

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

FAKULTA BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

KATEDRA BIOMEDICÍNSKÉ TECHNIKY

**Meta-analýza studií na diagnostickou přesnost
A meta-analysis of studies on diagnostic accuracy**

DIPLOMOVÁ PRÁCE

Studijní program: Biomedicínská a klinická technika (BMKT)

Studijní obor: Systémová integrace procesů ve zdravotnictví (SIPZ)

Vedoucí práce: Ing. Vojtěch Kamenský

Bc. Monika Čapková

18. května 2017, Kladno

Katedra biomedicínské techniky

Akademický rok: 2016/2017

Z a d á n í d i p l o m o v é p r á c e

Student: **Bc. Monika Čapková**
Studijní obor: Systémová integrace procesů ve zdravotnictví
Téma: **Meta-analýza studií na diagnostickou přesnost**
Téma anglicky: A meta-analysis of studies on diagnostic accuracy

Z á s a d y p r o v y p r a c o v á n í :

Cílem diplomové práce je zpracování meta-analýzy se zaměřením na diagnostickou přesnost. V práci analyzujte současné postupy při hodnocení studií na diagnostickou přesnost v ČR a ve světě. Zaměřte se na problematiku hodnocení heterogenity/homogenity studií na diagnostickou přesnost a možností řešení této problematiky. Následně vybrané metody využijte při zpracování dat v rámci meta-analýzy. Meta-analýzu zaměřte na diagnostickou přesnost vybrané zdravotnické technologie při diagnostice roztroušené sklerózy. Dále v práci zhodnoťte výhody a nevýhody a limitace tohoto typu meta-analýzy.

Seznam odborné literatury:

- [1] M.M.G. Leeflang, Systematic reviews and meta-analyses of diagnostic test accuracy., Clinical Microbiology and Infection, ročník 20, číslo 2, 2014
- [2] Goddman, C.S., HTA 101 - Introduction to health technology assessment, Virginia USA, 2004

Vedoucí: Ing. Vojtěch Kamenský

Zadání platné do: 20.08.2017

.....
vedoucí katedry / pracoviště

.....
děkan

V Kladně dne 20.02.2017

Abstrakt

Roztroušená skleróza je chronické zánětlivé onemocnění centrální nervové soustavy, projevující se demyelinizací nervových vláken. Pro včasnou diagnostiku a určení míry poškození mozku a míchy je nejčastěji využíváno vyšetření magnetickou rezonancí. Tato metoda je považována za efektivní screeningovou metodu. Ve většině studií je tato efektivita potvrzena poměrně vysokou diagnostickou přesností. Cílem práce je tuto diagnostickou přesnost ověřit pomocí meta-analýzy dostupných studií zaměřených na diagnostickou přesnost této metody.

Meta-analýza byla provedena dle standardního postupu. Tedy výběr studií na základě definovaných kritérií, ověření kvality metodologického postupu jednotlivých studií určených k meta-analýze, vlastní meta-analytické zpracování vstupních dat a správná interpretace výsledků.

Z celkem 625 studií bylo vybráno 12 splňujících určená kritéria. Těchto 12 studií obsahovalo celkem 23 hodnotících kritérií testu, tedy bylo možné použít 23 údajů o senzitivitě a specificitě, použitelných v následné meta-analýze. Pomocí QUADAS 2 byla zhodnocena metodologická kvalita studií. A následně prostřednictvím statistických programů Open-Meta[Analyst] a Meta-DiSc byly vypočteny souhrnné hodnoty senzitivity 0,76 v intervalu (0,68; 0,82), specificity 0,65 v intervalu (0,57; 0,72), *NLR* 0,28 v intervalu (0,18; 0,42), *PLR* 2,00 v intervalu (1,65; 2,42) a *DOR* 6,15 v intervalu (3,88; 9,74). Dále byl určen Spearmanův korelační koeficient 0,68 a na základě SROC křivky určen parametr *AUC* 0,75 se směrodatnou odchylkou 0,04.

Na základě meta-analýzou určených agregovaných hodnot bylo přes patrnou heterogenitu studií zhodnoceno použití magnetické rezonance pro diagnózu roztroušené sklerózy jako metoda s dobrou diagnostickou přesností.

Klíčová slova: Magnetická rezonance, roztroušená skleróza, diagnostická přesnost, senzitivita, specificita, meta-analýza.

Abstract

Multiple sclerosis is a chronic inflammatory disease of the central nervous system manifested by demyelination of the nerve fibers. Magnetic resonance imaging is most commonly used for early diagnosis and determination of brain and spinal cord damage. This method is considered as effective screening method. In most studies, this effectiveness is confirmed by a relatively high diagnostic accuracy. The aim of the thesis is to verify this diagnostic accuracy by meta-analysis of available studies focusing on the diagnostic accuracy of this method.

The meta-analysis was performed according to the standard procedure. Thus, the selection of studies based on defined criteria, the verification of the methodological quality of individual studies for meta-analysis, the own meta-analytical processing of the input data and the interpretation of the results.

From total of 625 studies was selected 12 studies meeting the specified criteria. These 12 studies included a total of 23 evaluation criteria for the test, so it was possible to use 23 sensitivity and specificity data to be used in subsequent meta-analysis. Using QUADAS 2, the methodological quality of studies was evaluated. Then summarized values were calculated using the OpenMeta[Analyst] and Meta-DiSc statistical programs and sensitivity was 0,76 in the interval(0,68; 0,82), specificity was 0,65 in the interval (0,57; 0,72); NLR 0,28 in the interval (0,18; 0,42); PLR 2,00 in the interval (1,65; 2,42) and DOR 6,15 in the interval (3,88; 9,74). Then a Spearman correlation coefficient of 0,68 was determined and the AUC 0,75 with a standard deviation of 0.04 was determined based on the SROC curve.

Based on the meta-analysis of the determined aggregate values, despite the apparent heterogeneity of studies, the use of magnetic resonance imaging for the diagnosis of multiple sclerosis as a method of good diagnostic accuracy was evaluated.

Key words: Magnetic Resonance Imaging, multiple sclerosis, diagnostic accuracy, sensitivity, specificity, meta-analysis.

PODĚKOVÁNÍ

Chtěla bych na tomto místě poděkovat vedoucímu diplomové práce panu Ing. Vojtěchovi Kamenskému za odborné vedení této práce, poskytnutí zásadních studijních materiálů a za jeho ochotnou při řešení dané problematiky.

PROHLÁŠENÍ

Prohlašuji, že jsem diplomovou práci s názvem:

META-ANALÝZA STUDÍ NA DIAGNOSTICKOU PŘESNOST

vypracovala samostatně a použila k tomu úplný výčet citací použitých pramenů, které uvádím v seznamu přiloženém k závěrečné zprávě.

Nemám závažný důvod proti užití tohoto školního díla ve smyslu §60 Zákona č.121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Kladně dne 18. května 2017

Podpis:

Obsah

Úvod	1
1 Přehled současného stavu	2
1.1 Meta-analýza	2
1.2 Diagnostická přesnost	4
1.2.1 Specificita a senzitivita	4
1.2.2 Prediktivní hodnota	5
1.2.3 Pravděpodobnostní poměr (<i>LR</i>)	6
1.2.4 ROC křivka	7
1.3 Meta-analýza DTA studií	8
1.3.1 Rozdíl mezi meta-analýzou intervenčních a DTA studií	8
1.3.2 Doporučené postupy pro tvorbu meta-analýzy DTA studií	9
1.4 Roztroušená skleróza	15
1.5 Diagnostika RS pomocí magnetické rezonance	16
2 Metody	18
2.1 Definice výzkumné otázky	18
2.2 Výběr studií	18
2.3 PRISMA diagram	19
2.4 Hodnocení metodologické kvality studií	19
2.5 Extrakce dat	21
2.6 Analýza dat	22
3 Výsledky	23
3.1 Výběr studií	23
3.2 QUADAS 2	26
3.3 Meta-analýza	28
3.3.1 <i>NLR</i>	30
3.3.2 <i>PLR</i>	32
3.3.3 Senzitivita	34
3.3.4 Specificita	36
3.3.5 <i>DOR</i>	38
3.3.6 SROC křivka	40
4 Diskuze	42

Závěr	45
Reference	46
Příloha I	50
Struktura souborů na příložením CD	50

Seznam obrázků

1.1	Příklady ROC křivek	8
1.2	MRI snímky mozku s RS	17
3.1	Prisma diagram	24
3.2	QUADAS 2 graf	26
3.3	<i>NLR forest</i> -graf	30
3.4	<i>NLR</i> kumulativní <i>forest</i> -graf	31
3.5	<i>PLR forest</i> -graf	32
3.6	<i>PLR</i> kumulativní <i>forest</i> -graf	33
3.7	<i>Forest</i> -graf senzitivity	34
3.8	Senzitivita kumulativní <i>forest</i> -graf	35
3.9	<i>Forest</i> -graf specificita	36
3.10	Specificita kumulativní <i>forest</i> -graf	37
3.11	<i>DOR forest</i> -graf	38
3.12	<i>DOR</i> kumulativní <i>forest</i> -graf	39
3.13	SROC symetrická analýza	40
3.14	SROC symetrická analýza včetně intervalů spolehlivosti	41

Seznam tabulek

1.1	2x2 tabulka senzitivity a specificity	4
1.2	Vztah mezi AUC a diagnostickou přesností	7
1.3	Revidovaná McDonaldova diagnostická kritéria z roku 2010	16
2.1	Výpočet dat v 2x2 tabulce senzitivity a specificity	21
3.1	Studie rozřazené podle kritérií výběru pacientů	25
3.2	QUADAS 2	27
3.3	Použitá data jednotlivých studií	28
3.4	Výsledné hodnoty NLR , PLR , senzitivity, specificity a DOR	29

Seznam zkratek

<i>AUC</i>	plocha pod křivkou (<i>Area under the curve</i>)
CNS	centrální nervový systém
CT	výpočetní tomografie (<i>Computed Tomography</i>)
DIS	diseminace v prostoru
DIT	diseminace v čase
<i>DOR</i>	poměr diagnostických šancí (<i>Diagnostic odds ratio</i>)
DTA	přesnost diagnostických testů (<i>Diagnostic Test Accuracy</i>)
<i>FN</i>	falešně negativní
<i>FP</i>	falešně pozitivní
HSROC	hierarchický souhrnný ROC (<i>Hierarchical Summary Receiver Operating Characterctics</i>)
<i>LR</i>	pravděpodobnostní poměr
MR	magnetická rezonance
<i>NLR</i>	negativní pravděpodobnostní poměr
<i>NPV</i>	negativní prediktivní hodnota
<i>PLR</i>	pozitivní pravděpodobnostní poměr
PP RS	primárně progresivní roztroušená skleróza
<i>PPV</i>	pozitivní prediktivní hodnota
QUADAS	hodnocení kvality studií na diagnostickou přesnost (<i>Quality Assessment of Diagnostic Accuracy Studies</i>)
ROC	<i>Receiver Operating Characteristic</i>
RR RS	relaps-remitentní roztroušená skleróza
RS	roztroušená skleróza
SP RS	sekundárně progresivní roztroušená skleróza
SROC	souhrnná ROC křivka (<i>Summary Receiver Operating Characterctics</i>)
<i>TN</i>	skutečně negativní
<i>TP</i>	skutečně pozitivní

Úvod

Roztroušená skleróza (RS) je chronické zánětlivé onemocnění postihující mozek a míchu. Jedná se o nejčastější neurologickou příčinu invalidity lidí v produktivním věku s rostoucí prevalencí v populaci v ČR 1/1000 obyvatel. Touto chorobou trpí dvakrát častěji ženy než muži. Příčinou onemocnění je průnik buněk vlastního imunitního systému přes poničenou hematoencefalickou bariéru do mozku a míchy. V centrální nervovém systému poté aktivované imunitní buňky útočí proti tkáním vlastního těla a to konkrétně proti myelinu. Myelin je bílkovinné povahy a zajišťuje správnou funkci neuronů. Poškození myelinu může vést až k destrukci neuronů a k následnému postižení nemocného. [21]

Pro diagnostiku RS se nejčastěji používá vyšetření mozku a míchy pomocí magnetické rezonance (MR) a průkaz oligoklonálních pásů v mozkomíšním moku. Magnetická rezonance má neocenitelný přínos jak v diagnostice, tak i ve sledování průběhu onemocnění a účinnosti léčby. Průběh onemocnění má několik fází, ale na léčbu dobře reaguje jen pacient v počátečních fázích onemocnění, kdy demyelinizace převažuje nad neurodegenerací. Proto je zásadní diagnostikovat RS včas a začít co možná nejrychleji s vhodnou léčbou. Pro rychlou a přesnou diagnostiku RS jsou nezbytné diagnostické testy s vysokou diagnostickou přesností. Ve světě existuje několik studií potvrzující poměrně dobrou diagnostickou přesnost MR na rozdíl od vyšetření pomocí CT, které nedokáže odhalit takový počet lézí jako magnetická rezonance. [21, 25]

Cílem práce je zhodnotit prostřednictvím meta-analýzy vhodného souboru studií diagnostickou přesnost metody MR vyšetření, především se zaměřením na hodnoty senzitivity a specificity.

Kapitola 1

Přehled současného stavu

V následující kapitole je charakterizovaná meta-analýza a diagnostická přesnost. Dále jsou zde popsány jednotlivé pojmy vztahující se k diagnostické přesnosti jako je senzitivita, specifita, prediktivní hodnota, pravděpodobnostní poměr a ROC (*Receiver Operating Characteristics*). V další kapitole jsou uvedeny doporučené postupy pro tvorbu meta-analýzy DTA (*Diagnostic Test Accuracy*) studií. Následující část se zabývá roztroušenou sklerózou a použitím magnetické rezonance při tomto onemocnění.

1.1 Meta-analýza

Meta-analýza využívá statistických postupů s cílem shrnout výsledky ze dvou či více empirických studií, které jsou zaměřené na stejný nebo obdobným problémem. Meta-analýza se zaměřuje na velikost účinků (*effect size*), zjištěných v jednotlivých studiích. K důležité formě vědeckého sdělení patří tzv. přehledy (*review*), které je možné klasifikovat do čtyř skupin. První typ přehledů má za úkol identifikovat a posuzovat nové trendy ve vědě. Druhá skupina využívá empirickou evidenci k vysvětlení nebo hodnocení určité teorie a také k navržení teorie nové. Třetí druh přehledů třídí a hodnotí znalosti z rozdílných oblastí výzkumu. Čtvrtý typ přehledů se nazývá integrativní typ a řadí se do něj i meta-analýza. Tento druh přehledů usiluje o zobecnění výsledků výzkumu z množiny studií o specifické vědecké otázce. [1, 2]

Meta-analýza má formu systematického přehledu, který aplikuje vědecké strategie za účelem omezení vzniku systematických chyb při shromažďování, kritickém hodnocení a syntéze všech relevantních studií k danému tématu. Využitím meta-analýzy lze získat dva typy poznatků. Prvním typem jsou poznatky získané z primárních studií, např. testování hypotézy účinnosti léku, a druhým typem poznatky vycházející z použití meta-analytické metody. Druhá kategorie vychází z variability vlastností jednotlivých primárních studií. Na základě této variability je možné označit tzv. moderující proměnné, které mají vypovídající charakter účinnosti terapie. [1, 2]

Stejně jako empirický výzkum je důležité vytvářet meta-analýzu systematicky v několika etapách. První etapou je **definice problému**. Je nutné jasně definovat výzkumnou otázku, parametry, které ve studiích budeme porovnávat, a vytvořit hypotézu, kterou budeme meta-analýzou ověřovat. [1, 2]

Druhou velmi důležitou etapou meta-analýzy je **vyhledávání a selekce primárních studií**. Pro vhodnou selekci studií by měla být stanovena kritéria výběru, např. omezení se pouze na randomizované kontrolované studie publikované v impaktovaných časopisech. Správnou selekcí se minimalizuje výsledkové zkreslení vyplývající z heterogenity použitých studií. Mohou se totiž lišit např. díky rozdílným podmínkám, při kterých byla primární studie prováděna, nebo odlišným výběrem testovaných subjektů. Výběrová kritéria primárních studií musí být součástí závěrečné zprávy meta-analytické studie. [1, 2]

Časově nejnáročnější etapou je **vytváření matice dat** extrahovaných z vyselektovaných studií. Při extrakci posuzovaných parametrů je nutné uvažovat také potenciální ovlivnění dat faktory nerelevantními danému problému. [1, 2]

Po vytvoření matice dat přistupujeme k **agregaci výsledků jednotlivých studií**. To znamená, že z jednotlivých zpráv se snažíme agregovat odhad průměrného účinku, který by reprezentoval všechny studie. Agregaci provádíme pomocí statistických metod, které se řídí zvoleným modelem. Jde o modely „pevných efektů“, které očekávají, že rozdělení účinků v jednotlivých studiích mají stejný průměr, ale rozdílné rozptyly, nebo modely „náhodných efektů“. Modely náhodných efektů jsou zobecněným modelem pevných efektů, jelikož nepředpokládají stejný průměr, ale pouze normální rozdělení, čímž vzniká dvoustupňový model zahrnující míru nehomogenity mezi studii. [1, 2]

Při zjištění nehomogenity výsledků zkoumáme její příčiny a provádíme **analýzu citlivosti**, kdy porovnávat např. vliv kvality studie na velikost účinku a snažíme se charakterizovat podskupiny výsledků, kde má sledovaný vztah danou účinnost. [1, 2]

Závěrečnou fází jsou **interpretace a prezentace výsledků**. Zde dochází ke zhodnocení, do jaké míry jsou zjištěné efekty závislé na sledovaných faktorech, a ke zanalyzování možných publikačních šumů, které by měly být potlačeny právě vhodnými statistickými metodami v těchto kvantitativně pojatých přehledech. [1, 2]

1.2 Diagnostická přesnost

Diagnostické testy se používají k identifikaci přítomnosti či nepřítomnosti určitého onemocnění (obecněji určitého cílového stavu) za účelem zvolení vhodného postupu léčby. Diagnostická přesnost se vztahuje ke schopnosti testu rozlišovat pacienty s onemocněním a bez něj. Mezi příklady diagnostických testů patří využití zobrazovacích a biochemických metod, různá patologická a psychologická vyšetření, pozorování symptomů a klinické hodnocení. Vzhledem k požadavkům pro zlepšení rychlosti, snadnosti provedení, zvýšení bezpečnosti pacientů, přesnosti a snížení nákladů se neustále vyvíjejí nové diagnostické testy. [3, 4]

