

České vysoké učení technické v Praze  
Fakulta biomedicínského inženýrství

# DIPLOMOVÁ PRÁCE

Květen 2016

Alexandra Petrlíková



**ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE**

---

**FAKULTA BIOMEDICÍNSKÉHO INŽENÝRSTVÍ**

**Katedra biomedicínské techniky**

Název diplomové práce:

**PACIENT – specifické modely pro  
rozhodování o léčbě**

Studijní program: Biomedicínská a klinická technika

Studijní obor: Systémová integrace procesů ve zdravotnictví

Autor diplomové práce: Bc. Alexandra Petrlíková

Vedoucí diplomové práce: Ing. Vojtěch Kamenský

---

**Kladno 2016**

Katedra biomedicínské techniky

Akademický rok: 2015/2016

## Zadání diplomové práce

Student: **Bc. Alexandra Petrlíková**  
Studijní obor: Systémová integrace procesů ve zdravotnictví  
Téma: **Pacient-specifické modely pro rozhodování o léčbě**  
Téma anglicky: Patient-specific models for decision making in health care

### Zásady pro vypracování:

Cílem diplomové práce je využití pacient-specifických modelů pro rozhodování o léčbě pacienta. Analyzujte současný stav využití pacient-specifických modelů ve světě a ČR. Zhodnoťte využití modelů pro predikci klinických a nákladových výstupů. Určete vhodné diagnózy k modelování a na příkladech pneumonie aplikujte model. Využijte metody modelování a pro zhodnocení ekonomických nákladů zvolte vhodnou metodu HTA. Zhodnoťte využití modelů pro predikci klinických a nákladových výstupů. Na základě získaných informací zhodnoťte výhody a nevýhody použití modelů a odhadněte ekonomické přínosy.

### Seznam odborné literatury:

- [1] Goddman, C.S., HTA 101 - Introduction to health technology assessment, Virginia USA, 2004
- [2] Tappenden, Paul et al., Using whole disease modeling to inform resource allocation decisions: economic evaluation of a clinical guideline for colorectal cancer using a single model, Value in health, ročník 16, číslo 4, 2013
- [3] Visweswaran, Shyam et al., Learning patient-specific predictive models from clinical data, Journal of biomedical informatics, ročník 43, číslo 5, 2010

Vedoucí: Ing. Vojtěch Kamenský

Zadání platné do: 20.08.2017

.....  
vedoucí katedry / pracoviště

.....  
děkan

V Kladně dne 01.12.2015



## **PROHLÁŠENÍ**

Prohlašuji, že jsem diplomovou práci na téma „Pacient – specifické modely pro rozhodování o léčbě“ vypracovala samostatně. Veškerou použitou literaturu a podkladové materiály uvádím v příloženém seznamu literatury.

V Kladně 20. května 2016

.....

Bc. Alexandra Petříková

## **Poděkování**

Ráda bych v této části poděkovala především vedoucímu diplomové práce, Ing. Vojtěchu Kamenskému, který mě po celý čas práce vedl a motivoval. Dále děkuji primáři MUDr. Tomáši Dvořákovi za poskytnutí dat pro zpracování praktické části a především děkuji MUDr. Janu Fulíkovi za konzultace na medicínské části diplomové práce.

Název diplomové práce:

Pacient – Specifické modely pro rozhodování o léčbě

### **Abstrakt:**

V případě zkvalitňování zdravotnické péče se dostáváme i ke klinickému rozhodování o péči. U mnoha diagnóz se lékaři často v ambulanci rozhodují, zda léčit pacienta ambulantně nebo hospitalizací. Léčba hospitalizací znamená vyšší náklady pro zdravotnické zařízení a celý systém zdravotnictví, v každém případě je nutno zohlednit prospěch nemocného. Dalšími přidruženými náklady je také delší pracovní neschopnost pacienta po případné hospitalizaci. Tyto náklady jsou však u konkrétních diagnóz těžko spočítatelné, protože jsou velmi individuální. V diplomové práci jsou uvedeny rozdíly mezi prediktivními modely populačně-širokými a pacient-specifickými. Pacient-specifické modely mají výhodu použití. Využívají se u konkrétního případu pacienta a nejsou zprůměrované, jako modely populačně široké. U vybrané diagnózy pneumonie byly určeny důležité proměnné při rozhodování o léčbě ve zdravotnických zařízeních a dále byly analyzované získané výsledky. Vybraný vzorek respondentů byl získán v Klaudiánově nemocnici v Mladé Boleslavi. Pro sestavení klasifikačních a regresních stromů byly využity proměnné, které zahrnovaly osobní i rodinnou anamnézu, přidružené choroby a konkrétní příznaky při vstupu do ambulance. V práci byly dále zanalyzovány výsledky zabývající se počtem vykázaných bodů pojišťovně a jejich závislost na typu léčby a počtu dnů hospitalizace. Grafické znázornění prediktivních modelů ve formě klasifikačních a regresních stromů je vhodná volba pro praktické využití v ambulanci při podpoře rozhodovacího procesu. Tyto výsledky jsou prvotní podporou pro vznik prediktivních pacient-specifických modelů na větším vzorku pacientů a tím zkvalitnění predikce klinických výsledků.

### **Klíčová slova:**

pacient-specifické modely, populačně-široký model, pneumonie, klasifikační a regresní stromy

**Master's Thesis title:**

Patient – specific models for decision making in Health care

**Abstract:**

In regards of improving the medical care we are in close connection with decision making about necessary care. In many hospital cases the doctors need to decide if to take care of the problem immediately or to hospitalize the patient. Admitting the patient means higher costs for the hospital but we cannot forget about patient's health either. In addition to this fact are the costs of the longer sick leave. These costs cannot be generalized as they differ in every case. In my Thesis you see two different predicting modules: Population-wide and Patient-Specific. Patient Specific one is better in the way of usage. We can use it in particular case and it is not generalized as the second module. In the selected diagnose of pneumonia I have set important factors in decision making about urgent cases and analyzed collected results. The responding group comes from Kladiánova Nemocnice in Mladá Boleslav. The source for my findings and creation of classification and regression trees were based on personal and family anamnesis, additional symptoms and current symptom at the admission. Furthermore, in the Thesis is work with results of the totals sent to the insurance company and what is the connection to way of treatment and number of day spent in hospital. The graphs of both models classification and regression trees are useful for hospital usage. My findings support creation of predicting modules for patient specific way of treatment in high numbers of patients and therefore better prediction of results.

**Key words:**

patient-specific model, population-wide model, pneumonia, classification and regression trees



# Obsah

Úvod.....	11
1 Přehled současného stavu .....	12
1.1 Specifické modely pro predikci .....	12
1.2 Příklady využití specifických modelů s diagnózou komunitní pneumonie .....	15
1.2.1 Predikce nepříznivého průběhu a vážných komplikací u pacientů s komunitní pneumonií.....	16
1.2.2 Pravidlo predikce pro identifikování pacientů s komunitní pneumonií s nízkým rizikem .....	19
1.2.3 Model rozhodující cesty pro specifické modelování péče o pacienty .....	20
1.2.4 Pacient-specifické modely pro předvídání výsledků u pacientů s komunitní pneumonií.....	22
1.2.5 Posuzování závažnosti pro predikci mortality u hospitalizovaných pacientů s komunitní pneumonií.....	25
1.3 Komunitní pneumonie .....	26
2 Metody .....	31
2.1 Bayesovská statistika .....	31
2.2 Klasifikační a regresní stromy (CART).....	32
2.2.1 Klasifikační strom - rozhodovací.....	33
2.2.2 Regresní strom .....	35
2.2.1 Algoritmus PSDP – BA .....	35
2.2.3 Algoritmus PSDP - IG .....	36
2.3 Markovův model.....	36
3 Výsledky .....	38
3.1 Získaná data .....	38
3.2 Klasifikační a regresní strom (CART).....	45
3.2.1 Výsledky – Klasifikační strom .....	45
3.2.2 Výsledky – Regresní stromy.....	48
3.3 Zhodnocení využití modelů pro predikci klinických a nákladových výstupů. ...	58
4 Diskuze .....	60
Závěr .....	63
Reference .....	63
Seznam obrázků.....	69

Seznam tabulek .....	70
Seznam příloh: .....	71

# Úvod

V každodenní praxi lékaři v ambulancích rozhodují o druhu léčby mnoha onemocnění. To zahrnuje určitou míru nejistoty, stanovení diagnózy a nasazení adekvátní léčby. Nejistota je v tomto případě dvojnásobná, může nastat správné stanovení diagnózy a nasazení neadekvátní léčby nebo dokonce špatně stanovená diagnóza a neadekvátní léčba. Predikcí mohou být vážné komplikace a odhalení vážně nemocných pacientů. Čím přesněji je diagnóza stanovena, tím je dosaženo lepšího výsledku s ohledem na zdraví pacienta i optimalizace vynaložených finančních prostředků celého zdravotnického systému.

V současné době existují dva typy prediktivních modelů. První - populačně-široký model patří mezi nejvíce používané. Počítá s retrospektivními daty, které zprůměruje a výsledná předpověď se dá použít téměř ve všech budoucích případech. Naopak Pacient-specifické modely mají své použití na konkrétních případech a dají se využít individuálně. Jedná se o optimalizovaný statistický model, který pracuje s daty z minulosti a zohledňuje dané proměnné, které zahrnují konkrétní příznaky dané diagnózy pacienta. I malé zlepšení predikce vývoje onemocnění, dokáže zkvalitnit klinickou péči a potenciálně šetřit náklady.

U mnoha diagnóz řeší lékař zásadní otázku postupu léčení, zda bude léčba ambulantní anebo je nutná hospitalizace. V případě pneumonie je léčba perorálními antibiotiky v domácí péči mnohem méně nákladná, ale mnohdy i méně efektivní, zvláště u těžkých infekcí. Stále se jedná o zdraví pacienta, a proto při vytvoření prediktivních modelů počítají autoři při svých studiích s nejhorsími možnými výsledky. Mezi takovéto výsledky patří závažné komplikace spojené s nemocí, potřeba hospitalizace na oddělení JIP či dokonce úmrtí. Tyto výsledky vždy dáváme jako konečný bod statistických modelů. Do studií jsou také zahrnuti pacienti, kteří byli v minulosti léčení neúspěšně. To zahrnuje delší dobu hospitalizace, prodloužení pracovní neschopnosti nebo úmrtí v důsledku pneumonie.

Pacient-specifické modely mají velký potenciál k podpoře a zlepšení diagnostiky. Díky těmto modelům je možné optimalizovat především klinickou péči a to díky předpovědi výsledků a zvolení vhodné terapie. V některých nemocnicích v zahraničí se staly tyto modely naprostým standardem péče. Užití v České republice nebylo doposud ve studiích vyhledáno. Kladné ohlasy v zahraničí byly inspirací pro tuto diplomovou práci.

Cílem diplomové práce je využití pacient-specifických modelů pro rozhodování o léčbě pacienta. V práci se zabývám také analýzou současného stavu využití prediktivních modelů ve světě a u nás. Na základě analýzy byla určena vhodná diagnóza pro aplikaci těchto modelů a na konkrétním vzorku byly tyto metody aplikovány.

# 1 Přehled současného stavu

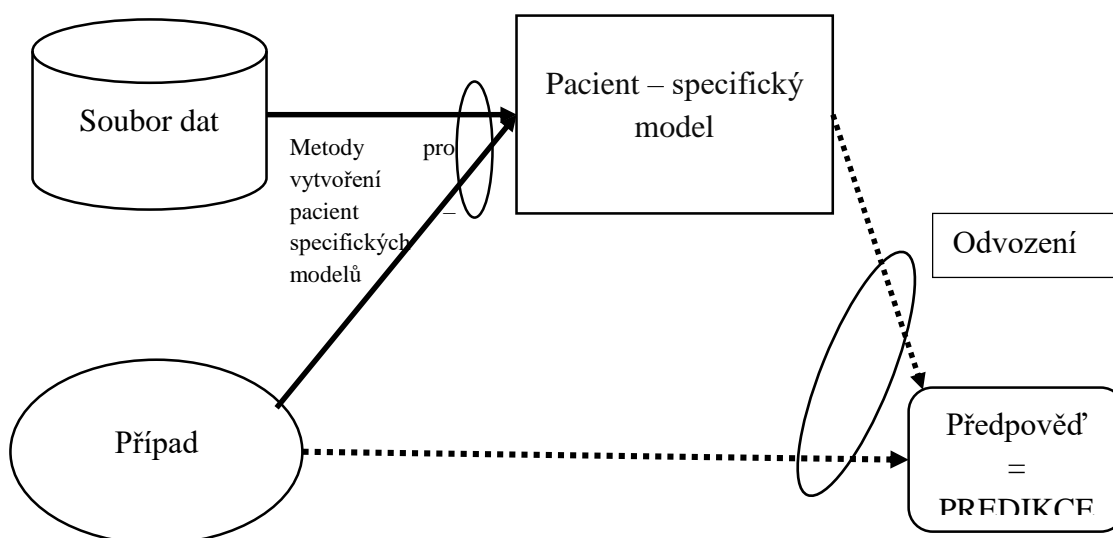
Pacient-specifické modely mají uplatnění především v zahraničí. Ve většině studií se používají modely populačně-široké. V literatuře nebylo dohledáno použití pacient-specifických modelů u nás v České republice. Pro zhodnocení současného stavu a výběr vhodné diagnózy pro zpracování praktické části byla provedena analýza současného stavu zahraničních studií.

## 1.1 Specifické modely pro predikci

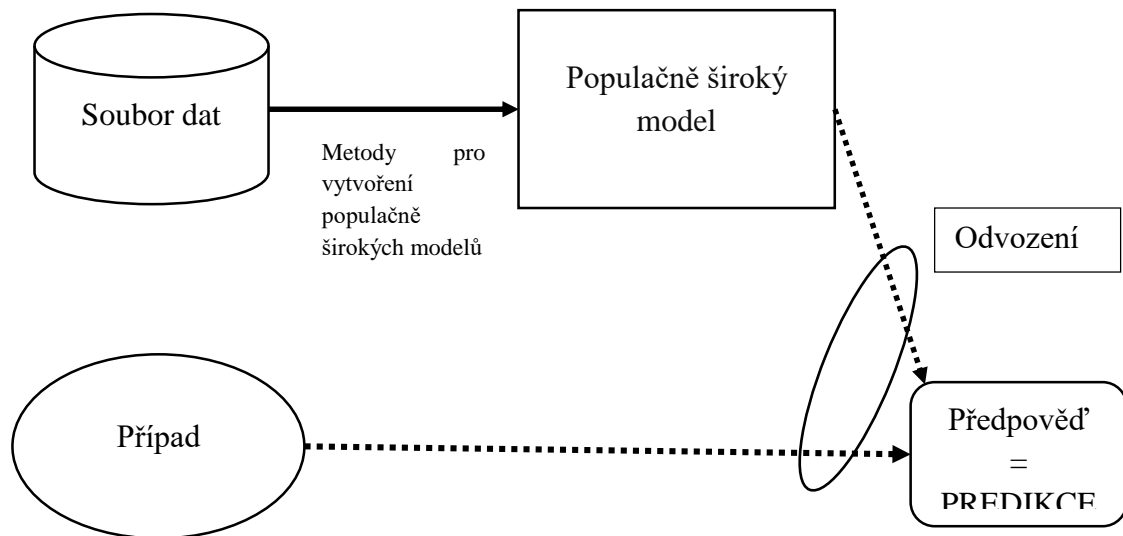
Predikce výsledků ovlivňuje rozhodnutí o dané léčbě pacienta. Ambulantní léčba je méně nákladná pro zdravotnictví, ale určujícím faktorem je zdraví pacienta. Čím přesněji je odhadnuta pacientova diagnóza, tím je rychleji nasazena optimální léčba. Použitím pacient-specifických modelů je umožněno predikovat riziko – vážné komplikace nebo úmrtí. Vytvořením predikčních modelů hodnotících stav a prognózu vede k efektivnějšímu využívání finančních prostředků a lepší péči o pacienta. Pacient – specifické modely jsou sestavovány tak, aby je bylo možné využít na konkrétních případech pacienta ve srovnání s běžně používanými plošnými modely [1,2,3,4,5].

Jak již bylo naznačeno, v prediktivním modelování se setkáváme se dvěma typy modelů:

- Pacient-specifický model: specifický pro konkrétní rysy daného případu pacienta, specializuje se na modelovou strukturu a parametry na základě známých vlastností, viz obrázek 1 [1,2,3].
- Populačně-široký model: čerpá z datové sady konkrétních případů pacientů, aplikován na celou populaci v budoucích případech, viz obrázek 2 [1,2,3].



Obrázek 1: Pacient-specifický model [1]



Obrázek 2: Populačně široký model [1]

Pacient-specifické modely jsou sestaveny tak, aby je bylo možné využít ve zvláštních případech v zájmu pacienta ve srovnání s běžně používanými plošnými modely. Ty jsou konstruovány tak, aby je bylo možné použít průměrně ve všech případech. Optimalizace obou modelů dosáhneme tím, že specializujeme modelovou strukturu a parametry na základě známých vlastností pacienta. Výhodou použití pacient-specifických modelů je také to, že dokáží počítat i s chybějícími daty [1,2,6].

Při sestavování prediktivních modelů se využívají data z minulosti u konkrétní diagnózy. Zohledňují se symptomy, laboratorní výsledky, fyzikální vyšetření a jiné funkce související s konkrétní diagnózou. Konstrukce samotného prediktivního modelu využívá veškeré dostupné poznatky. Konkrétní paradigma zahrnuje mnoho přístupů, které vyúsťují v různé modely. Modely se liší v daných proměnných, především v síle interakce mezi zahrnutými proměnnými, v předpovědích pro daný případ a také v identifikování podmnožiny dat dané skupiny [1,2,6].

Pacient-specifický model může mít větší sílu predikce a to z důvodu využití zvláštních charakteristik v daném případě pacienta, bývá jednodušší než model populačně-široký, tím tak poskytovat stručnější vysvětlení. Konstrukce těchto modelů je specifičtější a výpočetně rychlejší. Další výhodou modelu je schopnost ignorace chybějících dat. Dokáže pracovat pouze s dostupnými daty u konkrétního pacienta [1,6,7].

Tabulka 1: Přehled studií o použití pacient-specifických modelů

č.	Studie	Autoři	Rok	Zaměření
1	Patient-specific models of cardiac biomechanics [5].	Adarsh Krishnamurthy Christopher T. Villongco	2012	Srdeční biomechanika
2	Toward patient-specific myocardial models of the heart [8]	Sermesant M, Peyrat JM	2008	Infarkt myokardu
3	Patient-Specific Modeling of Cardiovascular Mechanics [9].	C. A. Taylor	2009	Kardiovaskulární mechanika
4	Determination of patient-specific multi-joint kinematic models through two-level optimization [10].	Reinbolt JA, Schutte JF	2005	Pohybové modely v ortopedii
5	Characterization of cerebral aneurysms for assessing risk of rupture by using patient-specific computational hemodynamics models [11].	Cebral JR, Castro MA	2005	Prasknutí mozkových výdutí
6	Design of Optimal Treatments for Neuromusculoskeletal Disorders using Patient-Specific Multibody Dynamic Models [12].	Benjamin J. Fregly	2009	Neurologické choroby ovlivňující svalový skelet člověka
7	Patient-Specific Models for Predicting the Outcomes of Patients with Community Acquired Pneumonia [2].	Shyam Visweswaran, Gregory F. Cooper	2005	Rozhodnutí o léčbě komunitní pneumonie
8	Decision Path Models for Patient-Specific Modeling of Patient Outcomes [1].	Antonio Ferreira, Gregory F. Cooper, Shyam Visweswaran,	2013	Rozhodnutí o léčbě komunitní pneumonie
9	Prediction Rule to Identify Low-Risk Patients with Community-Acquired Pneumonia [13].	Michael J. Fine, Thomas E. Auble	1997	Rozhodnutí o léčbě komunitní pneumonie
10	Predicting dire outcomes of patients with community acquired pneumonia [14].	Georgy F. Cooper, Vijoy Abraham	2005	Rozhodnutí o léčbě komunitní pneumonie

V tabulce je přehled studií, které používala pacient-specifické modely. Pod tímto pojmem náleží řada pojmů, mezi ně patří také tvorba počítačových anatomických modelů pro přesnější diagnostiku a plánování léčby. Autoři ve studii č. 1 na základě retrospektivních dat popsali metody pro vytvoření konkrétních modelů srdce pacienta a jeho specifické selhávání. Modely slouží především k odhadu oběhového algoritmu s využitím dat získaných z echokardiografie. Autoři ve studii č. 4 také zhodnotili na základě dat z minulosti pohybové modely pacienta pro řešení problému v ortopedii a následně rehabilitaci. Optimalizovaný model byl vytvořen s cílem minimalizování chyb značkování trajektorie v experimentální sféře. Autoři v článku č. 6 se zabývali poruchami nervosvalového systému člověka, mezi ně patří následky cévní mozkové příhody, osteoartritidy a paraplegie. Jedná se o následky, které významně snižují kvalitu života, protože ovlivňují mobilitu. Pomocí pacient-specifických modelů autoři sestavili dynamický model, který slouží jako základ pro plánování léčebného procesu a rehabilitace [5,10,12].

V práci je řešen typ pacient-specifických modelů pro predikci výstupů léčby, kde pro praktickou část byla vybrána diagnóza komunitní pneumonie a rozhodovacím problémem modelování byla predikce faktorů ovlivňující pravděpodobnost hospitalizace.

## **1.2 Příklady využití specifických modelů s diagnózou komunitní pneumonie**

Jak již bylo zmiňováno, pacient-specifické modely se využívají u konkrétní diagnózy a pracují s konkrétními daty daného pacienta. Následující vybrané příklady studií se zabývají diagnózou komunitní pneumonie. Jedná se o velmi vážný klinický stav, kdy je pro lékaře velmi důležitý aktuální stav, ve kterém se pacient nachází, přidružené choroby, výsledky laboratorních a fyzikálních testů. Autoři ve svých studiích předvídají vždy nejhorší možné následky, mezi které patří především úmrtí do 30 ti dnů od hospitalizace, závažné komplikace spojené s možnou hospitalizací na oddělení JIP nebo sepsí [13,14,16,15,17].

Posouzení pravděpodobného klinického průběhu stavu pacienta s komunitní pneumonií může velmi výrazně ovlivnit rozhodování lékařů o tom, zda bude pacient léčen ve zdravotnickém zařízení nebo pouze ambulantně. Rozhodování o léčbě také ovlivňuje využívání finančních zdrojů (celého zdravotnického systému) a lůžkovou kapacitu zdravotnického zařízení. Hlavním cílem je vždy zdraví pacienta, poskytování kvalitní a bezpečné zdravotní péče, ale také vhodné využití zdrojů zdravotnického systému [13,14,16,15,17].

### 1.2.1 Predikce nepříznivého průběhu a vážných komplikací u pacientů s komunitní pneumonií

Mezi autory této studie z roku 2005 patří Georgy F. Cooper a Vijoy Abraham. Tito autoři se zabývají pacient-specifickými modely. Studie popisuje vyhodnocení statistických metod, které predikují vážné komplikace u pacientů s komunitní pneumonií. Zahrnuje zejména vážné komplikace, jako je úmrtí do 30 dnů od stanovení diagnózy, hospitalizace na jednotce intenzivní péče (JIP), respirační selhání či srdeční zástava [15].

Zmiňovaná studie vychází především z výsledků předchozích let, kdy v USA byla každoročně diagnostikována komunitní pneumonie u 4,4 miliónů pacientů, z nichž 1,2 miliónů pacientů muselo být hospitalizováno. Komunitní pneumonie měla na svědomí celkem 64 milionů dnů omezené aktivity pacientů a 39 miliónů dnů pracovní neschopnosti. Celkové náklady finančních prostředků vynaložených v USA na léčbu komunitní pneumonie byly odhadnuty na 9 miliard dolarů ročně [15].

