Charles University

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Improvement of risk adjustment for health insurance companies in the Czech Republic - compensation of costs of patients with renal failure

Author: **Bc. Magdalena Škodová**

Supervisor: **PhDr. Jana Votápková Ph.D.**

Academic Year: **2019/2020**

Declaration of Authorship

The author hereby declares that she compiled this thesis independently; using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, July 27, 2020

Signature

Acknowledgments

I would like to thank to my supervisor PhDr. Jana Votápková Ph.D. for her dedicated supervision, guidance, and comments, which largely contributed to shaping of the thesis and considerably influenced the topic.

I also wish to thank to MUDr. Pavel Hroboň, M.S. for assistance with the choice of the topic, provision of the dataset and for being always helpful and inspirational throughout my studies. I am also grateful to Mgr. Petra Kučová for help and suggestions that improved the quality of the thesis. Special thanks go to MUDr. Alena Svobodová for her professional insight on complex medical issues and much appreciated advice.

Last but not least I wish to thank to my parents and brothers, who supported me with their unconditional love, patience, and faith throughout my whole studies and who made my studies possible in the first place.

Abstract

Risk adjustment models are used to predict health care costs of insurees and represent an important part of mechanisms for redistribution of funds among insurance companies. In the Czech Republic, pharmacy-based cost groups (PCGs) were introduced into the risk adjustment model in 2018, reflecting the costs of chronic diseases in addition to age and gender. The thesis reviews the model for the most expensive chronic disease – renal failure. Using the sample of General Health Insurance fund (GHI) insurees reported with typical health care consumption for kidney disease in years 2015-2018, we tested the current model and subsequently modified the classification criteria for PCG "renal failure". The classification based on the number of dialysis procedures proved to be much better indicator of costs than the currently used consumption of typical drugs. The incorporation of dialysis-based approach into the PCG model improved the explained variation from 26 % to 49 %, and the predictive power increased substantially. The study suggests improvements of the Czech risk adjustment model and proposes a fairer fund redistribution among insurance companies, while no additional data collection is needed.

JEL Classification	I13, I18
Keywords	Fund redistribution, PCG model, renal failure, risk adjustment

Abstrakt

Model pro rizikovou úpravu predikuje náklady na zdravotní péči pojištěnců a představuje jednu z nejdůležitějších součástí mechanismu přerozdělení financí mezi zdravotními pojišťovnami. V roce 2018 byl v České republice zaveden tzv. PCG model, který kromě věku a pohlaví bere v potaz také náklady na chronická onemocnění. Tato diplomová práce ověřuje uvedený model z pohledu nejdražší chronické nemoci – renálního selhání. Za použití vzorku pojištěnců VZP, kteří měli vykázanou spotřebu zdravotní péče typickou pro onemocnění ledvin v letech 2015-2018, jsme testovali současnou verzi modelu a dodatečně stanovili další klasifikační kritéria pro zařazení pojištěnců do PCG „renální selhání“. Klasifikace, která používá pro zařazení výkony dialýzy, se ukázala jako lepší indikátor budoucích nákladů oproti doposud používané spotřebě typických léčiv. Použití přístupu založeného na výkonech dialýzy zlepšilo koeficient determinace z 26 % na 49 % a schopnost predikce nákladů se rovněž významně zlepšila. Předkládaná studie tak může napomoci ke zlepšení rizikové úpravy nákladů v ČR a přispět ke spravedlivějšímu přerozdělení financí mezi zdravotními pojišťovnami, přičemž veškerá potřebná data jsou k dispozici.

Klasifikace	I13, I18
Klíčová slova	PCG model, přerozdělení pojistného, renální selhání, úprava rizika

Bibliographic Record

ŠKODOVÁ, Magdalena. *Improvement of risk adjustment for health insurance companies in the Czech Republic - compensation of costs of patients with renal failure.* Prague, 2020. 67 pages. Master’s thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies. Supervisor PhDr. Jana Votápková Ph.D.

Contents

List of Tables

List of Figures

Acronyms

ACR Albumin-to-Creatinine Ratio

AER Albumin Excretion Rate

ADG Ambulatory Diagnostic Group

APD Automated Peritoneal Dialysis

ATC Anatomical Therapeutic Chemical

CDS Chronic Disease Score

CKD Chronic Kidney Disease

CKF Chronic Kidney/renal Failure

CPD Continuous Peritoneal Dialysis

CR the Czech Republic

CRI Chronic Renal Insufficiency

CSN Czech Society of Nephrology

DCG Diagnostic Cost Group

DDD Defined Daily Doses

eGFR estimated Glomerular Filtration Rate

ESRD End Stage Renal Disease

GDP Gross Domestic Product

GFR Glomerular Filtration Rate

GHI General Health Insurance fund (VZP -Všeobecná zdravotní pojišťovna)

GLM Generalized Linear Model

HD Hemodialysis

LR Likelihood ratio

MAPE Mean Absolute Prediction Error

MARE Mean Absolute Relative Error

MoH Ministry of Health

MPE Mean Prediction Error

OECD The Organisation for Economic Cooperation and Development

OLS Ordinary Least Squares

OOP Out-of-pocket Payments

PCG Pharmacy-based Cost Group

PD Peritoneal Dialysis

PDD Prescribed Daily Doses

Pmp Per million population

REN PCG group “renal failure”

RRT Renal Replacement Therapy

SHI Statutory Health Insurance

SÚKL Institute for Drug Control (Státní ústav pro kontrolu léčiv)

WHO World Health Organization

WLS Weighted Least Squares

Master's Thesis Proposal

Author:	Bc. Magdalena Škodová
Supervisor:	PhDr. Jana Votápková Ph.D.
Defense Planned:	September 2020

Proposed Topic:

Improvement of fund redistribution to health insurance companies in the Czech Republic - compensation of costs of patients with renal failure

Motivation:

All over the world, there has been a rising pressure on health care systems, where the resources are limited, and therefore efficient allocation is a necessity. In most countries, health care demand is rising, partly due to ageing of the population, which is usually associated with higher occurrence of chronic diseases. Most chronic diseases are not curable and therefore chronically ill people rise costs for health systems. Health care reforms are aimed to improve the efficiency of the fund redistribution that should reflect the individual health care needs as close as possible. For this purpose, prediction models are being used to estimate the expenditure of each health insurance company and thus to indicate how funds should be redistributed. Such models use available information about insurees to predict their costs. Its simpler version might be solely a demographic model which was used also in the Czech Republic until 2018, when new model of risk assessment was implemented.

The pharmacy-based cost group (PCG) model, which is used in the Czech Republic since 1.1.2018, uses both demographic data and chronic diseases. As it was shown in many studies, incorporating such information in the model substantially improves its predictive power (Lamers & Vliet, 2003, Hájíčková, 2015). PCG model has been successfully implemented also in other countries, such as Slovakia or the Netherlands. In fact, the model currently used in the Czech Republic is based on the PCG model that was first implemented in the Netherlands. For the PCG model to work, information about individual drug consumption of patients is crucial. This information is being routinely reported to insurance companies who cover the costs of treatment. The drug consumption is usually highly specific for each of the chronic diseases and thus insurees can be distributed into several groups according to the disease identified, their age and sex. Hereby, specific groups with sufficiently homogeneous expenditures are obtained and the costs for each insurance company is estimated. Funds are then redistributed to individual insurance companies fairly.

In my thesis, I would like to replicate the PCG model that is currently used in the Czech Republic and address the issues of its sensitivity and efficiency. Since PCG model is based on a linear regression, I would like to conduct sensitivity analysis and demonstrate, how precisely the model predicts the real expenditures. The thesis will be focused on one specific group of chronically ill patients, namely patients with renal failure (this corresponds to group REN used in the PCG model).

Chronic kidney disease or renal failure is a condition when the kidneys are not able to work properly (e.g. because of diabetes or high blood pressure) gradually leading to loss of its function. In later stages, the kidneys are not able to clean the blood from wastes anymore and the person is obliged to visit hospital several times a week for special treatment – dialysis, that basically replaces the kidney's function. Unless a new functioning kidney is transplanted, the patient is dependent on dialysis for the rest of his life. In spite of the fact, that renal failure is less usual than most of the other chronic diseases, it characterized by very high costs. In terms of the average costs for patient, renal failure is the most expensive among all chronic diseases (Dungl, Jandová, Kubů, Macháček, & Svoboda, 2017).
For the PCG model, consumption of typical drugs for treatment of renal failure is used, namely from anatomical therapeutic chemical (ATC) classes with codes B03X – antianemic preparations other than iron, vitamin B12 and folic acid, or V03AE – drugs for treatment of hyperkalemia and hyperphosphatemia. It should be noted, that these pharmaceuticals are only necessary complements to dialysis treatment, not the treatment of renal failure itself. For the patient to be classified as chronically ill, one has to take minimum of 181 defined daily doses (DDD) of drugs from ATC group B03X or V03AE (combination of both is also possible). However, there exists a reasonable belief that the PCG group is not defined precisely, since some of these drugs (e.g. erythropoietin) could be also used for treatment of other diseases. Therefore, the PCG methodology might mistakenly identify individuals, who do not suffer from renal failure. On the other hand, some part of the patients dependent on dialysis is believed to be omitted from the group. This might be caused by improvement of patient's condition, when less than 181 DDD of complementary drugs could be used. However, this is always temporary, since only kidney transplantation can cure renal failure definitively (Dungl et al., 2017). As a result, insurance companies might not be compensated sufficiently for these costly patients. The aim of my thesis would be to explore these inaccuracies and possibly suggest how the methodology and the PCG model could be improved to adequately compensate insurance companies for patients with renal failure.

Hypotheses:

Hypothesis #1: The PCG group REN contains patients who do not suffer from renal failure.
Hypothesis #2: The PCG group REN does not contain some patients with renal failure.
Hypothesis #3: If the PCG group REN was identified also based on the treatment procedures, not solely on drug consumption, the accuracy would improve significantly (both sensitivity and specificity).

Methodology:

For the purpose of this thesis, data from the General Health Insurance (Všeobecná zdravotní pojišťovna) Fund from the period 2012-2018 will be used. The dataset contains anonymized ID of insurees, demographic data (age, sex, place of living), data about procedures, drugs and diagnoses, and also on individual costs. Only insurees suffering from chronic kidney disease (either according to their diagnosis or drug consumption) would be included in the dataset.
The data will be used to identify the patients with renal failure according to the currently used version of the PCG model for the purpose of fund redistribution in the Czech Republic. The regression analysis will be conducted and used to estimate the risk indices that will later be used to predict the expenses. In the econometric

model, demographic characteristics will be included as independent variables as well as the dummy variables indicating the PCG group the patient belongs to. The threshold of 181 DDD of specific drugs will be used for identification of patients with renal failure. The model will be modified using different independent variables as well as their combinations. The predictive power of the model is usually demonstrated in terms of R-squared value (explained variation). Consequently, alternative identification of REN group will be employed by changing the limit of drug consumption required (less than 181 DDD) in order to verify the accuracy of the currently used version. Identification based on more criteria than solely drug consumption will be discussed as well. Finally, using the cost indices obtained in the regression model, predicted costs will be calculated and compared with the real costs.

Expected Contribution:

The challenge of the current health care systems is to find the best predictors of health expenditure to manage effectively the limited funds. While the PCG model was proved to be more precise when compared to a simple demographic model (Lamers & Vliet, 2003), it can still contain some inaccuracies and there is a space for improvement. My aim is to continue the research of Chochláčová (2018), who conducted similar analysis using data from Slovakia, focusing on patients with Hypercholesterolemia and Hájíčková (2015), who also explored the possibilities and benefits of PCG model even before it was implemented in the Czech Republic. My thesis will therefore extend their studies by using up-to-date data and by demonstrating the model on a different PCG group.
The results of the analysis will contribute to model improvement in a sense of its predictive power and accuracy of risk adjustment. Furthermore, the analysis might reveal inaccuracies in the definition of REN group, where some patients could be placed by mistake, while others are omitted. Failing to identify patients with renal failure using the drug consumption may point to further issues, such as inadequate prescription or use of drugs.

Outline:

1. Introduction
2. Literature review – models used in other countries, studies concerning the use of PCG model, its use in Czech and Slovakia
3. Methodology and data – description of the data and explanation of models used (demographic model, PCG model)
4. Econometric analysis – regression analysis, testing of different models and use of different identification criteria, costs estimation and comparison with real expenses
5. Results and discussion – discussion of results, suggestions of improvements of the current model
6. Conclusion

Core Bibliography:

1. Dungl, M., Jandová, P., Kubů, J., Macháček, T., & Svoboda, J. (2017). PCG v České republice - Historie, postupy, výsledky analýz. Monitor HC. Retrieved from https://monitorhc.cz/images/PCG_v_CR_working2_web.pdf
2. Hájíčková, T. (2015). The Pharmacy-based Cost Group Model: Application in the Czech Health Care System. 58 p. Master thesis. Charles University, Faculty

of Social Sciences, Institute of Economic Studies, Supervisor PhDr. Jana Votápková
3. Chochláčová, B. (2018) Improvement of Risk Adjustment Model to Strengthen the Competition among Health Insurance Companies in Slovakia. 65 pages. Master thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies. Supervisor MUDr. Ing. Daniel Hodyc, Ph.D.
4. Lamers, L. M., & Vliet, R. C. J. A. (2003). Health-based risk adjustment Improving the pharmacy-based cost group model to reduce gaming possibilities. The European Journal of Health Economics, 4, 107–114. https://doi.org/10.1007/s10198-002-0159-9
5. Lamers, L. M., & Vliet, R. C. J. A. (2004). The Pharmacy-based Cost Group model : validating and adjusting the classification of medications for chronic conditions to the Dutch situation. Health Policy, 68, 113–121. https://doi.org/10.1016/j.healthpol.2003.09.001
6. Van de Ven, W. P., Vliet, R. C. J. A., & Lamers, L. M. (2004). Health-adjusted premium subsidies in the netherlands. Health A ff airs, 23(3),45–55.

1 Introduction

The health care systems all over the world have gone through many reforms in last few decades as a reaction to the increasing demand for health care and rising health care costs. Most European countries introduced universal health coverage with mandatory contributions to establish solidarity and equal access to health care. In the reimbursement system, the fairness among health insurance funds is crucial for reaching the optimal efficiency and quality of provided services. For this purpose, the pooling of funds and risk adjusted redistribution have been introduced in many countries.

The pooling of collected funds is based on the idea that the healthy subsidize the sick and the young subsidize the old, in accordance with the principles of solidarity. However, in the absence of fair redistribution of pooled funds, the insurance companies with sicker population are undercompensated and might even face financial problems. Consequently, insurance companies may risk-select healthier and younger individuals since they represent lower financial risk. Although risk selection is not allowed in the Czech Republic, insurance companies might offer benefits to attract particular groups of insurees or refuse to contract with some providers. This is where risk adjustment mechanisms step in to reduce the risk selection incentives and to improve fairness in the health insurance system.

Risk adjustment mechanisms are utilized to predict the health care costs of specific groups of individuals as close as possible. First risk adjustment models implemented in European countries (including the Czech Republic) used demographic variables. As age and gender performed poorly in predicting the real costs and did not adequately reflect the health status, new risk adjusters have been discussed throughout the years. In the Netherlands, the pharmacy-based cost groups (PCGs) were first introduced in 2002, taking into account one of the most important drivers of health care costs – chronic diseases.

The incorporation of PCGs into the model largely increases its predictive power, as verified by multiple studies (Hájíčková, 2015; Huber et al., 2013; Lamers & Van Vliet, 2003). Following the Dutch example, PCGs were implemented in the Czech Republic in 2018, improving the model's performance considerably (Dungl et al., 2017). The Czech model specifies 25 PCGs and classifies individuals into the groups based on consumption of typical drugs for specific diseases (e.g. antidiabetics for diabetes). The

current classification threshold is set to 181 defined daily doses (DDDs), which corresponds roughly to half a year of recommended consumption. While for some diseases the pharmaceuticals specifically define given condition, for others they are less accurate and may omit some individuals who suffer from the disease. If the classification criteria omit expensive individuals, the insurance company will be undercompensated and will tend to avoid those patients.

"Renal failure" is a PCG group where the typical pharmaceuticals are only complementary to the treatment and the prescribed amounts vary substantially. Thus, we suspect the classification criteria based on pharmaceutical consumption to be inaccurate. Besides, renal failure is the most expensive chronic disease on average due to regular dialysis procedures, which substitute the kidney function and are crucial for keeping the patient alive. As opposed to drug prescriptions, the dialysis procedures are regular and highly specific for renal failure, hence they are more suitable for disease identification from the data.

The objective of this thesis is to revise and modify the existing classification criteria for the PCG renal failure and suggest improvements to the model currently used in the Czech Republic. The data are provided by the General Health Insurance fund (GHI) and consist of individuals reported with the drug consumption typical for renal failure or dialysis procedures in years 2015-2018. The thesis verifies the following hypotheses:

1. Currently used PCG model does not identify all individuals suffering from renal failure and omits expensive cases.

2. Identification of renal failure based on dialysis procedures captures more patients and reflects their costs better.

3. If dialysis procedures were incorporated into the model instead of drug consumption (or as its complement), the model's predictive power would increase substantially.

For the first two hypotheses, we identify patients based on the drug consumption and alternatively based on dialysis procedures, since we believe the procedures are able to predict the costs more precisely. The number of classified patients under both approaches and their costs are analysed and compared. Regarding the last hypothesis, we suggest various regression models based on the current risk adjustment methodology and modify the definition of the PCG "renal failure". The models are estimated using the Ordinary Least Squares (OLS) and their performance is compared

in terms of variation explained (R^2). Finally, the predictive power of the model is verified on real data. The estimates are used to calculate the cost predictions for year 2018, which are compared with the real costs.

The thesis is structured as follows: Chapter 2 presents basic concepts related to the topic, provides background on Czech health care system, and elaborates on the history of risk adjustment in the Czech Republic, ending with the description of the current PCG model. Chapter 3 is dedicated to literature review, where the first part covers risk adjustment methodologies and models used in previous studies, while the second part presents the chronic kidney disease, its worldwide prevalence, and related costs. Chapter 4 describes the data and their preliminary analysis, and explains methodology used in the empirical part. Chapter 5 presents the results of the analyses and comparison of models' performance. The discussion of possible issues and motivations for further research follows in Chapter 6. Chapter 7 summarizes the findings and contribution of the thesis.

2 Health insurance, pooling of funds, and risk adjustment

Unlike other regular goods (e.g. food), the consumption of health care is largely unpredictable both in its magnitude and timing over lifetime. Consequently, individuals are unable to prepare for future health care costs. As individuals are generally risk averse, they prefer to sacrifice a part of the present consumption in order to protect themselves against adverse events in the future (Smith & Witter, 2004). This is where the health insurance steps in to satisfy the needs for financial protection, as well as to solve the issues of equity in access to health care and overall health system efficiency.

Since 1990, many European countries introduced reforms to the structure of their healthcare systems. Regarding the health insurance, most policymakers worldwide have gradually moved towards compulsory universal health coverage (Kutzin et al., 2010). The system of health financing generally consists of three components: revenue collection, accumulation and management of resources, and their allocation (Mathauer et al., 2019). The collection of revenues in most European countries has the form of contributions (through general taxation, employer contributions, user charges, social insurance, health insurance premiums etc.) that are unrelated to the health status. In other words, charging premiums based on expected expenses is not allowed for insurers. In return, the financial coverage of standard health care package is guaranteed by law (Smith & Witter, 2004).

Revenues accumulation, referred to as pooling of funds is one of the most important characteristics of healthcare systems. Although the features of risk pooling differ among countries and their structure have gone through numerous reforms, the main goals are shared: pooling the risks together, redirecting the funds where needed, and improving the financial protection of the population (Kutzin et al., 2010). Risk pooling effectively balances the resources between rich and poor (especially when contributions are income-based), healthy and sick, young and old, and creates a level playing field in the access to health care (Mathauer et al., 2019).

The competition among funds providing health insurance is limited in countries where the risk rated premiums are not allowed. On the other hand, the adverse incentives of health funds to select healthier and therefore less expensive part of the population are

often an issue. Such practice, referred to as risk selection (also cream-skimming or cherry-picking), undermines the benefits of healthy competition among health funds, because instead of competing in the quality of provided services, the competition relies on the ability to attract the most profitable groups of insurees (Barros, 2003; Pilny et al., 2017). Even in systems where the health funds are obliged to accept all applicants, cherry-picking might take form of marketing strategies or benefit provisions favouring younger and healthier population.

When addressing the problem of risk selection, the fund allocation is crucial. An efficient redistribution should ensure that the funds are financially compensated for insuring sicker population, i.e. the fund allocation takes into account the risk profiles of subsequent fund members. Many countries have reacted to this challenge by employing risk adjustment (also called risk equalization) to calculate the expected expenses of pools and compensate them for the variation in the risk exposure.

2.1 Basic concepts – risk pooling and risk adjustment

The World Health Organisation describes the risk pooling as "the practice of bringing several risks together for insurance purposes in order to balance the consequences of the realization of each individual risk" (Smith & Witter, 2004). In the absence of risk pooling, all health costs would be born by the individual in relation with his/her clinical needs. Consequently, older and sicker population would have to bear the highest expenditure, which is inconsistent with the principles of solidarity. The risk pooling therefore ensures that the financial risk is shared among all pool members.