Měření diagnostické přesnosti souvisí s různými aspekty diagnostického postupu. Zatímco některá měření jsou používána k posouzení diskriminačních vlastností testu, jiné se používají k posouzení jeho prediktivní schopnosti. Diskriminační vlastnosti jsou většinou využívány pro zdravotnicko-politická rozhodnutí, naopak prediktivní opatření jsou velmi užitečná pro předvídání pravděpodobnosti onemocnění u testovaného jedince. Měření diagnostické přesnosti je velmi citlivé na charakteristiky populace, na které je hodnocen diagnostický test. Některá měření do značné míry závisí na prevalenci onemocnění, zatímco jiná jsou citlivá na spektrum onemocnění ve studované populaci. [3, 4]

V primárních studiích na diagnostickou přesnost je test zájmu (indexový test) porovnáván s existujícím diagnostickým testem (referenční test). Výsledky těchto dvou testů se posléze porovnávají za účelem vyhodnocení přesnosti indexového testu. Dva hlavní typy studií na diagnostickou přesnost jsou *case-control* studie a *cross-sectional*. V *case-control* studiích pacienti s daným onemocněním pocházejí z jiné populace než zdraví pacienti. V *cross-sectional* studiích všichni pacienti s podezřením na danou chorobu podstupují oba testy, jak indexový, tak i referenční test. [3, 4]

1.2.1 Specificita a senzitivita

Dokonalý diagnostický postup má potenciál zcela rozlišovat vzorky s a bez onemocnění. K tomuto rozlišení slouží hodnoty *cut-off*, nebo-li limity, jejichž cílem je rozdělit populaci na skupinu bez přítomnosti choroby a skupinu s přítomností choroby. Hodnoty perfektního testu, které jsou vyšší než *cut-off* hodnota jsou vždy pozitivní, zatímco pod *cut-off* hodnotou jsou vždy negativní. Bohužel, dokonalý test neexistuje, a proto diagnostické postupy mohou odhalit rozdíl mezi vzorky s a bez onemocnění pouze částečně. [5, 6]

Tabulka 1.1: 2×2 tabulka senzitivity a specificity

	subjekty s nemocí	subjekty bez nemoci
pozitivní	TP	FP
negativní	FN	TN

Hodnoty nad *cut-off* nesvědčí vždy o nemoci, protože u pacientů bez onemocnění může někdy dojít ke zvýšení hodnoty. Tyto zvýšené hodnoty určitého parametru se nazývají falešně pozitivní hodnoty (*FP*). Na druhé straně, hodnoty nižší než hodnota *cut-off* se nacházejí hlavně u jedinců bez onemocnění. Nicméně, někteří z nich nemocní být mohou. Tyto hodnoty jsou pak falešně negativní (*FN*). Proto *cut-off* dělí populaci vyšetřených pacientů s a bez onemocnění do čtyř podskupin:

1. skutečně pozitivní (*TP*) - pacienti nemocní s hodnotou parametru zájmu nad *cut-off*,
2. falešně pozitivní (*FP*) - pacienti bez nemoci s hodnotou parametru zájmu nad *cut-off*,
3. skutečně negativní (*TN*) – pacienti bez nemoci s hodnotou parametru zájmu pod *cut-off*,
4. falešně negativní (*FN*) – pacienti nemocní s hodnotou parametru zájmu pod *cut-off*.

Prvním krokem k výpočtu senzitivity a specificity je vytvoření tabulky 2×2 , v níž sloupce rozdělují skupiny pacientů dle zlatého řezu (podle referenční metody) a řádky znamenají kategorie podle výsledku testu (viz Tabulka 1.1).

Senzitivita se udává v procentech a vyjadřuje poměr mezi skutečně pozitivními a všemi nemocnými ve zkoumané skupině pacientů:

$$\text{senzitivita} = \frac{TP}{TP + FN} \quad (1.1)$$

V podstatě jde o schopnost metody odhalit u pacienta onemocnění. [5, 6]

Specifita narozdíl od senzitivity udává poměr mezi zdravými pacienty a celkovým počtem zdravých pacientů v souboru:

$$\text{specifita} = \frac{TN}{TN + FP} \quad (1.2)$$

Jinými slovy specifita určuje pravděpodobnost negativního výsledku testu u zdravého pacienta. [5, 6]

Ani senzitivita ani specifita nejsou ovlivněny výskytem choroby, což znamená, že výsledky z jedné studie by mohly být snadno přeneseny do jiného prostředí s odlišným výskytem tohoto onemocnění v populaci. Nicméně senzitivita a specifita se může výrazně lišit v závislosti na spektru onemocněních ve studované skupině. [5, 6]

1.2.2 Prediktivní hodnota

Pozitivní prediktivní hodnota (*PPV*) znamená pravděpodobnost, že pacient s pozitivním výsledkem testu bude opravdu nemocný. Proto *PPV* určuje podíl mezi správně pozitivními vůči všem pozitivním pacientům [5]:

$$PPV = \frac{TP}{TP + FP} \quad (1.3)$$

Naopak negativní prediktivní hodnota (NPV) vyjadřuje pravděpodobnost, že zdravý pacient bude mít negativní výsledek testu. Jinými slovy, NPV je podíl mezi skutečně negativními a celkovým počtem negativních pacientů [5].

$$NPV = \frac{TN}{TN + FN} \quad (1.4)$$

Na rozdíl od senzitivity a specificity, jsou prediktivní hodnoty do značné míry závislé na prevalenci onemocnění v populaci. Proto by hodnoty z jedné studie neměly být přenášeny do jiného prostředí s rozdílnou prevalencí. Prevalence ovlivňuje PPV a NPV opačným trendem. PPV stoupá, zatímco NPV klesá s nárůstem prevalence onemocnění v populaci. Ale PPV je ovlivňováno prevalencí výrazněji než hodnoty NPV . [5]

1.2.3 Pravděpodobnostní poměr (LR)

Poměr pravděpodobnosti je velmi užitečným měřítkem diagnostické přesnosti. Je definován jako poměr očekávaného výsledku testu s určitým onemocněním vůči pacientům bez onemocnění. Jako takový LR přímo spojuje pre-testovou a post-testovou pravděpodobnost onemocnění u konkrétního pacienta. Jinými slovy, LR nám říká, kolikrát je větší pravděpodobnost, že test odhalí onemocnění, než pravděpodobnost, že test vyhodnotí pacienta jako zdravého. Jestliže jsou obě pravděpodobnosti stejné $LR = 1$.

PLR (*Positive Likelihood Ratio*) je poměr pravděpodobnosti pozitivních výsledků testů, tedy jak moc je pravděpodobnější, že bude výsledek testu pozitivní u nemocného pacienta než u zdravého pacienta. Vyjádřit se dá jako:

$$PLR = \frac{\text{senzitivita}}{1 - \text{specificita}} = \frac{TP \cdot (FP + TN)}{FP \cdot (TP + FN)} \quad (1.5)$$

PLR je významným indikátorem pro potvrzení diagnózy. Vyšší PLR svědčí o větší schopnosti odhalit nemocného pacienta. Dobré diagnostické testy přispívající ke spolehlivé diagnóze mají PLR větší než 10. [5]

Naopak NLR (*Negative Likelihood Ratio*) je definovaný jako:

$$NLR = \frac{1 - \text{specificita}}{\text{senzitivita}} = \frac{FN \cdot (FP + TN)}{TN \cdot (TP + FN)} \quad (1.6)$$

a určuje poměr pravděpodobnosti negativního výsledku testu u nemocných pacientů oproti pacientům bez onemocnění. Obvykle bývá menší než 1, protože je mnohem méně pravděpodobné, že k negativnímu výsledku dojde u pacienta s onemocněním, než u pacienta zdravého. [5, 6]

NLR je dobrým indikátorem pro vyvrácení diagnózy. Dobré testy mívají NLR menší než 0,1. Nižší hodnota NLR znamená, že čím nižší je NLR , tím signifikantnější je přínos testu pro vyvrácení diagnózy. [5]

Vzhledem k tomu, že jsou PLR a PLR definované pomocí senzitivity a specificity, nezávisí na prevalenci onemocnění v populaci a jsou tedy použitelné i v jiné klinické praxi, pokud se nezmění definice onemocnění. [5]

1.2.4 ROC křivka

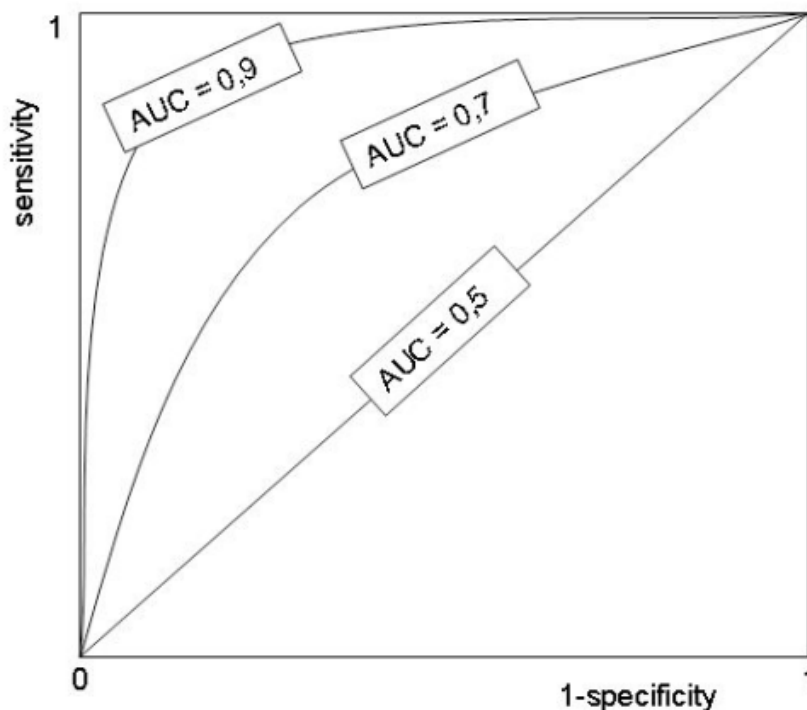
ROC (*Receiver Operating Characteristic*) křivka zobrazuje vztah mezi senzitivitou a specificitou pro různé *cut-off* hodnoty, kde specificita je vynášena na ose x a senzitivita na ose y . Sklon ROC křivky a plocha pod křivkou, která se označuje jako veličina AUC (*Area Under the Curve*), nám pomáhají odhadnout diskriminační sílu testu. Čím je křivka blíže levému hornímu rohu grafu a čím větší je plocha pod křivkou, tím vyšší je rozlišovací schopnost testu mezi nemocnými pacienty a pacienty bez onemocnění. Plocha pod křivkou může dosahovat hodnot mezi 0 a 1 a je ukazatelem kvality testu. Dokonalý diagnostický test má $AUC = 1$, naproti tomu málo diskriminující test se pohybuje kolem hodnot 0,5. [5] Obecně lze říci, že vztah mezi AUC a diagnostickou přesností platí, jak je popsáno v Tabulce 1.2:

Tabulka 1.2: Vztah mezi AUC a diagnostickou přesností [5]

AUC	Diagnostická přesnost
0,9–1,0	vynikající
0,8–0,9	velmi dobrá
0,7–0,8	dobrá
0,6–0,7	dostatečný
0,5–0,6	špatný
<0,5	neúčinný test

AUC je celosvětovým měřítkem diagnostické přesnosti. Ale neříká nám nic o jednotlivých parametrech, jako je senzitivita a specificita. Dva testy o stejné AUC mohou mít tyto parametry velmi rozdílné. AUC také nemá žádnou vypovídající hodnotu o přínosu pro potvrzení nebo vyvrácení diagnózy. Ale naopak srovnáním ploch pod křivkou lze odhadnout, který z použitých testů je vhodnější pro rozlišení zdravého subjektu od nemocného, nebo jiných podmínek zájmu. Je třeba upozornit, že tato srovnání by neměla být založena na vizuálním či intuitivním odhadu, ale na statistických testech vyhodnocující rozdíly mezi AUC na předem definované hladině významnosti. [5, 7]

Existují dva druhy ROC křivek, symetrická a asymetrická. Jejich použití závisí na variabilitě počtu pacientů s onemocněním a bez něj. ROC křivku lze popsat také parametrem Q^* , který značí bod protnutí ROC křivky s digonálou z levého horního rohu do pravého dolního rohu grafu. A podle definice je v tomto bodě hodnota senzitivity a specifity stejná. [6]



Obrázek 1.1: Příklady ROC křivek [5]

1.3 Meta-analýza DTA studií

Tato kapitola se zabývá rozdíly mezi meta-analýzou intervenčních studií a DTA (*Diagnostic Test Accuracy* – přesnost diagnostických testů) studii. Dále jsou v ní detailně popsány doporučené postupy při tvorbě meta-analýzy DTA studií, které jsou rozdělené na šest hlavních kroků.

1.3.1 Rozdíl mezi meta-analýzou intervenčních a DTA studií

Metody pro meta-analýzu intervenčních studií nemohou být snadno použity na DTA studie, protože k vyhodnocení přesnosti diagnostických testů je potřeba znalost dvou veličin a to specifity a senzitivity testu. Proto se v meta-analýze DTA studií používá statistické zpracování dvou rozdílných veličin na rozdíl od intervenčních studií, kde se hodnotí pouze jedna.

Meta-analýza DTA studií musí umožňovat porovnání specifity a senzitivity, ke kterému dochází u studií lišících se prahovou hodnotou definující pozitivitu a negativitu testu.

Ve výsledcích DTA studií je očekávána heterogenita, tudíž jsou požadovány náhodné efekty modelů k popisu variability mezi DTA studii. Hierarchické náhodné efekty modelů jsou vyžadovány u analýzy studií, které obsahují senzitivitu, specifitu a také heterogenitu v diagnostické přesnosti. [6]

1.3.2 Doporučené postupy pro tvorbu meta-analýzy DTA studií

Systematický přehled nebo meta-analýza přesnosti diagnostických testů má tři hlavní cíle:

1. získávat zobecňující souhrnné odhady přesnosti diagnostického testu,
2. identifikovat a poskytovat informace o faktorech, které ovlivňují DTA,
3. určovat oblasti pro další výzkum. [9]

Klíčovými procesy systematického přehledu a meta-analýzy lze rozdělit do šesti kroků, které jsou podrobně popsány níže.

Definování výzkumných otázek

Objasnění klíčových výzkumných otázek je nejdůležitějším krokem v celém procesu systematického přehledu. Důrazně se doporučuje, aby otázky byly tvořené podle strukturovaného rámce nazývaného formát PICO. [10] Je to akronym vzniklý spojením prvních písmen jednotlivých složek klinické otázky. Příčemž P (jako pacienti/účastníci/populace) znamená specifikování pacientů nebo populace, která má být zkoumána. Písmeno I jako intervence či indexový test, který bude hodnocen. Dále C jako komparátor, což je test nebo intervence, které mají být porovnány s indexovým testem. A nakonec O značí požadovaný výsledek (*outcomes*), jehož prostřednictvím bude intervence hodnocena. V některých případech se používají modifikované novější verze PICO(TS), které zahrnují všechny aspekty z PICO a navíc komponenty T (*time* - čas) a S (*setting* - prostředí).

Diagnostická přesnost se může lišit v závislosti na klinickém prostředí. Proto je důležité uvést klinické okolnosti nebo prostředí, ve kterém je diagnostický test použit. Vývoj kritérií pro určení, zda články v literatuře jsou vhodné pro systematický přehled a meta-analýzy, by se měl zaměřit na dva body: význam pro výzkumné otázky a metodická kvalita. Nejdůležitějším faktorem pro určení kritérií pro zahrnutí článků je shoda s výzkumnou otázkou. Zatímco kritéria pro vyloučení jsou založena na metodologické kvalitě. [8]

Systematické vyhledávání a výběr literatury

Účelem systematického vyhledávání je identifikovat co největší množství studií k danému tématu. Konečná strategie vyhledávání by měla být dostatečně podrobně popsána, aby bylo možné ji v případě potřeby reprodukovat. Obecný proces strategie vyhledávání zahrnuje tři fáze.

Fáze 1 *Rozvíjení vyhledávacích výrazů:*

- prohledat hlavní databáze, jako je například MEDLINE, EMBASE a další specifické databáze relevantní pro výzkumné otázky pomocí vhodných vyhledávacích termínů;
- identifikovat klíčové vyhledávací dotazy, které najdou studie relevantní pro výzkumné otázky;

- určit, které databáze by měly být prohledány;
- stanovit a zdokumentovat vyhledávací dotazy pro každou databázi.

Fáze 2 *Řízené vyhledávání*

- prohledat všechny databáze pomocí určených vyhledávacích výrazů;
- použití kritérií pro zařazení literatury, aby se určilo, který výzkum by měl být použit.

Fáze 3 *Rozšířené vyhledávání*

- hledat referenční seznamy a bibliografie všech zahrnutých studií pro nalezení dalších vhodných studií,
- hledat také v tzv. šedé literatuře, jako jsou disertace, abstrakty z konferencí, nebo výstupy z výzkumu vládního či soukromého sektoru, jsou-li k dispozici,
- konzultace s odborníky o řešeném problému. [8]

Hodnocení kvality studií

Strukturovaná metoda známá jako QUADAS (Quality Assessment of Diagnostic Accuracy Studies) slouží k posouzení metodologické kvality DTA studií. Tato metoda byla poprvé použita v roce 2003. [11] V roce 2011 byla metoda aktualizovaná na QUADAS-2 a byla obohacena o další nahromaděné důkazy, klinické zkušenosti, neoficiální zprávy a zpětné vazby od výzkumných pracovníků. [12] QUADAS-2 hodnotí kvalitu studií ve čtyřech oblastech:

- výběr pacientů,
- indexový test,
- referenční standard,
- *flow and timing* – průchod pacientů skrze studii (*flow*) a načasování indexového testu a referenčního standardu (*timing*). [12]

Každá oblast je posuzována z hlediska rizika zkreslení a první tři jsou také hodnoceny z hlediska obav ohledně použitelnosti. [12] Ke kategorizování rizika zkreslení se zavádějí tzv. signální otázky a dělí riziko na nízké, vysoké a nejasné. [12]

Extrakce dat a řízení

Při získávání dat se doporučuje využít strukturovanou extrakci dat, která obvykle obsahuje následující informace:

- charakteristiky studie – autoři, rok vydání, nemocnice nebo lékařské fakulty, rok nábory pacientů, velikost použitého vzorku, typy designu studií a diagnostická kritéria pro cílové onemocnění (referenční standard);

- demografické charakteristiky pacientů – věk, pohlaví;
- zobrazovací charakteristiky – typ zobrazujícího přístroje nebo techniky, zkušební protokoly;
- odhad přesnosti výsledků testů v diagnostické tabulce 2×2 , který obsahuje počet skutečně pozitivních (TP), falešných pozitiv (FP), falešně negativní výsledky (FN) a skutečně negativních (TN). Při použití diagnostické *cut-off* třeba uvést důvody pro výběr této hodnoty. Publikované studie často vykazují pouze závěrečný test parametrů přesnosti, jako je senzitivita a specificita, aniž by byly známy počty TP , FP , TN a FN . V těchto případech je nutné rekonstruovat 2×2 tabulku z dostupných informací, jako je např. celkový počet nemocných nebo celkový počet pacientů bez onemocnění. [8]

Analýza a syntéza dat

Syntéza dat z meta-analýzy DTA studií se liší od syntézy dat z terapeutických/intervenčních studií. Hlavním rozdílem je, že meta-analýza DTA studií se musí současně zabývat měřením specificity a senzitivity, zatímco terapeutická/intervenční meta-analýza se obvykle zabývá jediným měřeným parametrem. Proto musí meta-analýza DTA studií objasnit korelaci mezi specificitou a senzitivitou studie. [12]

Fáze 1. Deskriptivní statistika

První krok v syntéze dat zahrnuje výpočet deskriptivních statistik v každé primární studii. Tyto statistiky vycházejí ze senzitivity, specificity, PLR , NLR , a/nebo poměr diagnostických šancí (*diagnostic odds ratio*, tedy DOR) z diagnostických 2×2 tabulek jednotlivých studií.