Cílem této studie bylo především vyvinout co nejlepší model pro predikci vážných komplikací a nepříznivých následků, a tím odhadnout dopad na kvalitu zdravotní péče a náklady s tím spojené. Pro studii zvolili tři rozdílné zeměpisné oblasti a pět zdravotnických zařízení. Mezi ně patří *University of Pittsburgh Medical Center*, která patří mezi univerzitní fakultní nemocnici s kapacitou 942 lůžek. Další nemocnicí byla *St. Francis Medical Center* s kapacitou 427 lůžek. Ostatní nemocnice byly vybrány v Pittsburghu a v Bostonu, kde data z minulosti naznačovala velké rozdíly v predikci o typu léčby [15].

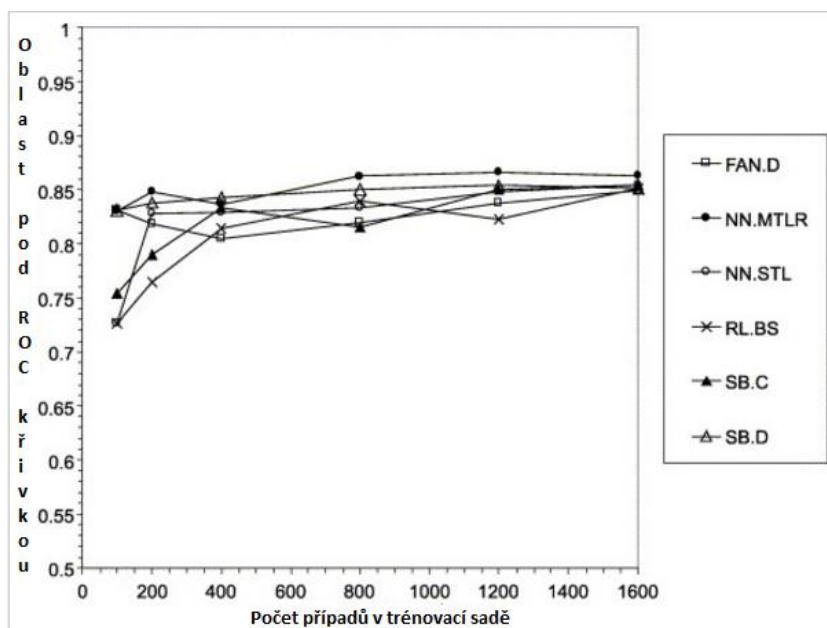
Mezi použité metody patřil jednoduchý Bayesův model, logistická regrese a model neuronové sítě. Data byla čerpána z PORT databáze, která byla sestavena v minulosti pomocí kohortních studií. Autoři vytvořili retrospektivní analýzu vlastního souboru, která se skládala z 1601 konkrétních případů pacientů, a sestavili 11 statistických modelů, které predikovaly vážné následky. Primárním cílem bylo zhodnotit úspěšnost léčby a provést rozbor komplikací (porovnávat koncové body) a vyvinout tak nejlepší model, který předpovídá nejzávažnější možné výsledky. Autoři studie dále vyhodnotili kvalitu zdravotní péče a dopady potřebných finančních nákladů za použití modelu neuronové sítě. V této studii bylo prokázáno, že není vhodné mít úmrtí jako hlavní prediktivní faktor, ale je důležité zahrnout i vážné komplikace. V následující tabulce 2 je 79 případů rozděleno dle výsledku v souboru dat a rozděleny do sedmi tříd. Pouze u 20 % pacientů mělo při výpočtech koncový bod úmrtí [15].



Tabulka 2: Predikce katastrofálních výsledků [15]

Třída	Závažné komplikace	Smrt	Přijetí na JIP	Počet případů pacientů
1.	Žádné	Žádná	Ano	27 (34%)
2.	Žádné	Ano	Žádný	16 (20%)
3.	Žádné	Ano	Ano	3 (4%)
4.	Ano	Žádná	Žádný	14 (18%)
5.	Ano	Žádná	Ano	4 (5%)
6.	Ano	Ano	Žádný	12 (15%)
7.	Ano	Ano	UES	3 (4%)

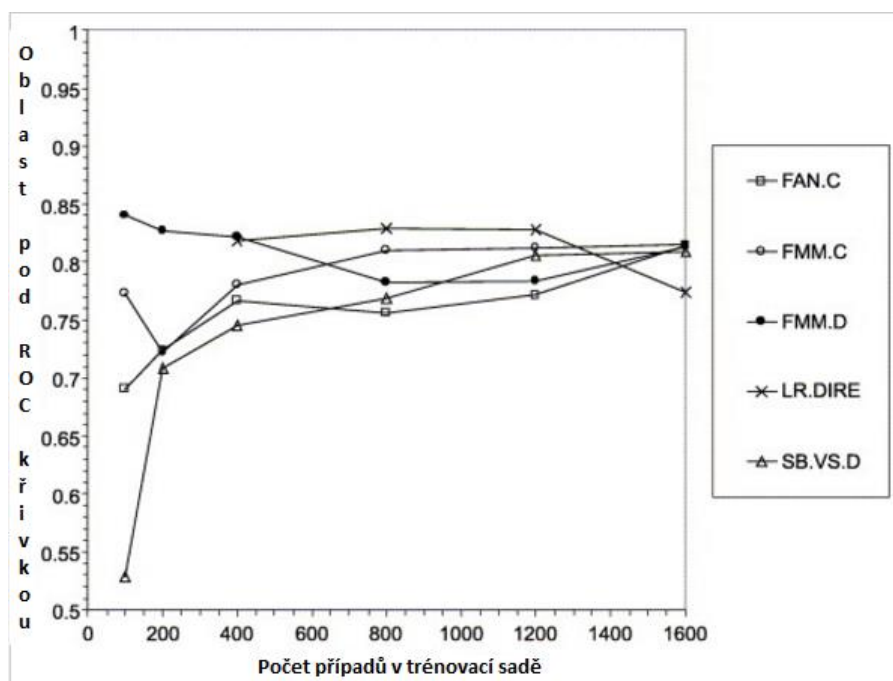
Na následujícím obrázku jsou vyobrazeny ROC křivky modelů, které se statisticky neliší na úrovni hladiny významnosti 0,05. Křivka ukazuje vztah mezi specificitou a senzitivitou daného testu [15].



Obrázek 3: Modely v oblasti ROC křivky s nejlepšími výsledky [15]

- FAN.D – spojení jednoduchého Bayesova modelu s konečnými modely, 158 diskretních proměnných
- NN.MTLR – model neuronové sítě
- NN.STL – model neuronové sítě, 196 proměnných z PORT databáze
- RL.BS – systém založený na pravidlech, použito 196 proměnných
- SB.C – jednoduchý Bayesův model, 161 proměnných, 35 spojitých
- SB.D – jednoduchý Bayesův model, 158 diskretních proměnných

Naopak na obrázku 4 můžeme vidět modely, které se v této studii naopak moc neosvědčily.



Obrázek 4: Modely v oblasti ROC křivky s nejhorsími výsledky [15]

- SB.VS.D – Bayesův model, 46 proměnných na základě vyhledávacích postupů
- FAN.C – model, který spojuje jednoduchý Bayesův model, obsahuje 161 spojitéch a diskretních proměnných
- FMM.C – konečné modely obsahující 161 spojitéch a diskretních proměnných
- FMM.D – konečné modely obsahující 158 diskretních
- LR.DIRE – model logistické regrese

Z celého počtu pacientů v souboru – 686 pacientů, zažilo celkem 79 pacientů katastrofální výsledek, z toho pouze 16 pacientů mělo za výsledek úmrtí, ostatní pacienti byli hospitalizováni na oddělení JIP. Závěrem této studie je, že není vždy důležité předvídat pouze úmrtí pacienta, ale podstatné je také počítat s možným orgánovým selháním nebo sepsí. Hospitalizace je spojená s vyššími náklady, protože je léčba spojena s podáváním intravenózních antibiotik. Pokud je pacient v ambulantní péči, nejsou náklady na perorální antibiotika tak vysoké [15].

## 1.2.2 Pravidlo predikce pro identifikování pacientů s komunitní pneumonií s nízkým rizikem

Mezi autory následující studie z roku 1997, která probíhala ve Spojených státech, patří Michael J. Fine a kol. Cílem této studie bylo snížení nákladů spojených s hospitalizací pacientů s komunitní pneumonií. Sestrojené objektivní modely mohou pomoci lékařům predikovat účinek zvoleného způsobu léčby. Kritéria pro indikaci léčby hospitalizace nejsou konzistentní. Pro sestavení modelů jsou rovněž důležitá data z minulých let. Hlavním účelem této studie bylo především vytvořit pravidlo pro stanovení prognózy, které by mnohem přesněji identifikovalo pacienty s komunitní pneumonií, kteří mají nízké riziko úmrtí do 30 dnů od stanovení diagnózy komunitní pneumonie [13].

Pro zařazení do této studie byla jasně stanovena pravidla. Pacienti museli mít nejméně 18 let, stanovenou diagnózu pneumonie dle Mezinárodní klasifikace onemocnění – MKN 10. Zároveň byli vyloučeni pacienti s osobní anamnézou, ve které se vyskytuje syndrom získané imunitní nedostatečnosti a také pacienti, kteří trpí onemocněním AIDS nebo jsou HIV pozitivní. Pacienti, u kterých byla diagnostikována pneumonie před více jak sedmi dny a nebyli hospitalizováni, taktéž nebyli zařazeni do této studie. V této studii bylo zároveň stanoveno pravidlo nejhoršího možného výsledku – úmrtí do 30 ti dnů od stanovení diagnózy. S tímto pravidlem také počítají i ostatní studie. V minulosti se používala šedesáti denní mortalita jako nejhorší výsledek [13].

Klinické profily pacientů byly rozděleny dle kohortních studií do 5 tříd, kde byli pacienti identifikováni dle demografických proměnných – věk, pohlaví, zdravotnické zařízení a dle fyzikálních nálezů – tepová frekvence, dechová frekvence, systolický krevní tlak, teplota a duševní stav [13].

- *Pacienti třídy I* – medián věku 35 – 37 let, žádné abnormality fyzikálního vyšetření.
- *Pacienti třídy II* – typicky střední věk 58 – 59 let, žádné abnormality fyzikálního vyšetření.
- *Pacienti třídy III* – střední věk až 72 let, alespoň jedna abnormalita fyzikálního vyšetření, radiografická abnormalita.
- *Pacienti třídy IV a V* – věk > 75 let, dělení IV a V především dle příslušných kategorií rizikových faktorů.

Výhodou prognostických modelů je především explicitní definování rozhodovacích ukazatelů pro snadné hodnocení příznaků již při přijetí pacienta, snadné rozřazení pacientů do tříd na základě informací získaných na počátku terapie. Zhodnocuje samotné fyzikální vyšetření, které umožňuje lékařům z části ušetřit prostředky za laboratorní testy. Pravdou je, že ne vždy se před samotnou hospitalizací laboratorní testy realizují a vyhodnotí. Často tyto testy nelze provádět ambulantně. Průběh komunitní pneumonie může být nekomplikovaný s krátkou dobou léčby nebo se mohou rozvinout komplikace, v některých případech i fatální. Prognóza se pohybuje od rychlého zotavení až k těm nejhorším možným výsledkům, jako je již zmiňované úmrtí. Výsledek studie také ukazuje, že pacienti mladší 50 let mohou být ve většině případů léčeni ambulantně, pokud se u nich nevyskytnou abnormality ve výsledcích

fyzikálního vyšetření. Vždy je totiž perorální antibiotická terapie a ambulantní léčba levnější. Důležité je poznamenat, že ve studii je uvažován předpoklad, že existuje velký podíl pacientů s nízkým rizikem hospitalizace [13].

Výsledkem je odvozené pravidlo pro predikci, které identifikuje pacienty s komunitní pneumonií s nízkým rizikem úmrtí a jiných vážných následků. Hlavním předpokladem je, že velká část pacientů s nízkým rizikem, kteří byli hospitalizováni, lze léčit ambulantně perorálními antibiotiky nebo pouze krátkou hospitalizací [13].

### 1.2.3 Model rozhodující cesty pro specifické modelování péče o pacienty

Tato studie byla publikována v listopadu 2013 a mezi autory patří Antonio Ferreira, Gregory F. Cooper a Shyam Visweswaran. Autoři v této studii dokazují, že pacient specifické modely jsou pro modelování lepší než modely populačně široké. Využívají především rozhodujícího stromu – Classification and regression trees (CART). Takové algoritmy jsou v praxi snadno aplikovatelné a přehledné. Porovnávají především algoritmus PSDP-BA a PSDP. Celkově lze říci, že pozorování vyvážené přesnosti s populačním algoritmem nejsou statisticky významně odlišné na úrovni 0,05 hladiny významnosti, od algoritmů PSDP-BA a PSDP-IG [1].

Soubor pacientů s diagnózou komunitní pneumonie zahrnoval tři zeměpisné oblasti: Pittsburgh, Boston a Halifax, Nové Skotsko. Kritéria pro zařazení pacientů do studie byla následující: pacienti museli být starší 18 ti let, musela být přítomnost minimálně dvou příznaků komunitní pneumonie a v neposlední řadě musel být doložen RTG snímek s nálezem. Dále v článku byly specifikovány vážné komplikace, mezi které patří stejně jako u již výše zmiňované studie [13] úmrtí do 30 ti dnů od přijetí, srdeční nebo respirační selhání a specifické těžké komplikace [1].

Tabulka 3: Porovnání použitých algoritmů – Pacient-specifický model [1]

Dataset	CART	PSDP-BA	PSDP-IG
pneumonie	<b>0,6133</b> [0.5644,0.6623]	0,6025 [0.5527,0.6523]	0,5542 [0.5244,0.5841]
sepsc-d	<b>0,6268</b> [0.5730,0.6805]	0,6158 [0.5671,0.6645]	0,6011 [0.5597,0.6424]
sepsc-s	0,6187 [0.5832,0.6542]	0,6200 [0.5884,0.6516]	<b>0,6355</b> [0.6043,0.6667]
srdeční selhání-d	0,5358 [0.5199,0.5518]	<b>0,5495</b> [0.5226,0.5763]	0,5232 [0.5104,0.5361]
srdeční selhání-c	0,5806 [0.5628,0.5984]	<b>0,5832</b> [0.5678,0.5987]	0,5660 [0.5564,0.5767]

Tabulka č. 3 uvádí vyváženou přesnost získanou ze tří algoritmů na daných pěti cílech. Ve výsledku studie hodnotí srovnatelnost algoritmů PSDP s CART, jelikož nejsou nijak významně statisticky rozdílné na hladině významnosti 0,05. Výrazně vyznačená vyvážená přesnost je u každého algoritmu ta nejvyšší průměrná [1].

Tabulka 4: Porovnání použitých algoritmů – Populačně široký model [1]

dataset	CART	PSDP-BA	PSDP-IG
pneumonie	0,6626 [0.6254,0.6999]	0,7062 [0.6769,0.7355]	<b>0,7437</b> [0.7166,0.7710]
seps-d	0,6670 [0.6172,0.7168]	<b>0,7539</b> [0.7191,0.7889]	0,7492 [0.7117,0.7867]
seps-s	0,6614 [0.6314,0.6915]	0,7142 [0.6881,0.7404]	<b>0,7263</b> [0.6995,0.7532]
srdeční selhání-d	0,6803 [0.6527,0.7080]	0,7182 [0.6962,0.7404]	<b>0,7235</b> [0.7020,0.7450]
srdeční selhání-c	0,6531 [0.6348,0.6715]	0,6635 [0.6477,0.6794]	<b>0,6963</b> [0.6811,0.7117]

Tabulka č. 4 hodnotí tři algoritmy dle modelů Populačně-širokých. Celkově lze říct, že pozorované hodnoty algoritmů se výrazněji liší a jsou významně odlišné na hladině významnosti 0,05. Jedinou výjimkou může být rozdíl u získaných hodnot v datové sadě srdeční selhání. Zvýrazněné hodnoty jsou ty nejvyšší. Studie dokázala, že pacient-specifické modely jsou přesnější oproti modelům populačně-širokým [1].

#### 1.2.4 Pacient-specifické modely pro předvídání výsledků u pacientů s komunitní pneumonií

Studie zabývající se pacient-specifickými modely pro predikci výsledků u pacientů s diagnózou komunitně získané pneumonie byla zveřejněna v roce 2005 a jejími autory jsou Shyam Visweswaran a Gregory F. Cooper. V této studii autoři zkoumali dva modely pacient-specifické a čtyři modely populačně-široké pro předvídání vážných komplikací. Jak již bylo několikrát zmiňováno, je předvídání vážných komplikací důležité pro rozhodnutí o typu léčby – hospitalizace nebo ambulantní léčba. Plošné modely jsou odlišné v tom, že se dají

použit v průměru ve všech budoucích případech, naopak modely pacient-specifické modely se indukují pro určitý případ pacienta [2].

Zlepšení přesnosti výsledku předpovídání má hlavní potenciál pro zlepšení zdravotnické péče a efektivnější využívání zdravotnických prostředků. Získaná komunitní pneumonie je významným zdravotním problémem, protože její následky mohou být vážné. K vypracování výsledků použili četné pravděpodobnostní modely založené na statistice. Mezi použité metody ve studii patří logistická regrese, neuronová síť, rozhodovací stromy a Bayesovská síť. Téměř v každém z případů jsou převzata data ze souboru pacientů s úmyslem jejich uplatnění na budoucích případech → populačně široký model. Oproti tomu stojí pacient-specifický model → konstruován na konkrétní případ pacienta [2].

Celkový počet pacientů s diagnózou komunitní pneumonie byl 2287. Data shromáždili pomocí prospektivní kohortní studie hospitalizovaných pacientů a pacientů léčených ambulantně. Studie probíhala v pěti nemocnicích ve třech geograficky rozdílných lokalitách – Pittsburgh, Boston a Halifax - Nové Skotsko. Pacienti museli splnit základní kritéria přijetí do této studie: věk alespoň 18 let, jeden nebo více příznaků související s komunitní pneumonií a RTG snímky s nálezy na plicích zhotoveny do 24hodin od přijetí do nemocnice [2].

Z celkového počtu 4002 splnilo požadavky 2287 pacientů, kteří byli zařazeni do této studie. Jedním z hlavních cílů projektu bylo stanovit kritéria, na základě kterých byla hodnocena prognóza pacientů s komunitní pneumonií. Vytvořit jakési vodítko, které by mohlo lékařům pomoci při rozhodování o typu léčby. Jako u předchozích výše zmiňovaných studií i tady byla sestrojena ROC křivka, která slouží k hodnocení klasifikačního testu (vztah mezi specificitou a senzitivitou). V následující tabulce vidíme výsledky různých použitých metod. Metoda pro výpočet chyby je uvedena ve třetím sloupci. V posledním sloupci jsou uvedeny počty pacientů, kteří mohli být léčeni ambulantně. V posledním řádku je uvedena skutečná míra chyby a skutečný počet pacientů léčených doma [2,18,19].

Tabulka 5: Provedené metody [2]

<b>Metoda</b>	<b>AUC s 95% intervalem spolehlivosti</b>	<b>Chyby a procento chyba s 95% interval spolehlivosti</b>	<b>Počet a procento (95% interval spolehlivosti) pacientů léčených doma</b>
Logistická regrese (LR)	0,741 [0,681, 0,802]	3 (1,8% [0,36%, 5,07%])	170 (24,8% [21,59%, 28,19%])
Neural Network (ANN)	0,828 [0,783, 0,873]	3 (1,7% [0,36%, 5,01%])	172 (25,1% [21,87%, 28,49%])
<i>k</i> -Nearest Neighbor ( <i>k</i> - NN)	0,787 [0,738, 0,837]	3 (1,5% [0,31%, 4,30%])	201 (29,3% [25,92%, 32,86%])
Jednoduché Bayes (SB)	0,850 [0,817, 0,883]	5 (1,3% [0,43%, 3,33%])	368 (53,6% [49,83%, 57,42%])
Modifikovaná LBR	0,861 [0,826, 0,896]	5 (1,2% [0,41%, 2,89%])	400 (58,3% [54,52%, 62,03%])
PSA	0,853 [0,818, 0,876]	5 (1,3% [0,42%, 2,97%])	390 (56,8% [53,05%, 60,59%])
Aktuální péče	-	5 (1,8% [0,58%, 4,12%])	280 (40,8% [37,11%, 44,60%])

V roce 1994 byly odhadované náklady na léčbu jednoho hospitalizovaného pacienta 7.517 dolarů v porovnání s náklady na ambulantně léčeného pacienta 421 dolarů. Pokud by následovala léčba dle doporučení upraveného modelu pro všechny pacienty s komunitní pneumonií, snížily by se náklady o 17,5%. To by vedlo k úsporám ve výši více než 1,5 miliardy dolarů ročně [2].

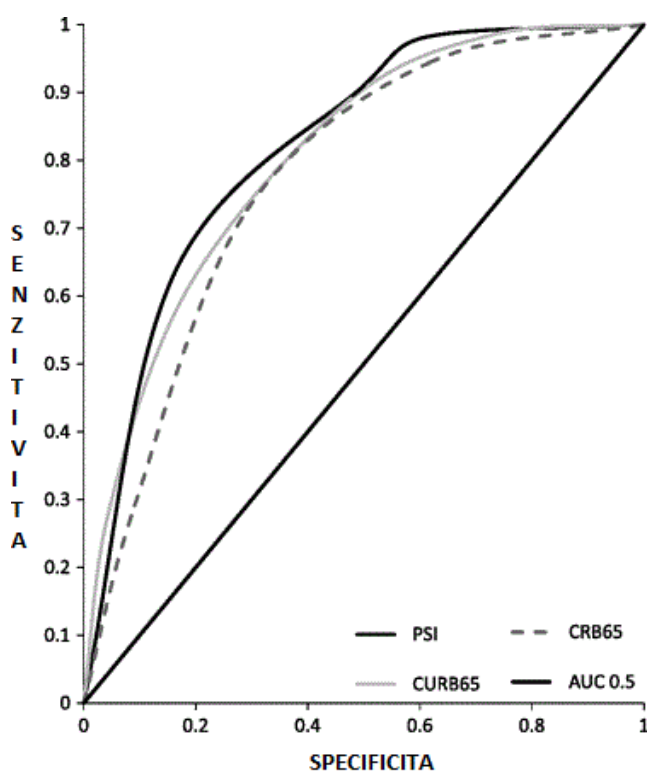


### 1.2.5 Posuzování závažnosti pro predikci mortality u hospitalizovaných pacientů s komunitní pneumonií

Článek byl uveřejněn v roce 2010 a mezi autory patří James D Chalmers, Aran Singanayagam a Ahsan R Akram. Studie probíhala v Edinburgu ve státě Texas - USA. K posuzování závažnosti komunitní pneumonie existují tři systémy, podle kterých se mohou lékaři rozhodovat o typu léčby. Mezi ně patří CURB65 – Britský systém, CRB65 – Německý systém a PSI – Americký systém. Mezinárodní skórovací systémy budou více popsány v následující kapitole [20].

Na základě údajů z minulých let bylo zjištěno, že úmrtnost se v daném státě pohybuje mezi 6 % a celých 14 % pacientů s diagnózou pneumonie musí být hospitalizováno na oddělení JIP. Klíčovým faktem, ale zůstává možné přeceňování nebo naopak podceňování úsudků lékařů. Stejně jako u výše zmiňovaných studií, hlavním cílem je zpřesnit prognózu a optimalizace finančních prostředků daného zdravotnického zařízení [20].