Smith & Witter (2004) distinguish between four basic approaches to risk pooling. The first approach does not use any risk pooling. Under this system, citizens meet their own health care costs and pay directly to the provider or, in case the insurance funds are present, the individuals pay risk premiums according to their perceived risk. The authors claim that such arrangement leads to dissatisfaction with the health care system, since most of the public health issues are neglected.

The other three mechanisms use risk pooling, differing in the number of pools and their interconnectedness. The unitary risk pool uses a single central pool where all the revenues from mandatory contributions are gathered and used later to cover the individual needs. While such arrangement effectively tackles the issue of cream-skimming and maximizes the pooling potential, it might induce excessive use of health care (moral hazard as well as supplier-induced demand). Moreover, the central risk pool might be particularly difficult to administer in larger countries.

The fragmented risk pooling assigns an individual into one of several pools in the system, based on the geographical location, the employment, individual characteristics or individual choice. Although the compulsory coverage results in pools with more diversity in terms of health risk sharing, the insurance contributions might be as well voluntary. The fragmented pooling is less difficult to administer, however, it introduces problems with variation in expenditure among different pools. Generally, the pools with older and sicker population (particularly with chronically ill individuals) are disadvantaged, since they bear higher costs. Unless some adjustment is made, the pure fragmented pooling may lead to cream-skimming and inequality among health funds.

Lastly, the integrated risk pools, usually accompanied with compulsory participation and free choice of health insurance fund, use transfers between the pools to ensure that the variation in risk exposures is reduced (Mathauer et al., 2019). Many countries implemented this system and developed various risk adjustment methods to predict the costs for health care and to compensate the insurance companies, accordingly (Schneider et al., 2008).

Although the health care costs are largely unpredictable, there exists number of factors that can be used as indicators for health care expenditure. The simplest risk adjustment schemes employ demographic indicators such as age, gender, and place of residence. More sophisticated models use individual information on the health status employing diagnoses-based or pharmacy-based indicators (Van Kleef & Van Vliet, 2012). The data used for risk adjustment should meet certain criteria: It should be feasible, robust against manipulation and easily applicable without excessive costs (Lamers & Van Vliet, 2004). A proper risk adjustment should reflect the variation in the risk exposure as precisely as possible and result in fairness among health insurance companies.

2.2 Risk adjustment in Czech healthcare

This chapter aims to set the context of Czech environment. Firstly, it presents the overall health care system in the Czech Republic (CR), the understanding of which is necessary for the following redistribution model. Subsequently, the evolution of the risk adjustment schemes in the CR is described, pointing out the main flaws of particular methodologies and incentives for their improvement.

2.2.1 Czech healthcare system and insurance policy

The Czech health care system is a statutory health insurance (SHI) system, based on compulsory contributions to health insurance funds. Also referred to as of Bismarckian type, the system is characterized by universality and a strong sense of social solidarity

(Alexa et al., 2015). The universality of access to healthcare is guaranteed by law. [Law on public health insurance (Zákon o veřejném zdravotním pojištění 48/1997 sb)].

In the Czech Republic, all citizens are obliged to pay contributions to the health insurance fund of their choice on monthly basis. The contributions are collected from the employers, employees, self-employed individuals, and people without taxable income, who are not paid for by the state. The part of the population which is economically inactive (e.g. students, retired, unemployed), is covered by the state contributions (Alexa et al., 2015).

Consequently, insurees are provided with a basic package of health care, that is covered by insurance. This package includes services such as inpatient and outpatient care, basic stomatologic procedures, rehabilitation and spa procedures, nursing and maternity care, screenings, vaccinations, and basic medical equipment (Zákon o veřejném zdravotním pojištění 48/1997 sb). Some procedures (e.g. in stomatology) and pharmaceuticals require cost sharing, i.e. out-of-pocket payments (OOP) by insurees. The mechanism of reimbursement and its regulation are subject to the Reimbursement Decree issued annually by the Ministry of Health.

The health insurance funds act as purchasers of health care and function as a quasi-public, self-governing, not-for-profit entities (Alexa et al., 2015). Originally, the General Health Insurance fund (GHI, Všeobecná zdravotní pojišťovna) was a single insurance fund operating in the Czech Republic since 1992, until other insurance companies joined the market. Currently, there are 7 health insurance funds in the Czech Republic, nevertheless, the GHI retained its dominant position by insuring approximately 57 % of population (as of 2018) (Cikrt, 2018). The main objective of insurance funds is to guarantee the provision of covered health care services, ensuring its local and time accessibility. For this purpose, the individual health funds contract with health care providers and negotiate the extent and the costs of covered services (Pelikánová, 2017). The competition between health insurance funds is limited, since the extent of benefit packages is determined by law and is considerably broad by its definition (Bryndová et al., 2019). As a result, individual funds differ only marginally in terms of contracted services, for example by offering bonuses to their members (e.g. contributions on sporting activities). Insurers are obliged to accept all applicants, hence any risk selection is prohibited (Alexa et al., 2015).

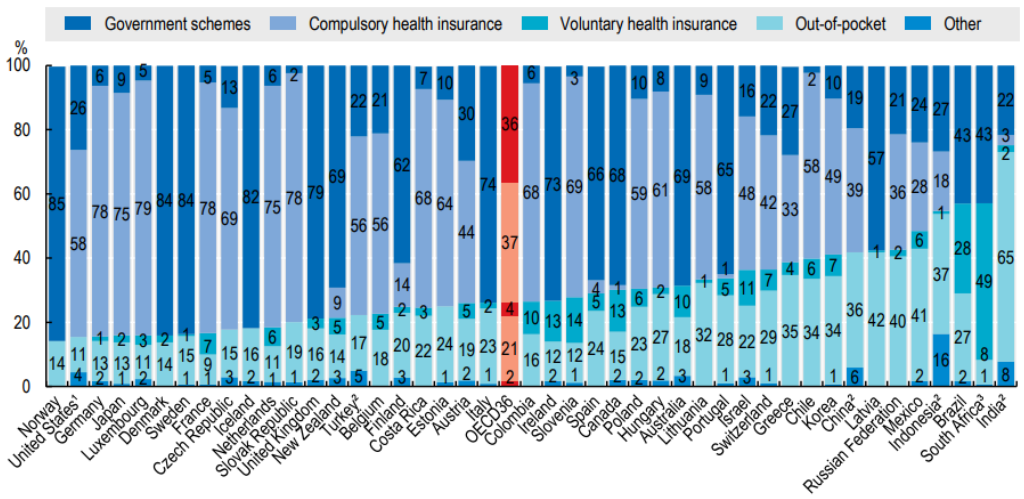
The health care providers can be distinguished according to their legal status and organization. Some of the medical units are organized as state entities (usually managed by the Ministry of Health), yet most of the units are of non-state character. These can be managed by regional or municipal authorities or by individuals, legal

entities and organizations (e.g. church). Regardless of ownership and legal form, medical entities negotiate contracts with health insurance funds and obtain reimbursement according to those agreements. The mechanism of reimbursement differs depending on the type of health care provider and health care provided (e.g. inpatient vs. outpatient care). The providers operating without contracts are paid directly by the patients without any claims for reimbursement (the only exception is necessary health care provided in case of emergency) (Pelikánová, 2017).

A crucial role is delegated to the Ministry of Health (MoH), which supervises the system, issues licences to health professionals, prepares legislation and policy agenda, and cooperates internationally. Moreover, the MoH manages several medical facilities and administers the State Institute for Drug Control (SÚKL). The regional authorities, which are by nature subordinate to the MoH, are in charge of registering local health care providers and managing own health care facilities (OECD/European Observatory on Health Systems and Policies, 2017).

SHI contributions are the main source of funding in the Czech healthcare system. According to the OECD (Health at Glance 2019), these accounted for 69 % of total sources in 2017. The rest of the sources consist of governmental schemes and OOP payments, which represent 13 % and 15 %, respectively. In comparison with other OECD countries, the Czech share of public sources on the total health expenditure is among the highest – 82 % vs. 73 % OECD average (see Figure 1). When related to the economy, the Czech Republic's overall health expenditure accounted for 7.5 % of GDP in 2018, which was slightly below OECD average of 8.8 % (Organization of Economic Cooperation and Development, 2019).

Figure 1: Health expenditure by the type of financing, 2017 (or nearest year)



Source: Organization of Economic Cooperation and Development, 2019

2.2.2 History of risk adjustment in the Czech Republic

The system based on compulsory SHI contributions does not automatically lead to a fair allocation of funds among health insurance companies. In the system of multiple health insurance funds, the structure of insurees among funds varies by their demographic characteristics and health status. If contributions were directly allocated to the insurance companies without any risk adjustment, the funds insuring less healthy patients would be disadvantaged and could face financial problems. Although the Czech insurance companies are not allowed to reject applicants, the risk selection might take form of marketing strategies aimed for selected groups (e.g. reimbursement of contraceptives, vitamins etc.) or the insurance companies might refuse to contract with some providers (Kutzin et al., 2010). Risk pooling and redistribution of collected funds according to risk adjustment models come as an adequate solution.

All premia collected from the insurees and the state contributions are pooled in one central fund administered by the GHI. Subsequently, the funds are redistributed to individual insurance companies with respect to the risk profile of their insurees and based on the predicted expenses. The predicted expenses are estimated using risk adjustment scheme which assigns risk indexes based on individuals' characteristics. This mechanism ensures relative fairness of the fund allocation and reduces attempts to attract only some groups of insurees (Chalupka, 2010).

The risk adjustment in the Czech Republic evolved dramatically in last decades. From the beginning of 90's when other health funds apart from the GHI entered the market, it was clear, that the revenues and expenses of individual funds would be unequal. Between 1993-1997 the insurance companies were still permitted to attract the insurees by offering additional benefits above the scope of the basic insurance package, such as travel insurance or wellness activities (Alexa et al., 2015). As a result, younger and more economically active part of the population frequently switched to new insurance companies who offered such services, which left the GHI in a disadvantaged position by insuring individuals with more complex health issues. In 1994, the first simple risk adjustment mechanism was implemented, taking into consideration the risk discrepancies of insurees (Chalupka, 2010). At that time, only 60 % of the collected funds and all state premia covering economically inactive population were subject to redistribution. The mechanism distinguished only two groups of insurees – those with age above 60, who were assigned the triple weight, and the rest of the population (Kutzin et al., 2010). The aim to improve the fund allocation was only partially fulfilled. Firstly, given that only two age groups were distinguished, the risk adjustment did little to reflect the variation in expenses between age groups. Secondly, the health

status within the age groups was not considered, although the insurees (e.g. above 60) could differ significantly as for their health condition (Bryndová et al., 2019).

Between 2004 and 2006, a demographic model was implemented with 100 % redistribution of collected premia, taking into account both age and gender of the insurees. The final version of the model accounted for 36 age/sex groups, where each group was assigned a specific risk index (Kutzin et al., 2010). Furthermore, the formula implemented in 2006 retrospectively compensated for extremely high costs. Approximately 10 % of collected funds was set aside for these purposes (Kutzin et al., 2010).

Despite the advantages of the demographic model, the employed risk adjusters still insufficiently captured the real health status of individuals. While indexes based on age and gender can explain some part of the variation in health care costs, they are unable to capture the variation within the same age/sex groups. Consequently, insurance funds with higher proportion of ill insurees (especially chronically ill) are always worse off even after accounting for the age/gender structure of the population (Chalupka, 2010).

To address these issues the new risk adjustment scheme has been discussed by the Government since 2010. Particularly the GHI supported the development of new risk adjustment, which would account for the health status, since they believed that their insurees were proportionally less healthy. Inspired by the models used in the Netherlands and later in Slovakia, the pharmacy-based cost groups (PCG) model was suggested, accounting for the most important drivers of the health care costs – chronic diseases. Moreover, given that the methodology uses drug consumption for classification to PCGs, the necessary data is already routinely collected by health insurance companies and does not require new data collection (Bryndová et al., 2019).

The final legislative proposal was submitted in 2016 and the PCG risk adjustment scheme came into force on 1.1.2018. This reform is expected to enhance the competition between the insurance funds, while additional compensation for chronically ill insurees should incentivize health insurance funds to offer more benefits to chronically ill individuals (Bryndová et al., 2019).

2.3 PCG model in the Czech Republic

The definition of the PCG model currently used in the Czech Republic is based on the Dutch model adopted in 2012. Redistribution of funds according to PCG methodology consists of two separate mechanisms with their own methods of calculation. The precise structure of the model, including all the coefficients set by the MoH is defined

in law on General health insurance premiums (Zákon České národní rady o pojistném na všeobecné zdravotní pojištění 592/1992 Sb.).

The first mechanism estimates risk indexes of individuals based on their age, gender, and occurrence of chronic diseases. As health care costs related to chronically ill patients account for approximately 80 % of total expenses on public health, the inclusion of PCGs substantially improves the predictive power of the model (Dungl et al., 2017). The second mechanism is intended to retrospectively compensate for extremely high costs and serves as a reinsurance tool. The funds for extreme cases are allocated ex post when real expenditures for given period are revealed. Retrospective risk sharing among funds reduces unexpected fluctuations in balances of health funds (Bryndová et al., 2019). The GHI is in charge of administration and supervision of central account which was established for these purposes.

Next subsections are presenting the detailed methodology of the model as specified by the law (Zákon České národní rady o pojistném na všeobecné zdravotní pojištění 592/1992 Sb.).

2.3.1 Demographic classification

Demographic classification is based on age and gender of insurees as of the first day of the month for which the funds are being redistributed. Currently, there exist 19 age groups for each gender, that is 38 groups in total. The list of age/gender groups with corresponding risk indexes computed for year 2018 is provided in

Table 1. The indexes carry information about the riskiness and expected costs of particular group. Intuitively, the higher the age, the higher the corresponding risk index[1], since the number of health complications and the probability of mortality are increasing. As indicated in many studies (Duncan et al., 2019; French et al., 2017) the highest expenses are usually incurred at the very end of the patient's life.

Table 1: Risk indexes for age/sex groups, 2018

Age	*Risk index - men*	*Risk index - women*
less than 1 year	0.7926	0.642
1-4 years	-0.5097	-0.5659
5-9 years	-0.5999	-0.6503
10-14 years	-0.616	-0.5818
15-19 years	-0.6427	-0.5095

[1] With the exception of newborns and babies younger than 1 year, who represent substantial health risk.

20-24 years	-0.7183	-0.5422
25-29 years	-0.7001	-0.4135
30-34 years	-0.6735	-0.359
35-39 years	-0.6448	-0.4212
40-44 years	-0.6051	-0.4667
45-49 years	-0.5357	-0.409
50-54 years	-0.4182	-0.3401
55-59 years	-0.2469	-0.2886
60-64 years	-0.0483	-0.2348
65-69 years	0.1832	-0.0784
70-74 years	0.4343	0.1191
75-79 years	0.5752	0.2726
80-84 years	0.6427	0.4432
85 years and more	0.7943	0.7461

Source: Zákon České národní rady o pojistném na všeobecné zdravotní pojištění 592/1992 Sb.

2.3.2 PCG classification

Patients are assigned into one of the PCGs based on their drug consumption. A patient suffering from a chronic disease uses specific medication for his or her illness. As chronic diseases are usually treated in the long-term or even for the rest of the patient's life, chronically ill individual is expected to consume a relatively stable amount of health care. In other words, if a patient consumes certain amount of pharmaceuticals that uniquely define the disease they suffer from, one can expect certain amount of health care that the patient will consume throughout the year (Dungl et al., 2017). The PCGs are constructed in a way that each group congregates patients with relatively homogenous health care needs and costs.

The specification of the right type of pharmaceuticals for each PCG and the threshold of their consumption to be reached are of utmost importance. Drugs are specified using Anatomical Therapeutic Chemical (ATC) coding. The threshold of consumption is defined in units of defined daily doses (DDDs), where 365 DDDs correspond to one year of daily usage of recommended doses. The threshold for PCG classification is specified by the MoH in the range of 121 and 365 DDD with regard to the number of expected individuals in each group and stability of the redistribution system. For year 2018 the threshold was set to 181 DDD for all PCG groups. Insurees may belong to more than one group, given that the exclusion conditions are met (e.g. DM2 cannot be combined with DM1 or diabetes with hypertension). The reclassification of patients is carried out on monthly basis.

Currently, the Czech model specifies 25 PCGs for which the corresponding indexes are estimated. As in the case of demographic factors, the indexes reflect the expected

effect of belonging to a group on health care costs. The index calculation relies on the methods of weighted least squares (WLS) (Bryndová et al., 2019). The list of PCGs and indexes calculated by the MoH for year 2018 is provided in Table 2.

Table 2: Risk indexes for PCGs, 2018

PCG code	*PCG group name*	*Risk index*
GLA	Glaucoma	0.2246
THY	Thyroid disorders	0.2533
PSY	Antipsychotics, Alzheimer's disease, treatment of addiction	1.9603
DEP	Treatment with antidepressants	0.8659
CHO	Hypercholesterolemia	0.2838
DMH	Diabetes with hypertension	1.0344
COP	Serious asthma, Chronic obstructive pulmonary disease	1.8142
AST	Asthma	0.8682
DM2	Diabetes mellitus type 2	0.4561
EPI	Epilepsy	1.3813
CRO	Crohn's disease, ulcerative colitis	0.9823
KVS	Heart disease	1.5601
TNF	Rheumatic diseases treated with TNF inhibitors	14.4966
REU	Rheumatic diseases treated otherwise than with TNF inhibitors	0.9963
PAR	Parkinson's disease	1.4167
DM1	Diabetes mellitus type 1	2.1692
TRA	Transplants	4.1426
CFP	Cystic fibrosis or disorder of pancreatic exocrine function	20.7391
CNS	Brain and spine disorders	10.1492
ONK	Malignancy	17.2183
HIV	HIV, AIDS	10.7017
REN	Renal failure	41.6000
RAS	Therapy with growth hormone	10.3981
HOR	Hormonal oncology	2.2946
NPP	Neuropathic pain	2.2671

Source: Zákon České národní rady o pojistném na všeobecné zdravotní pojištění 592/1992 Sb.

2.3.3 Risk index

Combining the demographic and PCG indexes as described above, each insuree is assigned with the final risk index, which is recalculated each month. The index quantifies the overall anticipated risk of the individual as for his health care costs covered by the health insurance. The final risk index is calculated as follows:

$$\begin{aligned}\textbf{\textit{Risk index}} = \mathbf{1} &+ risk\ index\ of\ demographic\ group\ the\ insuree\ belongs\ to \\ &+ sum\ of\ risk\ indexes\ of\ PCG\ groups\ the\ insuree\ belongs\ to \\ &+ correction\ for\ combination\ of\ groups^{2}\end{aligned}$$

Given that the average costs are known, multiplying them with the risk index gives the predicted expenses for the insuree in a given month (Dungl et al., 2017).

2.3.4 Reinsurance and reinsurance constant

If the real costs substantially exceed the risk adjusted predicted costs for that period, the reinsurance mechanism guarantees that the health insurance company will be compensated ex post. For this purpose, the reinsurance constant is calculated for each period. Reinsurance constant represents a threshold, that must be exceeded by additional costs in order to be subject to an ex post compensation. The reinsurance rules for retrospective compensation are as follows:

- If the real costs for a patient are higher than the sum of the reinsurance constant and the amount that was obtained based on the predictions, the health insurance fund has the right to be compensated for 80 % of the amount that exceeded this sum. The compensation must not exceed four-fold of the reinsurance constant.

- In case, that the real costs for a patient exceeded the sum of six-fold of the reinsurance constant and the amount that was obtained based on the predictions for that period, the health insurance fund has the right to be compensated for 95 % of the amount that exceeded this sum.

The calculation of risk indexes described in the previous chapters is also affected by the existence of the reinsurance. In fact, the reinsurance constant directly enters the WLS estimation.

---

[2] This component captures the correction for specific combinations of demographic and PCG risk groups given that these are specified (not specified for 2018).

3 Literature review

This chapter is dedicated to a brief literature review related to the main research questions of the thesis. The first subchapter provides an overview of research papers on risk adjustment methodologies and the performance of models considering different explanatory variables. The second subchapter introduces a chronic kidney disease and discusses its worldwide prevalence, health care costs and role in the risk adjustment.

3.1 Risk adjustment in the literature

During 1990s, risk adjustment schemes have been introduced in 11 European countries (Prinsze & Van Vliet, 2007). In last few decades, the risk equalization became a subject of the generous amount of research papers and the methods have evolved dramatically over years. The introduction of diagnostic cost groups (DCGs) in the US in 2000, and the development of pharmacy-based cost groups (PCGs) in the Netherlands in 2002, represented the core milestones in the modern risk adjustment methodology. Table 3 briefly summarizes the evolution of mechanisms in Israel, Germany, Switzerland, the Netherlands, and Slovakia. Although the currently employed methods are not perfect, many reforms have been done in these countries to address the issues of fair allocation and thus can be used as an illustration of different risk adjustment methods.