Veličina DOR je dfinována jako:

$$DOR = \frac{LR(+)}{LR(-)} = \frac{\text{senzitivita} \cdot \text{specificita}}{(1 - \text{senzitivita}) \cdot (1 - \text{specificita})} = \frac{TP \cdot TN}{FP \cdot FN} \quad (1.7)$$

a popisuje, kolikrát jsou vyšší šance získat pozitivní výsledek testu na nemocné osobě než na osobě bez onemocnění. [8]

Výpočtem deskriptivní statistiky lze odvodit popisný *forest*-graf a tzv. SROC-graf. *Forest*-graf zobrazuje na ose x výpočty deskriptivní statistiky na intervalu spolehlivosti 95% a na ose y jsou vynášeny identifikátory studií. Graf je často vykreslen tak, aby velikost datových bodů odrážela velikost vzorku každé studie, což má za následek monotónní rostoucí nebo klesající vzor. SROC-graf naopak zobrazuje souhrn bodů jednotlivých primárních studií (tj. graf senzitivity vůči falešně pozitivním výsledkům a jeho 95% interval spolehlivosti v ROC prostoru). [8]

Fáze 2. Posouzení heterogenity

Heterogenita se vztahuje k variabilitě celé studie. Heterogenita může být výsledkem náhody, chyby v analytické metodologii, a/nebo rozdílů v designu studie, protokolu, kritérií pro zařazení a vyřazení, a diagnostické prahové hodnoty. Je-li mezi studiemi objevena heterogenita, použití meta-analytické sdružování studií již není smysluplné, a proto se nedoporučuje. V takových případech by měla být systematická kontrola omezena na kvalitativní deskriptivní analýzu vybrané literatury. [8]

Jedním z nejdůležitějších zdrojů heterogenity mezi DTA studií je prahový účinek. Pro definování pozitivních výsledků testů může studie hodnotící přesnost testu použít stejný test, ale jiný práh (např. diagnostickou *cut-off* hodnotu nebo kritérium positivity). Z tohoto důvodu se přesnost diagnostického testu opírá o práh použitý ve studii pro klasifikaci výsledku testu jako pozitivní nebo negativní. Změny prahové hodnoty zvyšují senzitivitu a snižují specifitu a naopak. Korelace pozorovaná mezi senzitivitou a specificitou se nazývá prahový účinek. Proto je při posuzování DTA studií třeba vzít v úvahu senzitivitu a specifitu společně. Z meta-analytického hlediska, pokud studie používají různá kritéria pro definování výsledků testů, jejich souhrnná data se budou lišit v závislosti na použitém prahu, a to by mohlo přispívat k pozorované heterogenitě mezi výsledky studie. [8]

Zda prahový efekt je či není zaveden, může být posouzeno několika způsoby. Za prvé, můžeme zkontrolovat kumulativní *forest*-graf. Kumulativní *forest*-graf např. senzitivity a specificity, zobrazuje graf, ve kterém je pořadí nezávislých studií vytvořeno podle jednoho z parametrů (např. v sestupném pořadí hodnot senzitivity). Je-li použit prahový efekt, bude se senzitivita a specifita měnit inverzním způsobem. Proto celkové uspořádání kumulativního *forest*-grafu bude mít tvar V nebo tvar obráceného V. [13] Statisticky řečeno, na základě posouzení lineární korelace mezi senzitivitou a falešnou pozitivitou (tj. $1 - \text{specificita}$, s použitím Spearmanovy srovnávací analýzy) lze prahový efekt uvedený v přehledu považovat za podstatný, pokud existuje významná korelace s korelačním koeficientem 0,6 nebo vyšším. [13] Vykreslení SROC grafu je další způsob, jak zkoumat prahový účinek. Existuje-li prahový účinek, budou body v grafu ukazovat celkově zakřivené rozložení nezávislých studií v ROC prostoru (z levého dolního rohu do pravého horního rohu) a zároveň také konvexní do levého horního rohu grafu. Pokud je přítomný prahový účinek (a v případě, že korelace je pozitivní), použití ROC křivky pro shrnutí diagnostické přesnosti je vhodné. [14] SROC křivku lze získat pomocí bivariačního modelu či hierarchického souhrnného ROC (HSROC) modelu. [8]

V meta-analýze terapeutických/intervenčních studií, je heterogenita obvykle hodnocena pomocí Cochran Q testu nebo Higginsovy I^2 statistiky. [14] I^2 statistika nebo Cochran Q test jsou tradičně považovány za ukazatele značné heterogenity mezi výsledky studií, pokud je hodnota p v Q testu nižší než 0,10 nebo 0,20, případně I^2 statistika vychází vyšší než 50 %. Tyto statistiky mohou být také použity k prokázání přítomnosti a rozsahu heterogenity mezi senzitivitou a specificitou DTA studií.

Nicméně, protože Cochran Q-test a Higginsova I^2 statistika byly vyvinuty pro hodnocení heterogenity jedné výstupní proměnné, nemusí být v DTA analýze samostatně informativní nebo smysluplné, na rozdíl od případu terapeutických/intervenčních studií, a proto by měly být interpretovány obezřetně, protože neberou v úvahu prahový účinek. [8]

Fáze 3. *Kvantitativní (Meta-Analýza) nebo kvalitativní DATA Syntéza*

Obecně platí, že meta-analýza DTA studií vyplývá ze statistického sdružování výsledků primárních studií s cílem vytvářet souhrnné body, jako například souhrn senzitivity, specificity, *DOR*, nebo souhrnné přímky (např. ROC křivky). DTA studie často vykazují větší heterogenitu než terapeutické/intervenční studie z důvodu existence prahového efektu. Mnohé studie hodnotící DTA jsou retrospektivně observační, které obvykle mají relativně nízkou úroveň důkazů a potenciálně vysoké riziko zaujatosti. Proto, na rozdíl od klinických studií v oblasti terapeutických/intervenčních studií, je často obtížné kontrolovat zkreslení v DTA studií. Pokud je heterogenita vysoká mezi DTA studii, pouhé statistické sdružování jednotlivých výsledků studií může poskytnout nesmyslné odhady. Aby nedošlo k dosažení nepřesného výsledku hodnocení, je třeba se vyhnout nerozváznému použití meta-analytických sdružování studií. Namísto toho možné příčiny studie heterogenity je třeba prozkoumat a vyjasnit, buď pomocí analýzy podskupin, nebo meta-regresní analýzy. Pokud existuje značná heterogenita mezi primárními studii, které je obtížné vysvětlit nebo upravit, nemělo by se provádět shromažďování studijních výsledků s použitím meta-analýzy, ale měl by jednoduše omezit systematický přehled na kvalitativní deskriptivní syntézu. [8]

Fáze 4. *Posouzení publikačního zkreslení*

Stejně jako meta-analýza terapeutických/intervenčních studií, slučování výsledků DTA studií může být předmětem několika zdrojů předpojatosti. [16] Kromě toho, některé formy předpojatosti jsou specifické pro diagnostické testovací studie [17], jako je například spektrum předpojatosti [18], diferenciální nebo dílčího ověření předpojatosti [19], a/nebo zkreslení v důsledku volby *cut-off* hodnoty. [20]

U většiny meta-analýz dochází k publikačnímu zkreslení v důsledku nezveřejňování statisticky nevýznamných výsledků, což vede k přehnanému sdruženému odhadu v systematickém přehledu. Proto je důležité posoudit publikační zkreslení u každé DTA meta-analýzy. U meta-analýzy terapeutických/intervenčních studií je publikační zkreslení hodnoceno pomocí trychtýřového grafu, Beggova testu nebo Eggerova testu. Trychtýřový graf je grafické zobrazení souhrnných údajů v grafu ve tvaru nálevky, kde se hodnota výsledků studie vztahuje k ose x , zatímco přesnost nebo velikost je vynesena v ose y . V meta-analýze DTA studií je preferován *DOR* jako jediný ukazatel diagnostické přesnosti v případech možného publikačního zkreslení. Trychtýřový graf nebo Deeksův graf a jeho test lze použít tam, kde se zobrazuje přirozený logaritmus *DOR* a převrácená hodnota druhé odmocniny efektivní hodnoty velikosti vzorku na osu x a y . [8]

Prezentace výsledků k uveřejnění

Správné vykazování výsledků systematického přehledu a meta-analýzy je důležitou fází. Proces výběru relevantní literatury by měly být ve formě vývojového diagramu, který je obvykle první volbou v každém systematickém přehledu a meta-analýzách. Charakteristiky zahrnutých studií by měly být podrobně popsány a shrnuty v tabulce, která obsahuje údaje o indexovém testu, o účastnících, designu studie a referenčních standardech pro každou ze studií. Měly by být také poskytnuty výsledky hodnocení kvality získané pomocí nástroje QUADAS-2. Protože prezentace výsledků meta-analýzy bude záviset na použitých metodách, měl by být stanoven jasný popis rozhodovacího procesu, který vedl k výběru vhodné metodiky. *Forest*-grafy s pomocí souhrnných bodů by měly nabídnout užitečné vizuální shrnutí, které lze snadno interpretovat. Uvedení *TP*, *FP*, *FN* a *TN* jednotlivých studií ve formě tabulky nebo ve *forest*-grafu. je také užitečné pro čtenáře a recenzenty při interpretaci výsledků studie. [8]

Je-li provedena SROC analýza, měla by být příslušná křivka SROC uváděna podle statistického modelu použitého v meta-analýze. Je-li to možné, měly by být poskytnuty všechny odhadované hodnoty parametrů spolu s jejich odpovídajícími intervaly věrohodnosti. Měly by být rovněž prezentovány výsledky posouzení heterogenity. Interpretace těchto výsledků by měla být jasná a je třeba brát v úvahu potenciální účinek vyvolaný předpojatostí či heterogenitou. Výklad by měl také počítat s důsledky falešně negativních nebo falešně pozitivních výsledků. Pokud jde o metodickou a interpretační perspektivu, mělo by být jasně uvedeno potenciální omezení přehledu. [8]

1.4 Roztroušená skleróza

Roztroušenou sklerózu mozkomíšní (RS) lze definovat jako chronické zánětlivé onemocnění centrálního nervového systému (CNS) autoimunitní etiologie způsobující demyelinizaci neuronů. Demyelinizace vede k ztrátě myelinu a k destrukci axonů. Jde o nejčastější neurologickou příčinu invalidity lidí v produktivním věku. Roztroušenou sklerózou jsou více postiženy ženy a to v poměru 2:1. Prevalence roztroušené sklerózy v České republice je 1/1000 obyvatel. [21]

Etiologie RS je multifaktoriální jako u ostatních autoimunitních chorob. Faktory mající vliv na vznik RS jsou buď vnitřní jako je polygenně podmíněná náchylnost k autoimunitě a hormonální vlivy, nebo faktory zevního prostředí jako infekce, stres, nedostatek vitamínu D aj. Také zeměpisná šířka ovlivňuje četnost výskytu RS a to tak, že četnost onemocnění se zvyšuje směrem od rovníku k pólům. [22]

U většiny pacientů (zhruba 85 %) je prvním projevem onemocnění náhle vzniklá neurologická symptomatika. Zda se jedná o demyelinizační zánět v CNS podporuje nález na magnetické rezonanci (MR) a průkaz oligoklonálních páسů v likvoru bez korelátora v séru. [22]

V patofyziologii RS největší roli hrají T-lymfocyty, makrofágy a B-lymfocyty. Aktivované lymfocyty pronikají do CNS přes hematoencefalickou bariéru, která je pro ně normálně nepropustná. Lymfocyty přes bariéru proniknou pomocí adhezivních molekul. V bílé hmotě CNS se poté aktivované lymfocyty podílejí na poškození myelinu. [22]

Náhlé vzplanutí nemoci se nazývá ataka, která se definuje jako vznik nových nebo zhoršení stávajících symptomů. Tento stav musí trvat alespoň 24 hodin, aby byl považován za ataku. Pokud jsou tyto příznaky spojené s akutním hořčnatým stavem při infekci, nejedná se o ataku. [22]

Průběh onemocnění má více fází. V první fázi hraje hlavní roli aktivovaný imunitní systém. Tato fáze je charakterizovaná výskytem atak, které většinou úplně odezní a nazývá se relaps-remitentní RS (RR RS). Po několika letech RR RS postupně přechází do sekundárně progresivní RS (SP RS), kdy dochází k vyhasínání zánětu a převládá neurodegenerace a nárůst neurologického deficitu. U přibližně 15 % pacientů převažuje od počátku onemocnění neurodegenerace nad zánětem a ataky nejsou přítomny, přičemž se pacient zhoršuje od počátku onemocnění. Tato forma se nazývá primárně progresivní RS (PP RS). [21, 22]

Akutní ataka musí být co nejrychleji přeléčena a to intravenózním podáním kortikoidů. Dlouhodobá léčba se zaměřuje na omezení zánětlivé aktivity, snížení počtu a intenzity relapsů a zpomalení progresu onemocnění. První volbou léků u RR RS jsou léky modifikující průběh choroby. Tyto preparáty se podávají především injekčně a jeden schválený preparát perorálně. Pokud pacient na terapii dostatečně nereaguje, anebo trpí rychle progredující formou, poté lze léčbu eskalovat a využívají se léky druhé volby. Pokud selžou léky první i druhé volby může být indikována intenzivní imunoablace a léčba cytostatiky. [21, 22]

1.5 Diagnostika RS pomocí magnetické rezonance

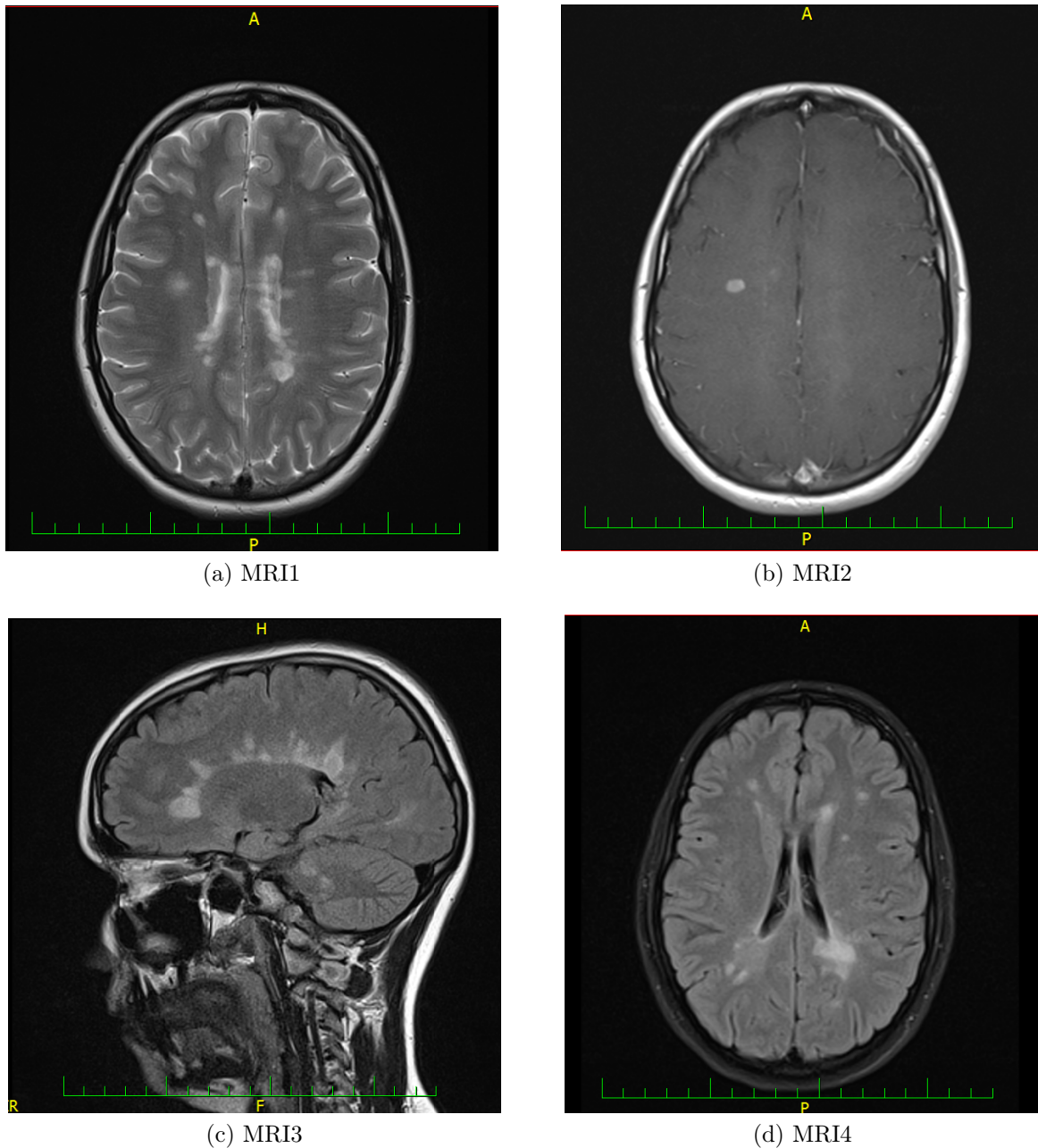
Magnetická rezonance (MR) pomáhá v diagnostice RS, stanovení prognózy i monitoraci průběhu nemoci a v současnosti představuje nejvíce využívaný biomarker. Nejnovější kritéria pro diagnostiku RS z roku 2011 umožňují diagnostikovat RS již po prvním vyšetření na MR. To vede k významnému zkrácení doby pro nasazení terapie. Nález na MR musí splňovat kritéria diseminace prosto v prostoru (DIS) a v čase (DIT). Za diseminaci v prostoru se považuje přítomnost alespoň 1 asymptomatické léze ve dvou typických lokacích ze 4 (juxtakortikálně, periventrikulárně, v oblasti zadní jámy nebo v míše). DIS byla v minulosti hodnocena na MR podle Barkhof–Tintoreových kritérií, která se vyznačovala vysokou specificitou i senzitivitou, ale byla hůře aplikovatelná. Z toho důvodu byla zavedena Swantonova kritéria, která byla postupně nahrazena častěji používanými McDonaldovými kritérii. Diseminace v čase je podle McDonaldových kritérií potvrzena přítomností enhancujících a neenhancujících ložisek v jednom MR vyšetření. Pokud kritéria nejsou splněna, je nutné provést s časovým odstupem další MR vyšetření. [23, 24]