Autoři zařadili pacienty do nízké, střední a vysoké rizikové kategorie dle případných rizik. Studie čerpala data z již proběhlých studií a na základě meta-analýzy, ukázala rozdíl v mezinárodních bodových systémech. Použitelnost bodového systému CRB především pro použití v ambulanci. Každý ze systémů stále predikuje 30 denní úmrtnost od hospitalizace pacienta [20].



Obrázek 5: Porovnání mezinárodních systémů [20]

Autoři v této studii vytvořili systematický přehled mezinárodních bodových systémů pro predikci mortality pacientů s komunitní pneumonií. Systémy CRB65 a CURB65 zahrnují pouze čtyři hodnoty z vyšetření pacientů na ambulanci oproti systému PSI, kde body se rozdělují téměř u 20 hodnot. Meta-analýzou dat bylo zjištěno, že zmiňované systémy se od sebe zásadně neliší. Lékaři se shodují na tom, že bodový systém PSI je velmi obtížné používat v praxi na ambulanci vzhledem k hodnocení většího množství dat. Autoři hodnotí systém PSI jako vhodný pro použití u pacientů s nízkým rizikem úmrtí [20].

### 1.3 Komunitní pneumonie

Komunitní pneumonie je nejčastější ze všech infekcí a nejzávažnější plicní onemocnění. Mortalita ani v době nejúčinnějších antibiotik neklesá nijak zásadně. Pneumonie je častou příčinou úmrtí v rozvinutých zemích. V České republice je roční incidence přes 100 000 případů, více než 20 000 osob je hospitalizovaných a kolem 3 000 pacientů na tuto diagnózu zemře. V celosvětovém měřítku řadíme pneumonii na třetí místo. Ročně na ni zemře tři až pět miliónů lidí [21,22].

Definice pneumonie → jedná se o akutní zánět v oblasti dýchacích bronchiolů a intersticia plic. Klinicky tento stav označujeme jako nález infiltrátu na skiagramu hrudníku. Dále jsou přítomny další příznaky infekce v oblasti respiračního traktu – kašel, dušnost, bolest na hrudi, horečka [4,22,23,24].

Definice komunitní pneumonie → epicentrem vzniku je běžný životní styk mimo prostředí zdravotnického zařízení a bez souvislosti k lékařským výkonům. Jsou způsobeny běžnými patogeny. Vedle komunitní pneumonie rozeznáváme právě i nozokomiální pneumonii, která je vázána na nemocniční prostředí [4,22,23,25].

Mezi nejčastější původce patří především *Streptococcus pneumoniae*, *Klebsiella pneumoniae*, *Escherichia coli*, *Mycoplasma pneumoniae*, *Chlamydia pneumoniae* (chlamydia) a *Haemophilus influenzae*, dále pak viry chřipky. U starších pacientů může být toto onemocnění souviset i s přidruženými nemocemi vyvolanými i jinými bakteriemi, mezi které patří například i *Escherichia coli* [22,23,25].

Klasifikace pneumonie [22]:

- LEHKÁ – nekomplikovaná, obvykle léčena ambulantně.
- ZÁVAŽNĚJŠÍ – stav, který je potřeba léčit za hospitalizace.
- TĚŽKÁ – vyžaduje vždy hospitalizaci na JIP nebo oddělení ARIM.

Existují tři systémy pro porovnávání závažnosti pneumonie a predikce mortality. Tyto systémy zohledňují aktuální údaje o současném stavu pacienta, jeho fyzikálního vyšetření a laboratorních výsledků. Jedním ze systémů je škála Britské hrudní společnosti, tj. systém CRUB. Tento systém hodnotí stav vědomí, hladinu močoviny, projevy dušnosti a hodnotu krevního tlaku. Při hodnocení se za každý tento znak přiděluje bod [23,25,26,27,28,29].

Tabulka 6: Britský systém určování rizika mortality – CURB [26]

C (confusion)	Zmatenost
U (urea)	Urea nad 7 mmol/l
R (respiratory rate)	Dechová frekvence nad 30/min
B (blood pressure)	Krevní tlak (90/60 torr)

Další systém, který existuje je německá škála CRB-65. Německý systém má výhodu, že eliminuje hodnoty z laboratorních vyšetření. Některé zdroje, ale uvádějí nepřesnost těchto dvou systémů. A to z důvodu, že přeceňují závažnost komplikací především u starších pacientů [23,25,26,27,28,29].

Tabulka 7: Německý systém CRB-65 [26]

C (confusion)	Zmatenost
Dechová frekvence nad 30/min	Dechová frekvence nad 30/min
B (blood pressure)	Krevní tlak (90/60 torr)
65	Věk nad 65 let

Systém Americké hrudní společnosti hodnotí závažnost stavu dle rizika mortality. Rozřazují nemocné do skupin dle počtu rizikových bodů. U pacientů do věku 50 ti let nejsou přidělovány žádné rizikové body. Jejich mortalita je velmi nízká [4,22,27,30].

Tabulka 8: Kritéria rizika dle systému Americké hrudní společnosti [30]

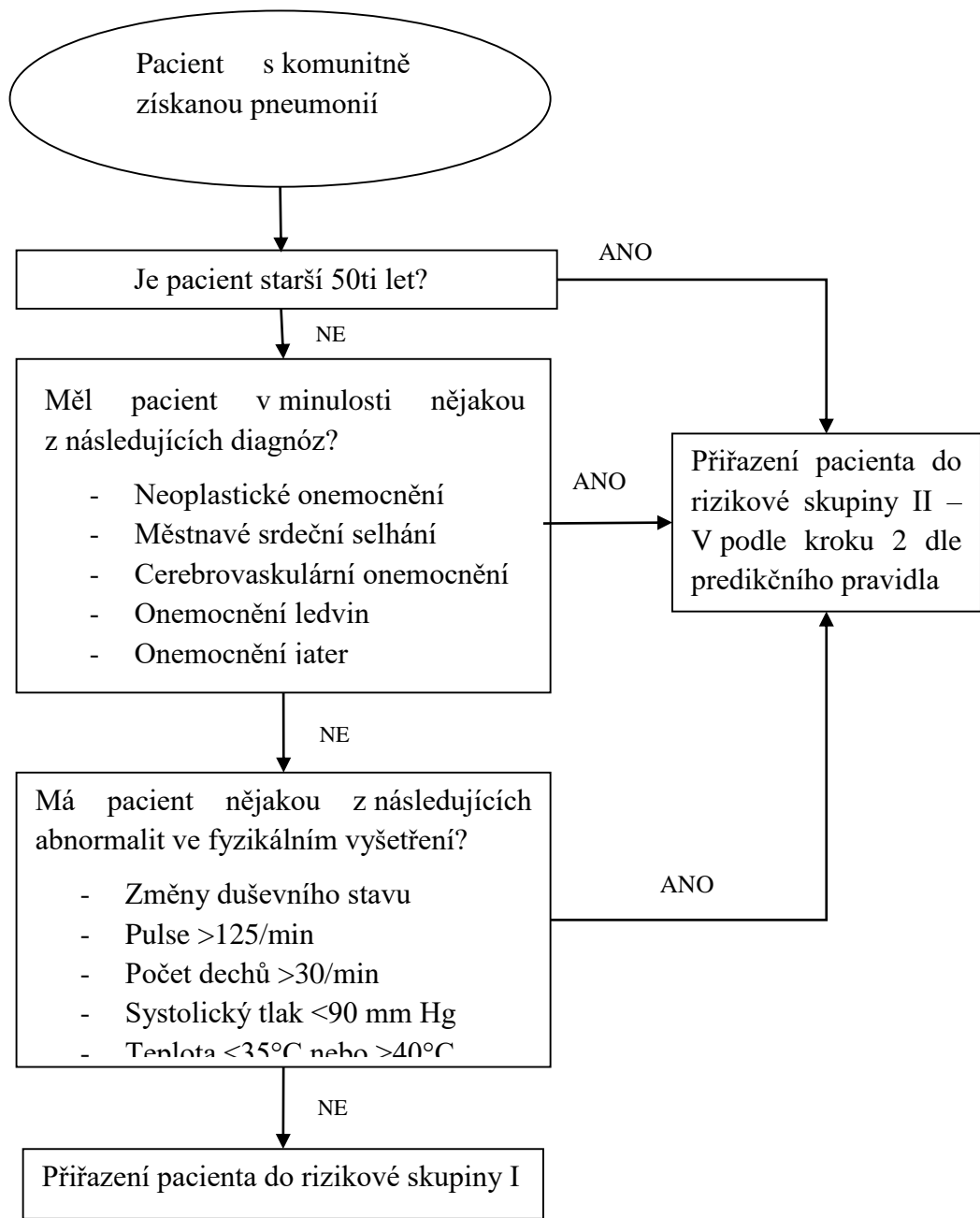
<b>Riziková kritéria</b>	<b>Body</b>
Muži	Věk
Ženy	Věk -10
Ústav sociální péče	Věk +10
<b>Přidružené emoci</b> <ul style="list-style-type: none"> <li>• Nádory +30</li> <li>• Hepatopatie +20</li> <li>• ICHS +10</li> <li>• CMP +10</li> <li>• Nefropatie +10</li> </ul>	
<b>Laboratorní a RTG nálezy:</b> <ul style="list-style-type: none"> <li>• pH pod 7,35 +30</li> <li>• urea nad 10,7 mmol/l +20</li> <li>• natrium pod 130 mmol/l +20</li> <li>• glykemie nad 13,9 mmol/l +10</li> <li>• hematokrit pod 30% +10</li> <li>• pO<sub>2</sub> pod 8 kPa +10</li> <li>• pleurální výpotek +10</li> </ul>	
<b>Fyzikální vyšetření:</b> <ul style="list-style-type: none"> <li>• alterace mentálního stavu +20</li> <li>• dechová frekvence nad 30 dechů/min +20</li> <li>• systolický TK pod 90 mm HG +20</li> <li>• teplota pod 35° C nebo nad 40°C +15</li> <li>• pulz nad 125/min +10</li> </ul>	

Tabulka 9: Rizikové skupiny dle systému Americké hrudní společnosti [30].

Skupina	Počet bodů	Mortalita	Doporučení typu léčby
I	< 50♂ / 55♀	0,1%	Ambulantně
II	< 70	0,6%	Ambulantně
III	71-90	2,8%	Ambulantně, event. Krátká hospitalizace
IV	91-130	8,2%	Hospitalizace
V	> 130	29,2%	Hospitalizace – JIP

Výhodou systému PSI je především to, že se jedná především o přesnější bodový systém, který zohledňuje aktuální stav pacienta, zároveň přihlíží na přidružené nemoci a zohledňuje také s laboratorními analýzy [4,27,30].

Kritéria k hospitalizaci pacientů nejsou konzistentní. Mimo objektivní hodnocení skórovacích systémů se lékaři často spoléhají na své subjektivní dojmy z klinického stavu pacienta. V případě, že se jedná o polymorbidního pacienta, se často rozhodnou pro hospitalizaci. Často se setkáváme s přeceňováním vážných komplikací spojená s pneumonií, z toho vyplývá hospitalizování pacientů s nízkými riziky. V medicíně je možné úplně vše. Objektivní modely za použití prognóz pro komunitní pneumonie by mohly pomoci lékařům vyhodnotit rizika u pacientů a zlepšit rozhodování o jejich hospitalizaci, případné ambulantní péči [4,13,20,22].



Obrázek 6: Odvozování predikčního pravidla – identifikace pacientů ve třídě rizik I [13]

Diagnostické hodnocení komunitní pneumonie je důležité pro posouzení závažnosti a pro management terapie, dále je důležité z hlediska prognózy a stanovení rizika vážných komplikací na základě vyšetření. Nejdůležitější léčebným opatřením je včasné podání antibiotika a ověření jeho účinnosti. Rychlost podání antimikrobiálních léčiv a správná detekce antibiotické terapie velmi ovlivňuje léčbu. Ověření účinnosti může být dokázáno snížením leukocytů, snížením C-reaktivního proteinu a zlepšením obrazu RTG [4,25,31].

## 2 Metody

Kapitola je věnována metodám pro vytvoření pacient-specifických modelů, které byly použity u výše zmiňovaných zahraničních studií. Statistické metody počítají s proměnnými týkající se zdravotního stavu pacienta. Výhodou je vytvoření modelů i při chybění některých dat. Počítají s retrospektivními daty vlastního souboru pacientů.

### 2.1 Bayesovská statistika

Do moderní statistiky patří také Bayesovská statistika. Pracuje s podmíněnými pravděpodobnostmi a zpřesňuje pravděpodobnost výchozích hypotéz. Podstata je v Bayesově větě. Klasická statistika pracuje tak, že stanoví pravděpodobnost určité události na základě známých faktů z minulosti. Bayesovskou statistiku použijeme tam, kde to není možné. Její použití proto cílíme tam, kde se pracuje s nejistotami, např. při diagnostickém testování [32,33,34].

Interpretace pravděpodobnosti Bayesovské statistiky může být chápána jako rozšířená výroková logika, která uvažuje s hypotézami. Takovou statistiku řadíme do kategorie důkazních pravděpodobností vyhodnocování hypotéz. Používá se u randomizovaných studií [32,33,34].

Bayesův teorém je výsledek podmíněných pravděpodobností, který mluví o dvou náhodných veličinách  $y$  a  $\theta$ . Jedná se o vzorec, který nám ukazuje, jak stávající řešení vyjádříme rozdělením pravděpodobností a modifikováním novými informacemi. Bayesova věta – udává, jak podmíněná pravděpodobnost určitého jevu souvisí s podmíněnou pravděpodobností opačnou [32,35,36,37].

Použitím rozdělených pravděpodobností vyjadřujeme naši nejistotu a tím určování skutečnosti z hlubších principů logického uvažování a racionálního chování. Určujeme kvalitativní předpoklady zaměnitelnosti binárních pozorování, tzn. jejich společnou distribuci. Předpokladem je vzájemná nezávislost podmíněného neznámého parametru  $\theta$ , kde  $\theta$  je distribuční a má omezující frekvenci [33,35,36,37,38].

**Bayesova věta :**

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \quad (2.1)$$

- $p(\cdot)$  ... značí rozdělení pravděpodobností
- $p(\cdot|\cdot)$  ... podmíněné rozdělení
- $y$  ... jev, který představuje data
- $\theta$  ... jev, který představuje parametry statistického modelu

$p(y|\theta)$  je podmíněná pravděpodobnost jevu  $y$  za předpokladu, že nastal jev  $\theta$

$p(\theta|y)$  je pravděpodobnost jevu  $\theta$  podmíněná výskytem jevu  $y$ . Oba jevy  $\theta$  a  $y$  jsou náhodné s pravděpodobnostmi  $P(y)$  a  $P(\theta)$ . Platí, že  $P(\theta) > 0$  [33,35,36,37,38].

Bayesovské statistické metody se většinou využívají v následujících situacích:

- Jednou z nich je, kdy nemá jinou možnost, než použít kvantitativní rozsudky z minulosti, z důvodů nedostatku žádných jiných údajů o aspektu modelu.
- Dalším případem může být také problém střední velikosti z více zdrojů důkazů, kde jsou hierarchické modely postavené na předpokladu dřívějších distribucí. Jejich parametry můžeme odhadnout z dat [33,35,36,37,38].

Bayesovský průměr je metoda odhadu střední hodnoty z populace v souladu s interpretací Bayesovské sítě. Střední hodnota je odhadnuta z některých nebo všech dat dostupného datového souboru. Ty jsou začleněny do výpočtu a tím se minimalizují odchylky, například v případě, že máme malou datovou sadu [32].

Použití Bayesovské metody při hodnocení zdravotnických technologií je přímé kvantitativní využití externích důkazů v plánování, analyzování, monitorování a interpretování. Pracuje na podkladě vícero studií a tak syntetizuje důkazy z více zdrojů [34,38,39].

Bayesova metoda – souhrn [39]:

- interpretuje data ze studií na základě externích důkazů a úsudků,
- věrohodnost hypotéz je založena na výsledcích diagnostického testu,
- využívá se u hodnocení zdravotnických technologií,
- je důležité kriticky hodnotit případové studie.

## 2.2 Klasifikační a regresní stromy (CART)

Rozhodující strom je prediktivní model, který dokáže zmapovat vyjádření o závěrech cílové hodnoty dané položky. Patří mezi přístupy prediktivního modelování. Klasifikační strom je model, kde lze cílovou hodnotu proměnné brát jako konečnou množinu hodnot. Regresní strom je takový model, kdy můžeme cílovou proměnnou exponovat do reálných čísel. Důležité je zmínit, že rozhodovací strom pouze popisuje data, nikoliv finální rozhodnutí. Bereme ho, jako vstup pro rozhodování. Cílem je vytvoření modelu, pomocí kterého můžeme předpovídat hodnotu cílové proměnné. Všechny funkce jsou omezeny časem diskrétní domény s cílem klasifikace [40,41,42].

Modely klasifikačních a regresních stromů se využívají pro konstrukci predikčních modelů z dat. Obecně máme jednu závisle proměnnou a celou sadu nezávisle proměnných. Modely znázorňují vzájemné vztahy mezi pozorovanými veličinami. Klasifikační stromy získáváme rozdělením datového prostoru a vytvoření jednoduchého přehledného systému, který předpovídá každou třídu dat. Jejich grafické prezentování v podobě rozhodujícího stromu je přehledné a vhodné především pro použití v praxi. Pro závislé proměnné se používají především klasifikační stromy spojitě pro neuspořádané hodnoty. Zohledňují také možnou chybu predikce. Zatímco regresní stromy jsou rovněž závislé proměnné, ale spojitě pro diskrétní hodnoty, jako chybu predikce měří druhou mocninu rozdílu mezi pozorovanými daty a vypočítanými daty [41,42,43].



Rekurzivním rozdělením datové sady z populačně širokého modelu lze sestavit model stromu. – klasifikační strom nebo regresní strom. Typ modelového stromu se určuje podle cílové proměnné, která je diskrétní nebo kontinuální. V tomto modelu listy představují cílové hodnoty a spojení listů představují větve, které vedou k cílovým hodnotám. Tyto algoritmy se obvykle provádějí výběrem prediktoru od shora dolů a každý další krok výběrem nejlepší následující proměnné z rozdělení datové sady. Pro výběr prediktorů je kritérium, které je zahrnuto do stromu a společná kritéria zahrnující Giniho index.

Giniho index je informace o přínosu a chybné klasifikaci. Rozhodovací stromy se používají především v lékařství, protože lze strom snadno interpretovat a realizovat v klinickém systému pro podporu rozhodování [1,40,41,42].

Klasifikační a regresní stromy mají výhodu všestranného komplexního využití téměř ve všech oborech nejen ve zdravotnictví, ale také například v ekologii a jiných oborech. Mezi nejnámější rozhodovací stromy patří klíče k určování živočichů či rostlin v biologii. Grafické vyjádření pomáhá k lepšímu porozumění. Díky jejich flexibilitě můžeme zahrnout širokou škálu typů proměnných. Nejen numerické, ale i kategorické a mají schopnost zpracovat chybějící hodnoty [41,42,43].

Modely se dají využít k [47]:

- **Segmentaci** – rozdělení rostlin na základě stejných vlastností do jednotlivých tříd
- **Stratifikaci** – rozvrstvení případů do tříd, např. malý, střední a velký výskyt
- **Predikci** – sestavení pravidel pro předvídaní událostí v budoucnu, např. předvídaní typu léčby u konkrétní diagnózy na základě známých příznaků
- **Identifikaci vztahů** – v rámci podskupin

### 2.2.1 Klasifikační strom - rozhodovací

Klasifikační strom řadíme mezi neparametrické metody, který klasifikuje vzorky do předem daného počtu tříd. Jedná se o lineární hierarchický klasifikátor a postup rozhodování, kdy výsledkem je roztrídění objektů do jednotlivých tříd na základě jejich vlastností [43,44,46,48].

Při vytváření stromů následuje metoda „*divide and conquer*“ (rozděl a podmaň). Data v sadě postupně rozdělujeme na menší podmnožiny a vytváříme tím uzly (větvení) stromu tak, aby v každé podmnožině převládaly vždy data z jedné třídy. Vždy postupujeme ze shora dolů k vytvoření stromu [43,44,46,48].

K vytvoření rozhodovacího stromu je důležitý výběr vhodného atributu = kořen stromu, rozdělení dat na podmnožiny dle hodnot kořene stromu, přidání uzlů pro každou podmnožinu. Celý datový soubor je obsažen právě v kořeni stromu. Pokud některý z uzlů nezahrnuje všechna data z jedné třídy → nová volba kořene stromu. Poslední část stromu jsou listy, které představují některou ze skupin [43,44,46,48].

Při jejich vytváření modelujeme závislost kategoriální proměnné na jedné či více nezávislých prediktorech. Prediktory jsou kategoriální a spojité. Cílem při sestavování

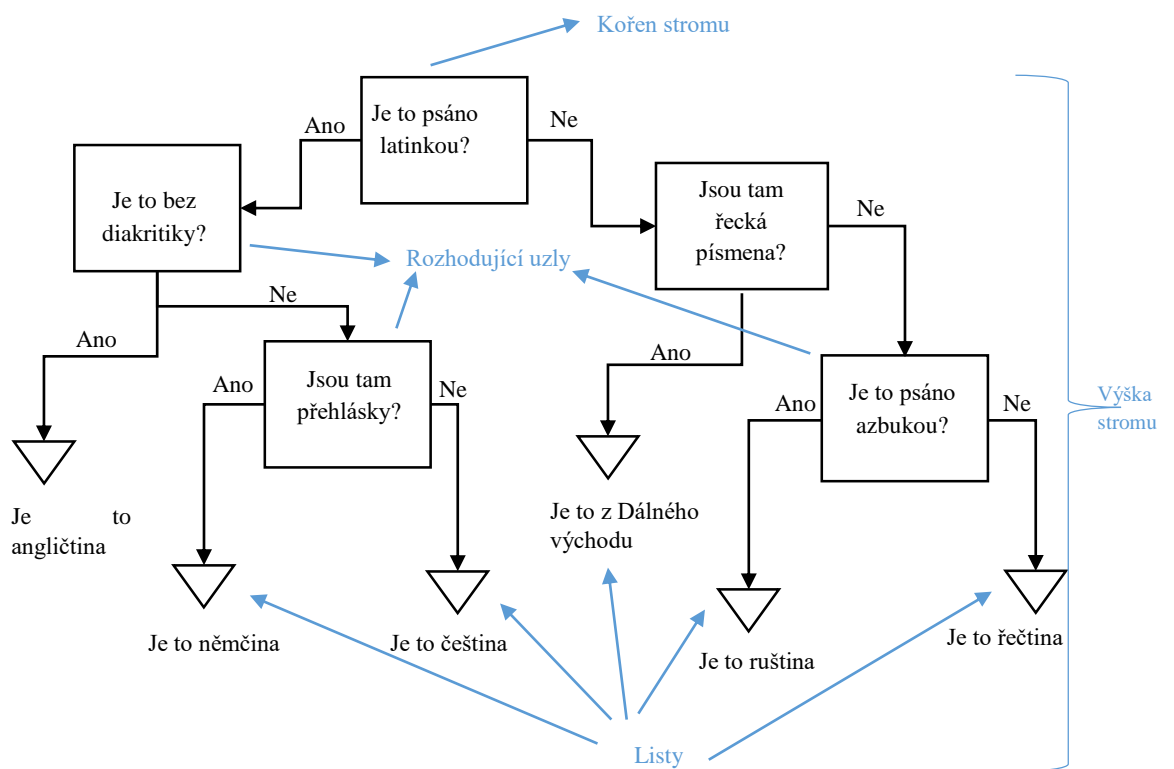
klasifikačního stromu je dosáhnout co nejmenší generalizační klasifikační chyby [43,44,46,48].

$$Rp(T) = P(d_T(X) \neq Y) \quad (2.4.1)$$

P ... sdružené rozdělení vektoru

X ... hodnoty prediktorů

Y ... hodnoty závislé proměnné,  $\{C_1 \dots C_k\}$ ,  $k > 2$   $\{C_1 \dots C_k\} \dots$  jednotlivé třídy



Obrázek 7: Příklad rozhodovacího stromu [1,2,40,41,42]

## 2.2.2 Regresní strom

Regresní strom je velmi podobný klasifikačnímu stromu jen s tím rozdílem, že každému uzlu je přiřazena reálná konstanta, tj. odhad kvantitativní závislé proměnné. Při jeho sestavení modelujeme závislost spojité proměnné na jedné či více nezávislých prediktorech, které jsou kategoriální a spojité [43,44,46,48].