Table 3: Evolution of risk adjustment in chosen countries

Country	*History of risk adjustment*	*Sources*
Israel	1995 – prospective payments based on age + retrospective payments for 5 severe diseases (including renal failure) 2010 – adding sex + peripheral status	(Shmueli, 2015) (Van de Ven et al., 2007)
Switzerland	1993 – retrospective payments based on age and sex 2011 – switching to prospective payments + adding hospital and nursing home stays 2020 – implementing PCGs in addition to age and sex	(Von Wyl & Beck, 2016) (Van de Ven et al., 2007) (Federal Office of Public Health, 2020)
Germany	1994 – age, gender, disability status as risk adjusters 2002 – adding Disease Management Program enrolments 2009 – introducing Hierarchical Morbidity Groups based on reported dialyses	(Pilny et al., 2017) (Wasem et al., 2018) (Ash et al., 2000) (Juhnke et al., 2016)

Netherlands	1990 – age, gender as risk adjusters	(Schneider et al., 2008)
	1995 – adding urbanization and source of income	(Prinsze & Van Vliet, 2007)
	2002 – introducing PCGs	(Van Kleef et al., 2018)
	2004 – introducing DCGs as a complement to PCGs and later adding retrospective compensation	
	2006-2017 – adding multiple risk adjusters, separating model types to somatic care, mental care, and OOP payments (see Appendix A 1 for the full list of risk adjusters used in somatic care in 2017)	
Slovakia	1995 – two age groups (below and above 60)	(Glova & Gavurová, 2013)
	1999 – switching to multiple sex/age groups	(Kutzin et al., 2010)
	2010 – adding economic activity/inactivity	(Health Policy Institute, 2014)
	2013 – implementing PCGs (24 groups)	

3.1.1 Development of risk adjusters

The efficiency of demographic models has been criticized from the beginning by many authors, who claimed that age and gender were insufficient predictors of the expected costs. Consequently, it was suggested to utilize risk adjusters related to the health status in addition to demographic factors. One of the first studies in this field was carried out by Newhouse et al. (1989), who tested model's efficiency after inclusion of multiple risk adjusters. The results of the study indicate that even after all relevant health-based indicators were incorporated in the model, the predictable variance explained reached a maximum of 30 %.

Inclusion of chronic diseases in the risk adjustment was firstly accomplished in the US in early 1990s, where so called Chronic Disease Score (CDS) was developed, using pharmaceutical information for disease classification (Von Korff et al., 1992). Later, Clark et al. (1995) revised the original CDS methodology and extended the range of drugs used for disease identification (29 groups in total). The revised CDS was compared with the simple demographic model, and additionally with a model using 34 ambulatory diagnostic groups (ADGs) based on outpatient diagnoses claims. Performing the regression analysis on the Group Health Cooperative of Puget Sound data from 1992, the explained variances (measured by means of R-squared) were equal to 3 %, 10 % and 8 % for the demographic model, the revised CDS model and the ADGs model, respectively. The combination of both CDS and ADGs improved the explanatory power of the model to 12 %.

Fishman et al. (2003) further developed the CDS by creating so called Rxmodel, which addressed the weaknesses and barriers of the original methodology. One of the main improvements was the expansion for children to reflect the special challenges of the drug prescription among pediatric population. Using the 1995-1996 data from large US

health maintenance organizations, the authors showed that Rxmodel is able to capture 8.7 % of cost variation, while demographic model only 3.5 %. For comparison, the models based on the diagnosis claims, using Ambulatory Clinical Groups and Hierarchical Coexisting Conditions managed to explain 10.2 % and 15.4 %, respectively. Employing the quintile analysis, the authors pointed out that all tested models performed similarly for the middle 60 % of the cost distribution.

The revised version of CDS was employed by Kuo et al. (2011), who used Taiwanese data of National Health Insurance from years 2006-2007. The authors identified 32 classes of chronic conditions based on the pharmaceutical consumption using WHO ATC classification. The resulting R-squared of the model using pharmacy-based metrics was among the highest with 30 % of variation explained compared to diagnoses-based morbidity measures (authors used Deyo's Charlson Comorbidity Index and Elixhauser's Index), none of which have exceeded 25 %.

In Europe, Huber et al. (2013) modified the CDS model to fit the Switzerland health care system. The authors defined 22 chronic conditions based on WHO ATC coding and estimated 3 different models using medical claims data from 2009-2010. The most expended model accounting for CDS, age, gender, language area and the type of health insurance plan, managed to explain 17.9 % of variance in health care costs. As opposed, the model without CDS explained 4.7 % (both estimated for individuals up to age 65).

The most important stream of literature originates in the Netherlands, where PCG model has been widely revised since its introduction in 2002. Lamers & Vliet (2003) firstly implemented 22 chronic conditions in addition to risk adjusters used before PCG implementation (i.e. age, sex, urbanization, type of insurance) and suggested improvements to reduce the gaming possibilities. The authors used different thresholds of prescriptions for PCG classification and different numbers of comorbidities allowed per person. The number of prescribed daily doses (PDDs) was set to at least 4 PDDs, 91 PDDs and 181 PDDs, where the model with at least 91 prescriptions had the best predictive power equal to 9.8 %. The model with unlimited number of conditions per person explained 9.9 % of the cost variation, which was better than models with one or two conditions per person. Demographic risk adjusters explained only 5 % of the variation. The authors also suggested using defined daily doses (DDDs) instead of number of prescriptions due to its better robustness to manipulation. Furthermore, the reduction of the number of PCGs by removing the diseases with low future costs was proposed as another strategy to prevent perverse incentives of sickness funds. Indeed, the number of groups in the Dutch risk equalization decreased to 13 in 2002 and later to 12 in 2004 (Prinsze & Van Vliet, 2007).

In 2004, DCGs were introduced in the Dutch risk adjustment. The combination of DCGs, PCGs, and demographic indicators reached the R-squared of 16.6 % (van de Ven et al., 2004). Van Veen et al. (2015) additionally included the diagnostic information from three prior years, which altogether managed to explain 24.8 % of the variation. The explanatory power of the current risk adjustment model utilized in the Netherlands (as of 2017) accounts for about 31 % (Van Kleef et al., 2018).

Available literature on PCG methodology in the Czech Republic is limited due to its recent implementation. One of the first researches has been accomplished by the Health Policy Institute (2014), that used the existing PCG model in Slovakia and tested its potential benefits in the Czech environment. The model included 23 chronic conditions following the Dutch example. Individuals were allowed to be classified into one PCG (the most expensive one) and the threshold of drug consumption was set to 181 DDDs. Using the data of Czech GHI from years 2009-2011, the PCG model was estimated to explain approximately 10.8 % of cost variance. This was a substantial improvement from the demographic model explaining only about 2.7 %.

The most detailed PCG analysis has been performed by the KlientPRO group, which significantly contributed to its implementation in the Czech Republic. As opposed to the model used in Slovakia, the authors proposed several modifications: Use of 25 PCGs, classification into more than one PCG, and lowered threshold of drug consumption. Moreover, the authors included an ex post compensation for extreme costs as described in the chapter on PCG model in the Czech Republic. Table 4 summarizes the performance of tested models. As can be seen, the model with the best predictive power allows for more PCGs, uses the threshold of 121 DDDs, adjusts for the combination of PCGs as well as for the combination of PCGs with demographic factors, and uses an ex post compensation (Dungl et al., 2017).

Table 4: Overview of PCG models and their performance, Czech data 2010-2011

Model	1	2	3	4	5
Prescription threshold (DDDs)	121	121	121	121	181
More PCGs per person	yes	yes	no	no	no
Correction for combination of two PCGs	yes	yes	no	no	no
Correction for combination of PCG with demographic group	yes	no	no	no	no
Ex post compensation	yes	yes	yes	no	no
R-squared	45,42%	45,09%	30,12%	19,81%	18,90%

Source: Dungl et al. (2017), edited

Lastly, it is worth to mention two diploma theses that contributed to the available literature on PCG models in Czechia and Slovakia. The thesis by Hájíčková (2015) showed the benefits of the PCG model over the demographic model used in the CR at that time. The author used the sample of 10 % of the Czech population in years 2011-2012. The predictive power largely improved from 2.03 % to 13.87 %, by which the appropriateness of PCG model was confirmed even before it was actually implemented in the CR.

The latter thesis by Chochláčová (2018) tested similar PCG model, however, with the use of Slovakian health insurance data. The author made several modifications to the current version, particularly focusing on the PCG group Hypercholesterolemia (CHO). Both lowering the prescription threshold (from 181 to 121 DDDs) and dividing the PCG group CHO into three subgroups according to the drug consumption, improved the model in terms of R-squared and profit/loss for the insurer.

3.1.2 Regression methodology

The majority of studies on risk adjustment (Clark et al., 1995; Dungl et al., 2017; Fishman et al., 2003; Prinsze & Van Vliet, 2007; Van Kleef & Van Vliet, 2012) employed ordinary least squares (OLS) or weighted least squares (WLS) in the regression analysis. Nevertheless, it is often argued that linear models do not fit the distribution of health care costs very well, since the distribution of expenditure is not normal, but rather skewed to the right with a heavy tail.

The main problem of skewness is that it typically produces heteroskedastic errors. To address this problem, various transformations of the model were suggested to better reflect the distributional properties. For example, Veazie et al. (2003) used the square root of the dependent variable, which resulted in the improvement of the model's precision. Other common practice is to use a log-transformation (Farley et al., 1996; Kuo et al., 2011), which is however problematic when zero-cost observations are included in the dataset. A common solution is to perform a two-part estimation, where the first equation is logit or probit used for all cases, and the second is least squares with log-transformation of all positive costs (Veazie et al., 2003). One of the main drawbacks of log-models is their interpretation which is not straightforward unless log-values are retransformed, which might induce a retransformation bias (Kuo et al., 2011). OLS or WLS are therefore the most frequently used methods.

Van Veen et al. (2015) suggest three main reasons in favour of OLS utilization on untransformed data. Firstly, the OLS provides unbiased and the most easily interpretable metrics compared to other models, which is a very important feature for

regulators and policy makers. Secondly, if the official methodology of risk adjustment used for redistribution purposes applies OLS (which is the case for the CR), it should be used as well when the model is being revised. Lastly, the authors point out that several studies have shown that with a large sample, the OLS perform similarly as other more complex models, such as general least squares or two-part estimation.

3.2 Chronic kidney disease

Chronic kidney disease (CKD) is a serious illness that can take on multiple forms according to its severity. The identification of the number of CKD patients in countries can be challenging, since the definition of the disease itself is ambiguous and the classification of its stages has changed over time. Generally, the prevalence of CKD tends to be lower in comparison with other chronic diseases, however, the disease is associated with very high costs, particularly in the case of end-stage renal failure. The number of patients suffering from kidney malfunction is recently increasing in most countries, including the Czech Republic. This creates a pressure on limited budgets of health care systems, particularly on insurance companies, that are responsible for the health care reimbursement.

In the Czech Republic, renal failure currently forms one of the groups in the PCG model used for fund redistribution. The risk adjustment should ensure that health insurance companies are well compensated for patients diagnosed with renal failure, that present substantial costs. Whether these costs are well predicted under the current system is a question to be investigated in the empirical part.

This chapter aims to provide the reader with the basic knowledge of the disease. Firstly, it introduces the main characteristics and forms of CKD. Subsequently, it presents the prevalence of the disease in the CR and worldwide, followed by a brief overview of its economic effects and role in the risk adjustment.

3.2.1 Renal failure as a chronic disease

Regarding the decreased function of kidneys, one can come across several different definitions. The most widely used term is the “chronic kidney disease” (CKD), which is defined as “abnormalities of kidney structure or function, present for more than 3 months, with implications for health” (National Kidney Foundation, 2013). Another term frequently used in Czech terminology is chronic renal insufficiency (CRI), which generally denotes that the kidney function has decreased (Ryšavá & Brejník, 2018). In the CR, the same condition might be also referred to as chronic kidney/renal failure, while in foreign literature “renal failure” usually represents the most serious kidney

malfunction, thus the last stage of the disease (National Kidney Foundation, 2013). In American literature, the term "end-stage renal (kidney) failure" also means that the patient is in the most serious condition, when renal replacement therapy (RRT) is a necessity (e.g. dialysis or transplantation) (Kramer et al., 2019). The chronic kidney disease should not be confused with acute kidney disease/injury or other non-progressive kidney insufficiency. While the symptoms might be similar, in acute condition the decrease in kidney function is always sudden and/or result of some injury (Ryšavá & Brejník, 2018).

Following widely spread definition from Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease[3] (National Kidney Foundation, 2013), the patient is considered to have the CKD if one of the markers depicted in Table 5 is present for more than 3 months. The glomerular filtration rate (GFR) is one of the basic classification markers for measuring the severity of kidney damage (Zima et al., 2014). Simply put, the GFR measures the amount of blood that is filtrated per unit of time (e.g. minute). The direct measure of GFR is rather complicated, therefore the estimated glomerular filtration rate (eGFR) is frequently employed instead (Viklický, 2013).

Table 5: Criteria of CKD (either of the following present for >3 months)

Markers of kidney damage	Albuminuria (AER ≥30mg/24 hours; ACR ≥30mg/g [≥3mg/mmol])
	Urine sediment abnormalities
	Electrolyte and other abnormalities due to tubular disorders
	Abnormalities detected by histology
	Structural abnormalities detected by imaging
	History of kidney transplantation
Decreased GFR	GFR <60ml/min/1.73 m2 (GFR categories G3a–G5)

Source: National Kidney Foundation (2013)

Currently used classification employs 3 different markers according to which the severity of CKD is evaluated: cause, GFR category and albuminuria category. The cause identification is based on other present diseases (e.g. diabetes), the location of damage within the kidney or presumed pathologic-anatomic findings. GFRs are classified into 5 categories (see Table 6), where GFR <60ml/min/1.73 m2 is already considered to mark decreased function of kidneys. Value of GFR less than 15ml/min/1.73 m2 signifies renal failure, in which case kidneys are not able to clean the blood from wastes and their function must be replaced by dialysis or transplant. Lastly, albuminuria is also classified into 3 categories according to albumin excretion

[3] Published by not-for-profit organization Kidney Disease Improving Global Outcomes (KDIGO) that was established by National Kidney Foundation (USA)

rate (AER) or alternatively using albumin-to-creatinine ratio (ACR) of the patient (see Table 7) (National Kidney Foundation, 2013).

Table 6: GFR categories

GFR category	GFR (ml/min/1.73 m2)	Terms
G1	≥90	Normal or high
G2	60-89	Mildly decreased
G3a	45-59	Mildly to moderate decreased
G3b	30-44	Moderately to severely decreased
G4	15-29	Severely decreased
G5	<15	Kidney failure

Source: National Kidney Foundation (2013)

Table 7: Albuminuria categories

Category	AER (mg/24 hours)	ACR (mg/mmol)	Terms
A1	<30	<3	Normal to mildly increased
A2	30-300	3-30	Moderately increased
A3	>300	>30	Severely increased

Source: National Kidney Foundation (2013)

In 75 % of cases the cause of the CKD is diabetes mellitus, arterial hypertension, or glomerular diseases. Less frequently, the CKD might be caused by genetical predispositions, return to dialysis after transplantation or other rare diseases (Ryšavá & Brejník, 2018). The risk of CKD and the pace of its progression also depend on factors such as age, gender, ethnicity, or family anamnesis. While these cannot be affected, others, such as hypertension, glycaemia, dietary or smoking habits are in hands of patients and their physicians. Appropriate conservative treatment can postpone the necessity of dialysis or kidney transplantation for several years. Besides, kidney diseases substantially increase the risk of infections and cardiovascular complications, that qualify as a major cause of mortality (Neild, 2017). Therefore, every patient should be sent to nephrological specialist in stage G3a at the latest (Zakiyanov et al., 2014).

The most serious cases (i.e. CKD stage G5) are subjects to renal replacement therapy (RRT), which can take a form of hemodialysis[4] (HD), peritoneal dialysis[5] (PD), or

[4] Hemodialysis - a treatment replacing function of kidneys by filtering wastes and water from the blood through needles placed into arm, helping to control blood pressure and balancing important minerals. (National Institute of Diabetes and Digestive and Kidney Diseases - NIDDK)

[5] Peritoneal dialysis - a treatment replacing function of kidneys by filtering the blood inside the body through catheter placed in patients' abdomen, or belly. (NIDDK)

transplantation. The choice of RRT has to take into account individual needs and patients have to be provided with all treatment possibilities (Ryšavá & Brejník, 2018). The most optimal treatment is kidney transplantation, ideally from a living donor. Unfortunately, this is not possible for all patients, mainly due to the lack of suitable donors. Therefore, dialysis is the best accessible option substituting kidney functions.

3.2.2 Epidemiology

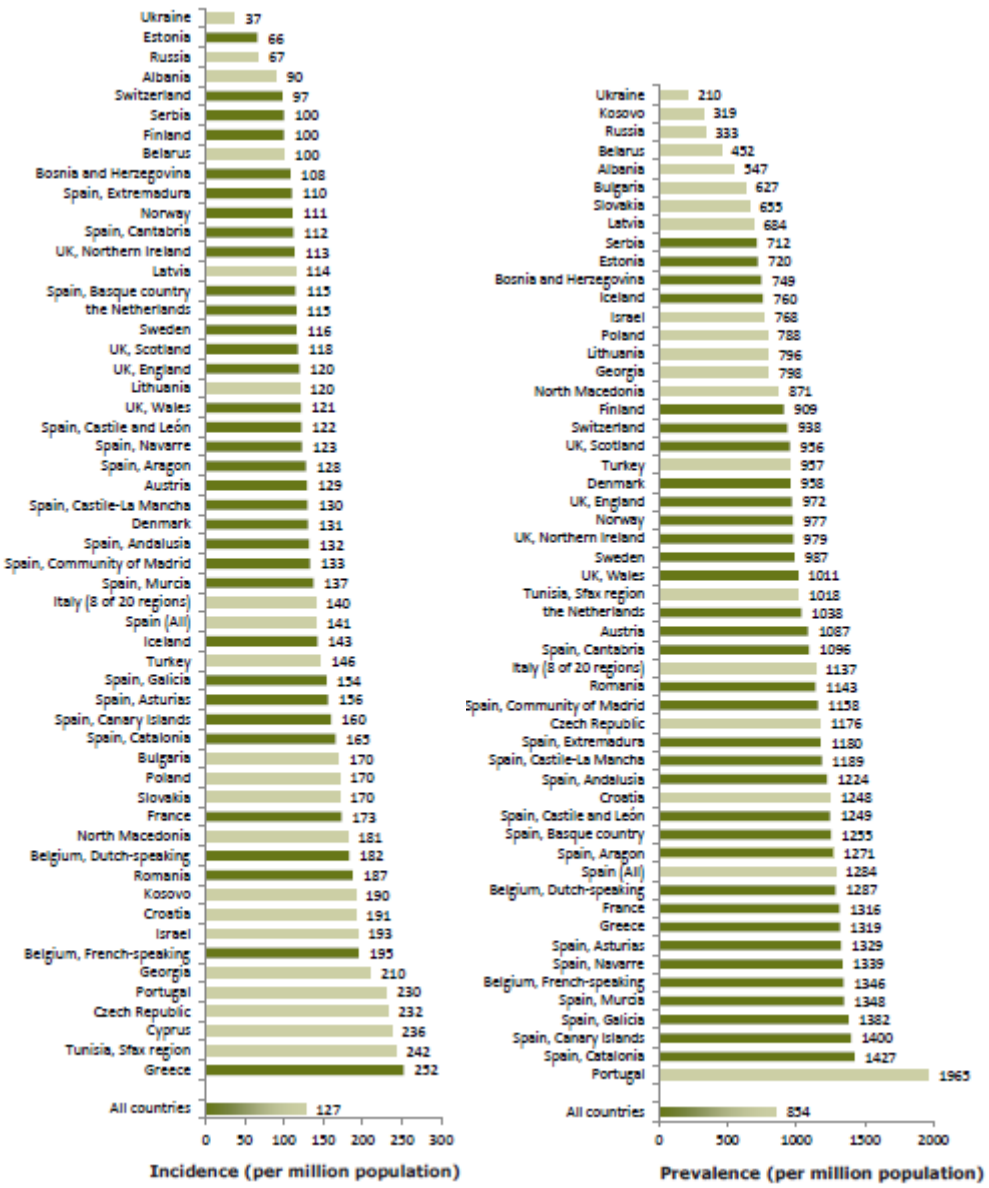
Prevalence and incidence of CKD have been increasing globally, partly as a result of population ageing and related expansion of diseases of civilization (hypertension, diabetes, cardiovascular diseases etc.). Decrease in kidney function presents the risk of progression into renal failure and increases the probability of death (Zakiyanov et al., 2014). Some of the health complications might be prevented or delayed, however, it requires proper management of the target population. Getting to know the prevalence of CKD population is thus highly important in the planning of disease management strategies.