Po zahájení terapie se přechází na tzv. monitorační protokol, který umožňuje v praxi vyhodnotit léze a vývoj atrofie. Magnetická rezonance byla zavedena do praxe v průběhu 80. let (v ČR v 90. letech) a má dominantní postavení ve všech těchto bodech - při stanovení diagnózy, odhadu prognózy a sledování aktivity onemocnění. [23]

Pro přítomnost RS jsou typické hyperintenzní ložiska na T2WI obrazech nebo sekvenci FLAIR (fluid attenuated inversion recovery = T2WI obraz s potlačením signálu likvoru. Dalším často se vyskytujícím a sledovaným parametrem u pacientů s RS na MR snímcích jsou ložiska na TWI obrazech barvící se po podání gadolinia (Gd), což je nejčastěji používaná kontrastní látka u RS. Vychytávání gadolinia ložisky je bezprostřední známkou aktivního zánětu, protože nově vzniklé ložisko má porušenou hemoencefalickou bariéru, kterou může kontrastní látka pronikat. Druhou velkou skupinou měření na MR, která umožňuje sledovat aktivitu RS, je měření mozkové atrofie, což je úbytek mozkové tkáně. [23]

Tabulka 1.3: Revidovaná McDonaldova diagnostická kritéria z roku 2010 [25]

Průkaz diseminace lézí v prostoru (DIS)	≥ 1 T2 léze nejméně ve dvou ze čtyř oblastí CNS; periventrikulární, juxtakortikální, infratenoriální nebo míšní. Pro demonstraci diseminace v prostoru není nutné, aby některá z lézí vychytávala gadolinium. Jestliže má pacient klinicky kmenový nebo míšní syndrom, symptomatické léze se nezapočítávají do počtu lézí.
Průkaz diseminace lézí v čase (DIT)	Nová T2 a/nebo gadolinium vychytávající léze na další MRI oproti prvnímu MRI skenu bez ohledu na načasování prvního skenu nebo současná přítomnost asymptomatických gadolinium vychytávajících lézí a nevychytávajících lézí v jakoukoli dobu.



Obrázek 1.2: Na obrázku 1.2a je zobrazen T2 snímek mozku 20 leté pacientky s rozsáhlým demyelinizačním nálezem odpovídajícím RS, na snímku 1.2b je zachycena aktivní léze téže pacientky na T1 snímku po podání kontrastní látky, obrázek 1.2c představuje T2 FLAIR snímek, kde jsou vidět tzv. Dawsonovy prsty typické pro diagnózu RS a obrázek 1.2d ukazuje T2 FLAIR snímek mozku dotyčné pacientky s RS.[autor]

Kapitola 2

Metody

Při tvorbě meta-analýzy bylo postupováno podle doporučených postupů pro tvorbu meta-analýzy DTA studií. Nejprve byla definována výzkumná otázka, podle které byla vybrána vhodná literatura. Studie vyhovující výzkumné otázce byly podrobeny hodnocení metodologické kvality QUADAS 2. A z těchto byla extrahována data, která byla následně zpracována v programu OpenMeta[Analyst], který umožňuje získání přehledných *forest*-grafů, a také v programu Meta-DiSc, jehož prostřednictvím lze generovat SROC křivky a informace o Spearmanově korelačním koeficientu.

2.1 Definice výzkumné otázky

Výzkumná otázka byla definována podle strukturovaného rámce PICO. Kde jako P (pacienti) byli zvoleni pacienti s podezřením na diagnózu roztroušenou sklerózu, pacienti již se stanovenou diagnózou a také pacienti s klinicky izolovaným syndromem. Indexový test (I) představovalo vyšetření pomocí MR. Jako komparátor (C) byla zvolena kritéria pro diagnózu RS. Výsledkem neboli *outcome* (O) byla stanovena diagnostická přesnost magnetické rezonance. Díky PICO byla výzkumná otázka meta-analýzy formulována jako stanovení diagnostické přesnosti magnetické rezonance při diagnóze roztroušené sklerózy.

2.2 Výběr studií

Systematické vyhledávání relevantní literatury probíhalo ve třech fázích, které postupně zvyšovaly počet vhodných studií pro tvorbu meta-analýzy.

Fáze 1: Rozvíjení vyhledávacích výrazů:

Primárně byly vhodné studie hledány pomocí databáze *Web of Science* a dále byly využity databáze *Medline* a *ScienceDirect*. Jako klíčová slova byla nejprve zvolena roztroušená skleróza (*Multiple Sclerosis*) a zároveň magnetická rezonance (*Magnetic Resonance*

Imaging, MRI). Postupně byla přidána specifická/senzitivita (*specificity/sensitivity*) jako další klíčové slovo.

Fáze 2: Řízené vyhledávání

Při screeningu literatury byly vytvořeny kritéria pro zahrnutí studií. Prvním kritériem bylo časové omezení a to na interval 2008–2017. Dalším kritériem bylo použití MR jako indexový test. Dále bylo důležité, aby studie obsahovaly dostatečná data pro tvorbu 2×2 tabulky nebo přímo obsahovaly údaje o specifitě a senzitivitě.

Fáze 3: Rozšířené vyhledávání

Ve třetí fázi bylo vyhledávání rozšířeno díky prohledávání seznamu referencí z vhodných studií a tím se zvýšil počet studií, které lze použít. Dále bylo využito doporučení databází podobných studií při nálezů relevantní studie.

2.3 PRISMA diagram

Prohlášení PRISMA bylo zveřejněno v roce 2009. Obsahuje kontrolní seznam a vývojový diagram. Pro účely této práce byl využit diagram PRISMA. PRISMA diagram zobrazuje tok informací v různých fázích systematického přehledu. Mapuje počet zaznamenaných, zahrnutých a vyloučených záznamů. PRISMA diagram se dělí na čtyři fáze, které byly pojmenovány: identifikace, *screening*, vhodnost a zahrnutí. [26]

2.4 Hodnocení metodologické kvality studií

Relevantní studie byly následně podrobeny hodnocení metodologické kvality DTA studií QUADAS a to konkrétně aktualizovanou verzí QUADAS 2. Všechny studie byly klasifikovány podle oficiálního formuláře QUADAS 2. Výsledky byly zaneseny do tabulky, kde bylo přehledně zaznamenáno riziko zkreslení čtyř aspektů studií a to výběr pacientů, indexový test, referenční standard a *flow and timing*. Riziko zkreslení bylo rozděleno na nízké, vysoké a nejasné. Kromě zkreslení byly u studií posuzovány obavy z hlediska použitelnosti, ale pouze u tří aspektů. Pro větší přehlednost byl vytvořen graf podle výsledků z tabulky. Tento graf udává počet studií s nízkým, vysokým nebo nejasným rizikem zkreslení a obavami týkajícími se použitelnosti.

Pro klasifikaci studií podle nástroje QUADAS 2 pro posouzení rizika zkreslení a použitelnosti bylo postupováno podle souboru následujících klasifikačních otázek:

DOMÉNA 1: Výběr pacientů

A Riziko zkreslení

Popište indexový test a jeho provedení a interpretaci:

- Byly interpretovány výsledky indexových testů bez znalosti výsledků referenčního standardu? [ANO/NE/Nejasné]
- Byla-li použita prahová hodnota, byla předem specifikována? [ANO/NE/Nejasné]
- Mohlo by provádění nebo interpretace indexového testu zavést zkreslení? [Riziko: Nízké/Vysoké/Nejasné]

B Obavy týkající se použitelnosti

- Existuje obava, že se indexová zkouška, její provedení nebo výklad neodpovídá otázce systematického přehledu? [Obavy: Nízké/Vysoké/Nejasné]

DOMÉNA 2: Indexové testy (pokud byl použit více než 1 indexový test, vyplňte pro každý test)

A Riziko zkreslení

Popište indexový test a jeho provedení a interpretaci:

- Byly interpretovány výsledky indexových testů bez znalosti výsledků referenčního standardu? [ANO/NE/Nejasné]
- Byla-li použita prahová hodnota, byla předem specifikována? [ANO/NE/Nejasné]
- Mohlo by provádění nebo interpretace indexového testu zavést zkreslení? [Riziko: Nízké/Vysoké/Nejasné]

B Obavy týkající se použitelnosti

- Existuje obava, že se indexová zkouška, její provedení nebo výklad neodpovídá otázce systematického přehledu? [Obavy: Nízké/Vysoké/Nejasné]

DOMÉNA 3: Referenční standard

A Riziko zkreslení

Popište referenční standard a jak byl proveden a interpretován:

- Je pravděpodobné, že referenční standard správně klasifikuje cílový stav? [ANO/NE/Nejasné]
- Byly interpretovány výsledky referenčního standardu bez znalosti výsledků indexových testů? [ANO/NE/Nejasné]
- Mohlo by provádění nebo interpretace referenčního standardu zavést zkreslení? [Riziko: Nízké/Vysoké/Nejasné]

B Obavy týkající se použitelnosti

- Existuje obava, že cílový stav, jak je definován v referenčním standardu, neodpovídá otázce systematického přehledu? [Obavy: Nízké/Vysoké/Nejasné]

DOMÉNA 4: *Flow and timing*

A Riziko zkreslení

Popište všechny pacienty, kteří nepodstoupili indexový test a/nebo referenční standard nebo kteří byli vyloučeni z tabulky 2×2 :

Popište časový interval a jakékoliv zásahy mezi indexovým testem a referenčním standardem:

- Existoval vhodný interval mezi indexovými testy a referenčními standardy? [ANO/NE/Nejasné]
- Podstoupili všichni pacienti referenční standard? [ANO/NE/Nejasné]
- Podstoupili pacienti ten samý referenční standard? [ANO/NE/Nejasné]
- Byli všichni pacienti zařazeni do analýzy? [ANO/NE/Nejasné]
- Mohl by *flow* pacientů zapříčinit zkreslení? [Riziko: Nízké/Vysoké/Nejasné] [12]

2.5 Extrakce dat

Ze studií byla získána data, která byla následně zaznamenána do tabulky. Mezi extrahovaná data patřila jména autorů, rok publikování, název studie a počet pacientů. Nejdůležitější bylo získat ze studií dostatek dat pro tvorbu 2×2 tabulky, ze které bylo dále zjištěno TP , TN , FP a FN . Ve většině studií místo těchto hodnot byla publikována již vypočtená specifická a senzitivita. U těchto studií bylo TP , TN , FP a FN zpětně určeno podle následujících vztahů.

Tabulka 2.1: Výpočet dat v 2×2 tabulce senzitivity a specifacity

	subjekty s nemocí	subjekty bez nemoci
pozitivní	$y_i(TP)$	$z_i(FP)$
negativní	$m_i - y_i(FN)$	$n_i - z_i(TN)$
celkem	m_i	n_i

Pro výpočet senzitivity a specifacity pak platí:

$$p_i = \text{senzitivita } i\text{-té studie} = \frac{y_i}{m_i}, \quad (2.1)$$

$$u_i = \text{falešně pozitivní poměr } i\text{-té studie} = \frac{z_i}{n_i}, \quad (2.2)$$

$$1 - u_i = \text{specifická } i\text{-té studie} = \frac{n_i - z_i}{n_i}. \quad (2.3)$$

Souhrnná data pro následnou meta-analýzu jsou uvedena v tabulce (3.3).

2.6 Analýza dat

Pro statistické zpracování dat byl použit program OpenMeta[Analyst], což je software s otevřeným zdrojovým kódem, takže je volně stažitelný i pro širokou veřejnost. Jedná se o program k provádění meta-analýz binárních, kontinuálních nebo diagnostických dat pomocí různých metod pevných a náhodných efektů včetně Bayesovy analýzy. Do tohoto programu byla zanesena data ze studií a to konkrétně autor studie, rok, TN , TP , FP a FN . Z těchto údajů program vypočetl specificitu, senzitivitu, NLR , PLR , DOR intervaly spolehlivosti a vygeneroval *forest*-grafy nezbytné ke grafickému zobrazení výsledků. [27]

Dále byl využit program Meta-DiSc, který sloužil ke kontrole výsledků, ke grafickému zobrazení SROC křivek a k výpočtu plochy pod SROC křivkou (AUC) a také k zjištění Spearmanovo koeficientu korelace. Meta-DiSc je volně přístupný software k tvorbě meta-analýz studií, které hodnotí diagnostické a screeningové testy. Program umožňuje zadávat data přímo nebo je lze importovat z textových souborů. [28]

Kapitola 3

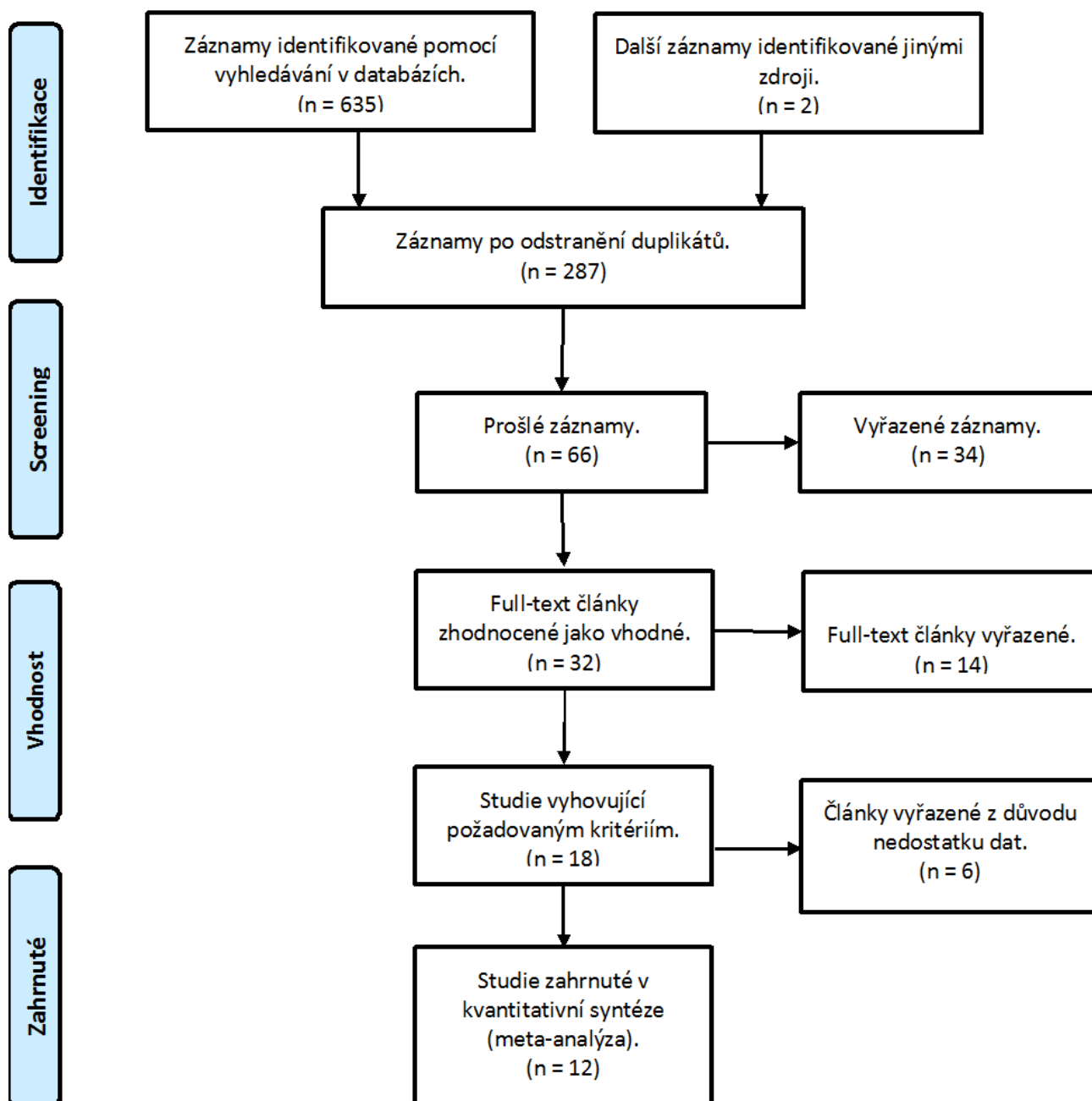
Výsledky

V následující kapitole jsou popsány jednotlivé výsledky. Nejprve je zde uveden postup výběru literatury, který je graficky znázorněn pomocí PRISMA diagramu. Relevantní studie jsou zaneseny v tabulce společně s údaji, které z článků vyplývají. Vhodné studie byly podrobeny hodnocení kvality a výsledky jsou popsány pomocí grafu a tabulky. V další části jsou výsledky meta-analýzy zobrazeny pomocí *forest*-grafů a souhrnné tabulky jednotlivých veličin. Kompletní shrnutí výsledků, jejich interpretace a význam jsou rozebrány v diskuzi.

3.1 Výběr studií

Postup selekce literatury byl znázorněn v PRISMA diagramu (viz obrázek 3.1) a rozdělen do čtyř fází a to do: identifikace, *screening*, vhodnosti a zahrnutí. Ve fázi identifikace bylo zjištěno 635 studií pomocí vyhledávání v databázích a 2 studie z dalších zdrojů. Po odstranění literatury, která se vyskytovala ve více databázích zbylo 287 studií. Po konkretizování klíčových slov zůstalo 63 studií, u kterých byl pročten abstrakt. Po *screeningu* abstraktů bylo zařazeno 32 studií, které splňovaly kritéria pro zařazení a byly dostupné ve full-textu. 18 studií splnilo kritéria pro zařazení. Z těchto článků bylo 6 vyřazeno, protože sice obsahovaly již vypočtenou specifitu a senzitivitu, ale nebylo možné v nich vyhledat dostatek údajů pro 2×2 tabulku. Ta byla ovšem nezbytná pro tvorbu meta-analýzy v programech OpenMeta[Analyst] a Meta-DiSc. Meta-analýza byla tedy finálně vytvořena celkem z 12 studií.