Cílem při sestavování regresního stromu je dosáhnout co nejmenší střední kvadratické chyby.

$$Rp(T) = Ep(Y - d_T(X))^2 \quad (2.4.2)$$

Střední kvadratická chyba s rostoucí velikostí stromu klesá nebo není rostoucí v trénovací sadě. V mnoha situacích v praxi chyba klesá jen do určité velikosti a s následným zvětšováním stromu opět roste [43,44,46,48].

Rozhodující cesta pro patientské modely je určena pro dobré předpovídání diskrétního cíle v zájmu klinického výsledku. U daného pacienta, u kterého známe rysy a vlastnosti, můžeme sestrojít rozhodovací cestu obsahující podmnožinu funkcí a tu pak aplikovat na případ pacienta a předpovědět cílovou hodnotu. Při stavbě rozhodovacího stromu ignorujeme chybějící funkce prediktorů. Obecně platí, že rozhodovací cesta sestrojena pro pacient-specifické modely neexistuje v modelech populačně širokých. Z toho vyplývá, že algoritmy sestrojené pro pacient-specifické modely jsou jednodušší a přesnější pro daný případ pacienta [1].

Příklady algoritmů pro sestavení rozhodovacích stromů::

- PSDP-BA – vyvážená přesnost,
- PSDP-IG – zisk informací.

### 2.2.1 Algoritmus PSDP – BA

Algoritmus, který využívá metodického vyhledávání modelového prostoru rozhodovacích cest. Ačkoliv je tento prostor modelu menší, než prostor u populačně širokého modelu, jedná se stále o vyčerpávající hledání z klinické datové sady. Metodický výzkumné hledání odpovídá hledání krok po kroku. Rozhodující cesta se skládá z kombinace různých funkcí, které jsou současně v testovacím případě a rozdělení pravděpodobností přes hodnoty cílové proměnné. Cílová proměnná se odhaduje z tréninkových sad, které splňují propojení funkcí. Používáme také Bayesův odhad pravděpodobnosti, který je ekvivalentní k velikosti vzorku. Parametr, který určuje, jak moc se vyskytuje vyrovnání v odhadu pravděpodobnostního parametru. Čím vyšší je ekvivalentní velikost vzorku, tím lepší je vyrovnání[1,6].

Pro daný testovací případ pacienta, kde chceme cílovou hodnotu předpovědět, začíná cesta modelem rozhodnutí bez prediktoru. Postupně přidáváme prediktory do cesty a při každém dalším kroku vyhodnocujeme zbývající prediktor pomocí vyvážené přesnosti. Vyvážená přesnost je definována jako aritmetický průměr, která získáme z citlivosti a přesnosti. Můžeme také použít průměrnou přesnost. Úvaha prediktoru, který je dočasně připojen k současné cestě a vyvážená přesnost je odhadnuta z tréninkových případů, které splňují dočasné prodloužení současné cesty. Konkrétněji, jeden z případů se vynechá a Bayesův odhad je použit k odhadu pravděpodobnosti parametrů ze zbývajících tréninkových případů. Cílová hodnota vynechaného tréninkového případu je předpovězena použitím pravděpodobnostního parametru. Proces se opakuje v takovém pořadí, že předpověď se získá pro každý jednotlivý případ, který splňuje požadavky rozhodovací cesty. Vyvážená přesnost pro prediktor se vypočítá dle následujícího vzorce [1,6].

$$Bcc = \frac{1}{2} (\textit{senzitivita} + \textit{specifičnost}) \quad (2.2)$$

Model s vyváženou přesností je přesnější, protože bere v úvahu přesnost každé hodnoty cíle zvlášť. Postupným krokováním se pomocí tohoto algoritmu cesta rozšiřuje do vyčerpání prediktorů, které splňují spojitost funkce nebo dokud nemají stejnou cílovou hodnotu [1,6].

### 2.2.3 Algoritmus PSDP - IG

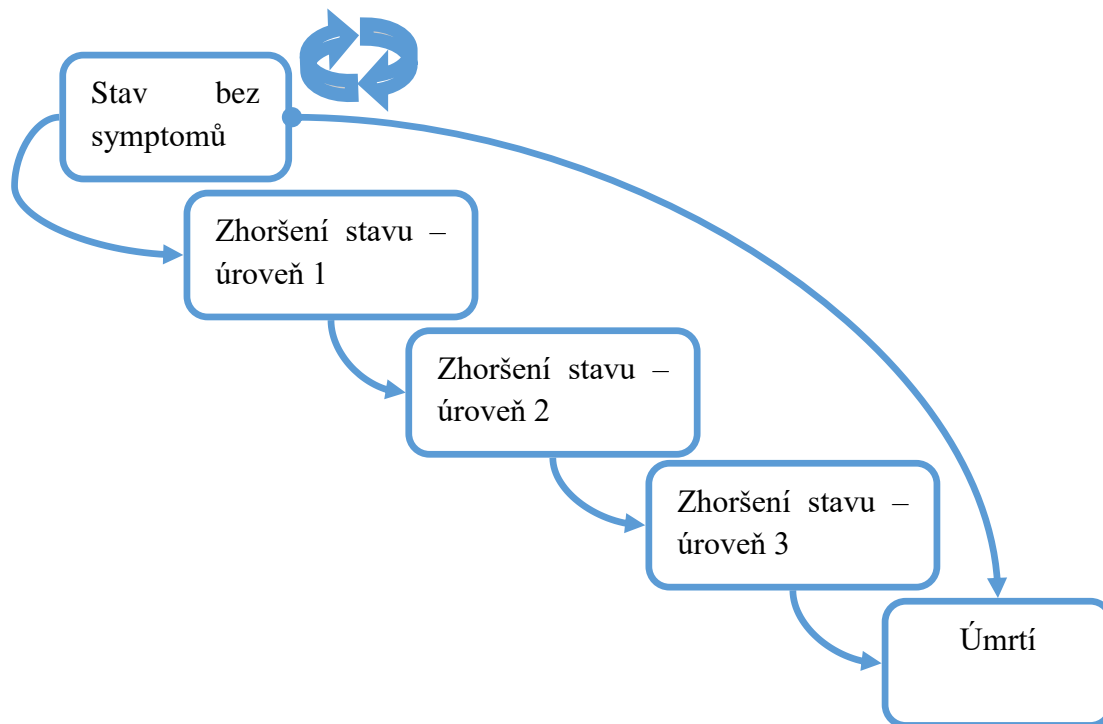
PSDP-IG algoritmus se využívá k informačnímu zisku v místě vyvážené přesnosti, kterou získáme z algoritmu PSDP – BA [1].

## 2.3 Markovův model

Statistický Markovův model je užitečný v plánování programů zdravotní péče, protože zahrnuje také riziko, které je kontinuální v průběhu času dané události. Používá se u složitějších klinických událostí pro přesnost. Pro jednodušší klinické události se obvykle používá právě rozhodující strom. Markovův model předpokládá, že se pacient nachází v jednom konečném zdravotním stavu. Všechny procesy v modelu jsou prezentovány jako přechody z jednoho stavu do druhého. Výhodou je jeho opakovatelnost událostí s časově závislou pravděpodobností. Předpokladem je, že stav budoucí závisí pouze na současném stavu a nebere ohled na sled událostí, které tomuto stavu předcházely [49,50,51,52].

Zásadní rozdíl Markova modelu od jiných statistických modelů používaných ve zdravotnictví je faktor času, který je spojován s pravděpodobností. Časová období se označují jako cykly, které jsou od sebe odděleny a doba jejich trvání závisí na onemocnění a jeho hodnocení, např. měsíční, roční aj. Hodnota nákladů je přiřazena ke každému konkrétnímu cyklu. Dobu, po kterou je pacient v neschopnosti (léčen hospitalizací nebo v domácí péči) je nástrojem pro výpočet očekávaných nákladů. Jak rychle je schopen pacient přejít z jednoho cyklu do druhého (pravděpodobnost přechodů) stanovuje využívání zdrojů. Můžeme hodnotit faktory spojené s chorobou a jejími vnějšími zásahy, které jsou s tím spojeny [51,52].

Při výstavbě Markova modelu je prvním krokem definování stavů onemocnění. Zajímají nás především zdravotní a ekonomické dopady. Jednotlivá stádia onemocnění se vzájemně vylučují a tím se sestavují dané cykly Markova modelu. Pacient nemůže mít více než jeden stav. Příklady u konkrétních diagnóz můžeme vidět třeba u diabetu nebo hypertenze [51,52,53].



Obrázek 8: Vzor Markova modelu ve zdravotní péči [52]

První fáze modelu je stav pacienta bez vážných symptomů, ale jsou známi faktory, které indikují danou chorobu. Značí stav, kdy pacient trpí chorobou, ale riziko úmrtí je nízké. Další cykly jsou ovlivněny progresí dané diagnózy směrem k nejhorším možným výsledkům na podkladě pravděpodobnosti a to je úmrtí. Součet pravděpodobností v každém cyklu se musí rovnat jedné a to z důvodu, že cyklus probíhá v konkrétním čase. Jednoduše řečeno, pravděpodobnost zůstává stejná v jeden okamžik a my odečítáme pravděpodobnost přechodu. Například v otázce střední délky života, řadíme váhu 1 = živý pacient a váhu 0 = zemřelý pacient. Průměrnou délku život vypočítáme provedením modelu o velkém počtu cyklu a sečteme jejich hodnoty [51,52].

Obecně mají Markovy modely velmi široké využití, včetně biologie, chemie, fyziky i matematice. Výhodou jsou pravděpodobnosti přechodů, které jsou řešeny maticí, kdy matice zohledňuje strávený čas a předpokládá výsledek přesně. Nevýhodou je použití při velkém počtu cyklů, kdy na zpracování je důležitá velká kapacita počítačů. Použití z hlediska zdravotní péče slouží jako analytický nástroj pro posouzení konečného výsledku z hlediska hospodářského [51,52].

## 3 Výsledky

Data pro zpracování praktické části byla získána v Klaudiánově nemocnici v Mladé Boleslavi. Pro nahlížení do nemocničního systému bylo nutností podepsat prohlášení o mlčenlivosti (viz.: příloha 1). Respondenty byli vybíráni z databáze podle kódu – mezinárodní klasifikace diagnóz – J18. Sesbírána byla jak data ohledně hospitalizovaných pacientů tak o pacientech léčených ambulantně. Všechna data jsou z roku 2015. Pro každého respondenta byla dokumentace analyzována zvlášť, aby byla získána data o anamnéze, fyzikálním vyšetření, přidružených chorobách a laboratorních výsledcích.

### 3.1 Získaná data

Kritéria pro výběr respondentů byla:

- Pacient musí být starší 18 ti let,
- mít jeden a více příznaků zápalu plic,
- mít RTG nález zápalu plic.

Zároveň byla určena i kritéria, kvůli kterým byly někteří respondenti vyřazeni.

Kritéria, která nebyla způsobilá pro zpracování praktické části: diagnóza AIDS nebo pozitivní protilátky proti HIV. Dále nebyli zařazeni pacienti, u kterých byl předpoklad plicního infiltrátu jiné etiologie (např. plicní edém nebo plicní embolie). Informace, které byly použity, představují první hodnoty, které jsou k dispozici lékařům při přijetí pacienta. Za vážnou komplikaci u pacienta byla považována situace, pokud došlo k úmrtí do 30 ti dnů od potvrzení pneumonie či došlo respirační nebo srdeční zástavě. Vážné komplikace budou dále popsány v práci.

Na základě analýzy současného stavu a výběru vhodných metod, byla vybrána data, které jsou bezprostředně potřeba na zpracování praktické části diplomové práce.

**Potřebná data:**

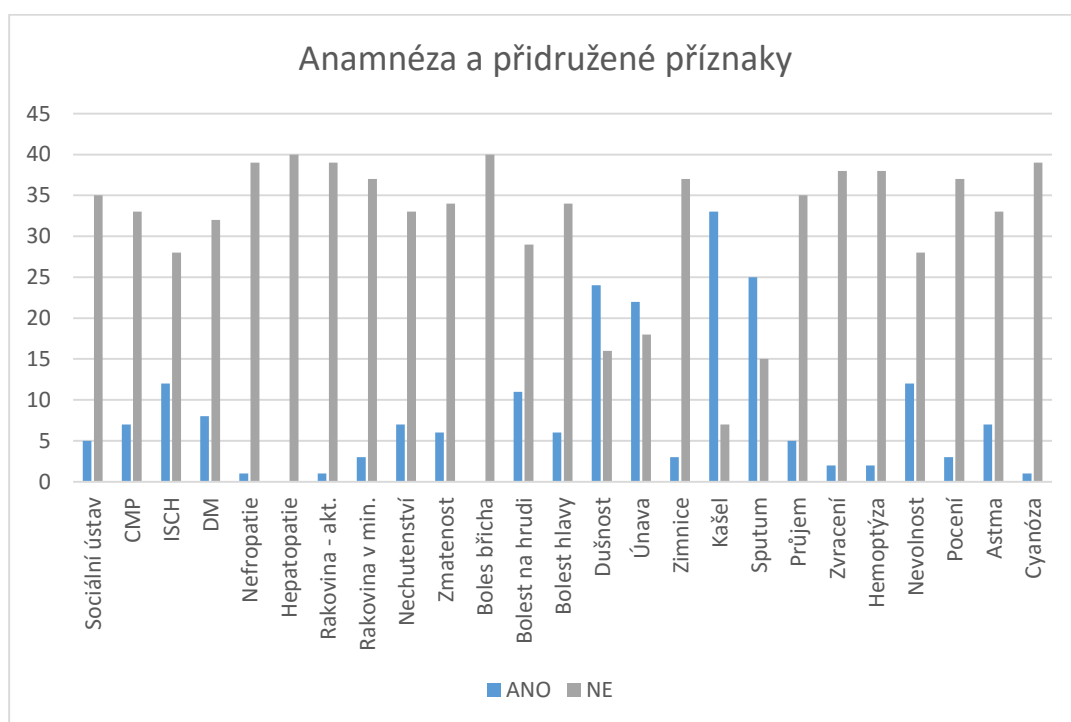
- věk pacienta,
- pohlaví,
- typ léčby – ambulantně/hospitalizací (počet dnů),
- vykázané body pojišťovně,
- sociální anamnéza (např. Ústav sociální péče),
- přidružené choroby – ICHS, CMP, pneumonie v minulosti, aktivní rakovina,
- současný stav pacienta – zmatenost, zimnice, ztráta chuti k jídlu, kašel, sputum, průjem, zvracení, projevy dušnosti aj.,
- fyzikální vyšetření – tepová frekvence, dechová frekvence, tělesná teplota, výška, hmotnost, systolický a diastolický tlak,
- laboratorní testy – leukocyty, hematokrit, pCO<sub>2</sub>, hodnota močoviny, C-reaktivní protein, natrium, glukóza, pH měřené.

Tabulka 10: Získaná data – věk, pohlaví, doba hospitalizace

Pohlaví	ŽENY	MUŽI
Počet respondentů	22	18
Průměrný věk	68	65
Věk - Max	91	89
Věk - Min	29	39
Průměrná doba hospitalizace	9	6
Doba hospitalizace - Max	15	14
Doba hospitalizace - Min	1	1

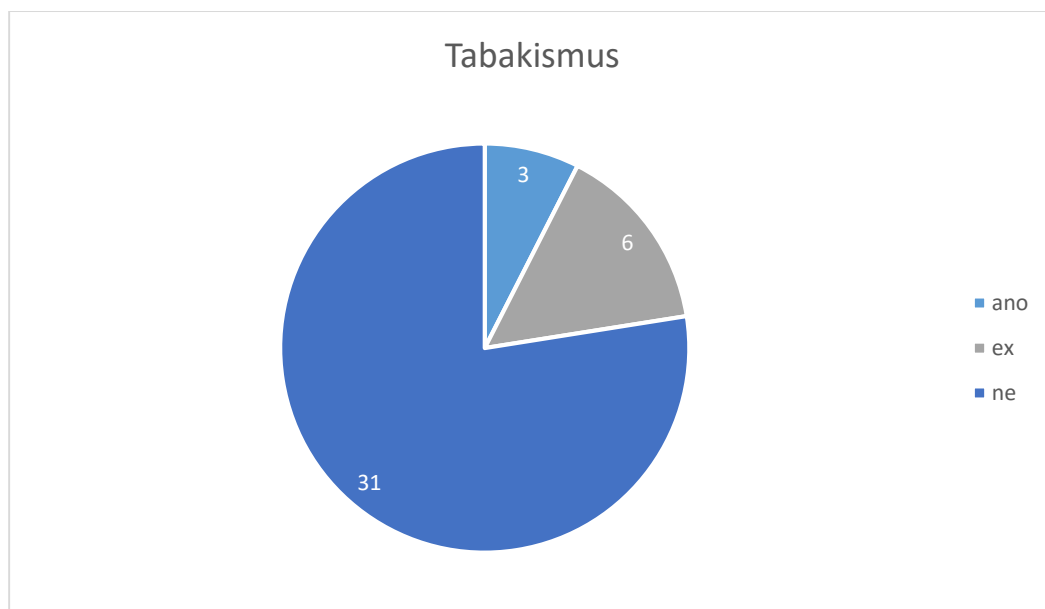
Celkový počet získaných respondentů z roku 2015 byl 40 pacientů. Ambulantně léčených bylo celkem 10 pacientů a hospitalizaci podstoupilo 31 pacientů. Jeden respondent ve stejném roce podstoupil ambulantní péči a následně byla nutnost k hospitalizaci. Vykázání bodů pojišťovně proběhlo dvakrát, proto součet respondentů typu léčby ambulantní a hospitalizace nesouhlasí. Nejstarší respondent je ve věku 91 let a nejmladší respondent ve věku 29 let. Průměrný věk všech 40 ti pacientů je 67 let.

Data, která byla sbírána o jednotlivých pacientech, začínaly u osobní anamnézy, pokračovaly přes rodinnou anamnézu, nejdůležitější část se týkala aktuálního stavu pacienta a v neposlední řadě se jednalo o data z fyzikálního vyšetření a laboratorních testů.



Obrázek 9: Graf - anamnéza a přidružené příznaky

Při vstupu pacienta do ambulance jsou pro lékaře velmi důležité celkový stav pacienta a osobní i rodinná anamnéza. Na obrázku 9 můžeme vidět, že u většiny respondentů skutečně převládaly příznaky spojené především s únavou, kašlem, produkcí sputa a dušností. Přidružené choroby, jako je cévní mozková příhoda, ischemická choroba srdeční a diabetes mellitus druhého typu se vyskytovala především u starších respondentů.

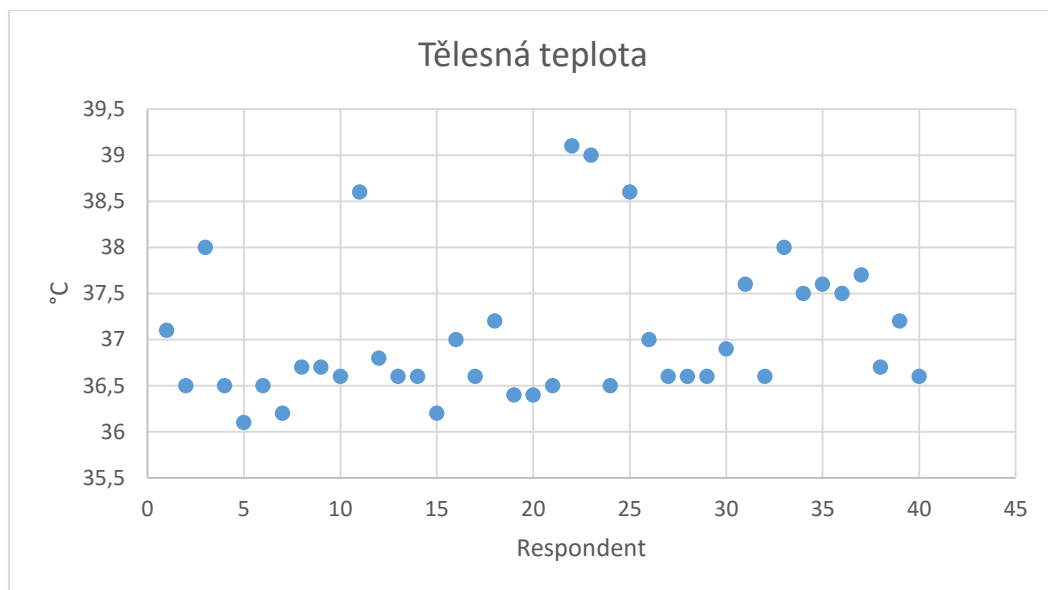


Obrázek 10: Graf – Tabakismus

Závislost na kouření byla v minulosti klasifikována jako onemocnění Světovou lékařskou organizací (WHO). Tabakismus ovlivňuje celý organismus a má vliv na choroby plic, srdce a cév. U drtivé většiny respondentů, tj. 31, ale kouření nebylo potvrzeno aktuálně, ani v minulosti. Na tomto grafu je vidět, že ne všechna plicní onemocnění vždy úzce souvisí s kouřením, jak si velká část veřejnosti myslí. Tím nemá být problematika závislosti na cigaretách zlehčována. Stále existují statistiky, že každý rok vlivem kouření zemře přes čtyři milióny lidí a čísla se rok od roku zvyšují.

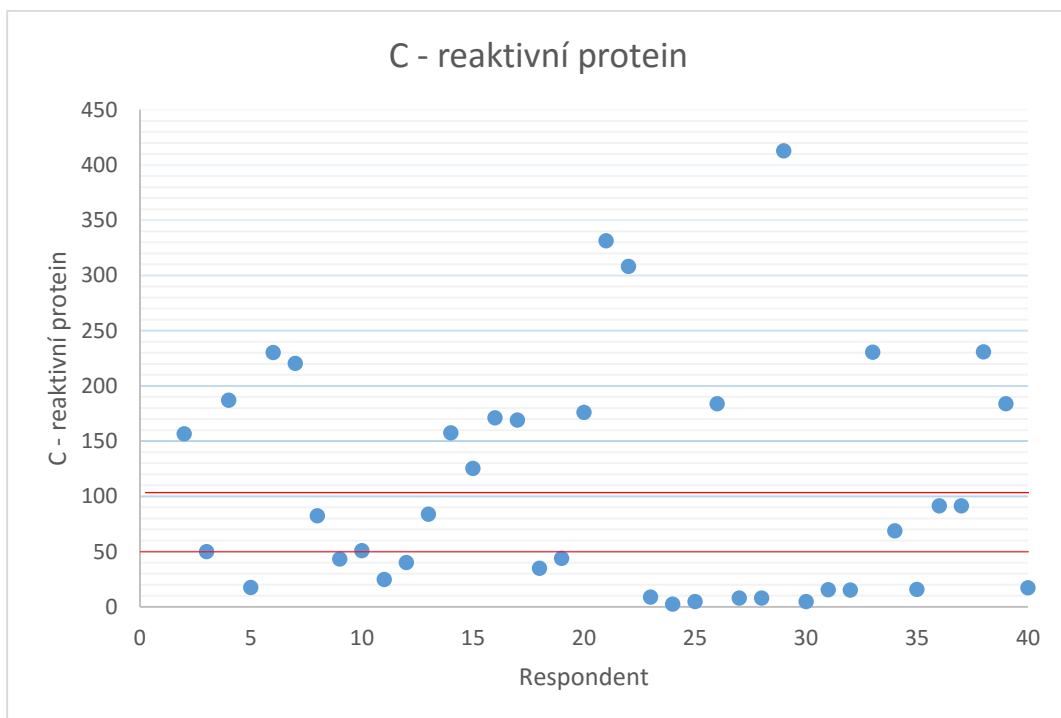
Dalším důvodem, proč není v našem vzorku větší část kuřáků, může být to, že většina takových pacientů trpí chronickou obstrukční plicní nemocí (CHOPN). Takové onemocnění je reverzibilní a je důsledkem závislosti na cigaretách nebo pracovním prostředím. Taková diagnóza je však v systému zdravotnického zařízení definována jinou diagnózou než pneumonií a to J44.1 [22,54].





