, several expansions of the model are conceivable. Among them are the incorporation of dynamic provider behavior and of further diseases as well as further medical services. Possibly an artificial “residual” medical problem could represent diseases that the model does not explicitly include.
REFERENCES


AUTHOR BIOGRAPHIES

PATRICK EINZINGER is researcher at dwh Simulation Services. He received his MSc. in Technical Mathematics from Vienna University of Technology. His main research interests are system dynamics, agent-based modeling and simulation, and applications to health services research. His email address is patrick.einzinger@dwh.at.
NIKI POPPER is CEO of dwh Simulation Services and teaches mathematical modelling and simulation at VUT (Vienna University of Technology). The mathematician has his expertise in modelling and simulation in health care and coordination of research projects. He coordinated various national research projects in this area, funded f.e. by the the Austrian Research Promotion Agency, the city of Vienna or VUT. He is specialist in methodological comparison and hybrid combination of models. He coordinated cooperative projects with partners like VUT, Medical University of Vienna, Ludwig Boltzmann Gesellschaft, Main Association of Austrian Social Security Institutions, UMIT Private University, Fraunhofer Gesellschaft. He was chair of various special sessions at international conferences and has published articles in the areas of theory of modelling & simulation and modelling & simulation in HTA/EBM/Health Economy. Niki Popper is editor of special issues of Simulation Notes Europe. His email address is niki.popper@dwh.at.

NINA PFEFFER works at the Main Association of the Austrian Social Security Institutions, in the Department for Evidence Based Health Care. As Deputy Head of Department, she is in particular responsible for the Unit of Health Economics. Her field of work includes payment methods for health services and health services research. Her email address is nina.pfeffer@hvb.sozvers.at.

REINHARD JUNG works at the Main Association of the Austrian Social Security Institutions, in the Department for Evidence Based Health Care. His field of work includes payment methods for health services, health services research and modeling approaches in health care. He is responsible for the development of various mathematic models, explaining phenomenons in the Austrian health care system. His email address is reinhard.jung@hvb.sozvers.at.

GOTTFRIED ENDEL works at the Main Association of the Austrian Social Security Institutions, in the Department for Evidence Based Health Care. He is responsible for the EBM-HTA team. His field of work includes HTA, health system research, e-health, patient safety and other questions related to the Austrian social system. He is lead partner representative in EUNetHTA, HTAi, iHEA, ISPOR, G-I-N member and board member in several Austrian associations. His email address is gottfried.endel@hvb.sozvers.at.

FELIX BREITENECKER studied “Applied Mathematics” and acts as professor for Mathematical Modelling and Simulation at Vienna University of Technology. He covers a broad research area, from mathematical modelling to simulator development, from DES via numerical mathematics to symbolic computation, from biomedical and mechanical simulation to process simulation. He is active in various simulation societies: president and past president of EUROSIM since 1992, board member and president of the German Simulation Society ASIM, member of INFORMS, SCS, etc. He has published about 350 scientific publications, and he is author/editor of 10 books. He is member of the editorial board of several journals and conference proceedings series, and since 1995 he is editor in chief of the journal Simulation Notes Europe (formerly Simulation News Europe). His email address is felix.breitenecker@tuwien.ac.at.