Worldwide, the prevalence of CKD of all stages (G1-G5) is estimated to amount to 13.4 % of population. The most frequent stage G3 accounts approximately for 7.6 % of population while the most serious stage G5 is minor – around 0.1 % (Ryšavá & Brejník, 2018). According to Brück et al. (2016), who investigated the occurrence of the CKD in 13 European countries based on 19 general population-based studies, the prevalence differs substantially across European populations. The study estimated the adjusted prevalence of the CKD of all stages to vary between 3.31 % (for Norway) and 17.3 % (for northeast Germany). When considering only more serious cases from G3 to G5, the prevalence in central Italy accounted for 1 % of population while in northeast Germany, it reached almost 6 %. The authors point out that the difference might be affected not only by true differences in population health but also as a matter of different methodologies used to collect the data in individual countries. The substantial part of the variation can be also ascribed to dissimilarities in lifestyle, environment, and regional health policies (Brück et al., 2016).

The most reliable data can be observed for cases where the CKD has already progressed to renal failure. These patients receive renal replacement therapy - either dialysis or kidney transplantation, both of which can be well monitored and documented. The epidemiology of RRT within Europe is described in the Annual Reports of European Renal Association – European Dialysis and Transplant Association (ERA-EDTA). The reports gather the data of individual national and regional renal registries in Europe and Mediterranean Sea bordering countries. The data from 52 renal registries from 37 countries were gathered in 2017 Annual Report (ERA-EDTA Registry, 2019).

According to this report the overall incidence of patients accepted for RRT in 2017 in selected countries was equal to 127 per million population (pmp). In the international comparison, as depicted on the left panel of Figure 2, the Czech Republic is among the highest with the incidence of 232 pmp. The right panel of Figure 2 illustrates the prevalence of patients with RRT in observed countries. The average prevalence is equal to 854 patients pmp (note that the CR is above the average). The detailed analysis showed that there is 60 % of men among patients on the RRT, 43 % of whom are older than 64 years (ERA-EDTA Registry, 2019).

Figure 2: Incidence and prevalence per mil. population by country/region, 2017



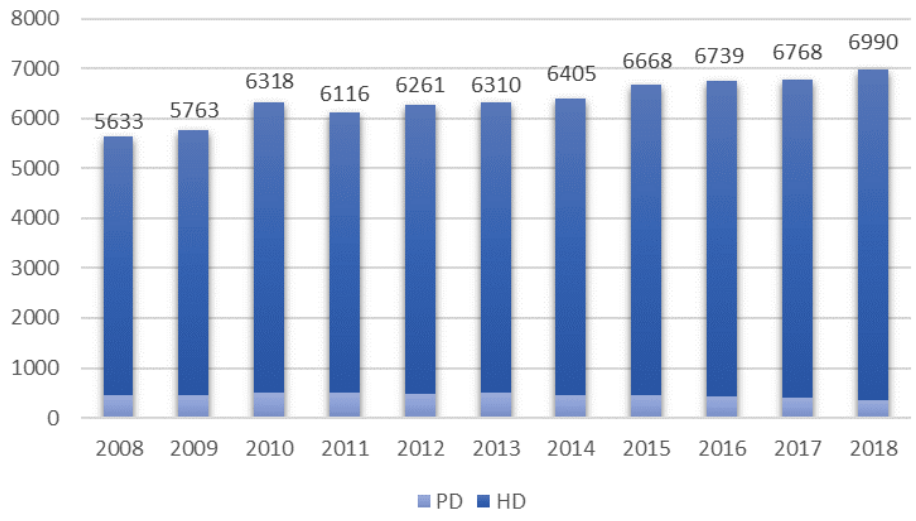
Note: Light bars – aggregated data; Dark bars – individual patient data

Source: ERA-EDTA Registry's Annual Report 2017 (ERA-EDTA Registry, 2019)

Currently, there is no population data capturing the prevalence of the CKD in the Czech Republic and there is a lack of studies researching this field. On the other hand, the data about RRT patients are well documented by the Czech Society of Nephrology (CSN) and published annually in statistical reports. Moreover, CSN administers the registry of dialysis patients providing another source of statistical information. In December 2018, there were 102 centres for dialysis where 6,990 patients (or 659 pmp) were treated. Most of those patients obtained hemodialysis, peritoneal dialysis accounted only for 5.1 %. Kidney transplant was performed in 508 cases (Czech Society of Nephrology, 2018). The information about the total amount of patients living with kidney transplant was missing in the 2018 report, however, in 2016, it amounted to 4,692 patients (Czech Society of Nephrology, 2016).

The number of patients on dialysis in Czechia has been increasing since 2008 (see Figure 3). Diabetes mellitus and hypertension represent the most common diagnoses in association with renal failure – in 2018, 46 % of patients on HD suffered from diabetes and 59 % from hypertension. Regarding the age structure of the patients, approximately 73 % and 58 % of patients on HD and PD, respectively, were older than 60 years (Czech Society of Nephrology, 2018). In 2018, 1,576 dialysis patients died, 45 % of whom in association with cardiovascular complications.

Figure 3: Number of patients on dialysis in the Czech Republic, 2008-2018



Notes: Patients who died during given year excluded; PD – peritoneal dialysis, HD – hemodialysis

Source: Annual Reports of Czech Society of Nephrology, retrieved from: www.nefrol.cz, edited

3.2.3 Costs

CKD patients represent high costs for health care systems, especially in the case of end-stage renal failure when dialysis or kidney transplant are necessary. In addition to costly procedures, CKD patients are often hospitalized for extensive periods of time and take numerous expensive medications (Da Silva et al., 2018). With the increasing number of dialysis patients and progression of the CKD in recent years, the financial impact is expected to become even more serious.

In their research, Da Silva et al. (2018) reviewed the studies investigating the cost burden related to the CKD in different countries all over the world, especially in the developing ones. The study indicates that CKD costs are increasing with the severity of the disease and with the need to initiate dialysis, while representing the highest burden for low and middle-income countries. The study discusses 37 articles, all of them focusing on the cost estimation of CKD, however using different methodology.

For example, in Sweden, in comparison with general population the costs related to HD and PD are 45 and 29-times higher, respectively. In Italy, the costs for a CKD patient before dialysis equal EUR 11,123, while the costs of a patient on dialysis amount to EUR 53,764. This demonstrates that if the progression of CKD to the end-stage renal failure could be prevented, the total expenditure would decrease substantially (Da Silva et al., 2018).

In the case of kidney transplantation, multiple studies (Axelrod et al., 2018; Helanterä et al., 2019; Jarl et al., 2018) showed that such surgery induces high costs in the year of operation, however, substantially reduces expenditure in the following years. Sijpkens et al. (2008) estimated the annual expenditure of the patient on dialysis to EUR 60,000, while the care for the patient one year after the kidney transplantation decreased to EUR 14,000.

The results are confirmed in the CR by Tichý (2015) who analysed the data of the GHI. Tichý (2015) shows that while in the year of surgery the kidney transplant presents higher costs compared to dialysis, the patient with transplant reduces the costs by 474,263 CZK in the following year. The investment thus returns approximately after 1.3 years. The positive effects are not solely financial; kidney transplant significantly improves patient's health, survival expectations and comfort (Ferrari, 2016). Tichý (2018) also revealed that while the number of patients on dialysis has been relatively stable in last years, their average costs have been increasing.

3.2.4 Renal failure in risk adjustment

The treatment options for the CKD vary substantially in their costs (e.g. conservative treatment vs. dialysis). In risk equalization models employing PCG methodology, the renal failure is usually represented by one PCG group, where the severity of the disease and the type of the treatment are not distinguished (since it is based solely on drug consumption). This might lead to undercompensation of severe cases requiring the most intensive health care or conversely, to overcompensation of less expensive cases. Consequently, the health funds could have incentives to avoid sicker patients or to reduce the expensive treatments, such as kidney transplantations.

In the models where DCGs are included, the insuree might be classified into both PCG and DCG representing renal failure. In the study by Prinsze & van Vliet (2007) the renal failure was reflected by one PCG and two DCGs, one of them being specifically for hemodialysis treatment. By including all of them, the estimated index decreased for each (compared to separate PCG and DCG model), as they were largely overlapping and therefore explaining the same variation. As the authors point out, the assignment into both PCG and DCG did not lead to substantial overcompensation, since such patients were sicker and costlier on average.

The research paper by Farley et al. (1996) addressed specifically the risk adjusted payments for Medicare end stage renal disease (ESRD) program. The authors developed a modified version of capitation payments which distinguished between the patients on dialysis and the patients with functioning graft. Additionally, the lump sum payments were designed to compensate for incremental costs incurred by the transplantation, the graft failure, or extremely high-cost individuals. The risk adjustment model included multiple variables, such as age, gender, years of renal failure, or presence of diabetes. The model managed to explain up to 25 % of the cost variation. The improvement of the capitation payment was believed to positively influence the treatment choices, address the specific needs of ESRD population, and protect the health funds against substantial financial risk.

Similarly, Levy et al. (2006) focused on the risk equalization among ESRD beneficiaries in the US Medicare. Since the authors believed that the treatment status is important for the cost prediction, they estimated three separate models for dialysis, transplant, and functioning graft. The models accounted for demographic variables, program eligibility, and diagnosis groups, following the risk equalization models employed in the US. The suggested model resulted in more accurate payments for ESRD patients compared to the demographic model, while enhancing incentives for the health care provision.

4 Methodology and data

This chapter provides a detailed overview of the provided datasets, performs preliminary analysis of costs, and describes the methodology used in the empirical part.

4.1 Data description

The dataset was provided by the General Health Insurance (GHI) company, which is responsible for the reimbursement of reported drugs and procedures on behalf of their insurees. The dataset captures years 2015-2018 and consists of patients who were treated with specific drugs and/or procedures related to chronic kidney disease (CKD). All individuals have been continuously insurees of the GHI throughout the observed period. In total, the dataset consists of 16,400 insurees with their medical records.

In 2018, the GHI was insuring approximately 5.9 million of individuals, which corresponds to the market share of 56.5 % (VZP ČR, 2018b). Although the GHI has a slightly different age profile compared to other insurance companies, Dungl et al. (2017) point out that the proportion of patients classified into PCGs is roughly the same for all insurance companies. Hence, we consider the sample to be representative of the whole population of the Czech Republic.

4.1.1 Insurees

Each insuree in the dataset is assigned a unique anonymized ID which merges all the medical records of the particular patient throughout the dataset. The demographic data about insurees contain gender, year of birth, and year of death, if any. Due to the sensitivity of the data, the place of residence was not provided. Although some studies use the place of residence in the PCG model to enhance its effectivity, the current model in the CR does not use this information and therefore we will not apply it, either.

The data is divided into separate years (2015-2018) and analysed individually. Patients who died during observed year are omitted from the analysis, since they might distort the comparability of annual costs among patients in two ways. Firstly, in case of death early in the year, the insuree incurs lower than average annual costs. As the exact date of death is not indicated in the data, an extrapolation to annual costs as employed by Hájíčková (2015) is not feasible. Secondly, some insurees might have incurred extremely high end-of-life costs before death (Duncan et al., 2019) and thus it is safer to exclude them in order to avoid inflation of the average costs.

4.1.2 Medical procedures

The dataset contains records of all dialysis procedures, the dates of their performance and quantities. There are six different codes for dialysis treatment as depicted in Table 8, appended with the total number of procedures performed in year 2017.

Table 8: Overview of dialysis codes and number of performed procedures, 2017

Code	*Name*	*Number of procedures*
18513	Continuous peritoneal dialysis (CPD)	46 591
18515	Automated peritoneal dialysis (APD)	40 934
18522	Chronic hemodialysis	115 716
18523	Chronic hemodialysis outside the dialysis centre	1 602
18530	Hemofiltration	276
18550	Hemodiafiltration	444 727

The last four codes correspond to hemodialysis (HD) with minor differences of the procedures (e.g. its duration and effectivity). Generally, hemodialysis cleans the blood through tubes inserted into veins. The blood is sent into a special machine that cleans it from waste and then sends it back to the blood circulation of patient. In most cases, hemodialysis is performed in hospital or specialized clinic with the frequency of 2-3 times a week. Hemodiafiltration (code 18550) is the most frequent hemodialysis method and is claimed to be the most efficient one (Nistor et al., 2015). The reimbursement of this procedure is limited in the CR by conditions imposed by insurance companies and thus it cannot be provided to all patients. The chronic hemodialysis (code 18522) is therefore the second most frequent procedure. Based on the individual's preference, the patient can be also educated to self-provide the dialysis at home (code 18523). According to dialysis specialists, the hemofiltration method (code 18530) is not used for chronic patients nowadays, since it is neither time effective, nor cost-efficient. Since hemofiltration claims are very rare in our dataset (only 276 procedures performed in 2017) and might have been reported by mistake, this code was omitted from the analysis.

Peritoneal dialysis (PD), including CPD and APD, is performed daily through a catheter permanently inserted into the abdomen of the patient. Although this type is less frequent than HD and is not suitable for all patients, its advantages may affect the patient's choice of treatment. First, it can be carried out at home by the patient, family or a trained nurse. It enables the patient to travel with all necessary equipment as opposed to HD, where a dialysis clinic must be searched in advance, making the travelling limited. Moreover, the dietary restrictions are lower for PD compared to HD. On the other hand, the patient can find the tube in the belly disturbing and there is a

certain possibility of infection development called peritonitis. Finally, not all patients feel comfortable to perform the procedure by themselves and thus they prefer to have it done by a professional instead. The difference between the two procedures (codes 18513 and 18515) is, that an automated peritoneal dialysis is performed during the night making the days of the patient dialysis-free as opposed to continuous peritoneal dialysis which typically takes place during the day. Nevertheless, the frequency of the continuous PD in the dataset is higher than the automated one.

The timing of the dialysis initiation and the choice of the method are not same for all. Instead of deciding solely on the rate of GFR, a subjective feeling of the patient, comorbidities, and symptoms such as nausea, fatigue, shortness of breath, or lack of appetite are also considered. According to the Canadian guideline (Nesrallah et al., 2014), the patient should be closely monitored when the eGFR declines under 15 mL/min per 1.73 m^2, however, the clinical indicators such as symptoms of uremia, fluid overload, hyperkalemia, or acidemia should be present to decide about the dialysis initiation (unless the eGFR drops below 6 mL/min per 1.73 m^2). The final decision also depends on the patient.

The beginning of the dialysis must be planned well in advance and the patient in so called predialysis period must be provided with all possible treatment options, including kidney transplant. In most cases, the decision between HD and PD depends on the personal preference. Many studies compared the impact of HD and PD on the health outcomes and mortality, but there is no clear evidence for the preference of one over another (Wong et al., 2018; Yang et al., 2015). While PD tends to be a cheaper option and some studies indicate that the patient satisfaction and the quality of life is higher (Jung et al., 2019; Sinnakirouchenan & Holley, 2011), HD is currently still favoured.

Looking at the frequency of the procedures, it can be indicated, that not all patients reached the amount corresponding to the whole year on dialysis. The reasons might be numerous: The patient could have started the dialysis for the first time during the observed year, or the treatment might have been intermitted after the health outcomes had substantially improved or worsened. It is also possible that the reported patient had acute kidney problems instead of chronic renal failure. Although acute kidney failure has its own reporting codes, it might have been wrongly reported as chronic kidney failure. We thus define a threshold of 40 HD or 90 PD a year, which both approximately correspond to the dialysis treatment for the period of 3 months (given that HD is received 3 times a week and PD daily). This will ensure that we end up with severe chronic patients, whose costs are high and comparable among each other. Moreover,

according to dialysis professionals, this should be the minimal observational period for determining how the patient reacts to the treatment. That is, during this period the patient might return to conservative treatment or die from serious complications.

Histograms of annual frequencies of hemodialysis and peritoneal dialysis for year 2017 are depicted in Figure 4 and Figure 5.

Figure 4: Histogram of hemodialysis frequency, 2017

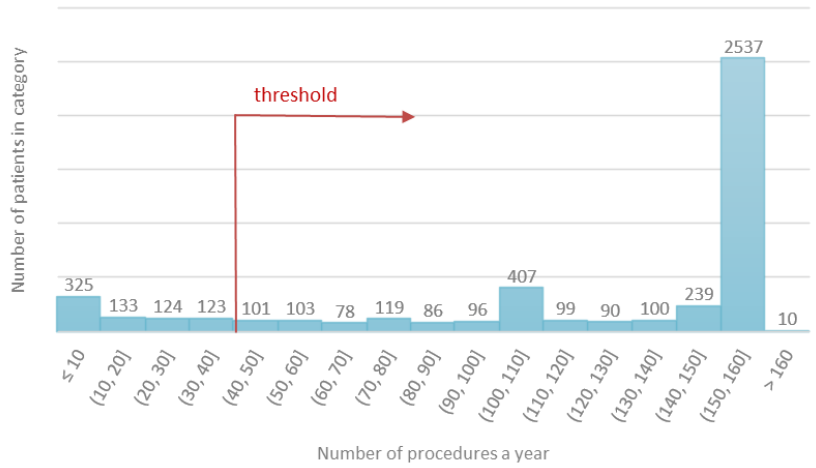
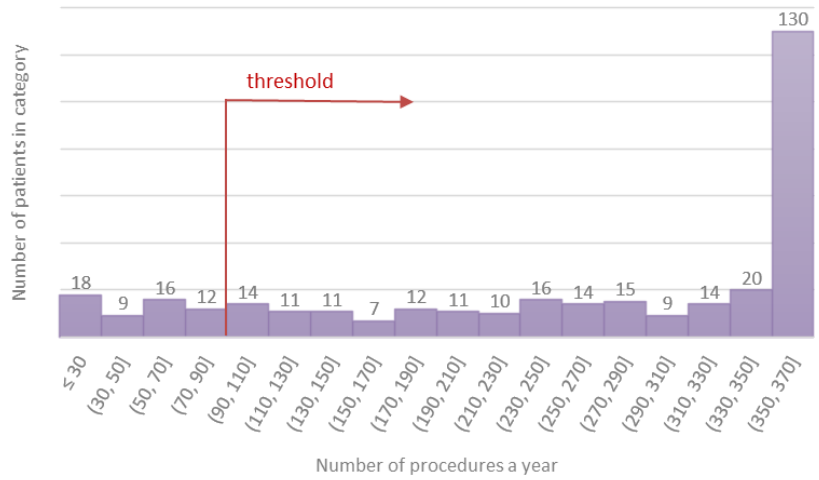


Figure 5: Histogram of peritoneal dialysis frequency, 2017

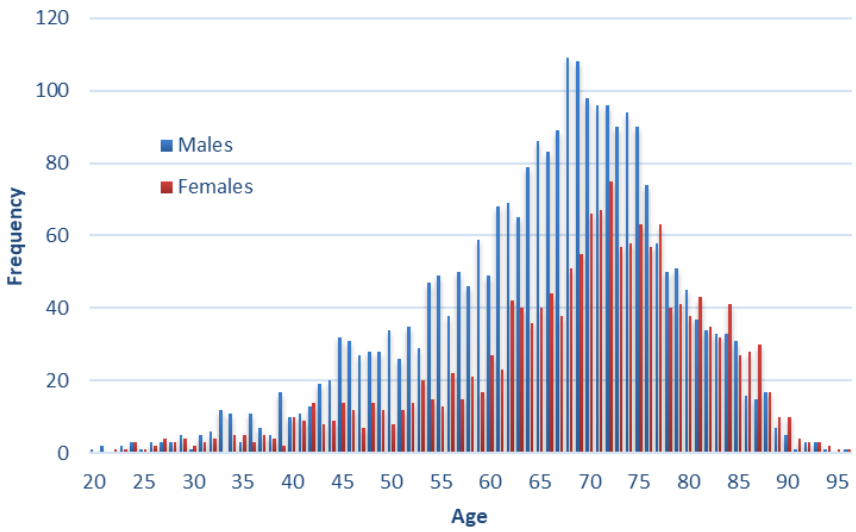


The standard frequency of hemodialysis is 3 times a week. Considering that a year has approximately 52 weeks, this corresponds to 150-160 procedures annually. As shown in Figure 4, the highest peak corresponds to this threshold. The second peak can be found at 100-110 procedures, which is consistent with receiving hemodialysis 2 times a week. Similarly, in Figure 5, the most frequent annual number of procedures falls into category 350-370, corresponding roughly to one dialysis a day. These could be the thresholds in case we wanted to limit the dataset solely to patients who received the

dialysis throughout the whole year. However, this would lead to the loss of considerable part of the dataset. Therefore, we decided to lower the threshold to include also the patients who received less than annual number of procedures, but still represent high cost burden for the insurance company.

Figure 6 shows the age structure of patients on dialysis in 2018 following the set threshold. The patients on dialysis are mostly represented by older population, since the disease is usually associated with high age and comorbidities. The frequency peak for men is reached at the age of 65-70, whereas for women somewhat later at 70-75. Renal failure is more frequent among men in almost all age categories except for the eldest population, being the consequence of higher age expectations for women.

Figure 6: The number of dialysis patients (PD ≥ 90 or HD ≥ 40) according to the age and gender, 2018



4.1.3 Pharmaceuticals

Chronic kidney disease is associated with many complications such as mineral and bone metabolism abnormalities, disbalance of blood pH, anaemia, neurological disorders, lowered immunity and more (Ryšavá & Brejník, 2018). Some of these can be regulated by proper dietary restrictions with limited intake of some minerals. Others are drug-compensated. In worse cases, renal replacement therapy is necessary, but even then, diet and pharmaceuticals remain important part of the treatment (Askar, 2015).