Obrázek 11: Graf – Tělesná teplota

Na základě teorie víme, že riziková skupina pacientů s komunitně získanou pneumonií podle Amerického hodnocení PSI s ohledem na pacientovu tělesnou teplotu, počítá s hodnotou pod 35°C a nad 40°C. Tyto hodnoty tělesné teploty nám rezonují systémovou infekci. Více riziková je velmi nízká teplota, která značí, že organismus se nedokáže více proti infekci bránit. Ani jeden respondent neměl rizikové hodnoty tělesné teploty. U všech pacientů se tělesná teplota pohybuje v rozmezí 36,1°C a 39,1°C.



Obrázek 12: Graf – C-reaktivní protein

Hodnoty C-reaktivního proteinu je pro lékaře ukazatelem imunitní odpovědi organismu. Patří mezi proteiny akutní fáze. Je to zánětlivá odpověď organismu na rozvíjející se poškození. Tvořen je játry a následně je vylučován do krve. Jednoduše řečeno nám hodnoty CRP ukazují přítomnost zánětu v organismu. V krvi známe hodnoty již po 6ti hodinách od zahájení zánětlivého procesu. Jeho jednotkou je mg/l a fyziologické hodnoty jsou v rozmezí 0,0 – 7,0mg/l. Dosahovat mohou až tisícinásobku [55].

- Koncentrace 10 – 50 mg/l – mírný zánět
- Koncentrace do 100 mg/l – vyšší stupeň zánětu
- Koncentrace nad 100 mg/l – těžký zánět

Na obrázku 12 můžeme vidět vyznačené hodnoty CRP u každého respondenta. Hranice od 10 mg/l do 100mg/l je vyznačena červenými přímkami. Většina respondentů léčená ambulantně je pod hranicí červené přímky. Na obrázku 12 je také vidět rozmanitost a velké skoky mezi hodnotami. Měření C-reaktivního proteinu je pro lékaře snadná nápověda, která probíhá ihned v ambulanci.

*Tabulka 11: Získaná data – Vykázané body, počet dnů hospitalizace a ambulantní péče*

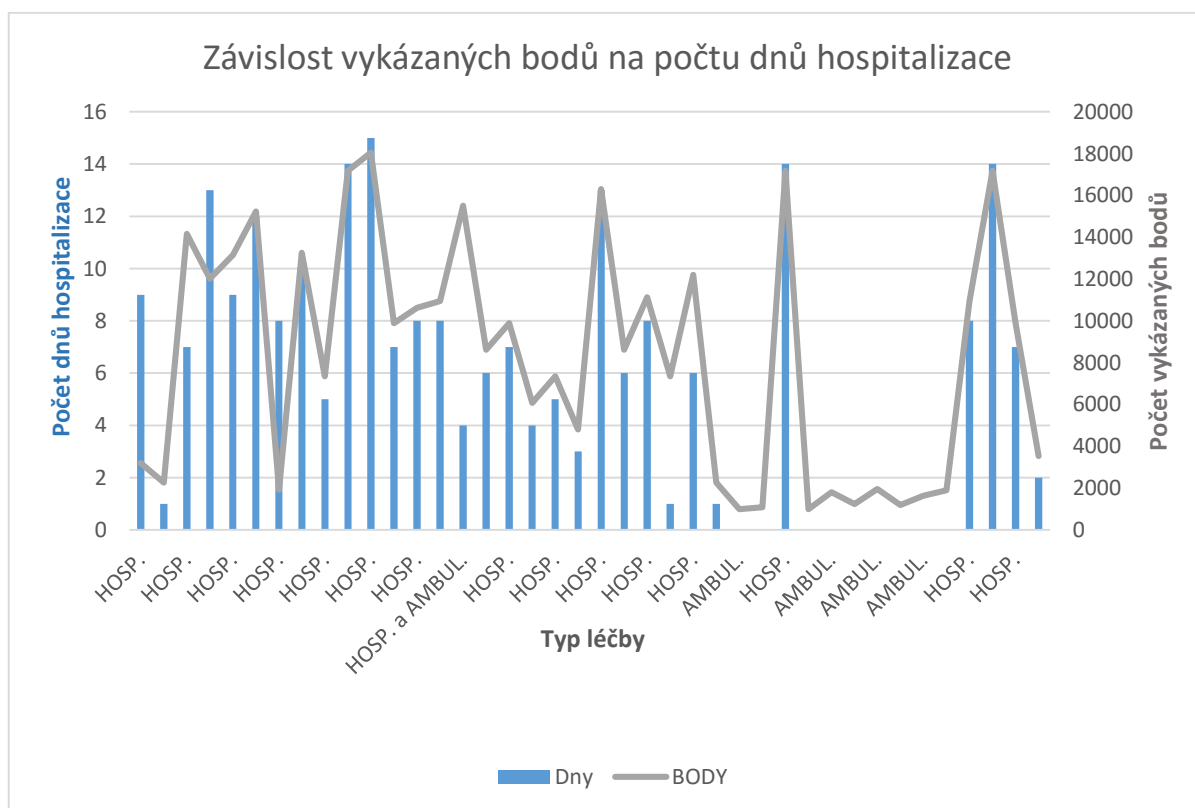
Typ léčby	AMBULATNÍ PÉČE	HOSPITALIZACE
Počet pacientů	10	31
Vykázané body	18115	312648
Průměr	1811,5	10085,42
Počet dnů hospitalizace	-	204
Průměr - hospitalizace	-	7,58
Max - body	5346	18047
Min - body	987	1924
Max - počet dnů hospitalizace	-	15
Min - počet dnů hospitalizace	-	1
Vykázané body v Kč (1,07Kč/bod)	19383,05	334533,36
Vykázané body v Kč (1,07Kč/bod) - průměr	<b>1938,305</b>	<b>10791,3994</b>
Vykázané body v Kč (1,07Kč/bod) - Min	1056,09	2058,68
Vykázané body v Kč (1,07Kč/bod) - Max	5720,22	19310,29
Medián	1432,5	10166
Medián v Kč	1590,075	11284,86

Směrodatná odchylka	1231,324	4675,68
Směrodatná odchylka v Kč	1366,769	5190

Stejně jako u tabulky 10 nám nesouhlasí součet pacientů typu léčby a to z důvodu, že jeden respondent byl zprvu léčen ambulantně a poté byla nutná jeho hospitalizace. Bylo potřebou takového respondenta zahrnout do tabulky dvakrát v rámci přehlednosti vykázaných bodů.

Tabulka 11 uvádí pozorovaná data. Na takové tabulce si může povšimnout nejdelší a nejkratší doby hospitalizace a zároveň poukazuje v rozdílu vykázaných bodů. Dále bylo v tabulce prokázáno, že nejnižší doba hospitalizace byla jeden den. V tomto případě je potřeba podotknout, že roli při léčbě hraje i pacientův souhlas. Stále mají pacienti možnost podepsání negativního reversu o propuštění z nemocnice. Případ odmítnutí hospitalizace byl v našem souboru dat pouze jeden. Zajímavostí je také nejvyšší cena za ambulantní péči, která byla 5.720,-. U tohoto respondenta byla v osobní anamnéze přítomna alergie na běžně používaná antibiotika. Pravděpodobně byla z toho důvodu nasazena dražší léčba.

Hodnota bodu je získána z vyhlášky o stanovení hodnot bodu, výše úhrad hrazených služeb a regulačního omezení pro rok 2015. Jedná se o vyhlášku č. 324/2014 stanovena Ministerstvem zdravotnictví podle §17 odst. 5 zákona č. 48/1997 Sb., o veřejném zdravotním pojištění a o změně a doplnění některých souvisejících zákonů, ve znění zákona č. 117/2006 Sb., zákona č. 245/2006 Sb., zákona č. 261/2007 Sb., zákona č. 298/2011 Sb. a zákona č. 369/2011 Sb. [56].



Obrázek 13: Graf - Závislost vykázaných bodů na počtu dnů hospitalizace

Obrázek 13 nám ukazuje počet vykázaných bodů dle počtu dnů hospitalizace. Počty vykázaných bodů nemohou být konzistentní z toho důvodu, že u každého respondenta se body vykazovaly za jiné úkony. Například respondent, který byl nejdříve léčen ambulantně a následně musel být hospitalizován má více vykázaných bodů.

U prvního respondenta můžeme pozorovat delší dobu hospitalizace a nižší počet vykázaných bodů oproti jiným respondentům. V tomto případě se jednalo o jednadevadesátiletou respondentku, která byla ležící, nekomunikující a zároveň byla umístěna v ústavu sociální péče. Zároveň nebyla orientována časem a místem, a také byla v minulosti již s diagnózou pneumonie hospitalizována, proto byla hospitalizace delší v důsledku pozorování. Velmi podobný případ byl i u respondentky č. 4, která měla demenci v progresi a více přidružených chorob. Zároveň byla také umístěna do ústavu sociální péče. U respondentky č. 25 můžeme pozorovat vyšší vykázaní bodů oproti délce trvání její hospitalizace. U této respondentky byl proveden test na albumin, který má vyšší vykazující hodnotu.

Na grafu také vidíme velmi podobné vykázaní bodů u respondentů, kteří byli léčeni ambulantně. Drobné odchylky se mohou lišit typem předepsaných antibiotik. Do grafu byl zahrnut celkový počet bodů, jak za ambulantní léčbu, tak za léčbu hospitalizací.

U všech respondentů byla bodově určena jejich riziková skupina na základě systému Americké hrudní společnosti PSI, která díky anamnestickým údajům, fyzikálnímu vyšetření, přidruženým chorobám a výsledkům laboratorních testů dokáže zhruba predikovat úmrtnost konkrétního pacienta. Výpočty u jednotlivých respondentů jsou umístěny v příloze.

Tabulka 12: Bodový systém PSI – Rizikové skupiny a typ léčby u respondentů

Skupina	Počet bodů	Mortalita	Doporučení typu léčby	Data - počet pacientů	Data - počet pacientů	
					A ambulantně	H hospitalizace
I	<50muži/55ženy	0,10%	Ambulantně	13	7	6
II	<70	0,60%	Ambulantně	4	1	3
III	71-90	2,80%	Ambulantně, event. krátká hospitalizace	6	1	5
IV	91-130	8,20%	Hospitalizace	11	0	11
V	>130	29,20%	Hospitalizace – JIP	6	0	6

Při přijetí pacienta a rozhodování o typu léčby je predikce velmi obtížná. Stále se jedná o zdraví pacienta a na možná rizika nikdy nesmíme zapomínat. Je důležité zmínit, že mnoho přidružených příznaků bodový systém Americké hrudní společnosti nepočítá a mnohdy jsou pro lékaře rozhodující. Mezi takové příznaky patří ztráta chuti k jídlu, bolesti na hrudi, dušnost, únava, zimnice, produkce sputa, průjem nebo zvracení, hemoptýza, pocení a cyanóza. V rozhodování lékařů hraje také roli prodělaný zápal plic v minulosti. Z laboratorních testů se lékaři také soustředí na hodnotu leukocytů v krvi a také na hodnotu C-reaktivního proteinu. Stále se jedná o hodnoty a příznaky, se kterými bodový systém nepočítá.

U respondenta z rizikové skupiny 3, u kterého byla zvolena léčba ambulantně, je důležité zmínit, že hospitalizaci odmítl, i když mu byla doporučována.

## 3.2 Klasifikační a regresní strom (CART)

Pro sestavení klasifikačních a regresních stromů byl použit matematický software specializovaný na statistiku – R verze 2.3, který je snadno k dispozici. Existují i jiné softwary, ale bohužel jsou licencované. Všechna získána data u 40ti respondentů byla vložena do přehledné tabulky Excel pro snadné nahrání do programu R. Proměnné byla zapsána ve zkratkách a jejich význam je v příloze č. 5.

Data nahrána pomocí příkazu:

```
> mydata <- read.csv("h:/ERKO.csv", header=TRUE, sep=";")
```

- *mydat* – značí uložená data
- *header=TRUE* – znamená, že hlavička je první řádek
- *sep=";"* – znamená, že data jsou oddělena středníkem

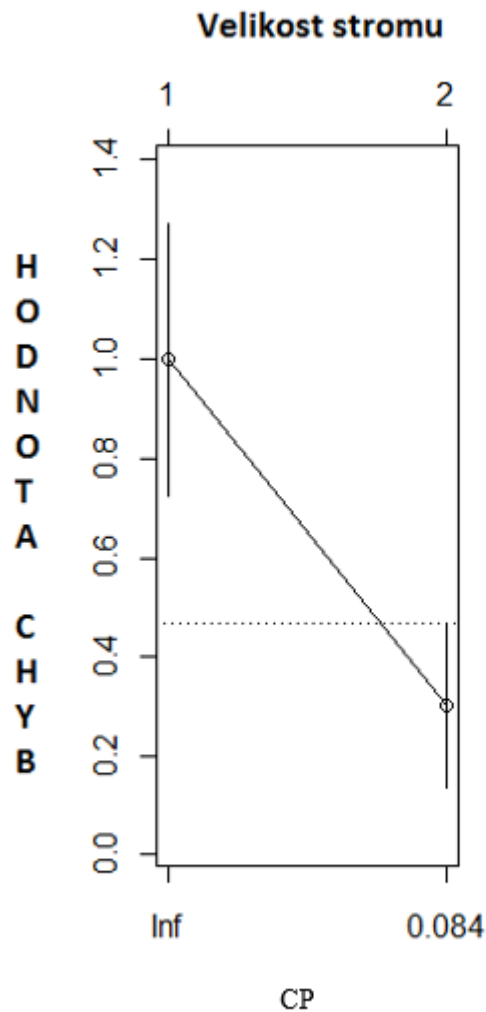
Na základě známé teorie víme, že proměnná kvalitativní slouží k sestavení klasifikačního stromu a kvantitativní proměnné slouží k sestavení regresního stromu.

### 3.2.1 Výsledky – Klasifikační strom

Klasifikační strom dokáže počítat také s chybějícími daty, které máme především u ambulantně léčených pacientů především u laboratorních testů, které nejsou z pravidla prováděny u pacientů s tímto typem léčby. Klasifikační strom také slouží k orientaci v datech. V souboru je mnoho vysvětlujících proměnných, které bylo třeba roztrždit do jednotlivých podmnožin. Celkový počet proměnných byl 34. První proměnná – TYP, značí typ léčby (ambulantní vs. hospitalizace) a byla zvolena jako cílová proměnná pro datový soubor.

Příkaz v R:

```
>fit<-  
rpart(TYP~VEK+ZEM+SU+CMF+ISCH+DM+NEF+HEP+RAK+RAKm+ZM+BB+BnH  
+BH+DU+UN+ZIM+KAS+SP+PR+ZV+VK+Nv+POC+AST+CYA+KUR+PLT+Na+Cl  
+TKS+TKD+TF+S,data=mydata,method="class")  
> plotcp(fit) ...příkaz, která vizualizuje křížovou validaci výsledků
```



Obrázek 14: Křížová validace – klasifikační strom

Křížové zhodnocení hovoří o rozhodnutí komplexnosti stromu. Datový soubor byl rozdělen na několik vzorků, ze kterých byly náhodně vybrány proměnné a z nich byl vytvořen klasifikační strom. Křivka je strmě klesající, což značí vyváženost predikce. V ideálních případech začne růst a jedno z pravidel hovoří o tom, že je potřeba vybrat takové dělení, které se jako první vyskytuje pod horizontální čarou. V grafu jsou zaneseny hodnoty chyb předpovědi typu léčby. Velikostí stromu se rozumí počet koncových větví → horní vodorovná osa. Přerušovaná čára uprostřed je referenční a prezentuje nejmenší hodnotu krosvalidované chyby. Hodnota  $cp$  = complexity parameter je parametr složitosti. V ideálních

případech začne růst a jedno z pravidel hovoří o tom, že je potřeba vybrat takové dělení, které se jako první vyskytuje pod horizontální čarou.

V našem případě byla velikost stromu 2, což znamená pouze jedno větvení. Parametr složitosti má hodnotu 0,084. Při sestavování stromů pro predikci jsou v ideálním případě vhodné složitější stromy minimálně o 5 ti větvení, aby byla předpověď odpovídající.

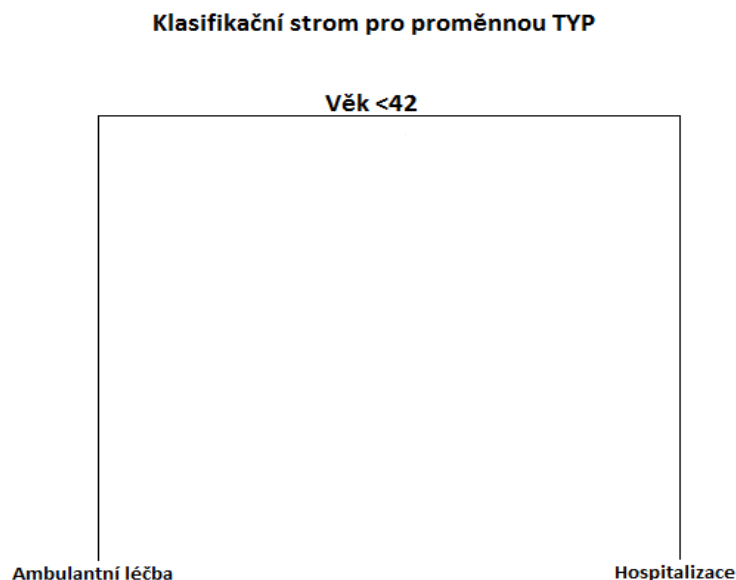
Při vytvoření stromu v program R byly zhodnoceny všechny potenciální proměnné a vybrány ty proměnné, které jsou variabilní nejvíce. V našem případě byla variabilita vyhodnocena pouze u proměnné věk.

```
>plot(pfit, uniform=TRUE,main="Pruned Classification Tree for TOP")
```

Teorie hovoří o tzv. ořezávání stromu, které namodelujeme pomocí výše uvedeného příkazu. Vzhledem k tomu, že variabilní proměnnou máme jen jednu, byl vytvořen strom pouze o kořenu a dvou listech, bez uzlů. Proto ořezávání stromu nemá význam a finální strom vypadá stejně. Diagram na obrázku 9 křížové validace nám také ukazuje, že pro další kleštění nebyl prostor. Jednoduše řečeno, výsledný strom je skutečné velikosti, aby ukázal správnou reprezentaci vloženého vzorku dat – proměnných. Nevariabilní proměnné nebyly dostatečně homogenní.

Následně byla vytvořena grafická a textová interpretace klasifikačního stromu pomocí následujících příkazů v programu R.

```
> plot(fit,uniform=TRUE,main="Classification Tree for TYP" ... grafická podoba stromu
```



Obrázek 15: Klasifikační strom

Vytvořená grafická podoba klasifikačního stromu nám ukazuje závislost proměnné VEK v predikci pro vysvětlovanou proměnnou. Mezní hodnota variabilní proměnné byla určena hodnotě 42 let. Klasifikační strom vždy čteme ze shora dolů a na uzlech si pokládáme klíčovou otázku. V tomto případě → „Je respondent mladší 42let?“. Pokud je odpověď na otázku ano, pokračujeme po levé větvi.

> *print(fit)* ...vytvoření textové podoby stromu

*n= 40*

*node), split, n, loss, yval, (yprob)*

*\* denotes terminal node*

1) *root 40 10 HOSP. (0.22500000 0.75000000 0.02500000)*

2) *VEK < 42 7 0 AMBUL. (1.00000000 0.00000000 0.00000000) \**

3) *VEK >= 42 33 3 HOSP. (0.06060606 0.90909091 0.03030303) \**

Závislost nebyla pravděpodobně nalezena z důvodu malého počtu respondentů, nesouvislosti proměnných a určení typu léčby, které probíhá především individuálně. Lékaři se vždy rozhodují na základě své odbornosti a zkušeností z minulosti.

### 3.2.2 Výsledky – Regresní stromy

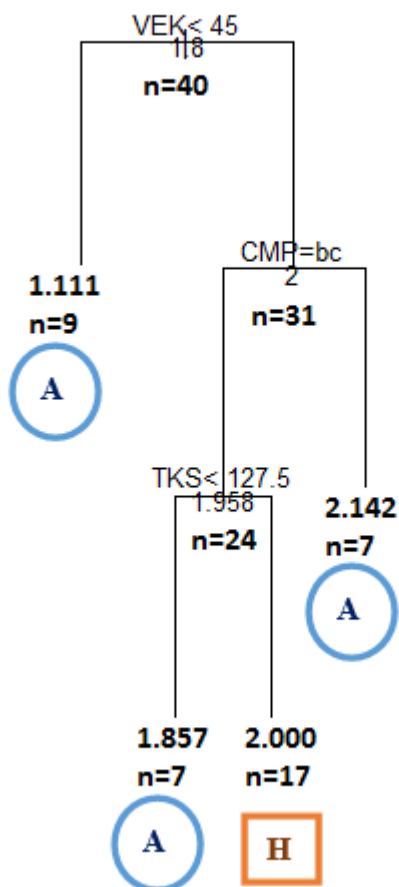
Stejně jako u Klasifikačního stromu i Regresní strom počítá s chybějícími daty, které máme především u ambulantně léčených pacientů. Příkaz pro zadání do softwaru R je téměř identický, jen změníme metodu „class“ na metodu „anova“. První proměnná – TYP nám opět značí typ léčby (ambulantní vs. hospitalizace). Typ léčby je pro nás cílová proměnná.

Příkaz v R:

```
> fit<-
rpart(TYP~VEK+ZEM+SU+CMP+ISCH+DM+NEF+HEP+RAK+RAKm+ZM+BB+BnH
+BH+DU+UN+ZIM+KAS+SP+PR+ZV+VK+NV+POC+AST+CYA+KUR+PLT+Na+Cl
+TKS+TKD+TF+S,data=mydata,method="anova")
> printcp(fit) ... ukáže výsledky
> plot(fit,uniform=TRUE,main="Regression Tree for TYP ") ... příkaz pro sestavení stromu
```



## Regresní strom pro proměnnou TYP



Obrázek 16: Regresní strom 1

Důležitá orientace po regresním stromu je taková, že směr je ze shora dolů. Na uzlech při dotazování je sled odpovědí závislý na předchozích otázkách. V případě, že odpověď je „ano“ pokračujeme po levé větvi. V našem případě tato situace nastává v prvním uzlu u proměnné CMP, pokud je odpověď ano, dostáváme se k dalšímu uzlu. U proměnné TKS postupujeme identicky. Pokud je systolický tlak menší než 127,5, pokračujeme opět po větvi levé.