Po selekci relevantních studií (12) byla vytvořena tabulka (viz tabulka 3.1) obsahující názvy všech vhodných studií, rok publikování, hodnoty specifity a senzitivity uvedené ve studiích a počet probandů. V tabulce jsou uvedena kritéria, ke kterým byla vztažena specifita a senzitivita. Některé studie obsahovaly více než jedno kritérium, proto výsledný počet zahrnutých hodnot specifity/senzitivity je 23.



Obrázek 3.1: Prisma diagram

Tabulka 3.1: Studie rozřazené podle kritérií výběru pacientů

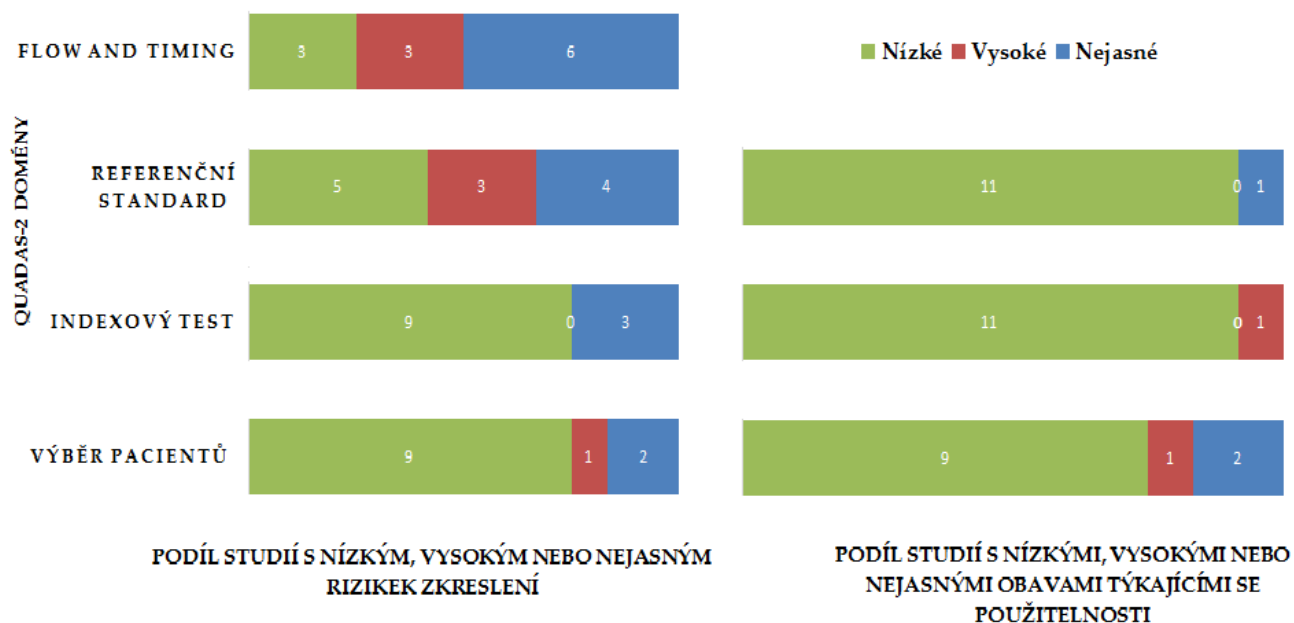
Název studie	Rok studie	Kritéria	Specifcita	Sensitivita	Počet pacientů
Gray matter MRI differentiates neuromyelitis optica from multiple sclerosis using random forest [29]	2016	x	0,72	0,77	144
Clinical 3-tesla FLAIR* MRI improves diagnostic accuracy in multiple sclerosis [30]	2016	T2-FLAIR FLAIR* T2-FLAIR + FLAIR*	0,88 0,87 0,73	0,91 0,90 0,98	68
Diagnostic Efficacy of Conventional MRI Pulse Sequences in the Detection of Lesions Causing Internuclear Ophthalmoplegia in Multiple Sclerosis Patients [31]	2014	T2 PD FLAIR	0,75 0,55 0,87	0,77 0,79 0,63	15
Accuracy of CSF and MRI criteria for dissemination in space in the diagnosis of multiple sclerosis [32]	2007	Barkhfova kritéria	0,88	0,74	58
Early MRI in optic neuritis: the risk for clinically definite multiple sclerosis [33]	2010	x	0,24	0,86	142
Magnetic Resonance Imaging Predictors of Conversion to Multiple Sclerosis in the BENEFIT Study [34]	2009	Barkhof–Tintoreova kritéria	0,36	0,64	468
Prediction of conversion from CIS to CDMS according to baseline MRI findings: comparison of revised McDonald criteria and Swanton modified criteria [35]	2009	McDonaldova kritéria Swantonova kritéria	0,53 0,60	1,00 1,00	64
Evaluation of the 2005 McDonald MRI criteria for dissemination in space in Afro-Caribbean patients with clinically isolated syndromes [36]	2009	McDonaldova kritéria 2001 McDonaldova kritéria 2005 Swantonova kritéria	0,68 0,66 0,63	0,36 0,39 0,39	66
Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis [37]	2008	Poserova kritéria	0,68	0,90	107
Accuracy of MRI criteria for dissemination in space for the diagnosis of multiple sclerosis in patients with clinically isolated syndromes [38]	2010	Barkhof–Tintoreova kritéria Swantonova kritéria	0,77 0,68	0,72 0,91	79
A search for new MRI criteria for dissemination in space in subjects with a clinically isolated syndrome [39]	2009	McDonaldova kritéria Swantonova kritéria	0,79 0,70	0,49 0,64	349
A Single, Early Magnetic Resonance Imaging Study in the Diagnosis of Multiple Sclerosis [40]	2009	Barkhof–Tintoreova kritéria Swantonova kritéria Frohmanova kritéria	0,67 0,60 0,50	0,64 0,77 0,84	250

3.2 QUADAS 2

Celkem dvanáct studií bylo podrobena hodnocení metodologické kvality QUADAS 2. Kvalita studií byla hodnocena ve čtyřech doménách a jak z hlediska zkraslení tak i z hlediska obav týkajících se použitelnosti. Z výsledků hodnocení byla vytvořena tabulka, kde bylo graficky rozlišeno riziko a obavy týkající se použitelnosti na nízké, vysoké a nejasné. Graf přehledně zobrazuje četnost studií s nízkým, vysokým a nejasným rizikem zkraslení a obavami týkajícími se použitelnosti.

V doméně *flow and timing* vykazovaly tři studie nízké riziko zkraslení, šest studií riziko nejasné a u tří studií hrozilo vysoké riziko zkraslení. Riziko zkraslení u referenčního standardu bylo nízké u 5 studií, nejasné u čtyř a vysoké u tří článků. Tři studie měly nejasné riziko zkraslení u indexového testu a devět studií riziko nízké. U indexového testu nehrozilo ani u jedné studie vysoké riziko zkraslení. Výběr pacientů mohl být zkraslen u jedné studie a to s vysokým rizikem. Nízké riziko u výběru pacientů vykazovalo devět studií a zbylé dvě měly nejasné riziko zkraslení u výběru pacientů.

Obavy týkající se použitelnosti byly hodnoceny ve třech doménách (vyloučení *flow and timing*). U referenčního standardu byly obavy nízké celkem u 11 studií a nejasné u zbývajících jedné. Obavy týkající se použitelnosti indexového testu byly nízké u 11 článků u jedné studie vysoké. U výběru pacientů obavy se nízké obavy vyskytovaly u 9, vysoké u 1 a nejasné u 2 studií.



Obrázek 3.2: QUADAS 2 graf

Tabulka 3.2: **QUADAS 2.** Symbol + značí nízké riziko, ? znamená nejasné riziko a - symbolizuje riziko vysoké.

Studie	Riziko zkreslení				Obavy týkající se použitelnosti		
	Výběr pacientů	Indexový test	Referenční standard	Flow and timing	Výběr pacientů	Indexový test	Referenční standard
ESHAGHI, A. [29]	+	+	?	?	?	+	+
GEORGE, ILENA C. [30]	+	+	-	?	+	+	+
MCNULTY, J. P. [31]	-	+	+	?	-	+	+
VILLAR, L. [32]	+	+	-	-	+	-	+
SWANTON, J. [33]	+	+	?	?	+	+	?
MORAAL, B. [34]	+	+	+	+	+	+	+
LO, CHUNG [35]	?	+	?	-	+	+	+
FISNIKU, L. [37]	+	+	-	+	+	+	+
DIAZ-SÁNCHEZ [38]	+	+	+	-	+	+	+
KORFEWEG, T. [39]	+	?	-	?	+	+	+
ROVIRA, A. [40]	+	?	+	?	+	+	+
CHAUSSEON, N. [36]	?	?	+	+	?	+	+

3.3 Meta-analýza

Tabulka 3.3: Použitá data jednotlivých studií

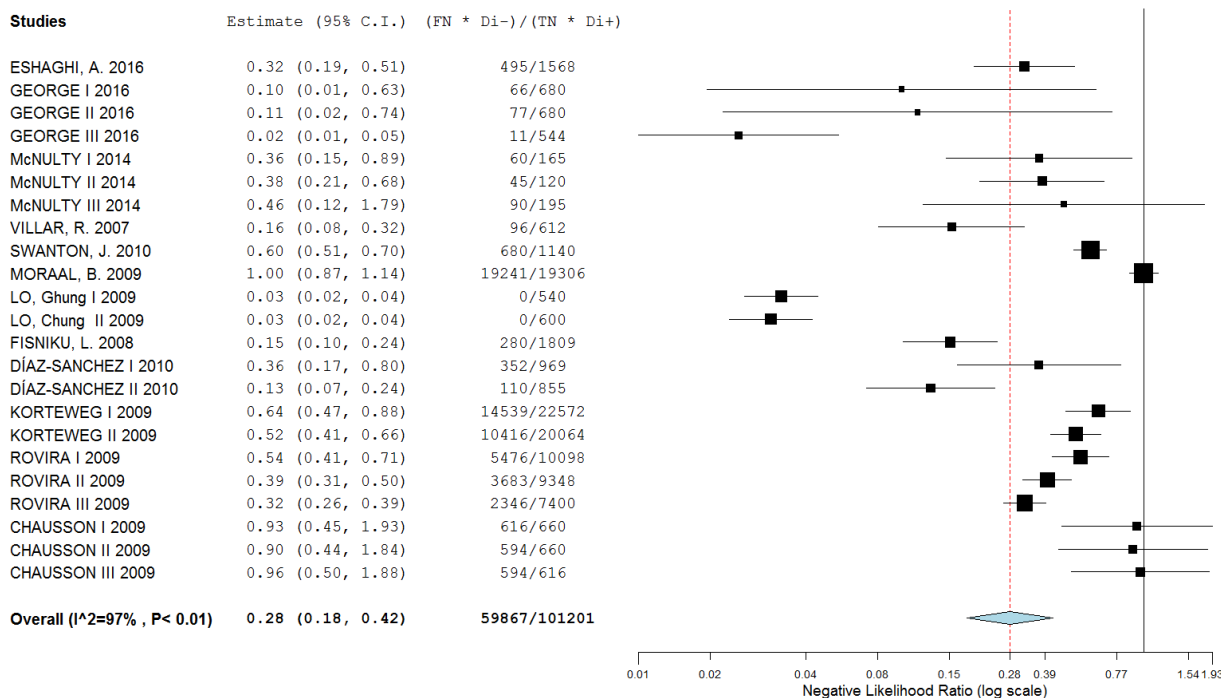
Autor	y_i (<i>TP</i>)	z_i (<i>FP</i>)	$m - y_i$ (<i>FN</i>)	$n - z_i$ (<i>TN</i>)	m	n	<i>Senz.</i>	<i>Spec.</i>
ESHAGHI, A. [29]	38	13	11	32	49	45	0.77	0.72
GEORGE, ILENA C. [30]	62	1	6	10	68	11	0.91	0.88
	61	1	7	10	68	11	0.90	0.87
	67	3	1	8	68	11	0.98	0.73
MCNULTY, J. P [31]	11	4	4	11	15	15	0.77	0.75
	12	7	3	8	15	15	0.79	0.55
	9	2	6	13	15	15	0.63	0.87
VILLAR, L. [32]	30	6	4	18	34	24	0.88	0.74
SWANTON, J. [33]	49	65	8	20	57	85	0.86	0.24
MORAAL, B. [34]	126	173	71	98	197	271	0.64	0.36
LO, CHUNG [35]	30	16	0	18	30	34	1.00	0.53
	30	14	0	20	30	34	1.00	0.60
FISNIKU, L. [37]	60	13	7	27	67	40	0.90	0.68
DÍAZ-SÁNCHEZ [38]	41	5	16	17	57	22	0.72	0.77
	52	7	5	15	57	22	0.91	0.68
KORTEWEG, T. [39]	65	46	67	171	132	217	0.49	0.79
	84	65	48	152	132	217	0.64	0.70
ROVIRA, A. [40]	65	49	37	99	102	148	0.64	0.67
	94	51	29	76	123	127	0.77	0.60
	125	52	23	50	148	102	0.84	0.50
CHAUSSON, N. [36]	16	7	28	15	44	22	0.36	0.68
	17	7	27	15	44	22	0.39	0.66
	17	8	27	14	44	22	0.39	0.63

V tabulce 3.3 jsou uvedeny jednotlivé studie reprezentované jménem autora. Jsou zde uvedeny hodnoty *TP*, *TF*, *FP* a *FN*, které byly buď zjištěny ze studií, nebo dopočteny podle hodnoty specifity a senzitivity. Další sloupce tabulky představuje množství pacientů s onemocněním m a bez něj n pro jednotlivé studie. Výchozími hodnotami pro následnou analýzu jsou specifity a senzitivity jednotlivých článků.

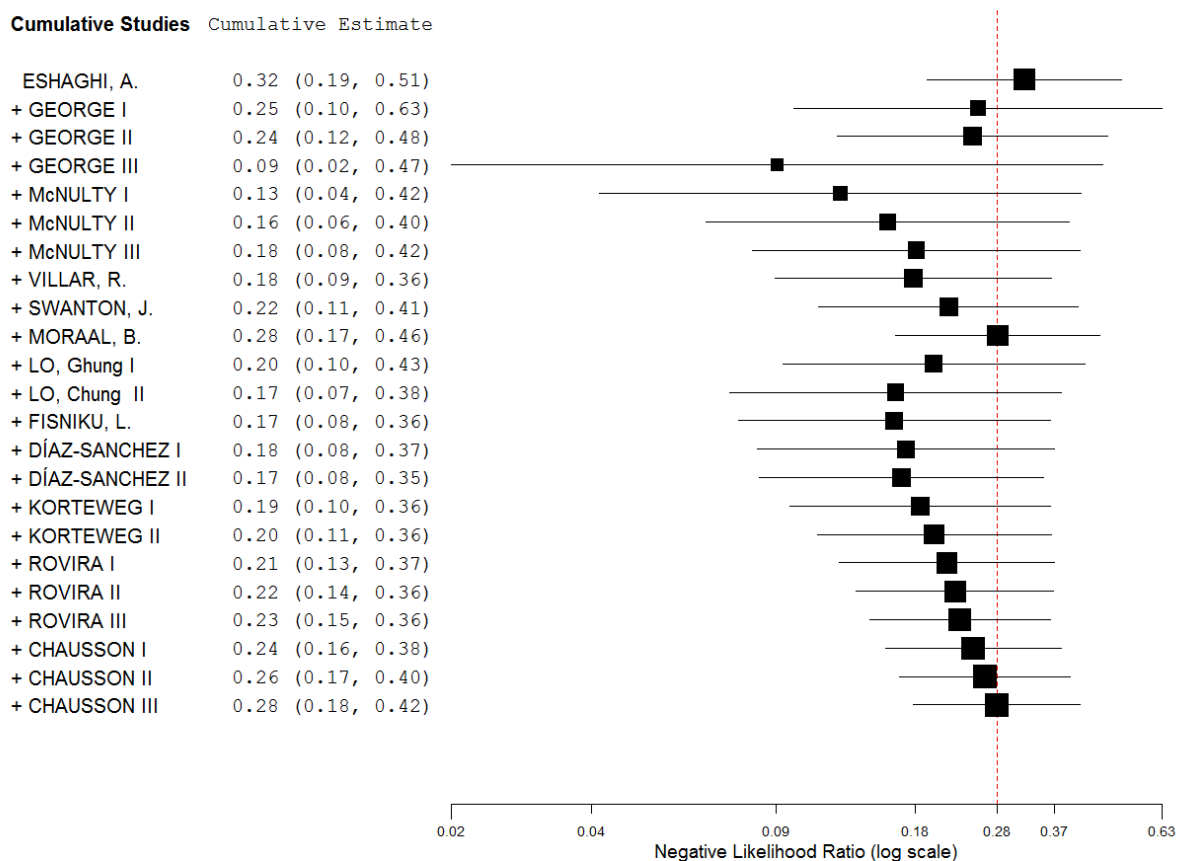
Negativní pravděpodobnostní poměr všech studií vyšel 0,28 s dolní mezí 0,18 a s horní mezí 0,42. Hodnota I^2 pro NLR dosahovala 97 % a p -hodnota heterogenity byla menší než 0,01. Souhrnná hodnota PLR byla 2,00 v intervalu (1,65; 2,42). Higinsova I^2 statistika pro PLR činí 82 % s p -hodnotou menší než 0,01. Souhrnná senzitivita studií dosahovala hodnoty 0,75 v intervalu (0,68; 0,82) s I^2 88 % při p -hodnotě menší než 0,01. Souhrnná specifická byla stanovena na 0,65 s horní mezí 0,57 a dolní mezí 0,72. I^2 pro specificku byla 87 % s p -hodnotou menší než 0,01. Agregovaný výsledek DOR byl určen na 6,15 v intervalu (3,88; 9,74) a hodnota I^2 byla 83 % s p -hodnotou menší než 0,01. Spearmanův korelační koeficient pro tuto meta-analýzu činí 0,68.

Tabulka 3.4: Výsledné hodnoty NLR , PLR , senzitivity, specificku a DOR . Horní a dolní mez značí 95% interval spolehlivosti.