The PCG methodology assigns the patient to a PCG based on the consumption of specific pharmaceuticals. In case of renal failure, the type and the number of drugs consumed varies substantially among patients. Therefore, the choice of the unique ATC group is rather complicated. Following the model from the Netherlands, the PCG

methodology in the CR uses two ATC groups for renal failure – B03X and V03AE. The threshold of drug consumption for the PCG classification is specified annually by the Ministry of Health and applies to the sum of the respective ATC groups (Vyhláška č. 114/2019 Sb.).

The dataset provided by the GHI captures the individual consumption of drugs belonging to one of these two ATC groups. Each observation in the dataset contains a unique drug code, the anonymized ID of the insuree, the date of the prescription and the amount prescribed expressed in the units of daily defined doses (DDD). For each insuree, the prescribed DDDs were summed to obtain the total annual consumption.

The ATC code B03X corresponds to antianemic preparations other than: iron, vitamin B12 and folic acid. All drugs in this group are based on erythropoietin and are used for treatment of anaemia. In some cases, patient's condition improves with the dialysis treatment and thus leads to lower drug usage. The problem with the use of B03X for identification of CKD is that these pharmaceuticals are also used for other conditions, e.g. for oncological diseases. The second ATC group V03AE stands for drugs for treatment of hyperkalemia and hyperphosphatemia (also called binders). Their prescription depends on the dietary habits of the patient and on the levels of potassium and phosphate in blood. The main advantage of this ATC group is that it is highly specific for CKD patients[6].

Classification of the insuree into the right PCG might fail in case that the drug consumption is insufficient or none. Figure 7 obtained from our dataset shows that almost half of the patients on dialysis had combination of both drug types, while as much as 10 % did not consume any of these drugs in 2017.

Figure 8 depicts the histogram of drug quantities (sum of B03X and V03AE) prescribed in 2017. A substantial part of the users did not reach the threshold of 181 DDDs (represented by the first two columns), nevertheless, 59 % of them already received dialysis. We believe that patients who are on chronic dialysis should be classified into the PCG, even in case they did not reach the pharmaceutical threshold. The reason is that these patients present high costs which are not going to be compensated to insurance companies under the current system.

[6] SÚKL, lists of prescriber and indication restrictions: http://www.sukl.cz/sukl/seznam-cen-a-uhrad-lp-pzlu-k-1-1-2020

Figure 7: The drug consumption of patients on dialysis, 2017

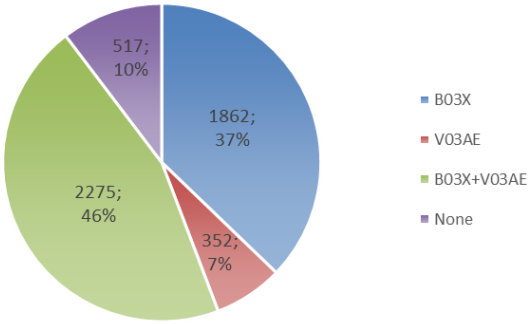
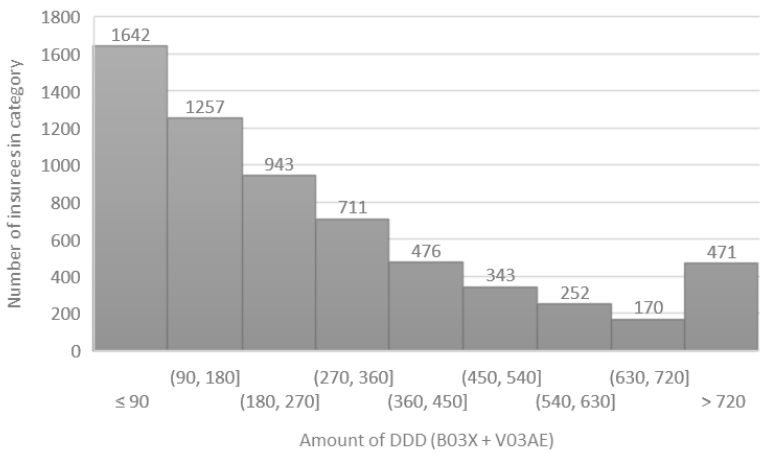


Figure 8: Histogram of drug consumption, ATC groups B03X + V03AE, 2017



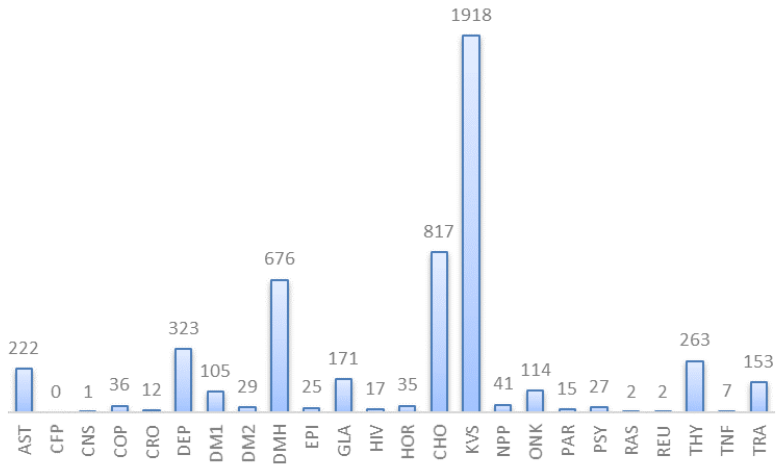
4.1.4 PCGs

For years 2017 and 2018, the dataset provides information about all PCGs the insurees were assigned to, based on their pharmaceutical consumption throughout the respective year. There are 25 different PCGs in total (see Appendix A 2 for list of PCGs with the classification criteria) including the group for renal failure (REN). The combination of chronic kidney failure with other chronic diseases is not an exception and their presence can further increase the health care costs. This is also reflected by the PCG model which allows for classification into more than one PCG.

The number of GHI insurees assigned to REN equals 3,255 and 3,231, for years 2017 and 2018, respectively. The most frequent comorbidities related to CKD are cardiovascular diseases (KVS), representing almost 60 % of patients in REN (see Figure 9). This is not surprising, since the risk of cardiovascular event increases with the decline in kidney function. Hypercholesterolemia (CHO) is the second most frequent disease, even though the association with the renal failure is not direct (Kaysen, 2007). The third frequent PCG is diabetes with hypertension (DMH), the

presence of which is observed in 20 % of REN patients. This follows from the fact that both diabetes and hypertension are among the most important risk factors for CKD development (Ryšavá & Brejník, 2018).

Figure 9: Frequency of PCGs for patients classified in REN, 2017



4.1.5 Costs and preliminary analysis

The dataset does not provide the real health care costs paid by the insurance company, but rather uses cost estimates. The estimates are based on the reported health care consumption of each insuree - inpatient and outpatient care, medical procedures, prescribed drugs, and other medical devices. The cost estimates are calculated as a multiplication of the number of points in different health care segments and their prices, as specified in the reimbursement decree published annually by the MoH. Although in some cases the real costs paid by the insurance company differ from the estimates, the Czech PCG model uses the same methodology for the cost valuation and thus it is suitable for our analyses of the model's efficiency.

Figure 10 depicts the histogram of costs for the sample in 2017. As argued in the Literature review, the costs are skewed to the right as most of individuals incur low to medium health care costs. The highest frequency occurs for the lowest cost group (below 100 thousand CZK), and it tends to decrease with higher costs. Table 9 provides the summary statistics for 2017 sample, which is later used in the regression analysis.

Figure 10: Histogram of annual costs, 2017 sample (skewness to the right)

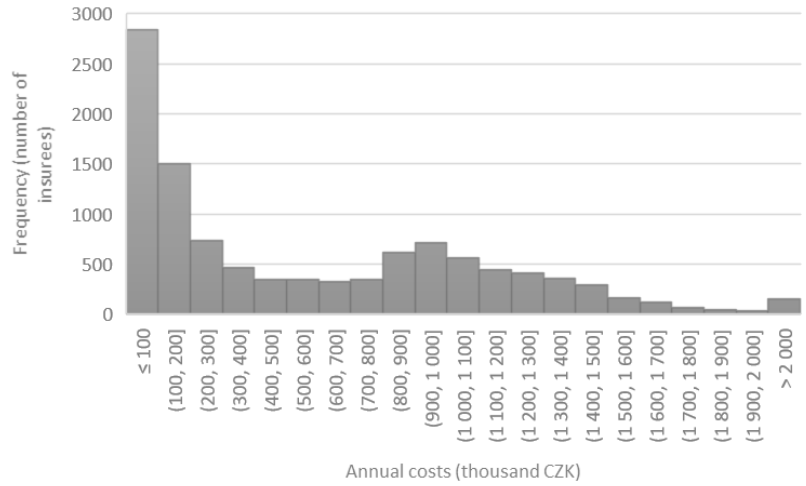
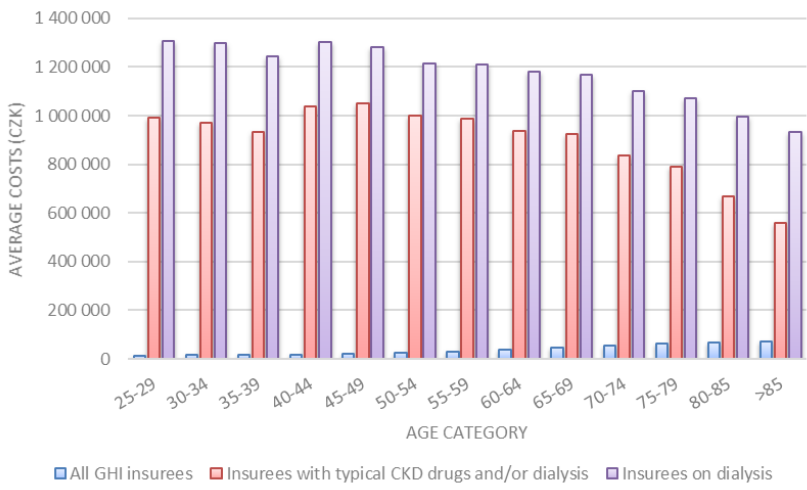


Table 9: Summary statistics of health care costs, 2017 sample

Min.	1st qu.	Median	Mean	3rd qu.	Max.
600	95 941	380 045	591 017	991 188	12 858 304

The mean costs calculated for insurees who consumed drugs and/or procedures typical for renal failure in 2018 are more than 30 times higher compared to the general population of GHI insurees (VZP ČR, 2018a). Figure 11 compares average health care costs of all GHI insurees with the average costs of CKD patients from our dataset (age groups below 25 were dropped due to unrepresentative number of CKD cases). Given that the general population of insurees contains also healthy individuals, the differences in costs are extremely large for all age groups. The third column depicts the patients on dialysis only, indicating that the average expenses are even greater for this group.

Figure 11: Comparison of mean costs for CKD and general GHI population, 2018

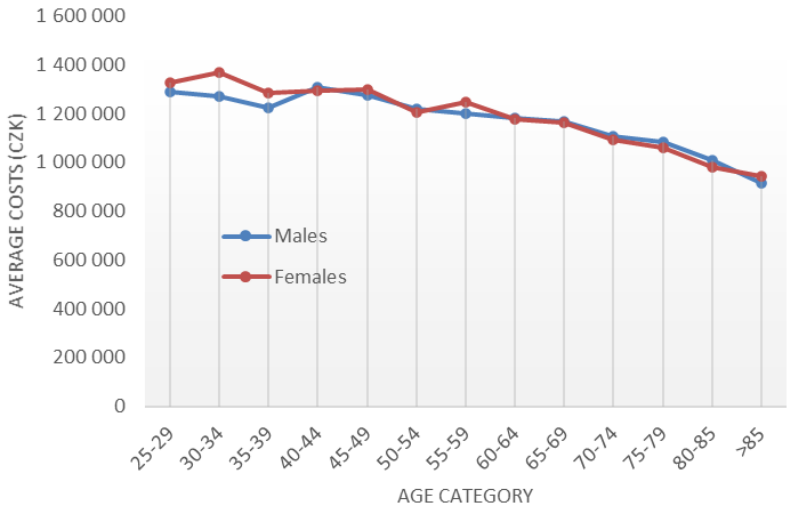


Source: Ročenka za rok 2018, VZP ČR (2018), and own data

Figure 11 also shows the relationship between costs and the age of insurees. While the average health care costs of the Czech population are generally increasing with the age (Bělohorský, 2018), the average expenses for CKD patients are decreasing in older age groups (approximately above the age of 70). Such trend seems counterintuitive since older patients usually experience more health complications, which is reflected by higher health care costs. The decrease in costs could possibly be the result of different approach to renal failure patients, where younger and more perspective individuals may obtain more intensive and cost demanding treatment. Other explanation could be that the most complicated patients tend to decease earlier, and thus only healthier CKD patients remain in higher age groups. Lastly, younger people are often eligible for a kidney transplant, in which case both the preparation and the surgery induce very high expenses.

Concerning the gender cost analysis, Figure 12 shows that for all age groups among dialysis patients, the average costs are more or less the same for both males and females. As suggested earlier, the costs are decreasing with the age for both genders.

Figure 12: Average costs of dialysis patients according to gender and age, 2018

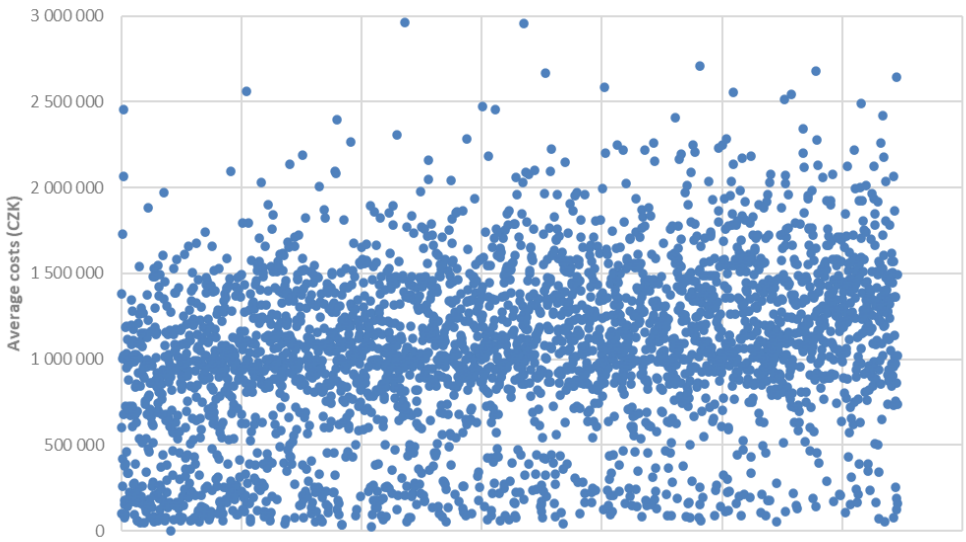


Comparing the costs among PCGs, according to Dungl et al. (2017) who used data from insurance companies for 2012, renal failure is the most expensive PCG regarding the average costs. On the other hand, given that renal failure is not as frequent as other chronic diseases (such as diabetes, hypercholesterolemia, or cardiovascular diseases), the total incurred costs are not among the highest.

When analysing the cost profile of REN PCG individually, Figure 13 shows that the expenses are mostly clustered around the mean, corresponding to 1.06 mil. CZK in 2018. However, it can be argued that the expenses lower than 0.5 mil. are frequent as

well, creating less evident cluster at lower cost levels. This reveals the cost heterogeneity of the PCG REN, possibly as a result of classification based solely on drug consumption, which reflects the expenses insufficiently. Indeed, the average cost of patient in REN who did not receive any dialysis treatment in 2018 is slightly below 0.5 mil. CZK. Based on this fact, the separation of REN group into two - patients with and without dialysis treatment – will be suggested in the methodological part.

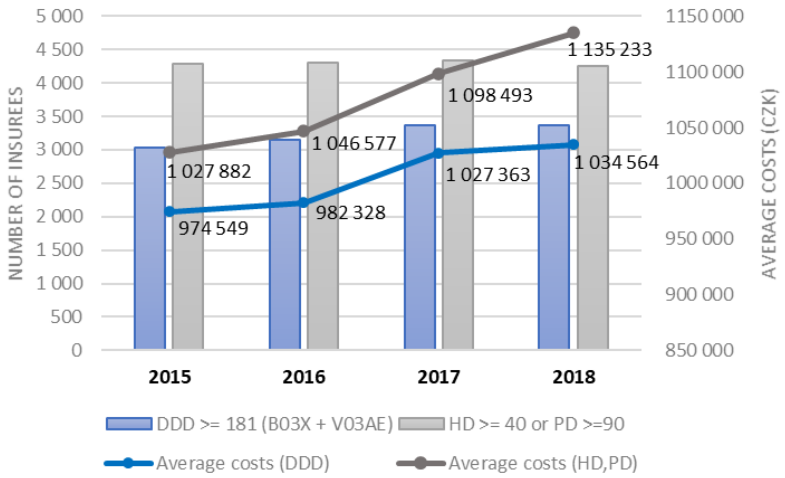
Figure 13: Cluster analysis of individual costs in the REN group (upper outliers not displayed), 2018



The extremely high costs are not exceptional, either. As many as 2.7 % of dialysis patients exceeded the total costs of 2 mil. CZK in 2018. The upper outliers might potentially inflate the average costs and make the prediction of expenses difficult. The problematic of outliers will be addressed later (see chapter on Heteroskedasticity and outliers in Methodology description).

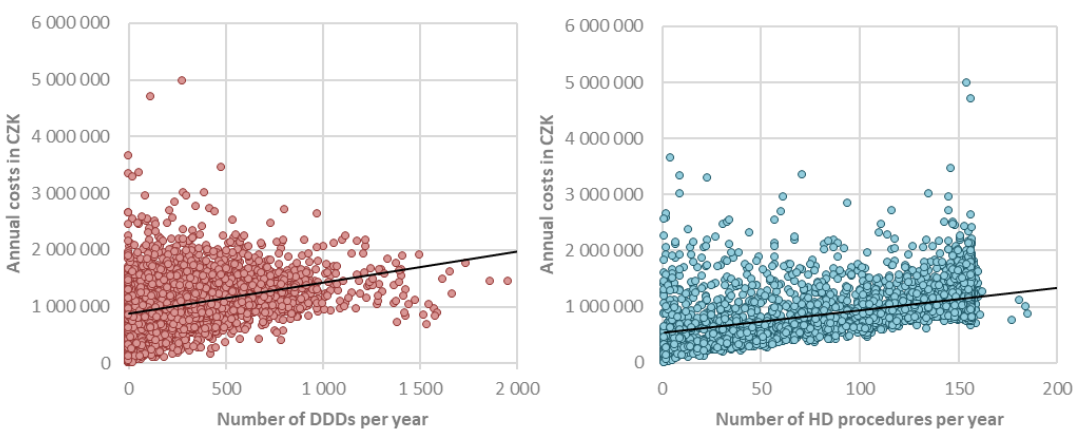
Figure 14 provides comparison of two classification methods that will be used in the analysis: The consumption of pharmaceuticals ≥ 181 DDD and the dialysis procedures with HD ≥ 40 or PD ≥ 90. So far, we presented mainly the statistics for the most recent years 2017 and 2018. Figure 14 shows the evolution of the number of insurees and their average costs during the whole period 2015-2018. Clearly, the size of both identified groups remains relatively stable throughout the years, while the average costs tend to increase in both cases. Note that the classification based on dialysis captures more insurees, while also being more expensive on average.

Figure 14: The evolution of CKD population and of average costs under different classification methods (drugs vs. dialysis procedures), 2015-2018



Additionally, Figure 15 inspects how the amount of drug consumption (typical for CKD) and the amount of dialysis procedures relate to the health care costs. Both graphs indicate a positive linear relationship, which implies that the use of either indicator have some justification in the cost prediction. Looking at the graphs, the dialysis procedures seem to follow the path more evidently. Both trends are significant at 1 % significance level as verified by a simple regression of costs on the respective health care consumption (drugs or dialysis). The regression using dialysis procedures, however, explains much more of the cost variation (41 % vs. 16 % measured by R^2).

Figure 15: The relationship of costs and health care consumption (upper outliers not displayed), 2017



4.2 Methodology description

The use of the OLS and WLS for the regression analysis was suggested in the Literature review, due to its advantageous features and easy interpretation. In case of WLS, the weights in the regression account for an individual eligibility for a given period, usually expressed as the number of months of enrolment. As already mentioned in the chapter on Data description, our dataset contains only individuals who were continuously enrolled at the GHI throughout the observed period; at the same time, we decided to exclude individuals, who passed away in a given year. Consequently, we can employ the OLS, since all insurees included in the dataset were enrolled for the whole year.

The analysis is carried out in 3 consecutive steps:

1. The model is estimated using the data on costs and risk adjusters from year t.
2. The coefficients and the R-squared are obtained for the model from year t.
3. The costs for year t+1 are predicted and compared with the real ones.

It is possible to use the coefficients calculated in the first step to compute the risk indexes as described in the chapter on PCG model in the Czech Republic. Nevertheless, we will not carry out the index computation, as we are interested in the models' performance rather than in the actual fund redistribution. Moreover, since we have the data from one insurance company only, we are not able to investigate the allocation of funds among different insurance companies.