Proměnné jsou rozdělovány v rozsahu svých hodnot tak, aby byly navzájem co nejvíce odlišné, ale vnitřně co nejpodobnější, vždy ve vztahu vysvětlované proměnné. V našem případě proměnná TYP. Výstupem regresních stromů je vždy číselná hodnota vyjadřující průměrnou hodnotu vysvětlované proměnné. Pro vytvoření větvení byla vybrána proměnná, která nejvíce maximalizuje dané kritérium (vysvětlovaná proměnná TYP). Větvení bylo ukončeno, pokud se hodnota vysvětlované proměnné v uvažovaném uzlu liší jen málo. V případě, že je směrodatná odchylka menší než 5%, větvení nepokračuje.

> *print(fit)* ... textová podoba regresního stromu

*n* = 40

*node*), *split*, *n*, *deviance*, *yval*

\* denotes terminal node

1) root 40 8.4000000 1.800000

2) *VEK* < 45 9 0.8888889 1.111111 \*

3) *VEK* >= 45 31 2.0000000 2.000000

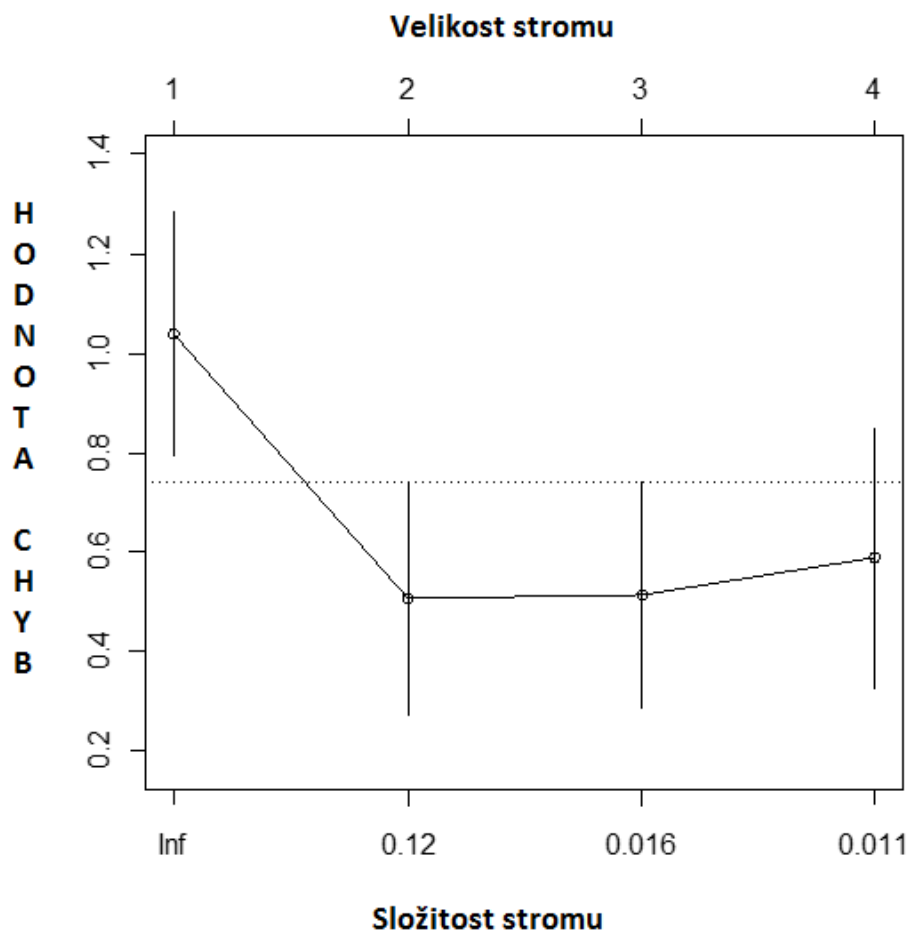
6) *CMP* = *ne,ne* 24 0.9583333 1.958333

12) *TKS* < 127.5 7 0.8571429 1.857143 \*

13) *TKS* >= 127.5 17 0.0000000 2.000000 \*

7) *CMP* = *ano* 7 0.8571429 2.142857 \*

> *plotcp(fit)* ... vizualizace křížové validace



Obrázek 17: Křížová validace – Regresní strom 1

Na diagramu můžeme vidět opět zanesení chyb v předpovídání typu léčby. Přerušovaná čára vyobrazuje hodnotu nejmenší křížové validace. Křivka strmě klesá až do velikosti stromu 2 a od velikosti stromu tři mírně stoupá. To nám vypovídá, že odhadnutá chyba roste

> `printcp(fit)` ... příkaz, ukáže výsledky souvisejícími s diagramem na obrázku 12

*Regression tree:*

```
rpart(formula = TYP ~ VEK + ZEM + SU + CMP + ISCH + DM + NEF +  
HEP + RAK + RAKm + ZM + BB + BnH + BH + DU + UN + ZIM + KAS +  
SP + PR + ZV + VK + NV + POC + AST + CYA + KUR + PLT + Na +  
Cl + TKS + TKD + TF + S, data = mydata, method = "anova")
```

*Variables actually used in tree construction:*

```
[1] CMP TKS VEK
```

*Root node error: 8.4/40 = 0.21*

*n= 40*

	<i>CP</i>	<i>nsplit</i>	<i>rel error</i>	<i>xerror</i>	<i>xstd</i>
1	0.656085	0	1.00000	1.05306	0.25465
2	0.021967	1	0.34392	0.39159	0.20754
3	0.012046	2	0.32195	0.41322	0.20659
4	0.010000	3	0.30990	0.44119	0.21855

Při zadání příkazu pro data související s předchozím diagramem si musíme dát pozor u následujících hodnot. Hodnoty *cp* jsou hodnoty na vodorovné ose (složitost stromu) a jsou odlišné, ale při prořezávání stromu musí být použity hodnoty z diagramu. Data „*nsplit*“ nám uvádějí počet koncových větví. Data „*rel.error*“ nám ukazují relativní pokles chyb v porovnání s nulovým modelem. Data „*xerror*“ nám ukazují hodnotu chyb. Regresní strom v našem případě našel větší variabilitu proměnných než v případě sestavení klasifikačního stromu. Již byly vytvořeny dva uzly, které mají větší vypoovídající hodnotu o predikci.

Při zadání datové souboru nám program R vybral ty nejlepší možné prediktory, tím jsou vybrány jeho podskupiny, které jsou schopny redukovat výchozí proměnnou (v našem případě proměnnou TYP – typ léčby). V případě, že chceme pozorovat i jiné závislosti, můžeme vidět výsledky díky následujícímu příkazu.

> summary (fit)

Primary splits:

VEK < 45 to the left, improve=0.6560847, (0 missing)

S splits as RRRRRLLRRRLLLLR, improve=0.4171627, (0 missing)

ISCH splits as RL, improve=0.1638322, (0 missing)

CMP splits as RLR, improve=0.1257440, (0 missing)

DM splits as RL, improve=0.1257440, (0 missing)

Surrogate splits:

S splits as RRRRRRRRRRRRLLLLR, agree=0.875, adj=0.444, (0 split)

DU splits as RLR, agree=0.800, adj=0.111, (0 split)

Primary splits:

CMP splits as RLL, improve=0.09226190, (0 missing)

DM splits as RL, improve=0.09226190, (0 missing)

TF < 95 to the left, improve=0.09226190, (0 missing)

Cl < 106.5 to the right, improve=0.08630952, (2 missing)

TKS < 127.5 to the left, improve=0.08423913, (0 missing)

Surrogate splits:

S splits as RLLRLLLLLLRLLRL, agree=0.935, adj=0.714, (0 split)

Na < 141.5 to the left, agree=0.871, adj=0.429, (0 split)

ISCH splits as RL, agree=0.839, adj=0.286, (0 split)

RAK splits as RL, agree=0.806, adj=0.143, (0 split)

PR splits as RL, agree=0.806, adj=0.143, (0 split)

Primary splits:

TKS < 127.5 to the left, improve=0.10559010, (0 missing)

VEK < 86.5 to the right, improve=0.08695652, (0 missing)

TKD < 62 to the left, improve=0.08695652, (0 missing)

S splits as -LL-LRRRRR-RL-R, improve=0.08695652, (0 missing)

Cl < 105.5 to the right, improve=0.06851120, (2 missing)

Surrogate splits:

KUR splits as LRR, agree=0.792, adj=0.286, (0 split)

TKD < 62 to the left, agree=0.792, adj=0.286, (0 split)

TF < 69 to the left, agree=0.792, adj=0.286, (0 split)

```
S splits as -LL-RRRRRR-RR-R, agree=0.792, adj=0.286, (0 split)
```

```
NEF splits as LR, agree=0.750, adj=0.143, (0 split)
```

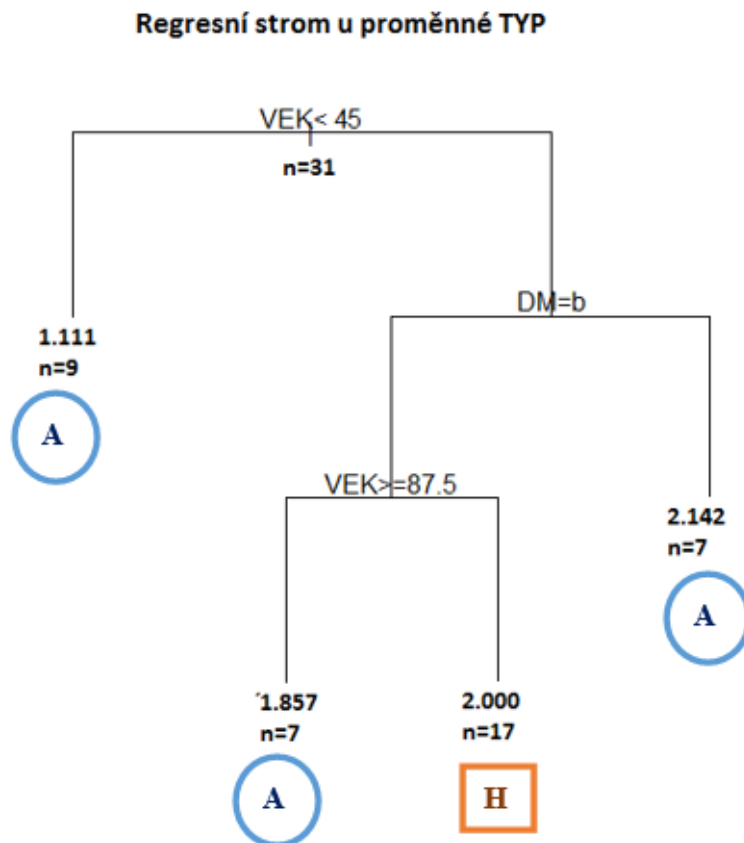
Výstup z výše zmiňovaného příkazu je velmi rozsáhlý, proto byla vybrána ta nejdůležitější kvůli zdůraznění daných propočtů. Můžeme tak pozorovat dělení, kdy jako první pravidlo pro rozdělení byla zvolena proměnná VEK (věk). Hodnota „agree“ nám ukazuje míru shody s předpovědí primárního pravidla. Hodnota „improve“ udává relativní kvalitu pravidel. Dále si můžeme povšimnout dělení i u jiných proměnných. Jak již bylo zmíněno, vidíme zde i propočty u proměnných s chybějícími daty, které vidíme především u proměnných Cl (chloridy).

Pro jednodušší grafické znázornění regresního stromu byla odstraněna jedna proměnná a tím bylo poukázáno i na jinou závislost na ostatních proměnných. Příkaz i postup je identický, jen chudší o proměnnou CMP.

```
>fit<-
```

```
rpart(TYP~VEK+ZEM+SU+ISCH+DM+NEF+HEP+RAK+RAKm+ZM+BB+BnH+BH+DU+UN+ZIM+KAS+SP+PR+ZV+VK+NV+POC+AST+CYA+KUR+PLT+Na+Cl+TKS+TKD+TF+S,data=mydata,method="anova")
```

```
> plot(fit,uniform=TRUE,main="Regression Tree for TYP ")
```



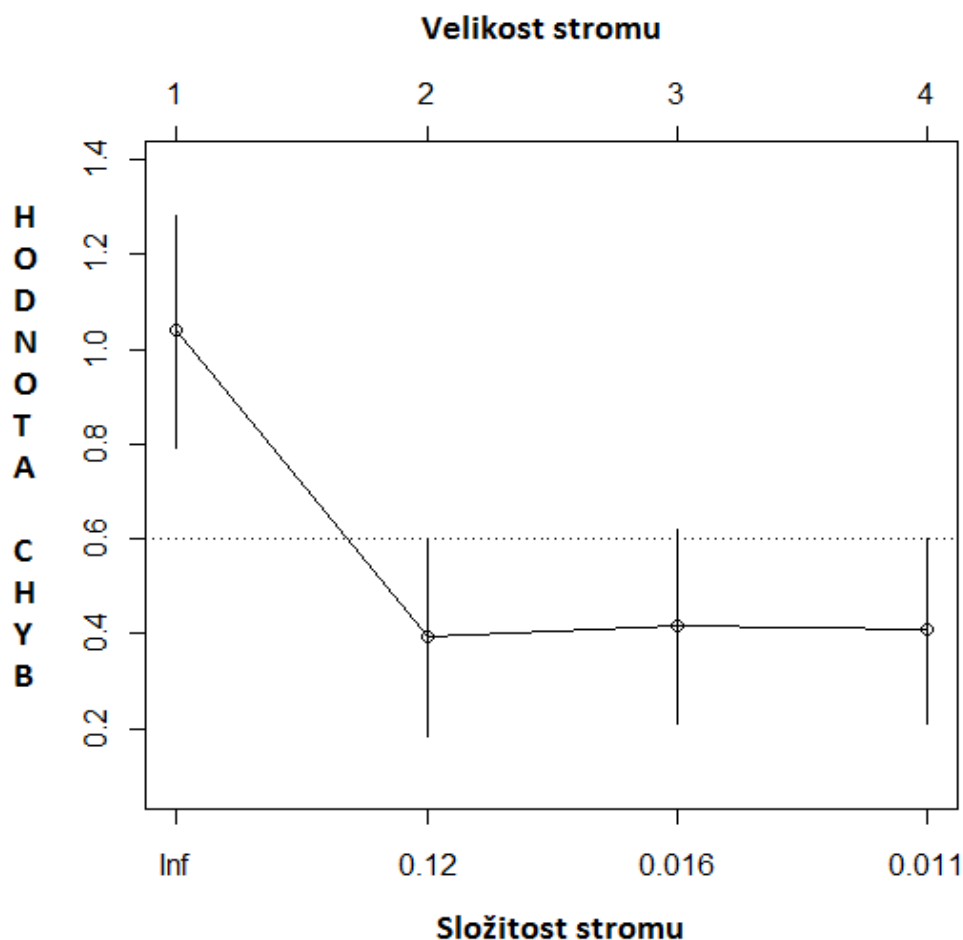
Obrázek 18: Regresní strom 2

Stejně jako u prvního modelování stromu, i v tomto případě byla nalezeny pouze dva uzly. Postup čtení je stejný jako výše zmiňovaný s tou výjimkou, že u druhého uzlu byl generován jiný směr. Pokud si na otázku přítomnosti proměnné DM (Diabetes mellitus II.typu) odpovíme „ne“, postupujeme doleva. Vizualizace „DM=b“ znamená otázku „není přítomna proměnná DM?“ → „ano není“, pokračuji po levé straně stromu. Dále můžeme vidět, že některé proměnné mohou mít závislost dvakrát. U druhého uzlu vidíme návrat k prozkoumání závislosti na proměnné VEK. Pokud je její hodnota větší nebo rovna 87,5, pokračuji po levé větvi.

Textovou podobu regresního stromu můžeme opět vidět pomocí následujícího příkazu.

```
> print(fit)
n= 40
node), split, n, deviance, yval
  * denotes terminal node
1) root 40 8.4000000 1.800000
 2) VEK< 45 9 0.8888889 1.111111 *
 3) VEK>=45 31 2.0000000 2.000000
 6) DM=ne 24 0.9583333 1.958333
 12) VEK>=87.5 7 0.8571429 1.857143 *
 13) VEK< 87.5 17 0.0000000 2.000000 *
 7) DM=ano 7 0.8571429 2.142857 *
```

```
> plotcp(fit) ... vizualizace křížové validace
```



Obrázek 19: Křížová validace – Regresní strom 2

Pro vyobrazení křížové validační chyby byl znovu vytvořen diagram, kde jsou zaneseny konkrétní chyby a počet vytvořených stromů. Křivka nám na počátku opět strmě klesá, pak je mírně vzrůstající, ale do velikosti třetího stromu je mírně klesající. Tím se snižuje její vypovídající hodnota.

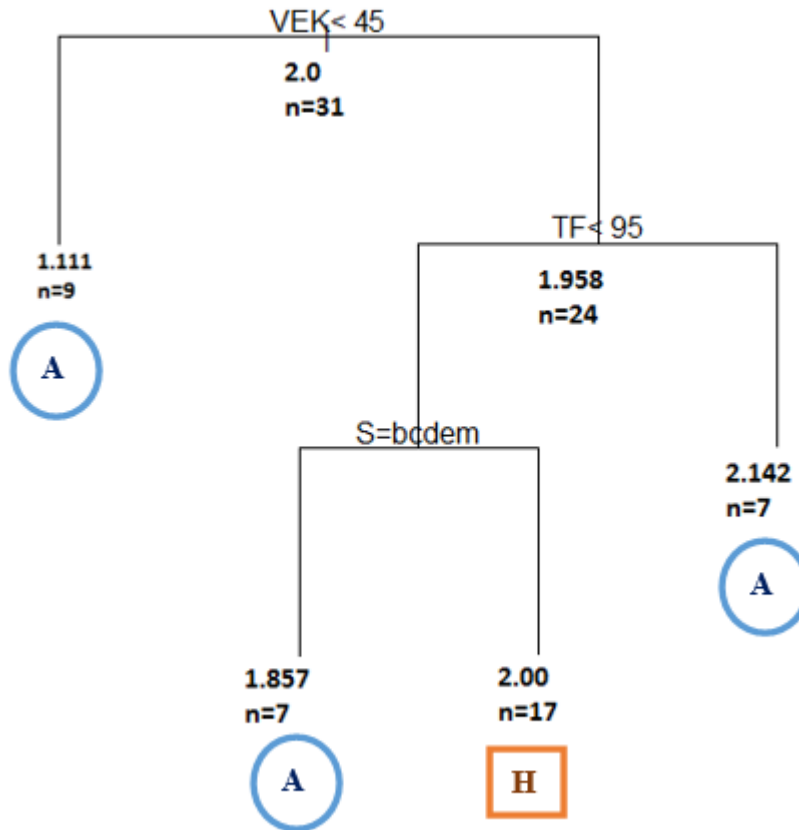
Vzhledem k tomu, že byla nalezena dělitelná proměnná jedna, a to DM, a následně se závislost vrátila k proměnné VEK, volím další znázornění s jejím vynecháním.

```
>fit<-
```

```
rpart(TYP~VEK+ZEM+SU+ISCH+NEF+HEP+RAK+RAKm+ZM+BB+BnH+BH+DU+
UN+ZIM+KAS+SP+PR+ZV+VK+NV+POC+AST+CYA+KUR+PLT+Na+Cl+TKS+TK
D+TF+S,data=mydata,method="anova")
```

```
> plot(fit,uniform=TRUE,main="Regression Tree for TYP ")
```

### Regresní strom u proměnné TYP



Obrázek 20: Regresní strom 3

I u třetího regresního stromu byla nalezena závislost u 2 proměnných. Čtení stromu v prvním uzlu se nijak neliší. Pokud je proměnná TF menší než 95, pokračujeme vlevo a dostáváme se k dalšímu uzlu s proměnnou S. U této proměnné byla závislost vypočítána jednotlivě, tzn.: že jednotlivé větve jsou bodově určeny. Jednotlivé body nám ukáže textová podoba regresního stromu.

```
> print(fit)
```

```
n = 40
```

```
node), split, n, deviance, yval
```

```
* denotes terminal node
```

```
1) root 40 8.4000000 1.800000
```

```
2) VEK < 45 9 0.8888889 1.111111 *
```

```
3) VEK >= 45 31 2.0000000 2.000000
```

```
6) TF < 95 24 0.9583333 1.958333
```



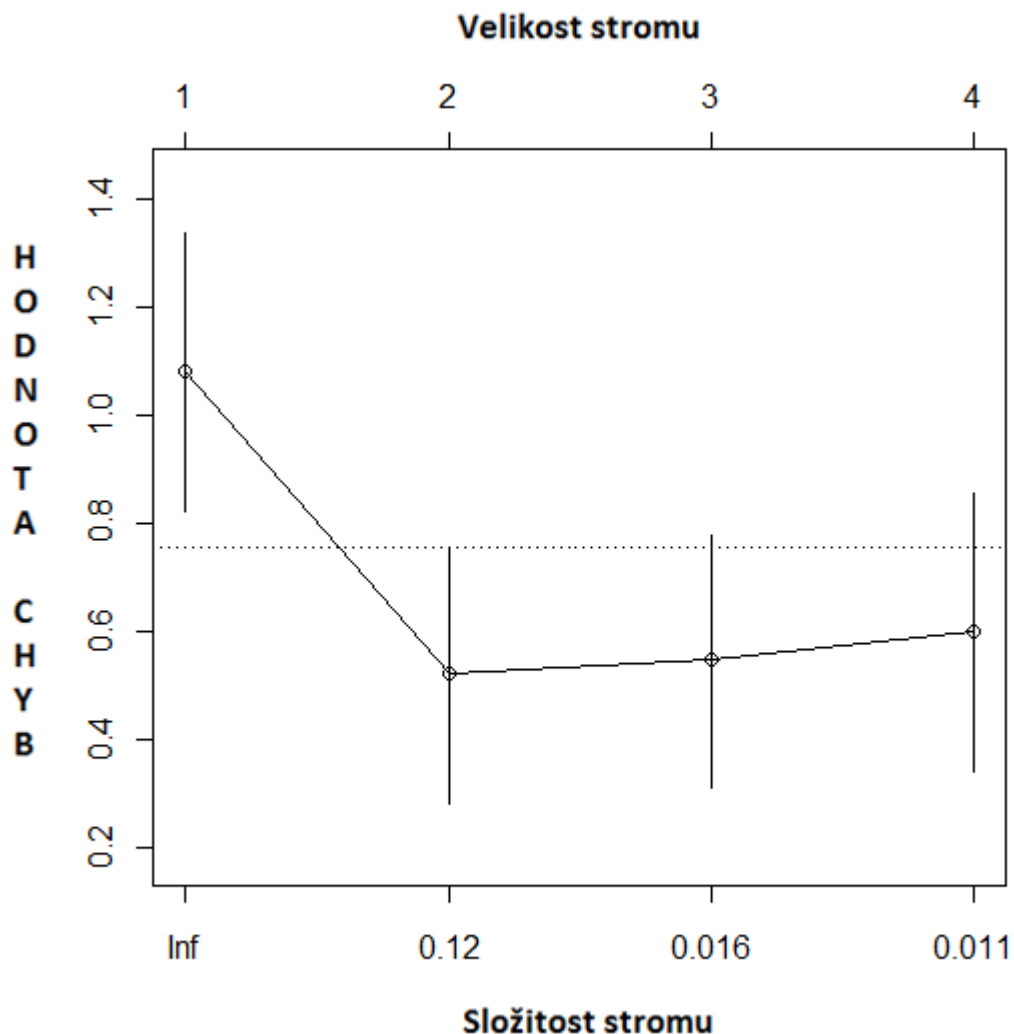
12)  $S=0,81,0,82,0,85,0,88,0,97$  7 0.8571429 1.857143 \*

13)  $S=0,89,0,92,0,93,0,94,0,95,0,96,0,98,0,99$  17 0.0000000 2.000000 \*

7)  $TF \geq 95$  7 0.8571429 2.142857 \*

Můžeme si povšimnout, že závislost nebyla určena rozmezí nízkých nebo vysokých hodnot. Levá větev bodů je vypsána v textové podobě stromu u bodu 13. Naopak pravá větev je vypsána u 12. Takové rozdělení nám ukazuje, že u proměnné S je závislost individuální u konkrétních pacientů. Nízké i vysoké hodnoty jsou v obou větvích. Nicméně vypovídají hodnotu křížové validace, vyobrazíme na diagramu pomocí následujícího příkazu.

```
> plotcp(fit)
```



Obrázek 21: Křížová validace – Regresní strom 3

Stejně jako u výše zmiňovaných příkladů, nám v první fázi křivka prudce klesá. Oproti regresnímu stromu 2 si můžeme všimnout, že po klesání nám křivka začne růst. Z toho vyplývá, že regresní strom 3 má větší vypovídající hodnotu, než strom 2.