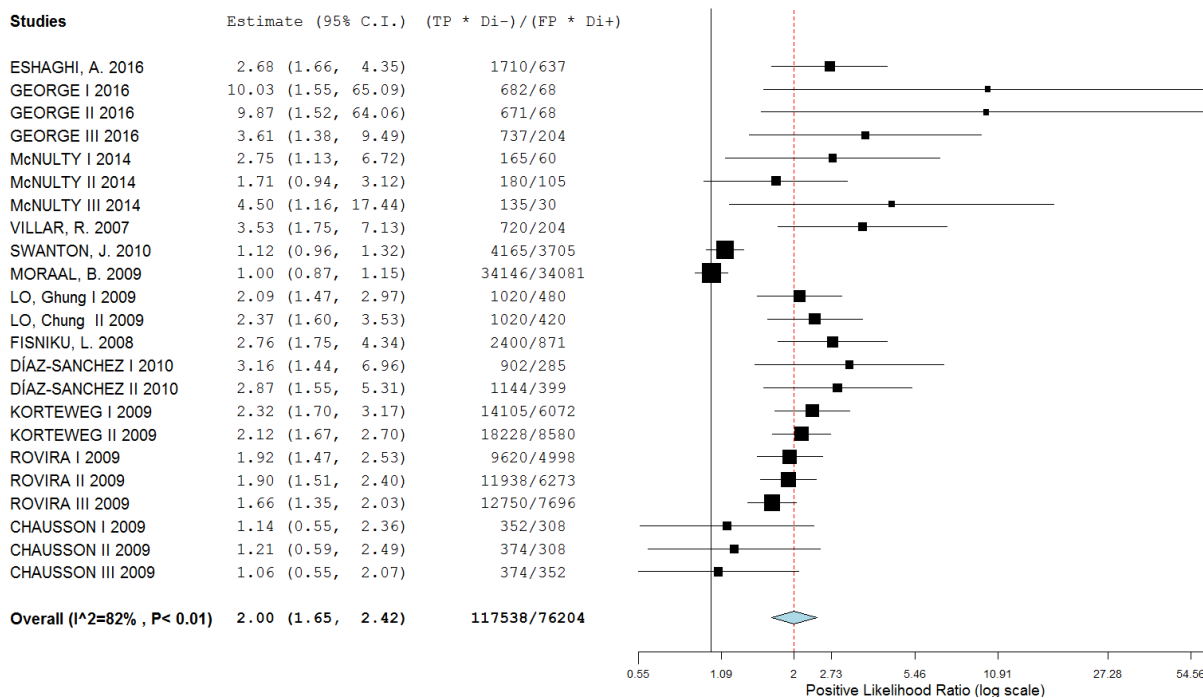
	souhrnná hodnota	dolní mez	horní mez	p -hodnota	p -hodnota heterogenity	I^2
NLR	0,28	0,18	0,42	< 0,01	< 0,01	97 %
PLR	2,00	1,65	2,42	< 0,01	< 0,01	82 %
Senzitivita	0,76	0,68	0,82	< 0,01	< 0,01	88 %
Specifická	0,65	0,57	0,72	< 0,01	< 0,01	87 %
DOR	6,15	3,88	9,74	< 0,01	< 0,01	83 %

3.3.1 *NLR*

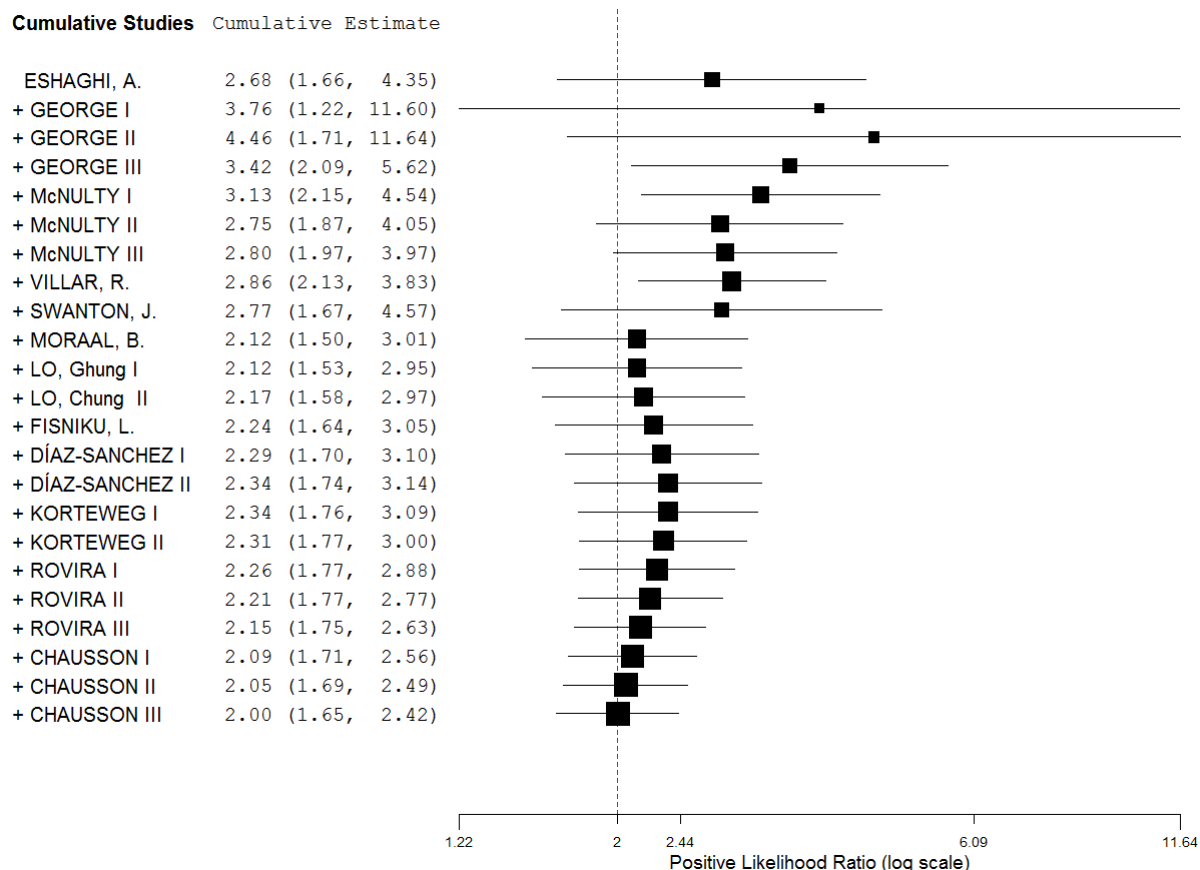
Obrázek 3.3: *NLR forest-graf*. Graf zobrazuje *forest-graf* negativního pravděpodobnostního poměru zahrnutých studií s 95% intervalem spolehlivosti. Na ose x jsou vyneseny hodnoty *NLR* v logaritmickém měřítku. Osa y odpovídá jednotlivým studiím. Červená kolmá osa vyznačena čárkovaně udává souhrnný *NLR* pro všechny studie, který nabývá hodnoty 0,28 s dolním odhadem 0,18 a horním odhadem 0,42 na daném intervalu spolehlivosti. Černá kolmice značí hodnotu 1. *NLR* studií se pohybuje od 0,02 (George III) do 1,00 (Moraal, B.). Největší rozptyl mají studie George I–III. Grafické znázornění *NLR* pro jednotlivé studie ve *forest-grafu* se liší velikostí v závislosti na počtu probandů. Higginsův I^2 test zde vychází 97 % s p -hodnotou $< 0,01$.



Obrázek 3.4: *NLR* kumulativní *forest*-graf. Graf zobrazuje kumulovaně propočítané průměry *NLR* ze všech předcházejících studií s 95% intervalem spolehlivosti. Všechny konfidenční intervaly studií se protínají s agregovanou hodnotou $NLR = 0,28$, proto žádná ze studií se významně statisticky neliší od agregované hodnoty *NLR* [2].

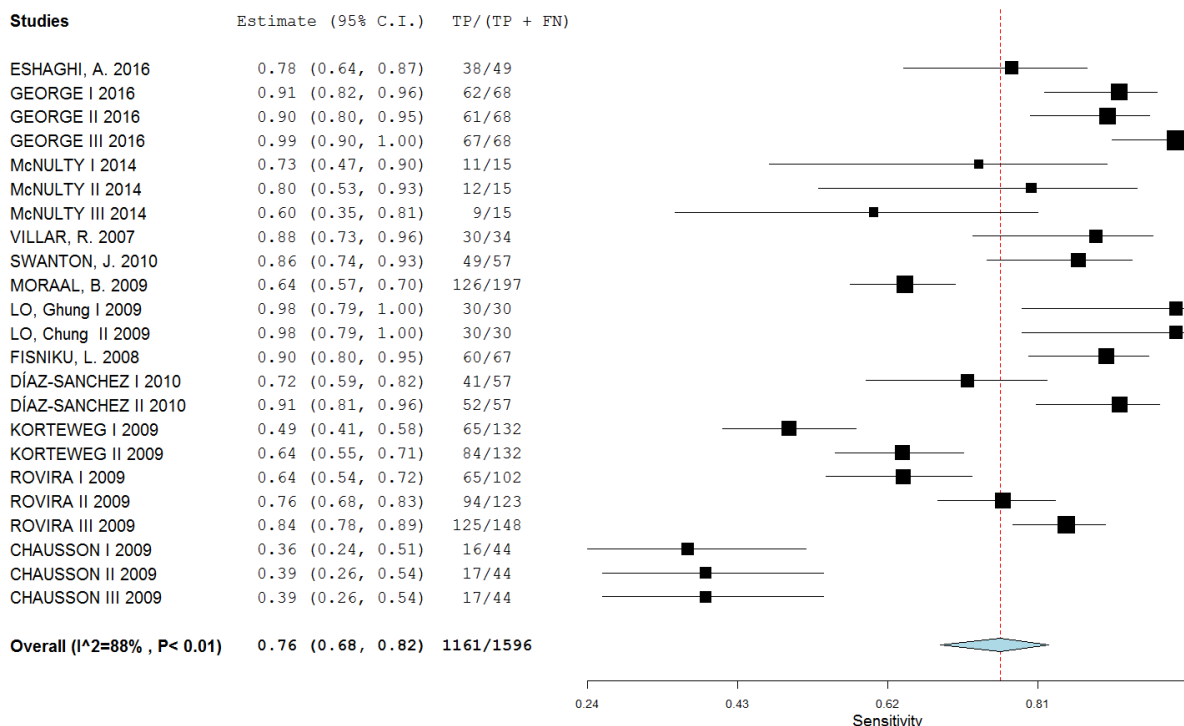
3.3.2 *PLR*

Obrázek 3.5: *PLR forest-graf*. Graf zobrazuje *forest-graf* pozitivního pravděpodobnostního poměru zahrnutých studií s 95% intervalem spolehlivosti. Na ose *x* jsou vyneseny hodnoty *PLR* v logaritmickém měřítku. Osa *y* odpovídá jednotlivým studiím. Červená kolmá osa vyznačená čárkovaně udává souhrnný *PLR* pro všechny studie, který nabývá hodnoty 2,00 s dolním odhadem 1,65 a horním odhadem 2,42 na daném intervalu spolehlivosti. Černá kolmice značí hodnotu 1. *PLR* studií se pohybuje od 1,00 (Moraal, B.) do 10,03 (George I). Největší rozptyl je na grafu vidět u George I–II a McNulty III. Grafické znázornění *PLR* pro jednotlivé studie ve *forest-grafu* se liší velikostí v závislosti na počtu probandů. Higginsův I^2 test zde vychází 82 % s *p*-hodnotou < 0,01.

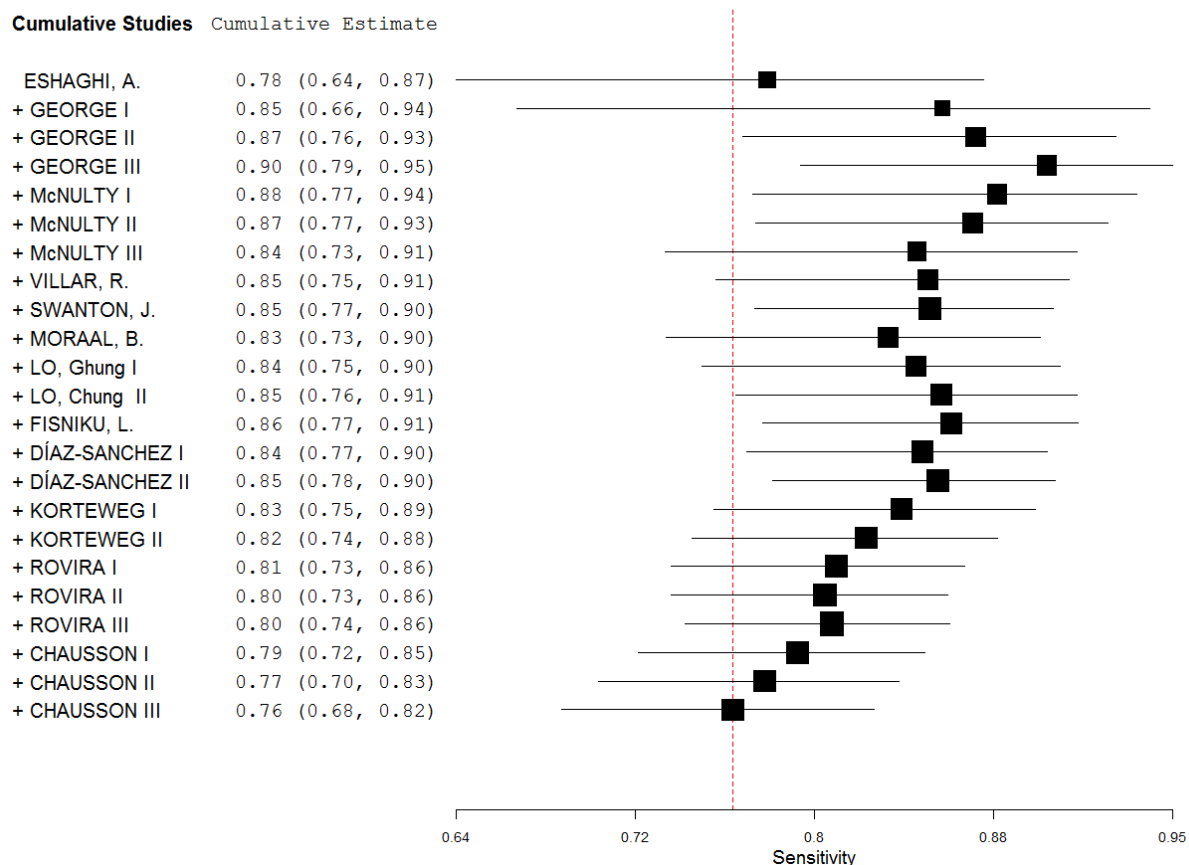


Obrázek 3.6: *PLR* kumulativníforest-graf. Graf zobrazuje kumulovaně propočítané průměry *PLR* ze všech předcházejících studií s 95% intervalem spolehlivosti. Konfidenční intervaly studií George III (kumulativní *PLR* = 3,42), McNulty I (kumulativní *PLR* = 3,13) a Villar, R. (kumulativní *PLR* = 2,86) se neprotínají s agregovanou hodnotou *PLR* = 2,00, tyto studie se významně statisticky liší od agregované hodnoty *PLR* [2].

3.3.3 Senzitivita

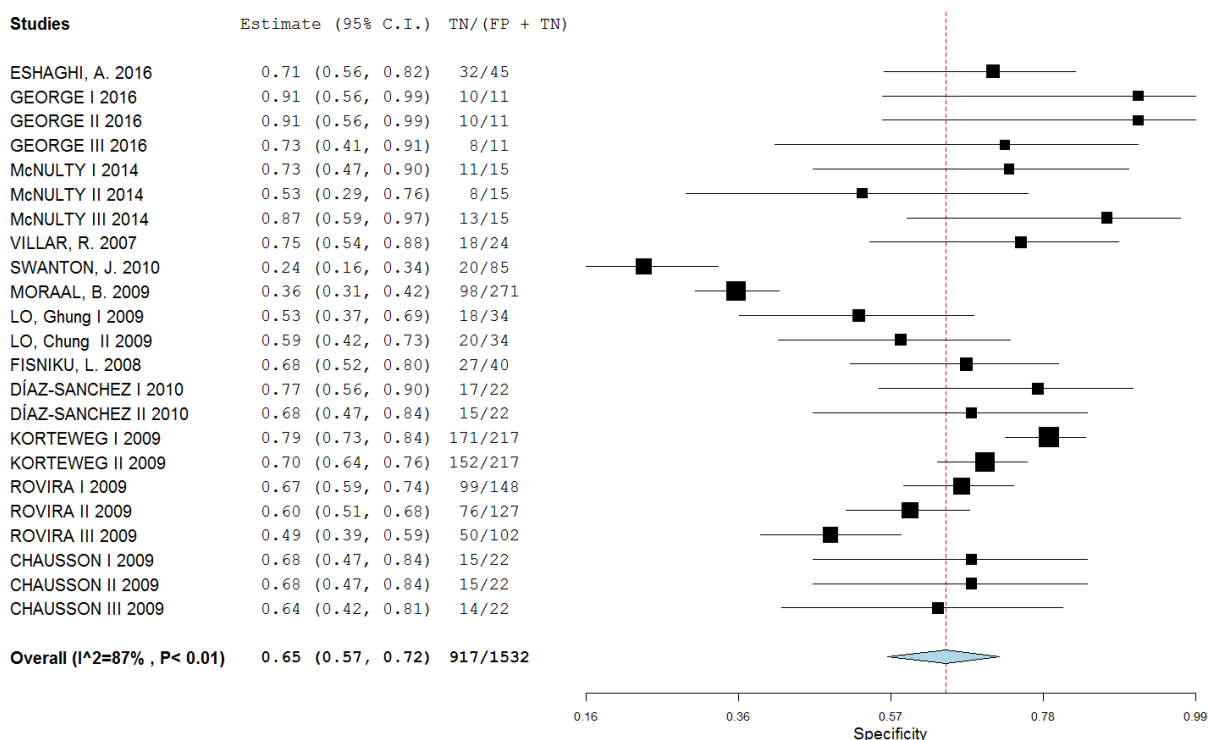


Obrázek 3.7: *Forest-graf senzitivity*. Graf zobrazuje *forest-graf* senzitivity zahrnutých studií s 95% intervalem spolehlivosti. Na ose x jsou vyneseny hodnoty senzitivity. Osa y odpovídá jednotlivým studiím. Červená kolmá osa vyznačená čárkovaně udává souhrnnou senzitivitu pro všechny studie, která nabývá hodnoty 0,76 s dolním odhadem 0,68 a horním odhadem 0,82 na daném intervalu spolehlivosti. Senzitivita studií se pohybuje od 0,36 (Chausson I) do 0,98 (LO, Chung). Studie Chausson I–III mají markantně nižší senzitivitu než studie zbývající. Naopak studie Rovira II dosahuje shodné meta-analýzou zjištěné senzitivity. Velký rozptyl je u studií McNulty I–III. Grafické znázornění senzitivity pro jednotlivé studie ve *forest-grafu* se liší velikostí v závislosti na počtu probandů. Higginsův I^2 test zde vychází 88 % s p -hodnotou $< 0,01$.

