

VYSOKÁ ŠKOLA POLYTECHNICKÁ JIHLAVA

Aplikovaná informatika

KOMPARATIVNÍ ANALÝZA NÁSTROJŮ UMĚLÉ  
INTELIGENCE PRO HODNOCENÍ SKELETÁLNÍCH  
RENTGENOVÝCH SNÍMKŮ

Bakalářská práce

Autor práce: Lukáš Brdíčko

Vedoucí práce: Mgr. Daniel Kvak

Jihlava 2026

# Vysoká škola polytechnická Jihlava

Tolstého 16, 586 01 Jihlava

## ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Autor práce: **Lukáš Brdíčko**

Studijní program: Aplikovaná informatika

Garant studijního programu: Ing. Lenka Kuklišová Pavelková, Ph.D.

Název práce: **Komparativní analýza nástrojů umělé inteligence pro hodnocení skeletálních rentgenových snímků**

Vedoucí práce: Mgr. Daniel Kvak, PhD.

Cíl práce: Cílem bakalářské práce je zhodnotit, jak čeští radiologové vnímají, přijímají a hodnotí nástroje umělé inteligence určené pro analýzu skeletálních rentgenových snímků. Práce bude mít kombinovaný design, zahrnující kvantitativní a kvalitativní část. V kvantitativní části budou vybrané systémy umělé inteligence (AI) otestovány na anonymizovaném souboru muskuloskeletálních rentgenových snímků. Hodnoceny budou výkonnostní metriky jako senzitivita, specificita či prediktivní hodnoty. Kvalitativní část bude postavena na dotazníkovém šetření mezi radiology, zaměřeném na důvěru k tzv. „black-box“ rozhodnutím, požadavky na vysvětlitelnost modelů, možnosti integrace do klinického workflow (např. PACS nebo EHR) a vnímané přínosy i rizika. Dotazníkové šetření bude doplněno o polostrukturované rozhovory, jejichž cílem bude identifikace hlavních bariér adopce AI a návrhů na jejich překonání. Na základě syntézy kvantitativních a kvalitativních výsledků práce nabídne doporučení pro efektivní a bezpečnou implementaci AI nástrojů v oblasti skeletální radiologie v českých zdravotnických zařízeních.

## Abstrakt

Bakalářská práce se zaměřuje na využití umělé inteligence v oblasti skeletální radiologie se zaměřením na hodnocení rentgenových snímků kostí. Cílem práce je provést komparativní analýzu vybraných komerčních nástrojů umělé inteligence určených pro detekci skeletálních patologií a zhodnotit postoje českých radiologů k jejich využití v klinické praxi.

Teoretická část shrnuje principy umělé inteligence ve zdravotnictví, specifika skeletálního zobrazování, typy AI nástrojů používaných v radiologii, přístupy k hodnocení jejich výkonnosti a faktory ovlivňující důvěru a akceptaci těchto systémů. Pozornost je věnována také regulačním aspektům, integraci do klinického workflow a etickým otázkám.

Praktická část je založena na kombinovaném výzkumném designu. Kvantitativní část porovnává výkonnost vybraných zdravotnických prostředků pomocí standardních diagnostických metrik na anonymizovaných rentgenových snímcích skeletu. Kvalitativní část vychází ze strukturovaného dotazníkového šetření mezi radiology. Výstupem práce jsou doporučení pro efektivní a odpovědnou implementaci AI nástrojů do klinické praxe.

## Klíčová slova

umělá inteligence; radiologie; rentgenový snímek skeletu; komparativní analýza; diagnostika

## Abstract

Bachelor's thesis focuses on the application of artificial intelligence in skeletal radiology, with emphasis on the evaluation of bone X-ray images. The aim of the thesis is to conduct a comparative analysis of selected commercial artificial intelligence tools designed for the detection of skeletal pathologies and to assess the attitudes of Czech radiologists towards their use in clinical practice.

The theoretical part summarizes the principles of artificial intelligence in healthcare, specifics of skeletal imaging, types of AI tools used in radiology, approaches to performance evaluation, and factors influencing trust and acceptance of these systems. Attention is also paid to regulatory aspects, clinical workflow integration, and ethical considerations.

The practical part is based on a mixed-methods research design. The quantitative part compares the performance of selected medical devices using standard diagnostic metrics on anonymized X-ray images of the skeleton. The qualitative part is based on a structured questionnaire survey among radiologists. The output of the thesis consists of recommendations for the effective and responsible implementation of AI tools into clinical practice.

## Keywords

artificial intelligence; radiology; X-ray image of the skeleton; comparative analysis; diagnostics

Prohlašuji, že předložená bakalářská práce je původní a zpracoval jsem ji samostatně. Prohlašuji, že citace použitých pramenů je úplná, že jsem v práci neporušil autorská práva (ve smyslu zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů, v platném znění, dále též „AZ“).

Byl jsem seznámen s tím, že na mou bakalářskou práci se plně vztahuje **AZ**, zejména § 60 (školní dílo).

Podle § 47b zákona o vysokých školách souhlasím se zveřejněním své práce podle Směrnice pro vedení, vypracování a zveřejňování závěrečných prací na VŠPJ, a to bez ohledu na výsledek obhajoby.

Beru na vědomí, že VŠPJ má právo na uzavření licenční smlouvy o užití mé bakalářské práce a prohlašuji, že **s o u h l a s í m** s případným užitím mé bakalářské práce (prodej, zapůjčení apod.).