The data from 2017 are utilized for the model's estimation and the predictions are made for year 2018, as these datasets are most up-to-date and include all indications of PCGs. The following subchapters present the suggested models and describe the methods used for their evaluation.

4.2.1 Model determination

Model 1

The first model is a simple demographic model using only age and gender for the cost prediction. The purpose is to show how the model performs without any health-based risk adjusters. The number of age/gender groups is markedly reduced compared to the official PCG methodology due to the low number of young insurees in the sample. For each gender, the age groups are formed in 5-year intervals, starting at the age of 35 (the first group captures age of 0-34) and ending at 85 (the last group captures age of 86 and more). Thus, there are 24 age/gender groups in total.

The regression model for year t is given as

Eq. 1: Model 1

$$h\_costs_{i,t} = \alpha + \beta_j age\_gender_{j,i,t} \qquad j = 1,\ldots,23$$

The dependent variable reflects costs in year t and the explanatory variables are 23 dummies representing the age/gender groups. For illustration, i[th] insuree being classified into group $j = 1$ means that the variable $age\_gender_{1,i,t}$ equals 1, while all other dummy variables $j = 2,\ldots,23$ equal 0. The number of included dummies accounts for 23 classes, since one class (the base group) must be omitted not to induce perfect collinearity. We decided to establish males at the age of 70-74 as the base group. The variable α represents the intercept.

The cost prediction for year t+1 follows Eq. 1, using the coefficients estimated in year t and the age/gender groups for year t+1; that is, the individual is reclassified in case that the additional year resulted in the age value belonging to different age group (for example if an individual was 34 in year t and reached 35 in year t+1).

Model 2

The second model includes PCGs in addition to age/gender groups from the Model 1. The same groups are used as displayed in Appendix A 2, except for the diseases with poor representation in the sample, thus 22 PCGs are included.

In this model, the definition of PCG REN is based solely on the consumption of typical drugs, following the official methodology. The threshold of consumption is firstly set to 181 DDDs, similarly to other PCGs (labelled as Model 2a). Additionally, we lower the drug consumption necessary for the classification into REN to 91 DDDs (labelled as Model 2b), which corresponds to the consumption of typical drugs for approximately 3 months, as proposed by Lamers & Vliet (2003), to test for the sensitivity of the model on modifications in classification criteria.

The regression model for year t is given as

Eq. 2: Model 2

$$h\_costs_{i,t} = \alpha + \beta_j age\_gender_{j,i,t} + \gamma_k PCG_{k,i,t}$$

$$j = 1,\ldots,23;\ k = 1,\ldots,22$$

Apart from the same variables included in Model 1, new 22 dummies reflecting the PCGs are incorporated. If an i[th] insuree belongs to the k[th] PCG, then $PCG_{k,i,t} = 1$, and

0 otherwise. Note that belonging into more than one PCG is possible. In this case, all 22 disease groups are included, since the base group is no PCG, i.e. when $PCG_{k,i,t} = 0$ for all k.

The cost prediction for year t+1 uses the estimated coefficients from Eq. 2 and age/gender groups for year t+1. As opposed to the age, which is well known in advance for the upcoming year, the classification into PCGs for year t+1 is unknown. Therefore, we use the PCG classification from year t as a proxy for the next years' cost prediction.

Model 3

The third model corresponds to the Model 2, as for the number of age/gender and disease groups. The difference lies in the specification of the REN group. Instead of drug consumption, the renal failure is identified based on the number of dialysis procedures in given year. As we justified in the Data description, the threshold for REN classification is set to HD ≥ 40 or PD ≥ 90. Therefore, instead of PCG REN, a new dialysis-based variable is introduced (called REN_dial), taking on value 1 in case the threshold is reached and 0 otherwise. All other PCGs are defined standardly using the drug consumption ≥ 181 DDDs.

The regression model for year t is given as

Eq. 3: Model 3

$$h\_costs_{i,t} = \alpha + \beta_j age\_gender_{j,i,t} + \gamma_k PCG_{k,i,t} + \delta REN\_dial_{i,t}$$

$$j = 1, \dots, 23;\ \mathrm{k} = 1, \dots, 21$$

The methodology of the cost prediction for year t+1 applies from the Model 2. Similarly as for the PCGs, the classification into REN for year t+1 is based on the data from year t (i.e. on the number of dialyses).

Model 4

Finally, both approaches to REN classification are combined in the last model. As discussed in the Data description, we believe that the CKD population consists of two types of patients with considerably different costs: The patients on dialysis and the patients on conservative treatment (using drugs only). The latter might be also referred to as predialysis treatment. Indeed, the data from 2018 suggest that the average costs of patients who exceeded the threshold of 181 DDDs, but did not obtain any dialysis procedures, equal less than half of the average costs for the whole REN group. This led us to define two different groups reflecting CKD patients:

1. Predialysis group: Drug consumption ≥ 181 DDDs and HD < 40 and PD < 90

2. Dialysis group: Drug consumption not considered; HD ≥ 40 or PD ≥ 90

The first group is represented by the new variable REN_pre and the second by REN_dial (same as in Model 3). We believe that the two defined groups are more cost homogenous and would lead to the model's improvement.

The regression model for year t is given as

Eq. 4: Model 4

$$h\_costs_{i,t} = \alpha + \beta_j age\_gend_{j,i,t} + \gamma_k PCG_{k,i,t} + \delta_1 REN\_pre_{i,t} + \delta_2 REN\_dial_{i,t}$$

$$j = 1, \dots, 23;\ k = 1, \dots, 21$$

4.2.2 Measures of model's performance

R-squared

R-squared (R^2) is the most common measure in the risk adjustment model's evaluation. The R^2 of regression, also referred to as the coefficient of determination, is interpreted as the fraction of the variation in dependent variable, that is explain by independent variables (Wooldridge, 2009). In econometrics, it is defined as

Eq. 5: R-squared

$$R^2 = \frac{SSE}{SST} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

SSE corresponds to the explained sum of squares and is computed as a sum of squared differences between the fitted values $\hat{y}_i$ and the sample average $\bar{y}$. Similarly, the total sum of squares (SST), uses the difference between the sample values y_i and the sample average $\bar{y}$. The value of R^2 is always between 0 and 1, as the SSE can never exceed the SST. Generally, the higher the value of R^2, the more variation in the dependent variable is explained and thus the better the choice of the explanatory variables. The R^2 is often interpreted as a percentage after multiplied by 100.

Nevertheless, we should be careful with the R^2 interpretation when comparing the performance of different models. From the general properties of R^2, it follows that the value tends to increase with every additional variable (or at least, it never decreases). This means that a small positive change in R^2 after new variable is added, does not have to necessarily indicate the model's improvement. To address this issue, the adjusted R-squared is sometimes used instead, although, it is not generally considered

a better measure than simple R^2 (Wooldridge, 2009). Both estimates are routinely reported by the econometric software along with the estimated coefficients and standard errors. Additionally, we compute another R^2 using the predicted costs and the actual costs for year t+1, to show how well the variance in the next year's costs is explained by our predictions.

Apart from the R^2, the models' accuracy is also verified by the means of the likelihood ratio (LR) test. The LR test compares the goodness of fit of two models – the restricted and the unrestricted, employing their log-likelihood functions. Under the null hypothesis, the dropped variables are not important in the model and thus the restricted model is more appropriate. Conversely, if H_0 is rejected (the LR test gives a significant result), the unrestricted model is significantly more accurate (Wooldridge, 2009).

MPE, MAPE, MARE

The forecasting power of the model can be assessed by three similar measures. Define C_pred_i as the costs predicted in year t for year t+1 for the ith insuree, C_act_i as the actual costs in year t+1 for the ith insuree, and n the number of insurees in the sample.

The Mean Prediction Error (MPE) is defined as

Eq. 6: MPE

$$MPE = \frac{\sum_{i=1}^{n}(C\_pred_i - C\_act_i)}{n} = \frac{C\_pred\_total}{n} - \frac{C\_act\_total}{n}$$

The MPE shows how well the model predicts the mean of the actual costs. The closer the MPE to zero (either in positive or negative values), the better the prediction.

The Mean Absolute Prediction Error (MAPE) is defined as

Eq. 7: MAPE

$$MAPE = \frac{\sum_{i=1}^{n}|C\_pred_i - C\_act_i|}{n}$$

The MAPE measures the average absolute difference between the predicted and actual costs. As the difference is always captured as a positive value, the predictions higher or lower than the actual costs do not cancel out, as opposed to the MPE (Fishman et al., 2003). In other words, the MPE allows for the losses to be offset by the profits resulting from inaccurate predictions, while the MAPE evaluates the average error (both positive and negative). Lower values indicate better forecasting performance.

The Mean Absolute Relative Error (MARE) is defined as

Eq. 8: MARE

$$MARE = \frac{\sum_{i=1}^{n} \frac{|C\_pred_i - C\_act_i|}{C\_act_i}}{n}$$

The MARE can be interpreted as the average absolute difference between the predicted and actual costs relative to the actual costs. It can be used to assess the relative importance of errors, which makes the comparison of models more intuitive (Hájíčková, 2015). The model's performance improves as the MARE approaches zero.

4.2.3 Heteroskedasticity and outliers

Heteroskedasticity is present whenever the variance of the explained variable is dependent on explanatory variables and is not constant. The heterogeneity in individual health care costs and the skewness of costs are expected to result in heteroskedasticity in our models. The presence of heteroskedasticity does not induce the bias of OLS estimators, nor has it an influence on R^2. On the contrary, it does affect the estimates of standard errors and the statistics based on it (e.g. t-statistics), which are therefore biased. To address this issue, we compute the heteroskedasticity-robust standard errors, which are asymptotically valid for any form of heteroskedasticity, provided we have a large enough sample (Wooldridge, 2009).

Extremely high costs occurring in few observations might negatively affect the estimation process and make the fit of the model less accurate. Since we want to see how well we can predict the future costs, we keep the outliers in the data for most of the analyses. Nevertheless, the outliers will be detected in the last section and the best performing model will be refitted without outliers to see whether it improves. For outlier's detection, the boxplot method is utilized. The boxplot standardly depicts the lower quartile (Q1) at 25th percentile, upper quartile (Q3) at 75th percentile and the median of the data. The middle 50 % of scores (the range between Q1 and Q3) is referred to as an Interquartile range (IQR). For the detection of outliers, the upper limit will be set as Q3 + 1.5*IQR and the observations larger than the upper limit will be dropped.

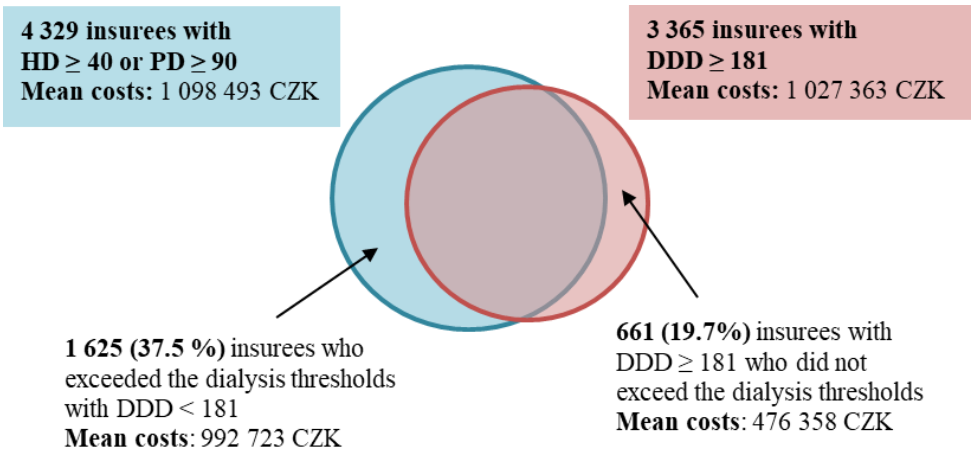
5 Results

5.1 Sensitivity of identification methods

The comparison of identification methods suggests that drug consumption is not a good indicator of renal disease, nor a suitable cost predictor. Figure 16 illustrates the overlap of groups identified based on drug consumption (181 DDDs) and dialysis procedures (40 HD or 90 PD) in 2017. The dialysis criteria capture not only higher number of patients, but also higher average expenses. As presented in the diagram, the individuals not captured with the use of dialysis criteria have average costs lower than 0.5 mil. CZK, while the patients omitted when using only drug consumption account for almost 1 mil. CZK in average costs.

Assuming the patients on dialysis for more than 3 months are truly the ones suffering from renal failure, we can observe the ability of drug consumption method to identify these individuals. Using the threshold of 181 DDDs, the sensitivity, defined as the probability of correct classification into REN when a patient truly suffers from renal failure, corresponds to 62.5 %. Alternatively, the specificity, defined as the probability of correct non-classification into REN when patient does not suffer from renal failure, accounts for about 90 %. This implies that while the current PCG methodology primarily does not include wrong individuals (as indicated by high specificity), it fails to capture a substantial part of renal failure patients (as indicated by low sensitivity).

Figure 16: The comparison of identification methods, 2017



5.2 Comparison of models

Table 10 summarizes the estimation results for five models suggested in the methodological part. Starting with Model 1, the variation of costs explained solely by demographic variables reaches less than 2 %. As anticipated, the inclusion of health-based indicators improves the explained variation substantially in remaining models, reaching as much as 49 % for the best fit. Clearly, the definition of the REN variable is crucial. In Model 2, where solely drug consumption is considered, the R^2 reaches 26 % for the threshold of 181 DDDs and increases to 30 % when threshold lowered to 91 DDDs. The largest improvement in R^2 is observed when the REN group becomes related to dialysis procedures instead of drug consumption; Model 3 manages to explain 48 % of the cost variation. The inclusion of additional variable capturing patients in predialysis (using drugs only) further improves the R^2, although by less than one percentage point. Note that since the last Model 4 combines both approaches, it captures all patients represented by the union of the two sets depicted in Figure 16.

Concerning the age/gender analysis, as the base group (excluded from variables) accounts for men at the age of 70-74, the other age/gender estimates are interpreted as additional costs compared to that group. For men, the interpretation of estimates is straightforward, since they represent only the age difference. For women, the coefficients reflect both age and gender difference when compared to the base group. The marginal costs for female gender can be viewed as the difference between the male and female coefficient estimated for the same age group.

The positive values for lower age groups indicate that younger insurees are more expensive on average. Conversely, all age groups higher than 75 have negative estimates for both genders, suggesting that health care costs decrease with age. The lowest costs in all models are assigned to the oldest individuals (> 84) for both genders. On the contrary, the highest average expenses are estimated at 50-54 for men (except for last two models, where second highest), while for women, at the age of 45-49.

In case we are interested in average costs for specific age/gender group, the intercept, which represents the average costs of the base group, and the corresponding estimated coefficient, are added up. For illustration, using Model 1 estimates, the average annual costs for men at the age of 80-84 are calculated as: (595,365 – 139,941) CZK = 455,424 CZK. Since Model 1 takes into account only age and gender, in remaining models, the estimates representing the PCG(s) the individual belongs to need to be added as well. Only in case the individual does not belong to any PCG, solely the intercept and the age/gender estimate are summed.

Table 10: Summary of results for suggested models, 2017

	Model 1	Model 2a	Model 2b	Model 3	Model 4
Intercept	595 365.25 ***	334 346.51 ***	281 555.91 ***	193 516.01 ***	171 603.33 ***
M < 35	18 808.97	93 087.76	99 033.44	135 333.76 *	138 632.31 *
M 35-39	48 109.97	90 119.25 *	81 828.10 *	34 312.51	46 977.23
M 40-44	67 746.18	99 754.65 **	114 863.94 **	51 660.91	56 758.35
M 45-49	140 386.70 ***	147 132.78 ***	153 583.14 ***	77 829.74 **	86 560.15 **
M 50-54	165 386.54 ***	160 380.72 ***	155 201.35 ***	101 378.62 *	103 571.97 **
M 55-59	124 487.39 ***	125 160.50 ***	124 384.47 ***	56 306.83 *	62 924.54 **
M 60-64	59 166.69 *	46 373.06 *	47 765.39 *	27 246.52	29 054.99
M 65-69	39 685.63	62 799.90 **	50 850.70 *	36 678.42 *	41 168.62 *
M 75-79	-9 583.03	-21 421.10	-34 538.61	-35 721.73 *	-37 795.29 **
M 80-84	-104 865.31 ***	-93 313.21 ***	-104 319.78 ***	-83 747.50 ***	-83 769.58 ***
M > 84	-163 004.62 ***	-150 670.93 ***	-154 343.38 ***	-95 022.90 ***	-96 313.59 ***
F < 35	-17 617.46	80 567.63	68 920.55	110 006.35 *	112 944.51 **
F 35-39	13 456.40	30 045.72	44 361.23	42 727.85	49 723.52
F 40-44	9 753.52	12 233.60	-36 883.12	53 322.07	52 383.55
F 45-49	151 020.11	164 052.18	151 509.71	149 020.47	156 536.26
F 50-54	60 226.96	72 878.93	79 874.45	106 742.26	107 394.26
F 55-59	48 610.54	70 204.11	80 248.64	81 580.21 *	87 913.71 *
F 60-64	4 686.52	-1 302.77	-11 183.40	19 717.06	20 304.55
F 65-69	-34 458.82	-20 746.69	-32 160.42	7 331.24	11 691.63
F 70-74	-33 216.49	-26 487.46	-36 547.77	2 035.93	930.35
F 75-79	-93 817.81 ***	-96 882.77 ***	-101 351.08 ***	-42 285.46 **	-45 079.33 **
F 80-84	-139 940.54 ***	-140 501.56 ***	-154 474.42 ***	-69 677.12 ***	-72 708.59 ***
F > 84	-195 802.00 ***	-184 883.78 ***	-207 572.91 ***	-116 610.99 ***	-117 150.93 ***
AST		32 752.54	27 777.24	36 736.05 *	34 255.64 *
COP		33 940.02	25 545.10	17 932.38	25 860.57
CRO		62 464.42	-27 022.33	-15 296.51	-4 810.02
DEP		78 562.52 ***	88 501.07 ***	68 629.58 ***	67 063.70 **
DM1		203 127.38 ***	178 688.67 ***	37 042.73	41 399.90 *
DM2		21 259.82	12 463.39	34 948.23	29 301.79
DMH		-58 579.31 ***	-53 840.14 ***	-5 772.16	878.14
EPI		38 433.95	31 567.46	-2 736.27	2 925.55
GLA		-11 108.75	-18 339.38	-1 641.29	-3 182.12
HIV		264 204.91 **	246 436.48 ***	172 228.49 *	173 961.91 *
HOR		-60 962.65 *	-31 688.11	-4 744.40	-3 664.64
CHO		-947.66	162.49	-957.41	1 644.00
KVS		150 001.63 ***	119 288.82 ***	21 277.82 *	17 309.67 *
NPP		-13 946.42	13 531.24	32 025.51	25 026.57
ONK		170 911.88 ***	133 473.23 ***	442 671.06 ***	408 898.29 ***
PAR		347.96	2 928.31	23 186.92	28 887.30
PSY		-72 090.17 *	-75 616.27 *	-24 871.65	-24 056.29
REU		-30 176.67	-19 328.44	48 230.32	69 387.75
THY		33 535.03	35 737.41	32 026.37 *	30 388.10
TNF		79 993.22	63 321.71	187 636.07 ***	140 467.53 **
TRA		-64 232.54 **	-55 452.61 *	87 395.58 ***	90 340.83 ***
REN_181		587 467.80 ***			
REN_91			600 776.35 ***		
REN_DIAL				863 290.44 ***	884 232.18 ***
REN_PRE					214 854.46 ***
N	10 937	10 937	10 937	10 937	10 937
R2	0.0197	0.2622	0.2958	0.4809	0.4874
Adjusted R2	0.0176	0.2591	0.2839	0.4788	0.4852

*Standard errors are heteroskedasticity robust. Significance: *** p < 0.001; ** p < 0.01; * p < 0.05.*

Apparently, not many age/gender variables are statistically significant. The two (for men) and three (for women) highest age/gender groups are negative and significant at 1 % level in all models, even though their magnitude changes. Additionally, the three age groups between 45 and 59 are significant for males in all models at 5 % level. The remaining groups are statistically insignificant (at 5 % level), or their significance fluctuates between the models. The loss of statistical significance in some cases can be ascribed to the fact that the additional health-based variables explain the cost variation better.

The inclusion of PCGs improves the model significantly, as suggested by both R^2 and adjusted R^2, and additionally supported by the likelihood-ratio (LR) test. The LR test is significant at all statistical levels for all models including PCGs, which rejects the null hypothesis that the restricted model (being Model 1) is the "true" model. Similarly to the age/gender variables, not all PCGs are statistically significant. In fact, only 5 PCGs apart from REN are significant in all models at 5 % level: Depression, Cardiovascular diseases, Oncological diseases, HIV and Transplantations. The importance of these conditions can be a result of their high representation in the population, as is the case of the former two diseases, or a consequence of very high average costs, as is the case of the latter three. All these PCGs have positive estimates (except for TRA which changes signs between models), oncological diseases being the most expensive. If all other insignificant PCGs were excluded, the model's performance would not deteriorate, as indicated by negligible decrease in R^2 (estimation not displayed) and approved by the insignificant result of LR test. These results apply solely to our limited sample, where some of the PCGs are underrepresented. In general population, all PCGs tend to be positive and significant (Chochláčová, 2018; Hájíčková, 2015).