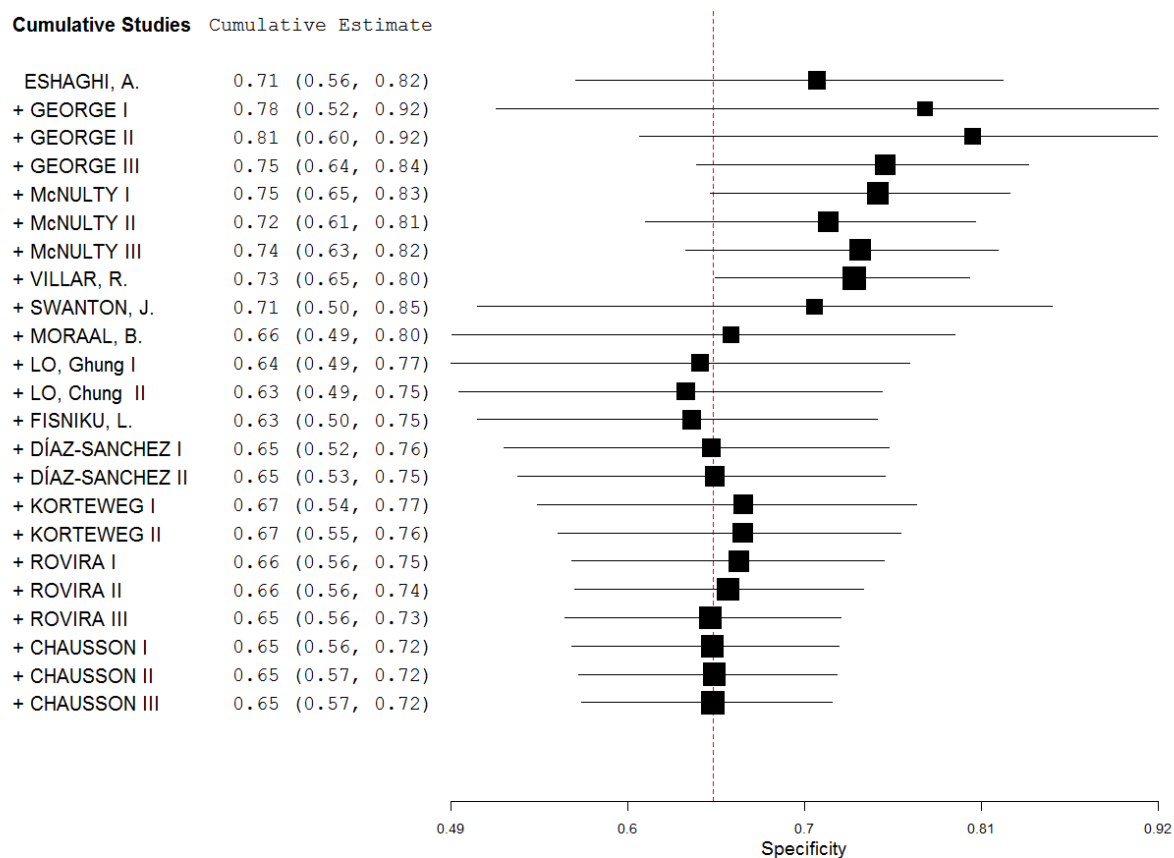


Obrázek 3.8: **Senzitivita kumulativní forest-graf.** Graf zobrazuje kumulovaně propočítané průměry senzitivity ze všech předcházejících studií s 95% intervalem spolehlivosti. Konfidenční intervaly studií George II (kumulativní senzitivita=0,87) George III (kumulativní senzitivita=0,90), McNulty I (kumulativní senzitivita=0,88), McNulty II (kumulativní senzitivita=0,87), Swanton, J. (kumulativní senzitivita=0,85), LO, Chung II (kumulativní senzitivita=0,85), Finisku, L. (kumulativní senzitivita=0,86), Díaz-Sanchez I (kumulativní senzitivita=0,84) a Díaz-Sanchez II (kumulativní senzitivita=0,85) se neprotínají s agregovanou hodnotou senzitivity=0,76, tyto studie se významně statisticky liší od agregované hodnoty senzitivity [2].

3.3.4 Specifická

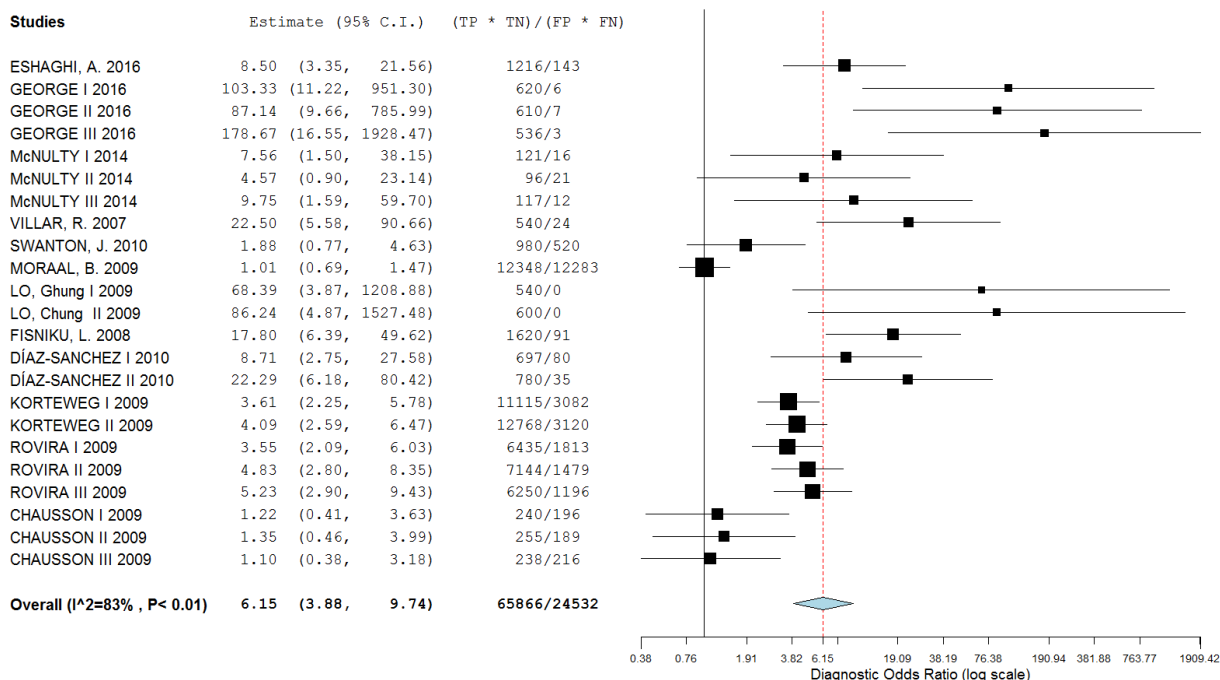


Obrázek 3.9: **Forest-graf specifická**. Graf zobrazuje *forest-graf* specifický zahrnutých studií s 95% intervalem spolehlivosti. Na ose x jsou vyneseny hodnoty specifický. Osa y odpovídá jednotlivým studiím. Červená kolmá osa vyznačena čárkovaně udává souhrnnou specificku pro všechny studie, která nabývá hodnoty 0,65 s dolním odhadem 0,57 a horním odhadem 0,72 na daném intervalu spolehlivosti. Specifická studií se pohybuje od 0,24 (Swanton, J.) do 0,91 (George I, II). U všech studií je téměř stejně velký rozptyl až na studii Moraal, B. a Swanton, J., které mají nejmenší rozptyl. Grafické znázornění specifický pro jednotlivé studie ve *forest-grafu* se liší velikostí v závislosti na počtu probandů. Higginsův I^2 test zde vychází 87 % s p -hodnotou $< 0,01$.

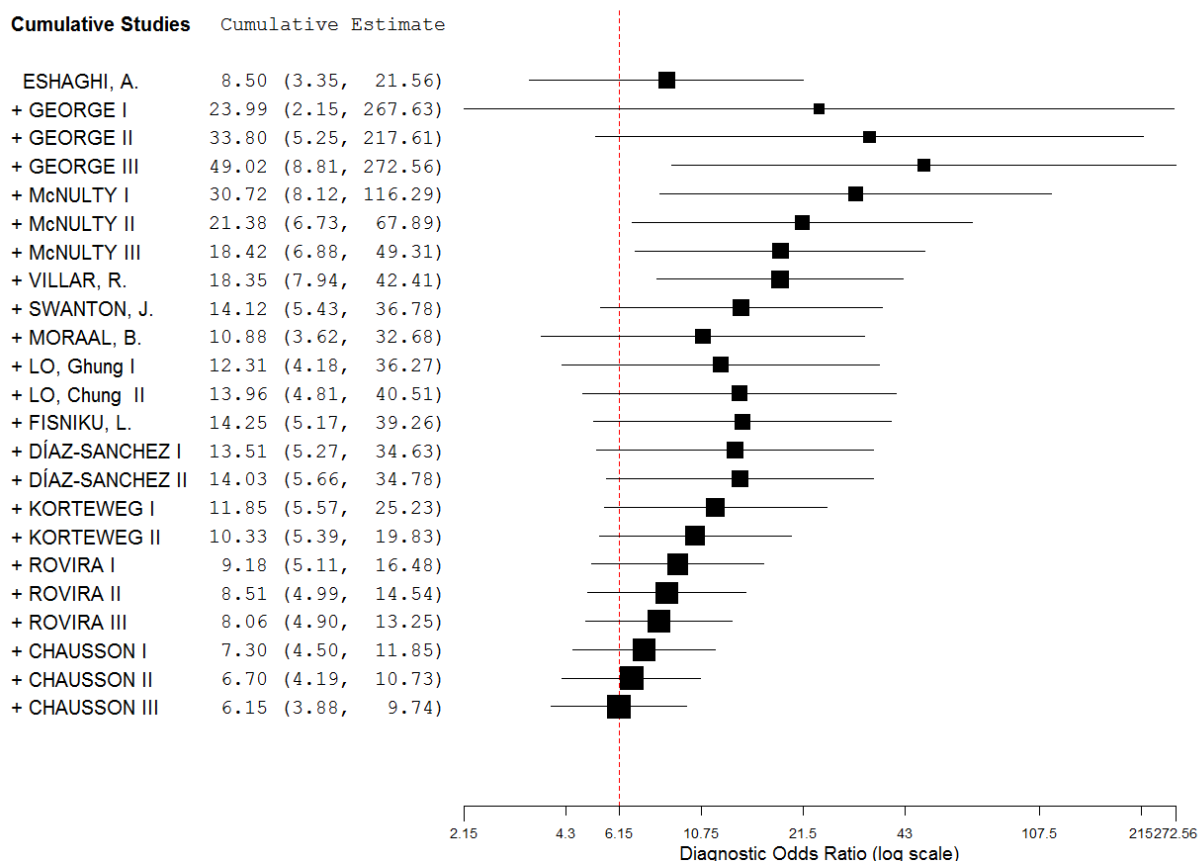


Obrázek 3.10: **Specificita kumulativní forest-graf.** Graf zobrazuje kumulovaně propočítané průměry specificity ze všech předcházejících studií s 95% intervalem spolehlivosti. Konfidenční interval studie Villar, R. (kumulativní specificita=0,73) se neprotíná s agregovanou hodnotou specificity=0,65, tato studie se významně liší od agregované hodnoty specificity [2].

3.3.5 DOR

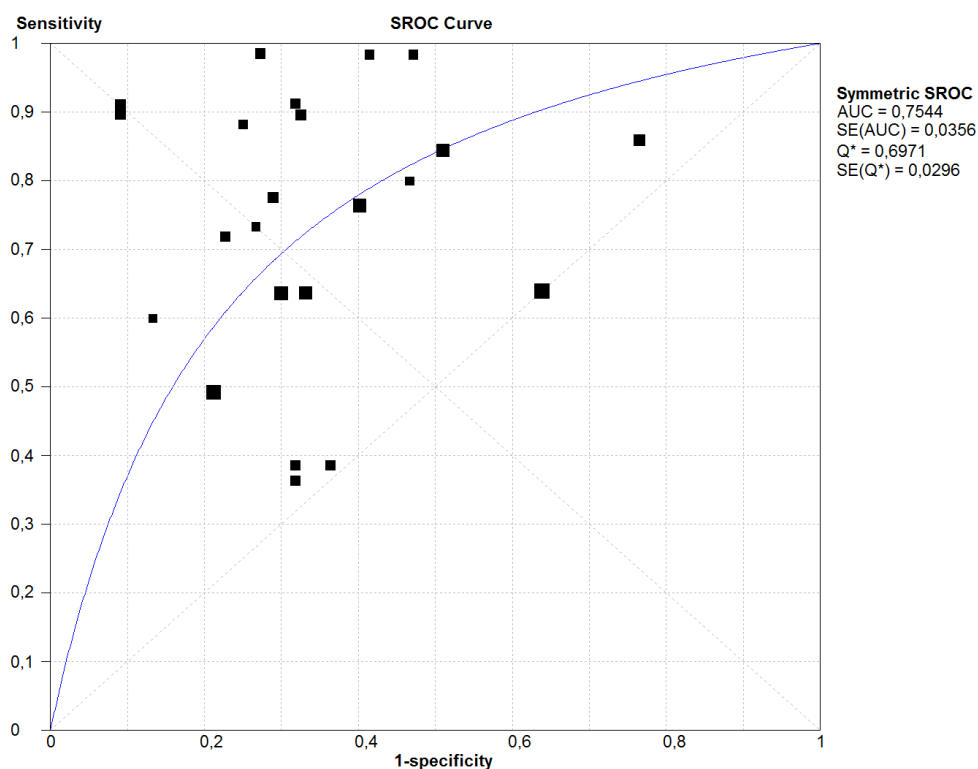


Obrázek 3.11: *DOR forest-graf*. Graf zobrazuje *forest-graf DOR* zahrnutých studií s 95% intervalem spolehlivosti. Na ose x jsou vyneseny hodnoty *DOR* v logaritmickém měřítku. Osa y odpovídá jednotlivým studiím. Červená kolmá osa vyznačená čárkovaně udává souhrnný *DOR* pro všechny studie, který nabývá hodnoty 6,15 s dolním odhadem 3,88 a horním odhadem 9,74 na daném intervalu spolehlivosti. Černá kolmice značí hodnotu 1. *DOR* zahrnutých studií se pohybuje od 1,01 (Moraal, B.) do 178,67 (George III). Největší rozptyl mají studie Lo, Chung I–II. Grafické znázornění *DOR* pro jednotlivé studie ve *forest-grafu* se liší velikostí v závislosti na počtu probandů. Higginsův I^2 test zde vychází 83 % s p -hodnotou $< 0,01$.



Obrázek 3.12: **DOR kumulativní forest-graf.** Graf zobrazuje kumulovaně propočítané průměry *DOR* ze všech předcházejících studií s 95% intervalem spolehlivosti. Konfidenční intervaly studií George III (kumulativní *DOR* = 49,02), McNulty I (kumulativní *DOR* = 30,72), McNulty II (kumulativní *DOR* = 21,38), McNulty III (kumulativní *DOR* = 18,42) a Villar, R. (kumulativní *DOR* = 18,35) se neprotínají s agregovanou hodnotou *DOR* = 6,15, tyto studie se významně liší od agregované hodnoty *DOR* [2].

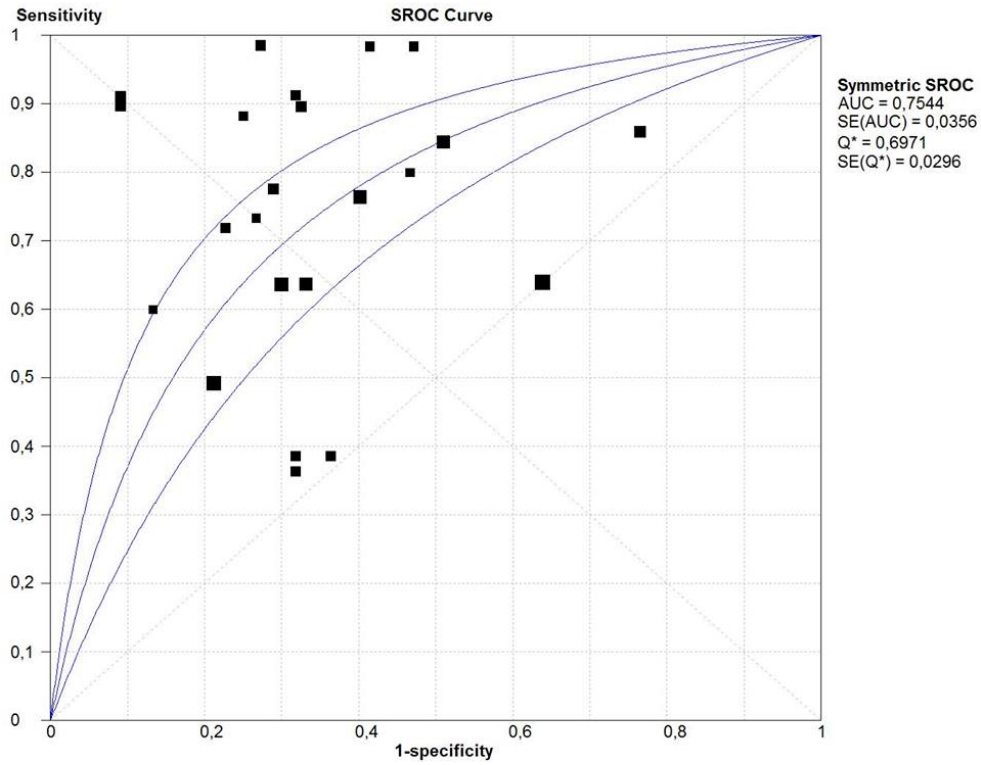
3.3.6 SROC křivka



Obrázek 3.13: SROC symetrická analýza

Průběh SROC křivky na obr. (3.13) poukazuje na dobrou diagnostickou přesnost testu svojí konvexností směrem k levému hornímu rohu. Po integraci jejího průběhu byla získána plocha pod křivkou, nebo-li AUC , jejíž hodnota byla stanovena na 0,75 se směrodatnou odchylkou 0,04. Parametr Q^* činí 0,70 se směrodatnou odchylkou 0,03 a označuje bod, ve kterém specifická i senzitivita nabývají shodné hodnoty, tedy 0,7.