Byl sestaven klasifikační strom a tři stromy regresní. Program určil závislost proměnných v uzlech a výstupní listy. Variabilní proměnná byla ve všech případech VEK (věk). U regresního stromu byla variabilita určena u dalších proměnných jako je CMP, TKS, DM, TF a S. Jsou to data týkající se přidružených chorob a fyzikálního vyšetření. Ovšem nejsou úzce spjaty s onemocněním pneumonie. Například proměnná S (saturace) byla hodnocena bodově, nikoliv intervalem. V obou rozdělených skupinách nebyly nalezeny nízké i vysoké hodnoty.

Nicméně z lékařského hlediska nemůžeme počítat tyto výsledky za adekvátní. Naopak nám rezonují velmi individuální péči o pacienty s komunitní pneumonií. Každý případ má svá specifika a onemocnění pneumonie je velmi nevyzpytatelné.

### **3.3 Zhodnocení využití modelů pro predikci klinických a nákladových výstupů.**

V klinickém rozhodování je důležitá správná prognóza. Pacient-specifické modely mají potenciál k snadné predikci u konkrétních diagnóz. Při modelování bylo zjištěno, že jednotlivé postupy se liší ošetřujícími lékaři a zároveň může být ovlivněno zvyklostmi daného zdravotnického zařízení. U pacientů s komunitní pneumonií je podstupováno riziko rychlého postupu onemocnění, kdy mohou nastat vážné komplikace, mezi které řadíme především respirační selhání. V našem souboru dat bylo potvrzeno, že největší váhu má aktuální stav pacienta. Vzhledem k tomu, že u ambulantně léčených pacientů nebyly výsledky laboratorních testů, nemohla být při modelování nalezena větší závislost pro možné sestavení klasifikačního a regresního stromu. Dále byla zjištěna delší doba hospitalizace u starších pacientů, kteří nemají dobré sociální zázemí. Právě pomocí pacient-specifických modelů by bylo možné lépe predikovat typ léčby u těchto pacientů. Pacient léčený ambulantně je méně zatěžující pro zdravotnický systém.

Na základě zahraničních výzkumů byly určeny koncové body pro sestavení modelů. Autor Georgy F. Cooper se ve svých článcích zabývá touto problematikou a na základě těchto článků bylo potvrzeno, že není vždy potřebné pro sestavení počítat s koncovým bodem úmrtí, ale právě vážné komplikace zahrnující možné orgánové selhání nebo sepsi. V našem souboru respondentů bylo potvrzeno, že převládá část s nízkým rizikem úmrtí a vážných následků [2,13]. Hospitalizace je spojena s vyššími náklady, protože je spojena s podáváním intravenózních antibiotik. Při shromažďování dat bylo obtížné pacienty správně vybrat. U některých respondentů byla některá velmi důležitá data chybějící, z toho důvodu nebyla při nalezena závislost a výsledný strom byl pouze o dvou listech.

Použití klasifikačního stromu má výhodu snadného použití v praxi. Díky jeho grafické vizualizaci není obtížná orientace po jednotlivých větvích. Výhodou je využití na budoucích případech. Pro jeho vytvoření je zapotřebí větší soubor respondentů a zajištění úplnosti dat u konkrétních pacientů.

Ekonomickým přínosem predikčních modelů je snížení zátěže pro zdravotnický systém vzhledem k solidárnímu systému pojištění v ČR. Je to především snížení finančních nákladů, ale také zohlednění obloženosti lůžek, lišící se zdravotnickým zařízením. Správná prognóza redukuje zbytečné hospitalizace a zároveň špatně odhalenou ambulantní léčbu,

která může vést v budoucnu k hospitalizaci. Lékařské rozhodování je spojené s dvojitou nejistotou. Může být správně odhalená prognóza, ale nasazena špatná léčba, nebo dojde ke špatnému stanovení prognózy a tím zároveň k nasazení neefektivní léčby.

V práci nebyla využita metoda HTA pro zhodnocení pro určení ekonomických nákladů z důvodů studie. Jedná se o vytváření predikčních modelů pro lepší podporu rozhodovacího procesu a metoda lze využít pro potenciální ušetření finančních prostředků. Avšak nebylo možné zakomponovat výsledky do některé z HTA analýz. Možnosti zhodnocení pomocí HTA analýz by byla po provedení rozsáhlejšího experimentu se dvěma větvemi pacientů - u jedné by byly využity pacient specifické modely a u druhé by byla použita současná běžná praxe. Následně by obě metody mohly být vyhodnoceny například pomocí analýzy minimalizace nákladů nebo analýzou nákladů a důsledků. Tento výzkum byl i z důvodů časové náročnosti a náročnosti na velikost vzorku pacientů z diplomové práce vynechán.

## 4 Diskuze

Při shromažďování dat bylo zjištěno, že jednotlivé postupy se liší mezi jednotlivými lékaři. Rozhodují se na základě své odbornosti a zkušeností z minulosti. Největší váhu rozhodování o léčbě měl aktuální stav pacienta a přidružené choroby. To se shoduje s autorem Georgy F. Cooper [15]. Diagnóza komunitní pneumonie může mít rychlý postup, kdy může mít nepříznivé dopady a tím jsou vážné komplikace, mezi které patří zejména respirační selhání. U starších pacientů velkou roli hrála také sociální anamnéza a to především to, zdali pacient žije sám, jaké je jeho rodinné zázemí nebo žije v ústavu sociální péče. Velkou podporou zdravotnictví je v dnešní době také tzv. Domácí péče, která poskytuje velmi kvalitní odbornou zdravotní a ošetrovatelskou službu. Zdravotnický pracovník, např. registrovaná zdravotní sestra nebo pečovatelka, dochází podle potřeby k pacientům přímo domů. Jejich velkou výhodou je také pozitivní vliv na psychiku pacienta, který zůstává v pohodlí domova. Domácí péče v některých případech dokáže do jisté míry nahradit hospitalizaci. Vždy je ale navržena lékařem dle stavu pacienta a s jeho souhlasem. Domácí péče má velký potenciál do budoucna.

Předpokladem aplikování prediktivních modelů bylo vytvořit grafické znázornění klasifikačních a regresních stromů pro snadné rozhodování o typu léčby u pacientů s komunitní pneumonií. Pomocí matematického softwaru R verze 2.3 byl strom namodelován. Byl použit vzorek proměnných, který obsahoval data týkající se osobní a rodinné anamnézy, přidružených onemocnění a příznaků u konkrétních pacientů. Cílová proměnná byla tedy zvolena TYP – typ léčby. Výhodou tohoto typu statistického modelování v našem případě bylo především to, že dokázal počítat také s chybějícími daty. Proměnné, týkající se dat laboratorních testů a fyzikálního měření, chyběly především u pacientů léčených ambulantně. Pokud se v praxi zvolí léčba perorálními antibiotiky, laboratorní testy se neprovádí.

V práci bylo také poukázáno na vykazování bodů pojišťovně, které není vždy konzistentní. Liší se provedenými testy a předepsanou léčbou. Každá léčebná terapie má jinou vykazující hodnotu. O našem vzorku můžeme říct, že průměrná cena hospitalizovaného pacienta s komunitní pneumonií je 10.791Kč, která také souvisí s průměrnou dobou hospitalizace osm dnů. Dále na tabulce 11 bylo prokázáno, že nejnižší doba hospitalizace byla jeden den. V tomto případě je potřeba podotknout, že roli při léčbě hraje i pacientův souhlas. Stále mají pacienti možnost podepsání negativního reversu o propuštění z nemocnice. Případ odmítnutí hospitalizace byl v našem souboru dat pouze jeden. Nejdelší doba hospitalizace byla ovlivněna stavem pacientky. V případě respondentky hospitalizované patnáct dní jsme se setkali s imobilitou a neorientovaností časem ani místem. Dále bylo v její anamnéze zahrnuto velké množství přidružených chorob. Zároveň byla pacientka umístěna v ústavu sociální péče.

Dále z pozorování získaných dat vyplývá, že pacient léčený ambulantně stojí zdravotnický systém v průměru 1.938Kč. Odchylky u respondentů jsou způsobeny rozdílnou perorální antibiotickou léčbou. Nejlevnější ambulantní léčba přijde zdravotnický systém na 1.056Kč. Naopak nejdražší léčba byla v našem souboru dat 5.720Kč. V případě respondenta s nejdražší ambulantní léčbou byla přítomna v osobní anamnéze alergie na typ antibiotikum,

kteří se vyskytovalo u většiny respondentů. Pravděpodobně bylo potřeba nasadit jinou, a možná právě dražší, léčbu.

Získaný datový soubor byl využit také pro pročet známého systému Americké hrudní společnosti – PSI [20]. Tento klasifikační systém pomáhá lékařům s predikcí léčby pacientů s diagnózou pneumonie. Zahrnuje data o přidružených chorobách, výsledcích laboratorních testů, fyzikálního vyšetření a RTG nálezů. Dle tabulky PSI je za každý nález u každého respondenta přiřazen určitý počet bodů. V práci byl tento systém využit. U skupiny I, která zahrnuje pacienty s velmi nízkým rizikem, se vyskytovali také respondenti, kteří byli ve finále hospitalizováni. Nutné podotknout, že všech šest takových respondentů byli starší osmdesáti let a měli více jak jednu přidruženou chorobu. Zůstávali v nemocnici v podstatě na pozorování. V tomto případě se také jednalo o nižší počet vykázaných bodů. Takže tito pacienti nejsou tolik zatěžující pro zdravotnický systém. Podobný případ se vyskytl i u skupiny II, u které systém stále doporučuje léčbu ambulantně. V našem souboru dat jsme měli opět větší počet respondentů, kteří byli léčeni hospitalizací. Jednalo se také o starší respondenty. V dalších skupinách pojednávajících o rizikovosti mortality naše data korespondovala s doporučením klasifikačního systému. Tento systém byl využit ve studii autora Antonio Ferreira [1].

Mým předpokladem bylo sestavení klasifikačního stromu usnadňující rozhodování lékařům v ambulancích u diagnózy pneumonie. V této části nebyl předpoklad potvrzen dle autora Ferreira Antonia [57]. Vytvoření klasifikačního stromu jsem modelovala v matematickém softwaru R verze 2.3. Výchozí proměnnou byl typ léčby a ostatní proměnné byly zvolené přidružené choroby, aktuální příznaky při příchodu do ambulance a také byla zohledněna osobní a rodinná anamnéza. Závislost program vyhodnotil však pouze na věku. Hodnotím, že pro sestavení klasifikačního stromu byl použit malý soubor respondentů. Na základě teorie vím, že rozhodování o léčbě u diagnózy pneumonie je velmi individuální. Největší roli hraje aktuální stav pacienta a popřípadě přidružené choroby.

Při sestavování regresního stromu bylo postupováno podobně jako u stromu klasifikačního. Cílová proměnná byla TYP a ostatní proměnné zahrnovala data týkající se osobní a rodinné anamnézy, přidružených chorob a příznaků v ambulanci. Podstatnou součástí byly také výsledky fyzikálního měření a laboratorních výsledků. Při modelaci byla nalezena závislost pouze u dvou proměnných. Jednou z nich byla cévní mozková příhoda v anamnéze a druhá proměnná se týkala fyzikálního měření – systolický tlak. Vizualizací křížové validace na diagramu byly zaneseny chyby v předpovídání typu léčby. V diagramu byla zanesena nejmenší hodnota křížové validace. Ve všech případech má křivka z počátku strmě klesající charakter a pak stabilní nebo mírně rostoucí. Rostoucí část nám ukazuje lepší míru predikce u dané proměnné.

Při prvním zadání všech proměnných byly vybrány nejlepší možné prediktory. U sumarizace bylo poukázáno na možnou závislost u jiných proměnných. Pro grafické znázornění jsme při modelování vyčlenily proměnnou, která v prvním případě byla v prvním uzlu. Ve druhém sestavení stromu nám diagram křížové validace ukázal, že míra predikce je nižší než u prvního případu.

V práci nebyla využita metoda HTA pro zhodnocení pro určení ekonomických nákladů z důvodů studie. Jedná se o vytváření predikčních modelů pro lepší podporu rozhodovacího procesu a metoda lze využít pro potenciální ušetření finančních prostředků. Avšak nebylo možné zakomponovat výsledky do některé z HTA analýz.

V našem systému mají lékaři prostor pro svobodné rozhodování na základě své odbornosti a praxe. Otázkou zůstává, mají prediktivní modely smysl aplikovat na naše zdravotnictví? Výsledky této diplomové práce jsou prvotní podporou pro vznik prediktivních pacient-specifických modelů na větším vzorku pacientů. Doporučením je také nashromáždit větší soubor dat v rámci výzkumného grantu za delší časovou jednotku. V takovém případě je důležité zahrnout více zdravotnických zařízení, aby byly vyřazeny zvyklosti jednoho pracoviště, a zároveň bude zajištěna geografická rozmanitost. Tím bude také zajištěno zohlednění profesního života jednotlivých respondentů.

## Závěr

V klinické praxi je velmi důležitá pacientova prognóza. Rozhodnutí lékařů se vždy opírá o jeho vlastní zkušenosti z minulosti, odbornosti a jeho subjektivního hodnocení pacienta. Predikce léčby je důležitá pro správnou a hlavně efektivní léčbu pacienta z důvodu vhodného čerpání finančních zdravotnických prostředků. Lékař predikuje výsledky za dvojnásobnou nejistoty. Může být správně stanovena diagnóza, ale aplikována špatná léčba, nebo dojde ke špatnému stanovení prognózy a tím i k nasazení neadekvátní léčby.

Mezi predikční modely patří pacient-specifické modely a populačně-široké modely. Pacient-specifické modely mají výhodu využití na budoucích případech, protože nejsou zprůměrované, jako populačně-široké modely. Zároveň nebylo v literatuře dohledáno použití v ČR. Pro modelaci byla vybrána diagnóza komunitní pneumonie.

Byla sesbírána datová sada pacientů v Klaudiánově nemocnici v Mladé Boleslavi. Celkový počet respondentů byl 40. Data u jednotlivých pacientů zahrnovala informace o osobní a rodinné anamnéze, přidružených chorobách, výsledky laboratorních testů, fyzikální vyšetření a příznaky. Pomocí těchto dat byla pro modelaci hledána závislost mezi jednotlivými daty. Předpokladem bylo vytvoření klasifikačního a regresního stromu. Bohužel pro nalezení vhodné proměnné byla datová sada malá a klasifikační strom byl vytvořen jen o dvou listech. V případě regresního stromu byla nalezena závislost, ale s ohledem na diagnózu neadekvátní. Na takové proměnné nelze brát ohled při stanovení prognózy.

Výsledky diplomové práce jsou prvotní podporou pro vznik prediktivních pacient-specifických modelů na větším vzorku pacientů. Doporučením je také nashromáždit větší soubor dat v rámci výzkumného grantu za delší časovou jednotku. V takovém případě je důležité zahrnout více zdravotnických zařízení, aby byly vyřazeny zvyklosti jednoho pracoviště, a zároveň bude zajištěna geografická rozmanitost. Tím bude také zajištěno zohlednění profesního života jednotlivých respondentů.

## Reference

- [1] FERREIRA, A. et al. Decision path models for patient-specific modeling of patient outcomes. In *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium* [online]. 2013. Vol. 2013, s. 413–21. Dostupné z internetu: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3900188&tool=pmcentrez&rendertype=abstract>>.
- [2] VISWESWARAN, S. - COOPER, G.F. Patient-Specific Models for Predicting the Outcomes of Patients with Community Acquired Pneumonia. In *{AMIA} Annual Symposium Proceedings* [online]. 2005. Vol. 2005, s. 759–763. Dostupné z internetu: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1560580/>>.
- [3] MARRIE, T.J.. et al. A controlled trial of a critical pathway for treatment of community-acquired pneumonia. In *Journal of the American Medical Association* [online]. 2000. Vol. 283, no. 6, s. 749–755. Dostupné z internetu: <<http://www.scopus.com/inward/record.url?eid=2-s2.0-0033974755&partnerID=40&md5=e5381cf551ffb2adec479d634c111f23>>.
- [4] METLAY, J.P. et al. Time course of symptom resolution in patients with community-acquired pneumonia. In *Respir Med* . 1998. Vol. 92, no. 9, s. 1137–1142. .
- [5] KRISHNAMURTHY, A. et al. Patient-specific models of cardiac biomechanics. In *Journal of Computational Physics* . 2013. Vol. 244, s. 4–21. .
- [6] VISWESWARAN, S. et al. Learning patient-specific predictive models from clinical data. In *Journal of Biomedical Informatics* [online]. 2010. Vol. 43, no. 5, s. 669–685. Dostupné z internetu: <<http://dx.doi.org/10.1016/j.jbi.2010.04.009>>.
- [7] TAPPENDEN, P. et al. Using whole disease modeling to inform resource allocation decisions: economic evaluation of a clinical guideline for colorectal cancer using a single model. In *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* [online]. 2013. Vol. 16, no. 4, s. 542–53. Dostupné z internetu: <<http://www.sciencedirect.com/science/article/pii/S1098301513000673>>.
- [8] SERMESANT, M. et al. 2008. .
- [9] TAYLOR, C. a - FIGUEROA, C. a Patient-specific modeling of cardiovascular mechanics. In *Annual review of biomedical engineering* . 2009. Vol. 11, no. June, s. 109–134. .
- [10] REINBOLT, J.A. et al. Determination of patient-specific multi-joint kinematic models through two-level optimization. In *Journal of Biomechanics* . 2005. Vol. 38, no. 3, s. 621–626. .
- [11] CEBRAL, J.R. et al. Characterization of cerebral aneurysms for assessing risk of rupture by using patient-specific computational hemodynamics models. In *American Journal of Neuroradiology* . 2005. Vol. 26, no. 10, s. 2550–2559. .
- [12] FREGLY, B.J. Design of Optimal Treatments for Neuromusculoskeletal Disorders using Patient-Specific Multibody Dynamic Models. In *International journal for computational vision and biomechanics* [online]. 2009. Vol. 2, no. 2, s. 145–155. Dostupné z internetu: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3141573&tool=pmcentrez&rendertype=abstract>>.



endertype=abstract>.

[13] FINE, M.J. et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. In *N Engl J Med* [online]. 1997. Vol. 336, no. 4, s. 243–50. Dostupné z internetu: <<http://www.ncbi.nlm.nih.gov/pubmed/8995086>>.

[14] COOPER, G.F. et al. Predicting dire outcomes of patients with community acquired pneumonia. In *Journal of Biomedical Informatics* . 2005. Vol. 38, no. 5, s. 347–366. .

[15] COOPER, G.F. et al. Predicting dire outcomes of patients with community acquired pneumonia. In *J Biomed Inform* [online]. 2005. Vol. 38, no. 5, s. 347–366. Dostupné z internetu: <<http://www.ncbi.nlm.nih.gov/pubmed/16198995>>.

[16] C.S., K. et al. [online]. .2013. Dostupné z internetu: <<http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed11&NEWS=N&AN=2013712321>>.

[17] FINE, M.J. et al. Processes and outcomes of care for patients with community-acquired pneumonia: results from the Pneumonia Patient Outcomes Research Team (PORT) cohort study. In *Arch Intern Med* [online]. 1999. Vol. 159, no. 9, s. 970–80. Dostupné z internetu: <<http://www.ncbi.nlm.nih.gov/pubmed/10326939>>.

[18] GREEN, D.G. - SWETS, J. a Signal detection theory and psychophysics. In *Society* [online]. 1966. Vol. 1, s. 521. Dostupné z internetu: <[http://andrei.gorea.free.fr/Teaching\\_fichiers/SDT and Psytchophysics.pdf](http://andrei.gorea.free.fr/Teaching_fichiers/SDT_and_Psychophysics.pdf)>.

[19] WIKISKRIPTA ROC křivka. In [online]. [cit. 2016-01-01]. Dostupné z internetu: <[http://www.wikiskripta.eu/index.php/ROC\\_k%C5%99ivka](http://www.wikiskripta.eu/index.php/ROC_k%C5%99ivka)>.

[20] CHALMERS, J.D. et al. Severity assessment tools for predicting mortality in hospitalised patients with community-acquired pneumonia. Systematic review and meta-analysis. In *Thorax* . 2010. Vol. 65, no. 10, s. 878–883. .

[21] FILE, T.M. et al. Diagnosis and treatment of community-acquired pneumonia. In *Lancet* . 2003. Vol. 73, no. 3, s. 1991–2001. .

[22] PROF. MUDR. VÍTĚZSLAV KOLEK, DRSC., PROF. MUDR. MILAN KOLÁŘ, PH.D., Mud.V. - KAŠÁK, DOC. MUDR. JIŘÍ BENEŠ, CSC., Mud.J.D. Klinické doporučení: Diagnostika a léčba komunitní pneumonie dospělých. In *Česká pneumologická a ftizeologická společnost ČLS JEP* . 2011. s. 1–15. .

[23] LUTFIYYA, M.N. et al. Diagnosis and treatment of community-acquired pneumonia. In *American Family Physician* . 2006. Vol. 73, no. 3, s. 442–450. .

[24] WUNDERINK, R.G. - WATERER, G.W. Community-Acquired Pneumonia. In *New England Journal of Medicine* [online]. 2014. Vol. 370, no. 6, s. 543–551. Dostupné z internetu:

<<http://www.nejm.org/doi/full/10.1056/NEJMcp1214869>\n<http://www.nejm.org/doi/pdf/10.1056/NEJMcp1214869>>.

[25] FILE, T.M. Community-acquired pneumonia. In *Lancet* [online]. 2003. Vol. 362, no. 9400, s. 1991–2001. Dostupné z internetu: <<http://www.sciencedirect.com/science/article/pii/S0140673603150210>>.

[26] GUEST, J.F. - MORRIS, A. Community-acquired pneumonia: the annual cost to the National Health Service in the UK. In *The European respiratory journal* [online]. 1997. Vol. 10, no. 7, s. 1530–4. Dostupné z internetu: <<http://www.ncbi.nlm.nih.gov/pubmed/9230242>>.