The variables capturing CKD patients are the most important, as they account for majority of cost variation. This can be demonstrated by fitting the model with REN_DIAL as the only independent variable, which still manages to explain 45 % of the variation in costs (see Appendix A 3 for estimation results). All specified REN variables are statistically significant at 0.1 % level and the estimates are positive and high in magnitude. The REN coefficient increases substantially with the introduction of dialysis-based approach, as dialysis procedures are able to capture more individuals with high costs. At the same time, the intercept, which corresponds to the average costs of individuals at the age of 70-74 not belonging to any specified disease group, decreases. The decline of intercept and some age/gender coefficients is a result of more expensive cases being reflected by the REN group(s) that better explain the variation.

Presenting the dialysis-based approach into last two models results in extensive improvement of model's performance – the variation explained by Model 3 is almost twice the variation explained by Model 2a. Additional variable REN_PRE in Model 4, accounting for the predialysis patients, seems to improve the model's performance only marginally, however, significantly according to the LR test between Model 3 and Model 4.

For illustration, consider a man at the age of 70-74 (base group) with the consumption of specific drugs for renal failure ≥ 181 DDDs. According to Model 2a, the expected average expenses are equal to 921,815 CZK. Now using the Model 4, if the man is on dialysis (classified into REN_DIAL), the expected costs are equal to 1,055,835 CZK. On the other hand, if the man is treated by drugs only (classified into REN_PRE), the expected costs amount to 386,457 CZK. This clearly demonstrates the difference between the two approaches and supports our hypothesis that solely drug consumption is not able to explain the cost heterogeneity in the CKD population.

5.3 Evaluation of predictive power

The predictive power of the model and the fit of the results on t+1 data are essential for model's evaluation. Firstly, the R-squared computed for t+1 represents how much variation in 2018 costs is captured by the predictions from 2017 data. Table 11 shows that the R^2 for predicted costs performs similarly as for the fitted models in 2017. For the first three models the variation explained is even higher, probably as a result of less outliers present in 2018 costs. The predicted R^2 increases with more advanced model, hence Model 4 is the best option with 48 % variation explained.

Table 11: R-squared: fitted model in t (2017) vs. prediction for t+1 (2018)

	Model 1	Model 2a	Model 2b	Model 3	Model 4
R-squared, t	0.0197	0.2622	0.2958	0.4809	0.4874
R-squared, t+1	0.0526	0.2759	0.3065	0.4730	0.4800

The measures of forecasting power depicted in Table 12 provide additional insight on how well the predicted costs match the actual 2018 costs. The MPE, corresponding to the average difference between the predicted and actual costs, indicates that all models underestimate the true costs. This could result in loss of insurance companies in case there is no ex-post compensation. According to the MPE, Model 3 performs the best with the lowest average loss equal to 118,755 CZK, however, we should not put too much weight on this measure. Less negative MPE might be a result of overestimation of other expenses, which cancelled out part of the underestimation, and generally, we want to avoid both.

The MAPE is considered a better measure since it captures the average size of the error, no matter if positive or negative. In this matter, the last two models appear to be the best predictors with the lowest MAPEs, Model 3 being slightly better.

Lastly, the MARE is the best measure for the comparison of models since it reflects the mean absolute error relatively to the actual costs. The MARE implies that Model 4 has the lowest relative error and thus the best forecasting potential.

Table 12: Measures of predictive performance

	Model 1	Model 2a	Model 2b	Model 3	Model 4
MPE	-119 322.5	-123 546.7	-124 423.9	-118 755.1	-119 387.8
MAPE	494 602.9	400 640.1	387 041.3	321 495	321 583
MARE	5.1071	3.4351	3.0781	2.1772	2.0893

Apart from the overall performance, we carried out more detailed analysis of specific cost groups using quantile approach. We analysed the fit of the predictions in 20 % of the most expensive and 20 % of the least expensive cases in 2018. As depicted in Table 13, the most expensive individuals are largely underestimated, in the worst case (Model 1) by as much as 986 thousand CZK on average. MAPEs are very close to MPEs in absolute values since the errors are usually negative for this part of the population. MAREs are smaller compared to the whole sample, because the errors are weighted by very high actual costs. On the contrary, the costs of the cheapest 20 % of population are overestimated by all models, as follows from the positive MPEs. MAREs are relatively high, as we weight the errors by low actual costs. Model 4 performs the best according to all measures, that is, it predicts the costs closest to the actual ones for both subsamples.

Table 13: MPE, MAPE and MARE for 80th and 20th percentile of 2018 costs

80th percentile	Model 1	Model 2a	Model 2b	Model 3	Model 4
MPE	-985 840.5	-791 771.6	-784 964.5	-663 555.7	-661 184.2
MAPE	985 840.5	792 485.1	785 114.3	664 839.7	662 305.8
MARE	0.5903	0.4599	0.4540	0.3698	0.3682

20th percentile	Model 1	Model 2a	Model 2b	Model 3	Model 4
MPE	510 956.5	343 171.1	318 517.4	189 749	182 781.1
MAPE	510 956.5	343 172.4	318 870.8	190 061.7	184 055.1
MARE	22.2573	14.6387	12.9002	8.8840	8.4042

5.4 Outliers' analysis

Lastly, we performed the boxplot analysis to identify the cost outliers and to evaluate their impact on the model's performance. The upper limit was set to 2.33 mil. CZK

corresponding to Q3 + 1.5*IQR (see Methodology description). Subsequently, 74 outliers exceeding the upper limit were detected, as depicted in Figure 17. The outliers were dropped from the 2017 data and the best performing Model 4 was refitted using the restricted sample (estimates provided in Appendix A 4). As expected, the R^2 for the fitted model increased extensively to 67.6 %, as the costs were much easier to fit in the absence of extreme values. Furthermore, the predictions explained 50 % of the cost variation in 2018, which is a better result compared to the complete sample (48 %).

Regarding the measures of predictive power (see Table 14), the overall average loss represented by the MPE is worse compared to all previous models. This is a consequence of omitting the outliers from the fitted model but keeping them in the actual data; the extreme costs in 2018 incur high losses against the predicted costs, resulting in large negative MPE. On the other hand, the MAPE and the MARE are lower than for the original models, implying that we managed to reduce the average (relative) absolute error. In the Czech risk adjustment, the outliers are usually not an issue, since the extreme costs are addressed by the ex-post compensation.

Figure 17: Boxplot with outliers, 2017

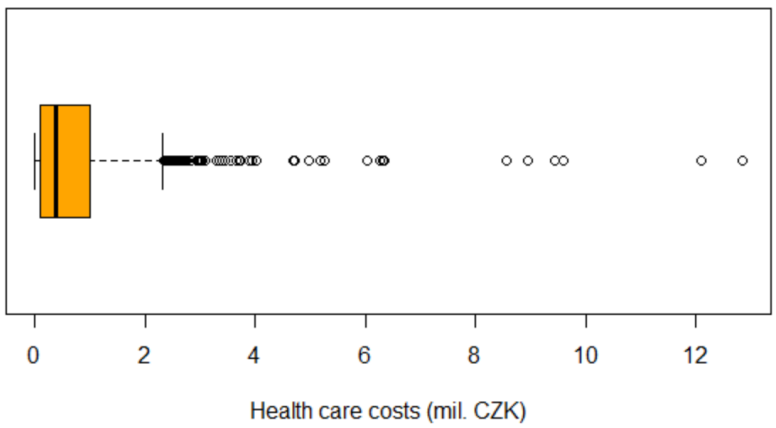


Table 14: Predictive performance of Model 4 without outliers

R-squared, t	0.6760
R-squared, t+1	0.5013
MPE	-141 209
MAPE	318 839
MARE	1.8225

6 Discussion

The results demonstrated a great potential of the PCG model for improvement from the view of the most expensive disease group – patients with renal failure. Firstly, lowering the drug consumption for classification of renal failure patient from 181 DDDs (currently used in the CR) to 91 DDDs improved the model's explanatory power. On the other hand, lower prescription threshold might include incidental drug users or even create a potential for manipulation of prescriptions as pointed out by Lamers & Vliet (2003). Thus, higher threshold, such as 121 DDDs, would be advised (in fact, 121 DDDs is currently a minimal threshold allowed in the Czech PCG model).

Next, we showed that the costs were reflected much better by the procedures than by the drug consumption and we verified that the introduction of dialysis procedures as an indicator for renal failure significantly improved the accuracy of predictions. Consequently, the risk indexes computed from the improved model could be more accurate and lead to better financial compensation of insurance companies for these patients. This is the most important outcome of our study since none of the studies on risk adjustment in the Czech Republic (nor abroad to our knowledge) considered reported procedures as a potential cost predictor. One of the limitations of this approach is its suitability only for diseases requiring typical and regular procedures, because the risk adjustment and the fund redistribution are carried out on monthly basis.

The data analysis showed that the average costs of patients on dialysis and patients treated solely by drugs differ substantially. We addressed this by establishing two separate REN groups reflecting the cost difference, by which we managed to further improve the model, particularly in terms of its forecasting performance. The main advantage of the methodology is that all necessary data are already being routinely collected by insurance companies and thus the model is easily applicable without additional data collection or extra costs.

The analysis additionally revealed that the expenses decrease for older CKD patients, which does not correspond to the trend of general population. Intuitively, older patients experience more health complications, which should result in higher health care costs. The reason for an opposite trend in case of renal failure might be the intensity of treatment which tends to diminish with high age, in order not to burden the patients with excessive amounts of drugs (that might be also contraindicated in combination with other diseases) and procedures. With regard to the patient's current prognosis, the

dialysis treatment can be reduced or even withdrawn, and instead, the patient obtains conservative treatment supported by palliative care (Fassett, 2014). On the contrary, younger patients are treated intensively, since the compensation of their condition and high life expectancy are of utmost importance. Another explanation of lower costs is the fact that older patients tend to adhere to dietary restrictions better (Kugler et al., 2005), and therefore there is a lower pharmaceutical intake (as also backed by the data). To address this specific trend in the context of risk adjustment and as a motivation for further research, we would suggest including the interaction of REN group with age as additional variable in the general PCG model.

Another suggestion for research extension would be to apply the presented models on the whole Czech population. The predictive power of the model for the general population would probably not improve so substantially as in our results, because the CKD population is minor. On the other hand, the condition presents very high expenses compared to other diseases and thus we believe the total profit/loss balance of insurance companies would be affected noticeably. Another step could be an introduction of the procedure approach into other currently used PCGs, which could improve their accuracy and overall model's performance.

The suggested methodology does not account for the ex-post compensation. The ex-post risk sharing is known to improve the overall model's performance and tackle the outliers' problem as suggested by Dungl et al. (2017). On the other hand, if ex-ante prediction was designed precisely enough, it would not be necessary to carry out the ex-post compensation in the first place. This is the case in the Netherlands, where the introduction of the most refined ex-ante model led to the mitigation of ex-post compensation (Van Kleef et al., 2018). As we managed to improve the accuracy of the predictions, the need for ex-post compensation reduced. Nevertheless, the outcomes revealed that even the best fitted model still underestimates the real costs, hence there is certainly a space for improvement.

The last issue worth mentioning is the use of OLS, which is not the most suitable option due to the substantial skewness of the data (as already discussed in the Literature review). While there has been suggested multiple modifications, such as log-transformation or the use of GLM, majority of authors pointed out the practical advantages of OLS, particularly its simplicity and easy interpretation, which are preferred in this field (Fishman et al., 2003; Van Veen et al., 2015). Furthermore, since the official PCG methodology in the CR still uses the method of least squares and our aim was to test its potential for improvement, the OLS was preserved in our models.

The potential heteroskedasticity stemming from the skewness was addressed by heteroskedasticity robust standard errors, although its presence does not influence the bias of coefficient estimates. The coefficients might be biased in case we omitted an important variable, which is at the same time correlated with some of the independent variables. For example, if income or socioeconomic status were correlated with the health care costs and at the same time with the age groups or PCGs in our model, it could induce the omitted variable bias. Nevertheless, as it is almost never possible to include all relevant variables, we rely on our current estimates to be at least consistent.

7 Conclusion

In 2018, new health-based risk adjusters were implemented in the Czech Republic following the Dutch example – the pharmacy-based cost groups (PCGs), which accounted for the costs of chronic diseases in addition to age and gender. The aim of this study was to revise the Czech PCG model, where we focused on one PCG only – renal failure (REN). The renal failure is the most expensive chronic disease and we believe that the classification criteria for the PCG are insufficient given the nature and the treatment of the disease. The PCG model for renal failure would thus benefit from methodology improvements resulting in fairer fund redistribution. The thesis suggests improvements which take advantage of current data availability, thus no additional costs are required.

The thesis shows that the consumption of typical drugs is not an accurate indicator of renal failure, nor a suitable cost predictor. Instead, we suggested employing the dialysis procedures, as they are unique for renal failure and present substantial costs for insurance companies. The dialysis procedures correctly identified not only more patients suffering from renal failure, but also those being more expensive on average. Although dialysis procedures do not capture the patients treated solely by drugs, we showed that those have substantially lower average costs than dialysis patients. In comparison, the drug consumption method captured only 62.5 % of patients on dialysis.

Five different models were tested using the sample of GHI insurees who were reported with any health care consumption typical for kidney disease throughout years 2015-2018. The OLS was employed following the methodology used in the Czech Republic. Firstly, we tested simple demographic model using only age/gender groups, which explained less than 2 % of variation in costs. The second model added PCGs (including renal failure) based on current threshold of 181 defined daily doses (DDDs) for PCG classification. The model explained 26 % of the cost variation.

In next three models, the definition of PCG for renal failure was modified. The threshold of drug consumption was lowered to 91 DDDs in the third model, in which the R^2 improved to 30 %. In the fourth model, we finally defined the REN variable based on number of dialysis procedures instead of drug consumption and the explained variation reached 48 %. Lastly, both approaches were combined, and two different REN groups were specified – patients treated only by drugs and patients on dialysis,

as these groups had substantially different cost profile (approximately 0.5 mil. vs 1 mil. CZK in average costs, respectively). The predictive power increased to 49 %.

Using the estimates from year 2017, the costs for year 2018 were predicted and compared with the real costs to evaluate the forecasting power of the models. As anticipated, the last two models explained best the variation in 2018 costs (47 – 48 %). Other measures of predictive power (MPE, MARE, MAPE) were also tested. Although all models' predictions on average underestimated the true costs, the models based on dialysis procedures were closest to the real values. Moreover, the quantile analysis revealed that the last model, which combined both drug consumption and dialysis procedures, performed best among the models for 20 % of the most expensive and 20 % of the least expensive cases in the sample.

The thesis contributes to current discussions about the accuracy of the PCG models, renal failure in particular, leading to fairer redistribution of funds among insurance companies. Although the renal failure patients form a minority among chronic diseases, they present extreme costs and therefore we believe that the total balance of insurance companies is affected considerably, if renal failure patients are identified inaccurately. Better financial compensation should motivate the insurance funds to offer better services or even to create specialized disease management programs, which could improve the clinical outcomes and the efficiency of chronic kidney disease management. As an extension of our research, we would suggest testing the new methodology with the complete version of PCG model using general population.

Apart from the renal failure, our results might enhance the discussion about the accuracy of other PCGs regarding their identification criteria, particularly in cases where the drug consumption does not match the costs very well. This may lead to future modifications in the Czech risk adjustment, either in terms of lowered drug consumption thresholds or including complementary indicators, such as procedures or diagnoses, as already used in other countries.

Bibliography

Alexa, J., Recka, L., Votapkova, J., van Ginneken, E., Spranger, A., & Wittenbecher, F. (2015). Czech Republic: health system review. *Health Systems in Transition*, *17*(1), 1–165.

Ash, A. S., Ellis, R. P., Pope, G. C., Ayanian, J. Z., Bates, D. W., Burstin, H., Iezzoni, L. I., MacKay, E., & Yu, W. (2000). Using diagnoses to describe populations and predict costs. *Health Care Financing Review*, *21*(3), 7–28.

Askar, A. M. (2015). Hyperphosphatemia: The hidden killer in chronic kidney disease. *Saudi Medical Journal*, *36*(1), 13–19. https://doi.org/10.15537/smj.2015.1.9843

Axelrod, D. A., Schnitzler, M. A., Xiao, H., Irish, W., Tuttle-Newhall, E., Chang, S. H., Kasiske, B. L., Alhamad, T., & Lentine, K. L. (2018). An economic assessment of contemporary kidney transplant practice. *American Journal of Transplantation*, *18*(5), 1168–1176. https://doi.org/10.1111/ajt.14702

Barros, P. P. (2003). Cream-skimming , incentives for efficiency and payment system. *Journal of Health Economics*, *22*, 419–443.

Bělohorský, A. (2018). The Methodology of Health Care , Long-term Care and Education Projections. *Methodological Compendium*. http://www.mfcr.cz/methodology

Brück, K., Stel, V. S., Gambaro, G., Hallan, S., Völzke, H., Ärnlöv, J., Kastarinen, M., Guessous, I., Vinhas, J., Stengel, B., Brenner, H., Chudek, J., Romundstad, S., Tomson, C., Gonzalez, A. O., Bello, A. K., Ferrieres, J., Palmieri, L., Browne, G., … Jager, K. J. (2016). CKD prevalence varies across the European general population. *Journal of the American Society of Nephrology*, *27*(7), 2135–2147. https://doi.org/10.1681/ASN.2015050542

Bryndová, L., Hroboň, P., & Tulejová, H. (2019). The 2018 risk-adjustment reform in the Czech Republic: Introducing Pharmacy-based Cost Groups and strengthening reinsurance. *Health Policy*, *123*(8), 700–705. https://doi.org/10.1016/j.healthpol.2019.05.017

Chalupka, R. (2010). Improving risk adjustment in the Czech Republic. *Prague Economic Papers*, *1*(3), 236–250. https://doi.org/10.18267/j.pep.374

Chochláčová, B. (2018). *Improvement of Risk Adjustment Model to Strengthen the Competition among Health Insurance Companies in Slovakia*. Charles University.

Cikrt, T. (2018). *Ekonomická analýza zdravotních pojišťoven v České republice* (pp. 1–40). https://www.zdravotnickydenik.cz

Clark, D. O., Von Korff, M., Saunders, K., Baluch, W. M., & Simon, G. E. (1995). A Chronic Disease Score with Empirically Derived Weights. *Medical Care*, *33*(8), 783–795.

Czech Society of Nephrology. (2016). *Statistická ročenka dialyzační léčby v České republice 2016*.

Czech Society of Nephrology. (2018). *Statistická ročenka dialyzační léčby v České republice 2018*.

Da Silva, G. B., De Oliveira, J. G. R., De Oliveira, M. R. B., De Souza Vieira, L. J. E., & Dias, E. R. (2018). Global costs attributed to chronic kidney disease: A systematic review. *Revista Da Associacao Medica Brasileira*, *64*(12), 1108–1116. https://doi.org/10.1590/1806-9282.64.12.1108

Duncan, I., Ahmed, T., Dove, H., & Maxwell, T. L. (2019). Medicare Cost at End of Life. *American Journal of Hospice and Palliative Medicine*, *36*(8), 705–710. https://doi.org/10.1177/1049909119836204

Dungl, M., Jandová, P., Jan, K., Macháček, T., & Svoboda, J. (2017). PCG v České republice Historie, postupy, výsledky analýz. *Monitor HC*, 1–160. www.monitorhc.cz

ERA-EDTA Registry. (2019). *ERA-EDTA Registry Annual Report 2017*. Amsterdam UMC.

Farley, D. O., Carter, G. M., Kallich, J. D., Lucas, T. W., & Spritzer, K. L. (1996). Modified capitation and treatment incentives for end stage renal disease. *Health Care Financing Review*, *17*(3), 129–142.

Fassett, R. G. (2014). Current and emerging treatment options for the elderly patient with chronic kidney disease. *Clinical Interventions in Aging*, *9*, 191–199. https://doi.org/10.2147/CIA.S39763

Federal Office of Public Health. (2020). *Health insurance: Risk equalization*. https://www.bag.admin.ch/bag/en/home/versicherungen/krankenversicherung/krankenversicherung-versicherer-aufsicht/risikoausgleich.html

Ferrari, P. (2016). Nurturing the benefits of pre-emptive kidney transplantation. *Nephrology Dialysis Transplantation*, *31*, 681–682. https://doi.org/10.1093/ndt/gfv383

Fishman, P. A., Goodman, M. J., Hornbrook, M. C., Meenan, R. T., Bachman, D. J., & Rosetti, M. C. O. (2003). Risk Adjustment Using Automated Ambulatory Pharmacy Data The RxRisk Model. *Medical Care*, *41*(1), 84–99.