Zaměříme-li se na průběh SROC křivky včetně průběhů intervalů spolehlivosti na obr. (3.14), je patrná konvexnost směrem k levému hornímu rohu i u obou těchto intervalů.



Obrázek 3.14: SROC symetrická analýza včetně intervalů spolehlivosti

Kapitola 4

Diskuze

V hodnocení metodologické kvality studií QUADAS 2 nejhůře dopadl referenční standard, kde u 4 studií hrozilo vysoké riziko zkreslení a u 3 studií riziko nejasné. Bylo to způsobeno především tím, že hodnotitelé výsledků studií již dopředu věděli hodnoty referenčního standardu a mohli tím být ovlivněni. Dalším velkým nedostatkem v metodologické kvalitě studií byla krátká doba sledování pacientů, která byla delší než 10 let pouze u 3 studií. Během kratší doby sledování mohou být opomenuty některé případy RS, které se mohou z CIS rozvinout do klinicky definované RS až po 10 letech a více. U výběru pacientů nebylo riziko zkreslení příliš velké, ale mohlo být způsobeno malým počtem zahrnutých pacientů. Při hodnocení obav týkajících se použitelnosti nebyly zjištěny žádné zásadní problémy, obavy byly u velké části studií ve všech 3 doménách nízké.

Nevýhodou meta-analýzy DTA studií je její větší náročnost než u meta-analýzy intervenčních studií. Meta-analýza DTA studií lze provést jen z primárních studií, které spolu významově souvisí, proto je nezbytné správně definovat kritéria pro zahrnutí/vyloučení studií. Při meta-analýze DTA studií se očekává vysoká heterogenita nebo specifická a senzitivita záleží na situaci, v které byl diagnostický test použit. Hodnoty může ovlivnit také prahový efekt. [6]

Výhodou této metody je shrnutí výsledků zahrnutých relevantních studií. Poskytuje odhady průměrné diagnostické přesnosti testu nebo testů, nejistoty tohoto průměru a variability výsledků studie v okolí odhadů. Použití meta-analýzy zpřístupní velké množství dat a je rychle asimilováno odbornou veřejností i zdravotnickou administrativou. Meta-analýza pomáhá pochopit zdánlivě protichůdné výsledky studií, neboť identifikuje, které rozdíly pravděpodobně budou skutečné, které lze vysvětlit náhodou a které lze vysvětlit známými rozdíly ve studijních charakteristikách. Vzhledem k tomu, že přesnost odhadů typicky vzrůstá s množstvím dat, meta-analýza může mít větší sílu k odhalení skutečných rozdílů v přesnosti testů než jednotlivé studie. A to může poskytnout přesnější odhady očekávané senzitivity a specifity. [6]

Meta-analýza DTA studií je limitována kvalitou a dostupností primárních studií, které jsou relevantní pro dané téma. Pro správné provedení meta-analýzy DTA studií je zapotřebí dostatečné množství studií, které obsahují vhodné spektrum probandů, srovnávají různé

testy, používají přísnou metodiku a jasně oznamují metody a zjištění. Druhým významným omezením meta-analýzy DTA studií je menší přehlednost interpretovaných výsledků. [4]

Pro studie na diagnostickou přesnost je typická vysoká heterogenita, která se projevuje i v této meta-analýze. [6] Při vytváření studií pro určení diagnostické přesnosti magnetické rezonance pro diagnózu roztroušené sklerózy hraje roli několik faktorů, které značně ovlivňují konzistenci dat napříč studiemi. Tyto faktory jsou proto zdrojem heterogenity pro výslednou meta-analýzu.

Jedním z faktorů je velká variabilita rozlišovací schopnosti použitých magnetických rezonancí. Existuje zde např. značný rozdíl mezi rozlišovací schopností při použití 1,5T MR a 3T MR, které byly ve studiích zastoupeny v porovnatelném počtu.

Dalším významným faktorem je subjektivita vnášená do studií při samotném odečítání lézí z pořízených snímků. Počty lézí byly nejčastěji nezávisle určovány dvěma či více zkušenými radiology. Díky lidskému faktoru je míra objektivity relativně snížena. Heterogenita může být podpořena i následným publikačním zkreslením studií.

Heterogenitu dat potvrzují i *forest*-grafy jednotlivých parametrů, u kterých si lze všimnout např. výrazného rozptylu mezi nejnižší a nejvyšší senzitivitou 0,62 a mezi nejnižší a nejvyšší specificitou 0,67. Hodnota Spearmanova koeficientu byla stanovena na 0,68, což poukazuje na přítomnost prahového efektu, tedy korelace mezi senzitivitou a specificitou. Z tohoto důvodu bylo možné pro určení diagnostické přesnosti použít SROC křivku. Z níž výpočtem hodnoty *AUC* lze určit do jaké kategorie diagnostické přesnosti tato metoda spadá.

Z výsledků je tedy patrné, že výsledná senzitivita 0,76 v intervalu (0,68; 0,82) a výsledná specificita 0,65 v intervalu (0,57; 0,72) potvrzují dobrou schopnost metody odhalit nemoc. Tvar SROC křivky z levého dolního rohu do pravého horního rohu s konvexním prohnutím směrem k levému hornímu rohu poukazuje na hodnoty *AUC* typické pro dobrou diagnostickou přesnost testu. Tato plocha pod křivkou tedy činí 0,75 se směrodatnou odchylkou 0,04.

Dalšími meta-analýzou získanými parametry jsou *PLR* a *NLR*. Hodnota *PLR* byla stanovena na 2,00 v intervalu (1,65; 2,42). Tato hodnota je nižší než standardně uváděná hodnota 10 u velmi spolehlivých testů pro odhalení nemocného pacienta. Je to dáno právě výrazným rozptylem *PLR* hodnot u jednotlivých studií.

Naopak hodnota *NLR* vychází 0,28 v intervalu (0,18; 0,42) a spadá do typických hodnot *NLR* pod 1,00. Tedy odpovídá to skutečnosti, že je pravděpodobnější získat negativní test u pacienta bez onemocnění než u pacienta s onemocněním.

Parametr *DOR* vychází 6,15 v intervalu (3,88; 9,74), což znamená, že pravděpodobnost získání pozitivního testu u nemocné osoby je $6,15\times$ vyšší než u zdravé.

Přes významné rozptyly minimálních a maximálních hodnot mezi jednotlivými studiemi spadá magnetická rezonance podle vytvořené meta-analýzy do kategorie dobrých diagnostických metod. Tento výsledek byl především ovlivněn četnějším zastoupením studií s vysokými hodnotami senzitivity a specificity.

Data z provedené meta-analýzy odpovídají výsledkům systematickému přehledu *Accuracy of diagnostic tests in multiple sclerosis – a systematic review*, kde specificita a senzi-

tivita vykazovala též vysokou heterogenitu, neboť hodnoty senzitivity se v přehledu pohybovaly od 34–100 % a specificity od 24–92 %. Hodnoty *DOR* jednotlivých studií se konfidenčními intervaly překrývají s vypočtenou agregovanou hodnotou této meta-analýzy. [41]

Druhý systematický přehled problematiky *Accuracy of magnetic resonance imaging for the diagnosis of multiple sclerosis: systematic review* využil stejné postupy, podobné vyřazovací metody a vhodné studie byly také podrobeny hodnocení kvality QUADAS. Senzitivita zahrnutých studií se pohybovala od 15 % do 100 % a specificita se pohybovala v rozmezí 12–99 %, ale většina studií se vyznačoval senzitivitou více než 75 % a specificitou nad 70 %. Z výsledných hodnot je zřejmá vysoká heterogenita mezi studii. Hodnocení QUADAS poukazuje na stejné problémy v kvalitě studií jako byly zjištěny v této práci. Zobrazení senzitivity/specificity v ROC prostoru vykazují stejné vlastnosti jako SROC analýza v této práci. Lze konstatovat, že meta-analýza v této práci udává výsledky srovnatelné s dostupnými systematickými přehledy dané problematiky. [42]

Ve světě je problematika diagnostické přesnosti poměrně často diskutované téma, což potvrzuje velké množství DTA studií. V České republice je toto téma opomíjeno a neexistuje žádná studie na diagnostickou přesnost MR u RS v prostředí ČR. Shrnující studie na diagnostickou přesnost MR při diagnóze roztroušené sklerózy mají formu systematického přehledu, zřejmě z důvodu vysoké heterogenity vstupních dat a obtížné interpretace výsledků. V ČR ani v zahraniční literatuře není k dispozici kompletní meta-analytické zpracování této problematiky.

Závěr

Meta-analýzou 12 studií, s 23 porovnatelnými hodnotami senzitivity a specificity, zaměřených na diagnostickou přesnost magnetické rezonance pro diagnózu roztroušené sklerózy byly určeny souhrnná senzitivita 0,76 s intervalem spolehlivosti (0,68; 0,82) a souhrnná specificita 0,65 s intervalem spolehlivosti (0,57; 0,72) na 95% hladině spolehlivosti. Tedy poměr nemocných pacientů s pozitivním testem vůči všem nemocným pacientům činí 76 %. A poměr zdravých pacientů s negativním testem vůči všem zdravým pacientům je 65 %. Hodnota *PLR* je podle meta-analýzy 2,00 v intervalu (1,65; 2,42) a hodnota *NLR* 0,28 v intervalu (0,18; 0,42). Tedy je 2× pravděpodobnější, že výsledek testu bude pozitivní u nemocného pacienta než u zdravého. A pravděpodobnost, že nemocný pacient dosáhne negativního testu než zdravý pacient je 28 %. Hodnota *DOR* byla určena na 6,15 v intervalu (3,88; 9,74). Je tedy 6,15× pravděpodobnější, že pozitivní výsledek bude získán na osobě s onemocněním než na osobě bez onemocnění.

Spearmanův test pro tato data vykazuje hodnotu 0,68, tudíž spolu s analýzou výsledných SROC křivek a určením hodnoty *AUC* (0,75 se směrodatnou odchylkou 0,04) můžeme tuto diagnostickou metodu zařadit do metod s dobrou diagnostickou přesností pro určení onemocnění roztroušenou sklerózou.

Literatura

- [1] GOODMAN, Clifford S. *Introduction to health technology assessment*. The Lewin Group. Virginia, USA, 2004.
- [2] HENDL, Jan. *Přehled statistických metod: analýza a metaanalýza dat*. 4., rozš. vyd. Praha: Portál, 2012. ISBN 978-80-262-0200-4.
- [3] CAMPBELL, Jared M., et al. *Diagnostic test accuracy: methods for systematic review and meta-analysis*. International journal of evidence-based healthcare, 2015, 13.3: 154-162.
- [4] LEEFLANG, M. M. G. *Systematic reviews and meta-analyses of diagnostic test accuracy*. Clinical Microbiology and Infection, 2014, 20.2: 105-113.
- [5] ŠIMUNDIĆ, Ana-Maria. *Measures of diagnostic accuracy: basic definitions*. EJIFCC, 2009, 19.4: 203.
- [6] MACASKILL, Petra, et al. *Cochrane handbook for systematic reviews of diagnostic test accuracy*. Version 0.9. 0. London: The Cochrane Collaboration, 2010.
- [7] ZVÁROVÁ, J. *Základy statistiky pro biomedicínské obory*. 2., dopl. vyd. Praha: Karolinum, 2011. 219 s. ISBN 978-80-246-1931-6.
- [8] KIM, Kyung Won, et al. *Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers-part I*. general guidance and tips. Korean journal of radiology, 2015, 16.6: 1175-1187.
- [9] IRWIG, Les, et al. *Guidelines for meta-analyses evaluating diagnostic tests*. Annals of internal medicine, 1994, 120.8: 667-676.
- [10] JAROŠOVÁ, Darja a Renáta ZELENÍKOVÁ. *Ošetrovatelství založené na důkazech: evidence based nursing*. Praha: Grada, 2014. Sestra (Grada). ISBN 978-80-247-5345-4.
- [11] WHITING, Penny, et al. *The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews*. BMC medical research methodology, 2003, 3.1: 25.
- [12] WHITING, Penny F., et al. *QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies*. Annals of internal medicine, 2011, 155.8: 529-536.

-
- [13] REITSMA, Johannes B., et al. *Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews*. Journal of clinical epidemiology, 2005, 58.10: 982-990.
- [14] CHAPPELL, F. M.; RAAB, G. M.; WARDLAW, J. M. *When are summary ROC curves appropriate for diagnostic meta-analyses?*. Statistics in medicine, 2009, 28.21: 2653-2668.
- [15] LEE, Juneyoung, et al. *Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers-part II*. statistical methods of meta-analysis. Korean journal of radiology, 2015, 16.6: 1188-1196.
- [16] SONG F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ. *Publication and related biases*. Health Technol Assess. 2000;4:1–115
- [17] TATSIONI, Athina, et al. *Challenges in systematic reviews of diagnostic technologies*. Annals of internal medicine, 2005, 142.12_Part_2: 1048-1055.
- [18] MULHERIN, Stephanie A.; MILLER, William C. *Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation*. Annals of Internal Medicine, 2002, 137.7: 598-602.
- [19] DE GROOT, Joris AH, et al. *Adjusting for partial verification or workup bias in meta-analyses of diagnostic accuracy studies*. American journal of epidemiology, 2012, 175.8: 847-853.
- [20] LEEFLANG, Mariska MG, et al. *Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: mechanisms, magnitude, and solutions*. Clinical Chemistry, 2008, 54.4: 729-737.
- [21] KRASULOVÁ, Eva; HAVRDOVÁ, Eva. *Současná léčba relaps-remitentní roztroušené sklerózy a perspektivy do budoucna*. Klinická farmakologie a farmacie, 2008, 22.1: 23-29.
- [22] MELUZÍNOVÁ, Eva. *Současné možnosti léčby roztroušené sklerózy*. Neurol. Pro praxi, 2010, 11.5: 307-311.
- [23] HORÁKOVÁ, D. Vaněčková M. *Význam magnetické rezonance ve sledování aktivity roztroušené sklerózy–pohled klinického neurologa*. Neurol. praxi, 2012, 13.6: 320-324.
- [24] POLMAN, Chris H., et al. *Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria*. Annals of neurology, 2011, 69.2: 292-302.
- [25] BENEŠOVÁ, Yvonne. *ROZTROUŠENÁ SKLERÓZA A JINÁ DEMYELINIZAČNÍ ONEMOCNĚNÍ: Diagnostika, léčba, diferenciální diagnostika*. Neurologická klinika Fakultní nemocnice Brno. Brno: Lékařská fakulta Masarykovy univerzity, 2014, 2014(1), 1-62.

-
- [26] MOHER, David, et al. *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement*. PLoS med, 2009, 6.7: e1000097.
- [27] BRODLEY, C. E., et al. Open meta-analyst. 2012.
- [28] ZAMORA, Javier, et al. *Meta-DiSc: a software for meta-analysis of test accuracy data*. BMC medical research methodology, 2006, 6.1: 31.
- [29] ESHAGHI, Arman, et al. *Gray matter MRI differentiates neuromyelitis optica from multiple sclerosis using random forest*. Neurology, 2016, 87.23: 2463-2470.
- [30] GEORGE, Ilena C., et al. *Clinical 3-tesla FLAIR* MRI improves diagnostic accuracy in multiple sclerosis*. Multiple Sclerosis Journal, 2016, 22.12: 1578-1586.
- [31] MCNULTY, J. P., et al. *Diagnostic efficacy of conventional MRI pulse sequences in the detection of lesions causing internuclear ophthalmoplegia in multiple sclerosis patients*. Clinical neuroradiology, 2015, 25.3: 233-239.
- [32] VILLAR, Luisa M., et al. *Accuracy of CSF and MRI criteria for dissemination in space in the diagnosis of multiple sclerosis*. Journal of the neurological sciences, 2008, 266.1: 34-37.
- [33] SWANTON, J. K., et al. *Early MRI in optic neuritis: the risk for clinically definite multiple sclerosis*. Multiple Sclerosis Journal, 2010, 16.2: 156-165.
- [34] MORAAL, Bastiaan, et al. *Magnetic resonance imaging predictors of conversion to multiple sclerosis in the BENEFIT study*. Archives of neurology, 2009, 66.11: 1345-1352.
- [35] LO, Chung-Ping, et al. *Prediction of conversion from clinically isolated syndrome to clinically definite multiple sclerosis according to baseline MRI findings: comparison of revised McDonald criteria and Swanton modified criteria*. Journal of Neurology, Neurosurgery and Psychiatry, 2009, 80.10: 1107-1109.
- [36] CHAUSSON, N., et al. *Evaluation of the 2005 McDonald MRI criteria for dissemination in space in Afro-Caribbean patients with clinically isolated syndromes*. European journal of neurology, 2009, 16.11: 1191-1196.
- [37] FISNIKU, L. K., et al. *Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis*. Brain, 2008, 131.3: 808-817.
- [38] DÍAZ-SÁNCHEZ, María, et al. *Accuracy of MRI criteria for dissemination in space for the diagnosis of multiple sclerosis in patients with clinically isolated syndromes*. Multiple Sclerosis Journal, 2010, 16.5: 576-580.
- [39] KORTEWEG, Tijmen, et al. *A search for new MRI criteria for dissemination in space in subjects with a clinically isolated syndrome*. European radiology, 2009, 19.9: 2244-2248.

-
- [40] ROVIRA, Alex, et al. *A single, early magnetic resonance imaging study in the diagnosis of multiple sclerosis*. Archives of Neurology, 2009, 66.5: 587-592.
- [41] SCHÄFFLER, N., et al. *Accuracy of diagnostic tests in multiple sclerosis—a systematic review*. Acta Neurologica Scandinavica, 2011, 124.3: 151-164.
- [42] WHITING, Penny, et al. *Accuracy of magnetic resonance imaging for the diagnosis of multiple sclerosis: systematic review*. BMJ, 2006, 332.7546: 875-884.

Příloha I

Struktura souborů na přiložením CD

1. **Diplomova_prace_Capkova.pdf**: kompletní práce v elektronické podobě.
2. **Abstract_cz.pdf**: Abstrakt v Českém jazyce
3. **Abstract_en.pdf**: Abstrakt v Anglickém jazyce
4. **Klicova_slova.pdf**: Klíčová slova
5. **ZadaniDP.pdf**: Zadání diplomové práce