- [27] FALCONE, M. et al. Performance of PSI, CURB-65, and SCAP scores in predicting the outcome of patients with community-acquired and healthcare-associated pneumonia. In *Internal and Emergency Medicine* . 2011. Vol. 6, no. 5, s. 431–436. .
- [28] BROWN, S.M. - DEAN, N.C. Defining and predicting severe community-acquired pneumonia. In *Current opinion in infectious diseases* . 2010. Vol. 23, no. 2, s. 158–164. .
- [29] GONZALEZ, C. et al. Predicting pneumonia mortality using CURB-65, PSI, and patient characteristics in patients presenting to the emergency department of a comprehensive cancer center. In *Cancer Medicine* . 2014. Vol. 3, no. 4, s. 962–970. .
- [30] GLEASON, P.P. et al. Associations between initial antimicrobial therapy and medical outcomes for hospitalized elderly patients with pneumonia. In *Archives of internal medicine* . 1999. Vol. 159, s. 2562–2572. .
- [31] OSTAPCHUK, M. et al. Community-acquired pneumonia in infants and children. In *American family physician* . 2004. Vol. 70, no. 5, s. 899–908. .
- [32] GOPALAKRISHNAN, V. et al. Bayesian rule learning for biomedical data mining. In *Bioinformatics (Oxford, England)* [online]. 2010. Vol. 26, no. 5, s. 668–75. Dostupné z internetu: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2852212&tool=pmcentrez&rendertype=abstract>>.
- [33] BAYES, T. - PRICE, R. An Essay towards Solving a Problem in the Doctrine of Chances. In *Philosophical Transactions of the Royal Society of London* [online]. 1763. Vol. 53, no. 472-477, s. 370–418. Dostupné z internetu: <<http://www.jstor.org/stable/10.2307/105721>\n<http://rstl.royalsocietypublishing.org/cgi/doi/10.1098/rstl.1763.0053>>.
- [34] KASS, R.E. - RAFTERY, A.E. Bayes Factors. In *Journal of the American Statistical Association* [online]. 1995. Vol. 90430, s. 773–795. Dostupné z internetu: <<http://amstat.tandfonline.com/loi/uasa20>\n<http://dx.doi.org/10.1080/01621459.1995.10476572>\n<http://amstat.tandfonline.com/page/terms-and-conditions>>.
- [35] BERNARDO - SMITH 2000. .
- [36] JACOBS, R. a. - KRUSCHKE, J.K. Bayesian learning theory applied to human cognition. In *Wiley Interdisciplinary Reviews: Cognitive Science* . 2011. Vol. 2, no. 1, s. 8–21. .
- [37] JAYNES, E.T. *Probability theory: the logic of science*. . 2005. 83-83 s. ISBN 0521592712.
- [38] OLSHAUSEN, B. Bayesian probability theory. In *The Redwood Center for Theoretical ...* [online]. 2004. s. 1–6. Dostupné z internetu: <<http://redwood.berkeley.edu/bruno/npb163/bayes.pdf>>.
- [39] ALLMARK, P. Bayes and health care research. In *Med.Health Care Philos.* [online]. 2004. Vol. 7, no. 1386-7423 (Print), s. 321–332. Dostupné z internetu: <[file:///c:/Article database/2004 Allmark Bayes and Health care.pdf](file:///c:/Article%20database/2004%20Allmark%20Bayes%20and%20Health%20care.pdf)>.
- [40] LEARNING, M. et al. Induction of Decision Trees. In *Expert Systems* . 2007. s. 81–106. .
- [41] MOISEN, G.G. Classification and Regression Trees. In *Encyclopedia of Ecology* . 2008. no. 2000, s. 582–588. .

- [42] ROKACH, L. - MAIMOM, O. *Data mining with decision trees: theory and applications* [online]. . 2007. 242 s. ISBN 9789812771711.
- [43] LOH, W.-Y. Classification and regression trees. In *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* [online]. 2011. Vol. 1, no. 1, s. 14–23. Dostupné z internetu: <<http://doi.wiley.com/10.1002/widm.8>>.
- [44] BREIMAN, L. et al. *Classification and Regression Trees*. . 1984. 368 s. ISBN 0412048418.
- [45] DE'ATH, G. - FABRICIUS, K.E. Classification and Regression Trees: a Powerful Yet Simple Technique for Ecological Data Analysis. In *Ecology* [online]. 2000. Vol. 81, no. 11, s. 3178–3192. Dostupné z internetu: <<http://www.esajournals.org/doi/abs/10.1890/0012-9658%282000%29081%5B3178%3ACARTAP%5D2.0.CO%3B2>>.
- [46] DE'ATH, G. - FABRICIUS, K.E. Classification and regression trees: A powerful yet simple technique for ecological data analysis. In *Ecology* [online]. 2000. Vol. 81, no. 11, s. 3178–3192. Dostupné z internetu: <[http://www.esajournals.org/doi/abs/10.1890/0012-9658\(2000\)081%5B3178:CARTAP%5D2.0.CO%3B2](http://www.esajournals.org/doi/abs/10.1890/0012-9658(2000)081%5B3178:CARTAP%5D2.0.CO%3B2)>n<Go to ISI>://000165384000018>.
- [47] PŘIBÁŇOVÁ, T. [online]. .Dostupné z internetu: <[http://samba.fsv.cuni.cz/~soukp6as/OLD/AKD/old/zaverecneprace/Pribanova\\_regresni\\_a\\_klasifikačni\\_stromy.docx](http://samba.fsv.cuni.cz/~soukp6as/OLD/AKD/old/zaverecneprace/Pribanova_regresni_a_klasifikačni_stromy.docx)>.
- [48] KUBOŠOVÁ, K. Pokročilé neparametrické metody. In [online]. Dostupné z internetu: <<http://www.iba.muni.cz/esf/res/file/bimat-prednasky/pokrocile-neparametricke-metody/PNM-01.pdf>>.
- [49] HASMAN, a. 1991. .
- [50] FINE, S. et al. The Hierarchical Hidden Markov Model : Analysis and Applications. In *Maschine Learning* . 1998. Vol. 32, no. 1, s. 41–62. .
- [51] SONNENBERG, F.A. - BECK, J.R. Markov models in medical decision making: a practical guide. In *Medical decision making* [online]. 1993. Vol. 13, no. 4, s. 322–38. Dostupné z internetu: <<http://www.ncbi.nlm.nih.gov/pubmed/8246705>\n<http://mdm.sagepub.com/content/13/4/322.short>\n<http://www.ncbi.nlm.nih.gov/pubmed/8246705>>.
- [52] SATO, R.C. - ZOUAIN, D.M. Modelos de Markov aplicados a saúde. In *Einstein* [online]. 2010. Vol. 8, no. 3 Pt 1, s. 376–379. Dostupné z internetu: <<http://bases.bireme.br/cgi-bin/wxislind.exe/iah/online/?IsisScript=iah/iah.xis&src=google&base=LILACS&lang=p&nextAction=lnk&exprSearch=561633&indexSearch=ID>>.
- [53] WOUTERSE, B. et al. Modeling the relationship between health and health care expenditures using a latent Markov model. In *Journal of Health Economics* . 2013. Vol. 32, no. 2, s. 423–439. .
- [54] P, P. et al. Vyent en leukotrien v kondenztu vydechovanho vzduchu u nemocnych s astmatem a CHOPN. In *Alergie* . 2002. Vol. 4, no. 1, s. 15–20. .
- [55] KOTAČKOVÁ, Mud.M.L. C-reaktivní protein. In [online]. 2016. Dostupné z internetu: <<http://www.toplekar.cz/laboratorni-hodnoty/c-reaktivni-protein.html?znak=C>>.
- [56] MINISTERSTVO, Z. Sbírka zákonů. In *Sbírka zákonů* [online]. 2014. s. 144. Dostupné z internetu: <[http://www.mzcr.cz/Legislativa/dokumenty/vyhlasaka-c324/2014-sb-o-stanoveni-hodnot-boduvyse-uhrad-hrazenych-sluzeb-a\\_9871\\_11.html](http://www.mzcr.cz/Legislativa/dokumenty/vyhlasaka-c324/2014-sb-o-stanoveni-hodnot-boduvyse-uhrad-hrazenych-sluzeb-a_9871_11.html)>.

[57] FERREIRA, A. et al. Decision path models for patient-specific modeling of patient outcomes. In *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium* [online]. 2013. Vol. 2013, s. 413–21. Dostupné z internetu: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3900188&tool=pmcentrez&rendertype=abstract>>.

## Seznam obrázků

Obrázek 1: Pacient-specifický model [1].....	12
Obrázek 2: Populačně široký model [1] .....	13
Obrázek 3: Modely v oblasti ROC křivky s nejlepšími výsledky [15].....	17
Obrázek 4: Modely v oblasti ROC křivky s nejhoršími výsledky [15] .....	18
Obrázek 5: Porovnání mezinárodních systémů [20].....	25
Obrázek 6: Odvozování predikčního pravidla – identifikace pacientů ve třídě rizik I [13]	30
Obrázek 7: Příklad rozhodovacího stromu [1,2,40,41,42].....	34
Obrázek 8: Vzor Markova modelu ve zdravotní péči [52] .....	37
Obrázek 9: Graf - anamnéza a přidružené příznaky .....	39
Obrázek 10: Graf – Tabakismus .....	40
Obrázek 11: Graf – Tělesná teplota .....	41
Obrázek 12: Graf – C-reaktivní protein.....	41
Obrázek 13: Graf - Závislost vykázaných bodů na počtu dnů hospitalizace.....	43
Obrázek 14: Křížová validace – klasifikační strom.....	46
Obrázek 15: Klasifikační strom .....	47
Obrázek 16: Regresní strom 1 .....	49
Obrázek 17: Křížová validace – Regresní strom 1 .....	50
Obrázek 18: Regresní strom 2 .....	53
Obrázek 19: Křížová validace – Regresní strom 2 .....	55
Obrázek 20: Regresní strom 3 .....	56
Obrázek 21: Křížová validace – Regresní strom 3 .....	57

## Seznam tabulek

Tabulka 1: Přehled studií o použití pacient-specifických modelů .....	14
Tabulka 3: Predikce katastrofálních výsledků [15] .....	17
Tabulka 4: Porovnání použitých algoritmů – Pacient-specifický model [1] .....	21
Tabulka 5: Porovnání použitých algoritmů – Populačně široký model [1] .....	22
Tabulka 6: Provedené metody [2] .....	24
Tabulka 7: Britský systém určování rizika mortality – CURB [26] .....	27
Tabulka 8: Německý systém CRB-65 [26] .....	27
Tabulka 9: Kritéria rizika dle systému Americké hrudní společnosti [30] .....	28
Tabulka 10: Rizikové skupiny dle systému Americké hrudní společnosti [30]. .....	29
Tabulka 11: Získaná data – věk, pohlaví, doba hospitalizace .....	39
Tabulka 12: Získaná data – Vykázané body, počet dnů hospitalizace a ambulantní péče ..	42
Tabulka 13: Bodový systém PSI – Rizikové skupiny a typ léčby u respondentů .....	44

## Seznam příloh:

Příloha 1: Prohlášení o mlčenlivosti .....	72
Příloha 2: bodové hodnocení dle Americké hrudní společnosti PSI – část 1 .....	73
Příloha 3: bodové hodnocení dle Americké hrudní společnosti PSI – část 2 .....	74
Příloha 4: bodové hodnocení dle Americké hrudní společnosti PSI – část 3 .....	75
Příloha 5: bodové hodnocení dle Americké hrudní společnosti PSI – část 4 .....	76
Příloha 6: Seznam zkratk (proměnných) použitých v příkazech v programu R .....	77



## Prohlášení o mlčenlivosti

Já níže podepsaný(á) \_\_\_\_\_, nar.: \_\_\_\_\_, se zavazuji zachovávat mlčenlivost o všech skutečnostech, o kterých jsem se dozvěděl(a) nebo se v budoucnu dozvím v souvislosti s výkonem mé odborné praxe v Oblastní nemocnici Mladá Boleslav, a.s., nemocnici Středočeského kraje, IČ: 272 56 456, se sídlem třída Václava Klementa 147, 293 01 Mladá Boleslav (dále jen „**nemocnice**“). Povinná mlčenlivost zahrnuje zejména skutečnosti související s poskytováním zdravotních služeb ve smyslu ustanovení § 51 zákona č. 372/2011 Sb., o zdravotních službách, ve znění pozdějších předpisů.

Beru na vědomí, že tato povinnost trvá i po skončení mé praxe v nemocnici.

Povinná mlčenlivost se vztahuje na jakoukoliv formu sdělování informací, např.: ústní sdělení, písemná zpráva, email, poskytnutí elektronických dat, vydání potvrzení, poskytnutí či nahlédnutí do zdravotnické dokumentace, umožnění přístupu k programovému vybavení apod. Charakter důvěrné informace má rovněž skutečnost, zda konkrétní osoba byla v nemocnici ošetřena či nikoliv.

Za porušení povinné mlčenlivosti se nepovažuje:

- a) sdělování údajů nebo jiných skutečností pro potřeby trestního řízení způsobem stanoveným právními předpisy upravujícími trestní řízení,
- b) sdělování údajů nebo jiných skutečností při plnění zákonem uložené povinnosti přezkazit nebo oznámit spáchání trestného činu,
- c) sdělování údajů nebo jiných skutečností v nezbytném rozsahu pro ochranu práv nemocnice v soudním, správním nebo rozhodčím řízení nebo sdělování skutečností soudu nebo jinému orgánu, je-li předmětem řízení spor mezi nemocnicí, popřípadě jejím zaměstnancem, a pacientem nebo jinou osobou uplatňující práva na náhradu škody nebo ochranu osobnosti v souvislosti s poskytováním zdravotních služeb.

Další podrobnosti stanoví provozní řád nemocnice, se kterým jsem se seznámil(a).

V Mladé Boleslavi dne \_\_\_\_\_

.....  
(podpis)



Príloha 2: bodové hodnotenie dle Americké hrudní spoločnosti PSI – časť 1

		1	2	3	4	5	6	7	8	9	10	11	12
<b>Muži</b>	věk	0	89	89	0	0	88	87	0	87	0	0	86
<b>Ženy</b>	věk - 10	81	0	0	79	78	0	0	77	0	76	76	0
<b>Ústav soc. péče</b>	věk + 10	10	0	10	0	0	10	0	0	0	0	0	0
<b>Nádory</b>	30	0	0	0	0	0	0	0	0	0	0	0	0
<b>Hepatopatie</b>	20	0	0	0	0	0	0	0	0	0	0	0	0
<b>ICHS</b>	10	0	10	10	10	10	0	0	0	10	10	0	10
<b>CMP</b>	10	0	0	10	10	0	0	0	0	0	0	0	10
<b>Nefropatie</b>	10	0	0	0	10	10	0	0	0	0	0	0	0
<b>pH pod 7,35</b>	30	0	0	30	0	0	0	0	0	0	0	0	30
<b>urea nad 10,7 mmol/l</b>	20	20	0	20	20	0	0	0	0	0	0	0	0
<b>natrium pod 130 mmol/l</b>	20	0	0	0	0	0	0	0	0	0	0	0	0
<b>glykemie nad 13,9 mmol/l</b>	10	0	0	0	0	0	0	0	0	0	0	0	10
<b>hematokrit pod 30%</b>	10	0	0	0	10	10	0	0	0	0	10	0	0
<b>pO2 pod 8kPa</b>	10	0	10	0	0	0	0	0	0	0	0	0	10
<b>pleurální výpotek</b>	10	0	10	0	0	0	0	0	0	0	0	0	0
<b>alterace mentálního stavu</b>	20	20	0	20	20	0	0	0	0	0	0	20	0
<b>dechová frekvence nad 30 dechů/min</b>	20	20	0	20	0	0	0	0	0	0	0	0	0
<b>systolický TK pod 90 mm HG</b>	20	0	0	0	20	0	0	0	0	0	0	0	0
<b>teplota pod 35° C nebo nad 40°C</b>	15	0	0	0	0	0	0	0	0	0	0	0	0
<b>pulz nad 125/min</b>	10	10	0	0	0	0	0	0	0	0	0	0	0
<b>SOUČET BODŮ</b>		161	119	209	179	108	98	87	77	97	96	96	156
<b>Typ léčby</b>		H	H	H	H	H	H	H	H	H	H	H	H

Legenda

I.skupina

II.skupina

III.skupina

IV. Skupina

V. Skupina

Příloha 3: bodové hodnocení dle Americké hrudní společnosti PSI – část 2

		13	14	15	16	17	18	19	20	21	22	23	24
<b>Muži</b>	věk	0	0	88	60	0	0	48	47	0	53	0	0
<b>Ženy</b>	věk - 10	75	75	0	0	43	41	0	0	46	0	53	53
<b>Ústav soc. péče</b>	věk + 10	0	0	0	0	0	0	0	0	0	0	0	0
<b>Nádory</b>	30	0	0	30	0	0	0	0	0	0	0	0	0
<b>Hepatopatie</b>	20	0	0	0	0	0	0	0	0	0	0	0	0
<b>ICHS</b>	10	10	10	10	0	0	0	0	0	0	0	0	0
<b>CMP</b>	10	10	10	10	0	0	0	0	0	0	0	0	0
<b>Nefropatie</b>	10	0	0	0	0	0	0	0	0	0	0	0	0
<b>pH pod 7,35</b>	30	0	0	0	0	0	0	0	0	0	0	0	0
<b>urea nad 10,7 mmol/l</b>	20	0	0	0	0	0	0	0	0	20	0	0	0
<b>natrium pod 130 mmol/l</b>	20	0	0	0	0	0	0	0	0	0	0	0	0
<b>glykemie nad 13,9 mmol/l</b>	10	0	0	10	0	0	0	0	0	0	0	0	0
<b>hematokrit pod 30%</b>	10	0	0	10	0	0	0	0	0	0	0	0	0
<b>pO2 pod 8kPa</b>	10	0	0	10	0	0	0	0	0	10	0	0	0
<b>pleurální výpotek</b>	10	0	0	0	0	0	0	0	0	0	0	0	0
<b>alterace mentálního stavu</b>	20	0	20	0	0	20	0	0	0	0	0	20	0
<b>dechová frekvence nad 30 dechů/min</b>	20	0	0	0	0	0	0	0	0	0	0	0	0
<b>systolický TK pod 90 mm HG</b>	20	0	0	0	0	0	0	0	0	0	0	0	0
<b>teplota pod 35° C nebo nad 40°C</b>	15	0	0	0	0	0	0	0	0	0	0	0	0
<b>pulz nad 125/min</b>	10	0	0	0	0	0	0	0	0	10	0	0	0
<b>SOUČET BODŮ</b>		95	115	168	60	63	41	48	47	86	53	73	53
<b>Typ léčby</b>		H	H	H	H	H	H	H	H	H	H	H	H

Legenda:

I.skupina
II.skupina
III.skupina
IV. Skupina
V. Skupina

Příloha 4: bodové hodnocení dle Americké hrudní společnosti PSI – část 3

		25	26	27	28	29	30	31	32	33	34	35	36
<b>Muži</b>	věk	0	62	0	43	43	41	39	39	0	0	0	0
<b>Ženy</b>	věk - 10	52	0	81	0	0	0	0	0	26	23	21	19
<b>Ústav soc. péče</b>	věk + 10	0	0	0	0	0	0	0	0	0	0	0	0
<b>Nádory</b>	30	0	0	0	0	0	0	0	0	0	0	0	0
<b>Hepatopatie</b>	20	0	0	0	0	0	0	0	0	0	0	0	0
<b>ICHS</b>	10	0	0	0	0	0	0	0	0	0	0	0	0
<b>CMP</b>	10	0	0	0	0	0	0	0	0	0	0	0	0
<b>Nefropatie</b>	10	0	0	0	0	0	0	0	0	0	0	0	0
<b>pH pod 7,35</b>	30	0	0	0	0	0	0	0	0	0	0	0	0
<b>urea nad 10,7 mmol/l</b>	20	0	0	0	0	0	0	0	0	0	0	0	0
<b>natrium pod 130 mmol/l</b>	20	0	0	0	0	0	0	0	0	0	0	0	0
<b>glykemie nad 13,9 mmol/l</b>	10	0	0	0	0	10	0	0	0	0	0	0	0
<b>hematokrit pod 30%</b>	10	0	0	0	0	0	0	0	0	0	0	0	0
<b>pO2 pod 8kPa</b>	10	0	0	0	0	0	0	0	0	0	0	0	0
<b>pleurální výpotek</b>	10	0	0	0	0	10	0	0	0	0	0	0	10
<b>alterace mentálního stavu</b>	20	0	20	0	0	0	0	0	0	0	0	0	0
<b>dechová frekvence nad 30 dechů/min</b>	20	0	0	0	0	0	0	20	0	0	0	0	0
<b>systolický TK pod 90 mm HG</b>	20	0	20	0	0	0	0	0	0	0	0	0	0
<b>teplota pod 35° C nebo nad 40°C</b>	15	0	0	0	0	0	0	0	0	0	0	0	0
<b>pulz nad 125/min</b>	10	0	0	0	0	0	0	0	0	0	0	0	0
<b>SOUČET BODŮ</b>		52	102	81	43	63	41	59	39	26	23	21	29
<b>Typ léčby</b>		H	H	A	A	H	A	A	A	A	A	A	A

I.skupina
II.skupina
III.skupina
IV. Skupina
V. Skupina

		37	38	39	40
<b>Muži</b>	<i>věk</i>	0	0	80	0
<b>Ženy</b>	<i>věk -</i> <i>10</i>	74	72	0	63
<b>Ústav soc. péče</b>	<i>věk +</i> <i>10</i>	0	0	0	0
<b>Nádory</b>	<i>30</i>	0	0	0	0
<b>Hepatopatie</b>	<i>20</i>	0	0	0	0
<b>ICHS</b>	<i>10</i>	0	10	10	0
<b>CMP</b>	<i>10</i>	0	10	0	0
<b>Nefropatie</b>	<i>10</i>	0	10	0	0
<b>pH pod 7,35</b>	<i>30</i>	0	0	0	0
<b>urea nad 10,7 mmol/l</b>	<i>20</i>	20	0	0	0
<b>natrium pod 130 mmol/l</b>	<i>20</i>	0	0	0	0
<b>glykemie nad 13,9 mmol/l</b>	<i>10</i>	10	10	10	10
<b>hematokrit pod 30%</b>	<i>10</i>	0	0	0	0
<b>pO2 pod 8kPa</b>	<i>10</i>	0	0	0	0
<b>pleurální výpotek</b>	<i>10</i>	0	0	0	0
<b>alterace mentálního stavu</b>	<i>20</i>	0	20	0	0
<b>dechová frekvence nad 30 dechů/min</b>	<i>20</i>	0	0	0	0
<b>systolický TK pod 90 mm HG</b>	<i>20</i>	0	0	0	0
<b>teplota pod 35° C nebo nad 40°C</b>	<i>15</i>	0	0	0	0
<b>pulz nad 125/min</b>	<i>10</i>	0	0	0	0
<b>SOUČET BODŮ</b>		<b>104</b>	<b>132</b>	<b>100</b>	<b>73</b>
<b>Typ léčby</b>		H	H	H	H

I.skupina

II.skupina

III.skupina

IV. Skupina

V. Skupina

*Příloha 6: Seznam zkratk (proměnných) použitých v příkazech v programu R*

SU - sociální ústav	AST - astma
CMP - cévní mozková příhoda	CYA - cyanóza
ISCH - ischemická choroba srdeční	KUR - kuřák
DM - diabetes mellitus II. typu	WBC - leukocyty, krevní obraz
NEF - nefropatie	RBC - erytrocyty, krevní obraz
HEP - hepatopatie	HCT - hematokrit, krevní obraz
RAK - rakovina aktivní	MCV - str. obj. erytr, krevní obraz
RAK <sub>m</sub> - rakovina v minulosti	MCH - barvivo eryt., krevní obraz
ZM - zmatenost	RDW - eryt. křivka, krevní obraz
BB - bolest břicha	PLT - trombocyty
BnH - bolest na hrudi	U - urea, močovina v krvi
BH - bolest hlavy	Krea - kreatinin
DU - dusnost	Gluk - glukóza
U - únava	K - kalium
ZIM - zimnice	Na - natrium
KAS - kašel	Cl - chloridy
SP - sputum	CRP - C
PR - průjem	TKS - tlak systolický
ZV - zvracení	TKD - tlak diastolický
VK - hemoptýza, vykašlávání krve	TF - tepová frekvence
NV - nevolnost	T - teplota
POC - pocení	S - saturace