French, E. B., McCauley, J., Aragon, M., Bakx, P., Chalkley, M., Chen, S. H., Christensen, B. J., Chuang, H., Côté-Sergent, A., Nardi, M. De, Fan, E., échevin, D., Geoffard, P. Y., Gastaldi-Ménager, C., Gørtz, M., Ibuka, Y., Jones, J. B., Kallestrup-Lamb, M., Karlsson, M., … Kelly, E. (2017). End-of-life medical spending in last twelve months of life is lower than previously reported. *Health Affairs*, *36*(7), 1211–1217. https://doi.org/10.1377/hlthaff.2017.0174

Glova, J., & Gavurová, B. (2013). Vybrané aspekty systémov zdravotnej starostlivosti. *GRANT Journal*. http://www.grantjournal.com/issue/0202/PDF/0202glova.pdf

Hájíčková, T. (2015). *The Pharmacy-based Cost Group Model: Application in the Czech Health Care System*. Charles University.

Health Policy Institute. (2014). *Výpočet indexov rizika podĺa PCG a odhad dopadov na zdravotné poisťovne*. http://www.hpi.sk/cdata/Documents/HPI-PCG_pre_VZP_2014.pdf

Helanterä, I., Isola, T., Lehtonen, T. K., Åberg, F., Lempinen, M., & Isoniemi, H. (2019). Association of clinical factors with the costs of kidney transplantation in the current era. *Annals of Transplantation*, *24*, 393–400. https://doi.org/10.12659/AOT.915352

Huber, C. A., Schneeweiss, S., Signorell, A., & Reich, O. (2013). Improved prediction of medical expenditures and health care utilization using an updated chronic disease score and claims data. *Journal of Clinical Epidemiology*, *66*(10), 1118–1127. https://doi.org/10.1016/j.jclinepi.2013.04.011

Jarl, J., Desatnik, P., Hansson, U. P., Prütz, K. G., & Gerdtham, U. G. (2018). Do kidney transplantations save money? A study using a before-after design and multiple register-based data from Sweden. *Clinical Kidney Journal*, *11*(2), 283–288. https://doi.org/10.1093/ckj/sfx088

Juhnke, C., Bethge, S., & Mühlbacher, A. C. (2016). A Review on Methods of Risk Adjustment and their Use in Integrated Healthcare Systems. *International Journal of Integrated Care*, *16*(4), 1–18. https://doi.org/10.5334/ijic.2500

Jung, H., Jeon, Y., Park, Y., Kim, Y. S., Kang, S., Yang, C. W., Kim, N., Choi, J., Cho, J., Park, S., Kim, C.-D., & Kim, Y.-L. (2019). Better Quality of Life of Peritoneal Dialysis compared to Hemodialysis over a Two-year Period after Dialysis Initiation. *Scientific Reports*, 1–10. https://doi.org/10.1038/s41598-019-46744-1

Kaysen, G. A. (2007). Hiperlipidemia in Chronic Kidney Disease. *The International Journal of Artificial Organs*, *30*(11), 987–992. https://doi.org/10.1177/039139880703001107

Kramer, A., Pippias, M., Noordzij, M., Stel, V. S., Andrusev, A. M., Aparicio-Madre, M. I., Arribas Monzón, F. E., Åsberg, A., Barbullushi, M., Beltrán, P., Bonthuis, M., Caskey, F. J., Castro De La Nuez, P., Cernevskis, H., De Meester, J., Finne, P., Golan, E., Heaf, J. G., Hemmelder, M. H., … Jager, K. J. (2019). The European Renal Association - European Dialysis and Transplant Association (ERA-EDTA) Registry Annual Report 2016: A summary. *Clinical Kidney Journal*, 1–19. https://doi.org/10.1093/ckj/sfz011

Kugler, C., Vlaminck, H., Haverich, A., & Maes, B. (2005). Nonadherence with diet and fluid restrictions among adults having hemodialysis. *Journal of Nursing Scholarship*, *37*(1), 25–29. https://doi.org/10.1111/j.1547-5069.2005.00009.x

Kuo, R. N., Dong, Y., Liu, J., Chang, C., Shau, W., & Lai, M. (2011). Predicting Healthcare Utilization Using a Pharmacy-based Metric With the WHO ' s Anatomic Therapeutic. *Medical Care*, *49*(11), 1031–1039.

Kutzin, J., Shishkin, S., Bryndová, L., Schneider, P., & Hroboň, P. (2010). Chapter 5 Reforms in the pooling of funds. In J. Kutzin, C. Cashin, & M. Jakab (Eds.), *Implementing Health Financing Reform Lessons from countries in transition* (pp. 119–153). World Health Organization 2010, on behalf of the European Observatory on Health Systems and Policies.

Lamers, L. M., & Van Vliet, R. C. J. A. (2003). Health-based risk adjustment: Improving the pharmacy-based cost group model to reduce gaming possibilities. *European Journal of Health Economics*, *4*(2), 107–114. https://doi.org/10.1007/s10198-002-0159-9

Lamers, L. M., & Van Vliet, R. C. J. A. (2004). The Pharmacy-based Cost Group model : validating and adjusting the classification of medications for chronic conditions to the Dutch situation. *Health Policy*, *68*, 113–121. https://doi.org/10.1016/j.healthpol.2003.09.001

Levy, J. M., Robst, J., & Ingber, M. J. (2006). Risk-adjustment system for the medicare capitated ESRD program. *Health Care Financing Review*, *27*(4), 53–69.

Mathauer, I., Saksena, P., & Kutzin, J. (2019). Pooling arrangements in health financing systems: A proposed classification. *International Journal for Equity in Health*, *18*(1), 1–11. https://doi.org/10.1186/s12939-019-1088-x

National Kidney Foundation. (2013). KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. *Kidney International Supplements*, *3*(1), 135–150. https://doi.org/10.3182/20140824-6-za-1003.01333

Neild, G. H. (2017). Life expectancy with chronic kidney disease: an educational review. *Pediatric Nephrology*, *32*(2), 243–248. https://doi.org/10.1007/s00467-016-3383-8

Nesrallah, G. E., Mustafa, R. A., Clark, W. F., Bass, A., Barnieh, L., Hemmelgarn, B. R., Klarenbach, S., Quinn, R. R., Hiremath, S., Ravani, P., Sood, M. M., & Moist, L. M. (2014). Canadian Society of Nephrology 2014 clinical practice guideline for timing the initiation of chronic dialysis. *Canadian Medical Association Journal*, *186*(2), 112–117. https://doi.org/10.1503/cmaj.130363

Newhouse, J. P., Manning, W. G., Keeler, E. B., & Sloss, E. M. (1989). Adjusting capitation rates using objective health measures and prior utilization. *Health Care Financing Review*, *10*(3), 41–54.

Nistor, I., Sc, P., Jc, C., Saglimbene, V., Vecchio, M., Covic, A., & Gfm, S. (2015). Haemodiafiltration, haemofiltration and haemodialysis for end- stage kidney disease (Review). *Cochrane Database of Systematic Reviews*, *5*. https://doi.org/10.1002/14651858.CD006258.pub2.www.cochranelibrary.com

OECD/European Observatory on Health Systems and Policies. (2017). *Czech Republic: Country Health Profile 2017*. OECD. https://doi.org/10.1787/9789264283336-en

Organization of Economic Cooperation and Development. (2019). *Health and Glance 2019*. https://www.eib.org/attachments/general/the_eib_at_a_glance_en.pdf

Pelikánová, M. (2017). *Systém veřejného zdravotního pojištění v ČR*. Charles Univerzity.

Pilny, A., Wübker, A., & Ziebarth, N. R. (2017). Introducing risk adjustment and free health plan choice in employer-based health insurance: Evidence from Germany. *Journal of Health Economics*, *56*(10870), 330–351. https://doi.org/10.1016/j.jhealeco.2017.03.009

Prinsze, F. J., & Van Vliet, R. C. J. A. (2007). Health-Based Risk Adjustment: Improving the Pharmacy-Based Cost Group Model by Adding Diagnostic Cost Groups. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, *44*(4), 469–480. https://doi.org/10.5034/inquiryjrnl_44.4.469

Ryšavá, R., & Brejník, P. (2018). *Doporučené diagnostické a terapeutické postupy pro všeobecné praktické lékaře Základy Nefrologie - Novelizace 2018*. Společnost všeobecného lékařství ČLS JEP.

Schneider, U., Ulrich, V., & Wille, E. (2008). Risk adjustment systems in health insurance markets in the US, Germany, Netherlands and Switzerland. In *CESifo DICE Report 3*.

Shmueli, A. (2015). On the calculation of the Israeli risk adjustment rates. *European Journal of Health Economics*, *16*(3), 271–277. https://doi.org/10.1007/s10198-014-0572-x

Sijpkens, Y. W. J., Berkhout-Byrne, N. C., & Rabelink, T. J. (2008). Optimal predialysis care. *Nephrology Dialysis Transplantation Plus*, *1*(Suppl 4), 7–13. https://doi.org/10.1093/ndtplus/sfn117

Sinnakirouchenan, R., & Holley, J. (2011). Peritoneal dialysis versus hemodialysis: risks, benefits, and access issues. *Advances in Chronic Kidney Disease*, *18*(6), 428–432. https://doi.org/10.1053/j.ackd.2011.09.001

Smith, P. C., & Witter, S. N. (2004). Risk Pooling in Health Care Financing : The Implications for Health System Performance. *Health, Nutrition, and Population Family (HNP) Discussion Paper*.

Tichý, O. (2015). *Analýza VZP: Pacient po transplantaci ledviny stojí ročně o půl milionu méně než na dialýze*. VZP Aktuality. https://www.vzp.cz/o-nas/aktuality/analyza-vzp-pacient-po-transplantaci-ledviny-stoji-rocne-o-pul-milionu-mene-nez-na-dialyze

Tichý, O. (2018). *Za dialýzu dá VZP 3,8 miliardy. Nově zaplatí lékařům za pomoc s rozšířením transplantací*. VZP Aktuality. https://www.vzp.cz/o-nas/aktuality/za-dialyzu-da-vzp-3-8-miliardy-nove-zaplati-lekarum-za-pomoc-s-rozsirenim-transplantaci

Van de Ven, Wynand P.M.M., Beck, K., Van de Voorde, C., Wasem, J., & Zmora, I. (2007). Risk adjustment and risk selection in Europe: 6 years later. *Health Policy*, *83*(2–3), 162–179. https://doi.org/10.1016/j.healthpol.2006.12.004

Van de Ven, Wynard P.M.M., Van Vliet, R. C. J. A., & Lamers, L. M. (2004). Health-Adjusted Premium Subsidies In The Netherlands. *Health Affairs*, *23*(3), 45–55. https://doi.org/10.1377/hlthaff.23.3.45

Van Kleef, R. C., Eijkenaar, F., Van Vliet, R. C. J. A., & Van de Ven, W. P. M. M. (2018). Health Plan Payment in the Netherlands. In T. G. McGuire & R. C. Van Kleef (Eds.), *Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets* (pp. 397–429). https://doi.org/10.1016/C2016-0-01299-8

Van Kleef, R. C., & Van Vliet, R. C. J. A. (2012). Improving risk equalization using multiple-year high cost as a health indicator. *Medical Care*, *50*(2), 140–144. https://doi.org/10.1097/MLR.0b013e31822ebf8b

Van Veen, S. H. C. M., Van Kleef, R. C., Van de Ven, W. P. M. M., & Van Vliet, R. C. J. A. (2015). Improving the prediction model used in risk equalization: cost and diagnostic information from multiple prior years. *European Journal of Health Economics*, *16*, 201–218. https://doi.org/10.1007/s10198-014-0567-7

Veazie, P. J., Manning, W. G., & Kane, R. L. (2003). Improving risk adjustment for medicare capitated reimbursement using nonlinear models. *Medical Care*, *41*(6), 741–752. https://doi.org/10.1097/01.MLR.0000065127.88685.7D

Viklický, O. (2013). Nová klasifikace chronických onemocnění ledvin. *Postgraduální Nefrologie*, *11*(1), 2–4.

Von Korff, M., Wagner, E. H., & Saunders, K. (1992). A chronic disease score from automated pharmacy data. *Journal of Clinical Epidemiology*, *45*(2), 197–203. https://doi.org/10.1016/0895-4356(92)90016-g

Von Wyl, V., & Beck, K. (2016). Do insurers respond to risk adjustment? A long-term, nationwide analysis from Switzerland. *European Journal of Health Economics*, *17*(2), 171–183. https://doi.org/10.1007/s10198-015-0669-x

Vyhláška č. 114/2019 Sb. Vyhláška o nastavitelných parametrech přerozdělování finančních prostředků veřejného zdravotního pojištění pro rok 2020. Retrieved from: https://www.zakonyprolidi.cz/cs/2019-114#p2

VZP ČR. (2018a). *Ročenka za rok 2018*. https://www.vzp.cz/o-nas/dokumenty/vyrocni-zpravy

VZP ČR. (2018b). *Zdravotně pojistný plán Všeobecné zdravotní pojišťovny České republiky pro rok 2019*. https://www.vzp.cz/o-nas/dokumenty/zdravotne-pojistne-plany

Wasem, J., Buchner, F., Lux, G., & Schillo, S. (2018). Health Plan Payment in Germany. In T. G. McGuire & R. C. van Kleef (Eds.), *Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets* (pp. 295–329). Elsevier. https://doi.org/10.1016/C2016-0-01299-8

Wong, B., Ravani, P., Oliver, M. J., Holroyd-leduc, J., Venturato, L., Garg, A. X., & Quinn, R. R. (2018). Comparison of Patient Survival Between Hemodialysis and Peritoneal Dialysis Among Patients Eligible for Both Modalities. *American Journal of Kidney Diseases*, *71*(3), 344–351. https://doi.org/10.1053/j.ajkd.2017.08.028

Wooldridge, J. M. (2009). *Introductory Econometrics: A Modern Approach* (4th ed.). South-Western, Cengage Learning.

Yang, F., Khin, L. W., Lau, T., Chua, H. R., Vathsala, A., Lee, E., & Luo, N. (2015). Hemodialysis versus peritoneal dialysis: A comparison of survival outcomes in south-east Asian patients with end-stage renal disease. *PLoS ONE*, *10*(10). https://doi.org/10.1371/journal.pone.0140195

Zakiyanov, O., Vachek, J., & Tesař, V. (2014). Chronické onemocnění ledvin - Novinky v klasifikaci a terapii. *Kardiologicka Revue*, *16*(1), 49–52.

Zákon č. 592/1992 Sb. Zákon České národní rady o pojistném na všeobecné zdravotní pojištění. Retrieved from: https://www.zakonyprolidi.cz/cs/1992-592#cast3

Zákon č. 48/1997 Sb. Zákon o veřejném zdravotním pojištění a o změně a doplnění některých souvisejících zákonů. Retrieved from: https://www.zakonyprolidi.cz/cs/1997-48

Zima, T., Racek, J., Tesař, V., Viklický, O., Teplan, V., Schuck, O., Janda, J., Friedecký, B., Kubíček, Z., Kratochvíla, J., Rajdl, D., Šálek, T., Kalousová, M., & Granátová, J. (2014). *Doporučení k diagnostice chronického onemocnění ledvin* (pp. 1–21). Česká nefrologická společnost ČLS JEP a Česká společnost klinické biochemie ČLS JEP.

Appendix A: Additional figures

Appendix A 1: The risk adjusters for somatic care in the Netherlands, 2017

Age and gender: 20 age groups for each gender with 5-year gaps.

Pharmacy-based cost groups (PCGs): 33 groups corresponding to chronic diseases; classification based on the specific amount of drug consumption (usually 180 daily defined doses).

Diagnoses-based cost groups (DCGs): 15 groups of specific diagnoses from the previous year.

Multiple-year high-cost groups (MHCGs): 7 groups based on excessive spending on somatic care in last 3 years (or 2 years in case of extreme costs).

Durable medical equipment cost groups (DMECGs): 10 groups reflecting the use of durable medical equipment for specific chronic conditions in the previous year.

Physiotherapy-diagnoses cost groups (PDCGs): 4 groups including physiotherapeutic diagnoses from previous year.

Home care spending: 7 groups reflecting the costs of home care in the previous year.

Geriatric rehabilitation care spending: 2 groups reflecting the costs of geriatric rehabilitation care in the previous year.

Yes/no morbidity interacted with age: 4 groups reflecting the interaction of age with specific PCGs or DCGs or other risk-adjusters.

Urbanization: 10 clusters aggregating villages, towns, or their parts.

Socioeconomic status interacted with age: 12 groups capturing total households' income interacted with age.

Source of income interacted with age: 23 groups based on income or education level interacted with age.

Source: Van Kleef et al. (2018)

Appendix A 2: List of PCGs and their classification criteria

PCG code	PCG group name	List of ATC classification groups	Exclusion criteria
GLA	Glaucoma	S01E	
THY	Thyroid disorders	H03A, H03B	
PSY	Antipsychotics, Alzheimer's disease, treatment of addiction	N05A excluding (N05AL03, N05AN01), N06DA, N06DX01, N07BB, N07BC51	
DEP	Treatment with antidepressants	N06A excluding (N06AA09, N06AX21)	not if in PSY
CHO	Hypercholesterolemia	C10 excluding (C10AC01, C10BX03)	not if in DM1, DM2, DMH
DMH	Diabetes with hypertension	A10 and simulatneously C02 excluding (C02KX, C02CA04), C03 excluding (C03CA01), C07, C08 excluding (C08CA06), C09	
COP	Serious asthma, Chronic obstructive pulmonary disease	R03AC18, R03AK03, R03BB	
AST	Asthma	R03 excluding (R03AC18, R03AK03, R03BB, R03CA02, R03BC01, R03CC02, R03CC13)	not if in COP
DM2	Diabetes mellitus type 2	A10	not if in DM1, DMH
EPI	Epilepsy	N03 excluding (N03AX12, N03AX16,	
CRO	Crohn's disease, ulcerative colitis	A07EA06, A07EC02	
KVS	Heart disease	C01A, C01B, C01D, C01EB15, C01EB17,	
TNF	Rheumatic diseases treated with TNF inhibitors	L04AA11, L04AA24, L04AB, L04AC	
REU	Rheumatic diseases treated otherwise than with TNF inhibitors	A07EC01, L01BA01, L04AA13, L04AX03, M01CB01, M01CC01, P01BA02	not if in TNF
PAR	Parkinson's disease	N04B	
DM1	Diabetes mellitus type 1	A10A	not if in DMH
TRA	Transplants	L04AA06, L04AA10, L04AA18, L04AC02, L04AD01, L04AD02, L04AX01	
CFP	Cystic fibrosis or disorder of pancreatic exocrine function	J01GB01 , J01XB01, R05CB13	
CNS	Brain and spine disorders	L03AB07, L03AB08, L03AX13, L04AA23, M03BX01, M03BX02	
ONK	Malignancy	L01 excluding (L01BA01), L03AA, L03AC01,	
HIV	HIV, AIDS	J05AE, J05AF excluding (J05AF08, J05AF10, J05AF11), J05AG, J05AR, J05AX excluding (J05AX05)	
REN	Renal failure	B03X, V03AE	
RAS	Therapy with growth hormone	H01AC01, H01AC03	
HOR	Hormonal oncology	L02	
NPP	Neuropathic pain	N01BX04, N03AX12, N03AX16	

Source: Zákon České národní rady o pojistném na všeobecné zdravotní pojištění 592/1992 Sb.

Appendix A 3: Summary of results for model with REN_DIAL as the only explanatory variable

Intercept	258 561 ***
REN_DIAL	839 933 ***
N	10 937
R2	0.4472
Adjusted R2	0.4471

*Significance: *** p < 0.001; ** p < 0.01; * p < 0.05*

Appendix A 4: Summary of results for Model 4 without outliers, 2017

	Model 4 /outliers		
Intercept	155 308.49 ***	REN_DIAL	896 706.23 ***
M < 35	58 217.97 *	REN_PRE	180 446.73 ***
M 35-39	58 335.86 *	AST	25 950.04 *
M 40-44	48 305.61	COP	33 518.16
M 45-49	63 061.73 **	CRO	18 011.35
M 50-54	36 677.82	DEP	42 279.94 ***
M 55-59	40 916.31 *	DM1	50 190.76 **
M 60-64	10 263.33	DM2	49 825.05
M 65-69	36 324.10 **	DMH	10 671.22
M 75-79	-32 442.91 *	EPI	26 696.98
M 80-84	-73 229.06 ***	GLA	6 765.27
M > 84	-84 675.23 ***	HIV	159 791.18 **
F < 35	69 377.51 *	HOR	19 413.08
F 35-39	64 223.74	CHO	-2 762.74
F 40-44	67 605.53 *	KVS	21 128.88 ***
F 45-49	46 653.92	NPP	54 173.82 *
F 50-54	37 078.25	ONK	404 776.71 ***
F 55-59	17 439.86	PAR	37 589.92
F 60-64	30 840.32	PSY	-6 675.51
F 65-69	12 827.19	REU	26 097.54
F 70-74	-4 597.29	THY	22 950.07 *
F 75-79	-35 429.49 **	TNF	190 283.95 ***
F 80-84	-58 557.83 ***	TRA	85 510.73 ***
F > 84	-103 452.13 ***		
N	10 869		
R2	0.6760		

*Standard errors are heteroskedasticity robust. Significance: *** p < 0.001; ** p < 0.01; * p < 0.05*