Jsem si vědom toho, že užít své bakalářské práce či poskytnout licenci k jejímu využití mohu jen se souhlasem VŠPJ, která má právo ode mě požadovat přiměřený příspěvek na úhradu nákladů, vynaložených vysokou školou na vytvoření díla (až do jejich skutečné výše), z výdělku dosaženého v souvislosti s užitím díla či poskytnutím licence.

V Jihlavě dne 24. března 2026

.....

Podpis studenta

## Poděkování

*Rád bych poděkoval vedoucímu bakalářské práce Mgr. Danielu Kvakovi za odborné vedení, průběžnou podporu a cenné podněty při zpracování práce. Dále děkuji MUDr. Jakubu Dandárovi za cenné rady v oblasti zdravotnické terminologie a za pomoc při zajištění relevantních kontaktů na radiology pro realizaci dotazníkového šetření. Poděkování patří také PaedDr. PhDr. Michalu Matějkovi, DBA, Ph.D., za pomoc se stylistickou úpravou práce. Děkuji společnosti Carebot s.r.o. za součinnost při zajištění datasetu z Fakultní nemocnice Olomouc. Poděkování náleží rovněž Fakultní nemocnici Olomouc za poskytnutí podkladů pro zpracování výzkumné části práce. Zvláštní poděkování patří MUDr. Jakubu Olejkovi za zpracování kazuistik využitých pro výzkumné účely. Děkuji také všem respondentům, kteří se zapojili do dotazníkového šetření a přispěli tak k realizaci výzkumné části práce. V neposlední řadě mimořádně děkuji své rodině za podporu, důvěru a zázemí po celou dobu zpracování bakalářské práce.*

# Obsah

<b>Seznam obrázků .....</b>	<b>7</b>
<b>Seznam tabulek .....</b>	<b>8</b>
<b>Seznam zkratk .....</b>	<b>9</b>
<b>Úvod .....</b>	<b>11</b>
<b>1 Teoretická část.....</b>	<b>12</b>
1.1 Umělá inteligence ve zdravotnictví a radiologii .....	12
1.2 Skeletální zobrazování a typické diagnostické úlohy .....	14
1.3 Nástroje umělé inteligence pro hodnocení skeletálních rentgenových snímků .....	15
1.4 Výkonové metriky AI (senzitivita, specificita, věrohodnostní poměry) .....	17
1.5 Vnímání a důvěra radiologů v technologie AI – vysvětlitelnost a etika.....	19
1.6 Přehled certifikovaných AI nástrojů v radiologii .....	21
<b>2 Metodika výzkumu .....</b>	<b>22</b>
2.1 Výzkumný design a teoretická východiska zvoleného přístupu .....	22
2.2 Výběr, charakteristika a zaslepení testovaných AI nástrojů.....	22
2.3 Protokol pro extrakci dat, anonymizace a strukturace datového souboru .....	23
2.4 Kvantitativní vyhodnocení výkonnosti a definice referenčního standardu.....	24
2.5 Nástroje a postupy pro sběr kvalitativních dat .....	24
<b>3 Výsledky a analýza dat .....</b>	<b>26</b>
3.1 Kvantitativní část – výsledky testování AI nástrojů .....	26
3.2 Kvalitativní část – postoje radiologů a vnímání AI nástrojů .....	33
3.3 Komparace a syntéza výsledků .....	38
<b>4 Diskuse .....</b>	<b>41</b>
4.1 Interpretace výsledků ve vztahu k odborné literatuře .....	41
4.2 Hlavní bariéry a přínosy implementace AI v české skeletální radiologii.....	42
4.3 Možnosti integrace AI do klinického workflow .....	43
<b>Závěr .....</b>	<b>46</b>
<b>Seznam použité literatury.....</b>	<b>47</b>
<b>Přílohy.....</b>	<b>51</b>

## Seznam obrázků

Obr. 1: Výstup AI systému Carebot Bones při analýze rentgenového snímku ruky. Detekce fraktur je označena červenými rámečky s popiskem „FRA“; systém zároveň vyhodnotil celkové riziko jako vysoké. ....	17
Obr. 2: Rentgenový snímek pravého kolenního kloubu s depresí laterálního tibiálního plata... ..	29
Obr. 3: Rentgenový snímek pravé nohy s frakturami přednoží.....	30
Obr. 4: Rentgenový snímek levé nohy se stresovou frakturou II. metatarzu .....	31
Obr. 5: Rentgenový snímek levého loketního kloubu s frakturou hlavičky radia .....	32
Obr. 6: Délka klinické praxe respondentů v oboru radiologie .....	33
Obr. 7: Dosavadní praktická zkušenost respondentů s AI nástroji .....	34
Obr. 8: Vliv vizualizace (XAI) na důvěru v diagnostický výsledek.....	35
Obr. 9: Preferovaný způsob zařazení AI do klinického workflow .....	35
Obr. 10: Požadavky na integraci AI do nativního prohlížeče PACS .....	36
Obr. 11: Největší aktuální bariéry pro plošné nasazení AI v ČR.....	37
Obr. 12: Názor respondentů na nesení právní odpovědnosti za chybu AI .....	37

## Seznam tabulek

Tab. 1: Seznam výkonových metrik .....	18
Tab. 2: Struktura dotazníkového šetření .....	25
Tab. 3: Distribuce vyšetřovaných anatomických oblastí.....	26
Tab. 4: Výsledky diagnostické přesnosti – systém AI 1.....	27
Tab. 5: Výsledky diagnostické přesnosti – systém AI 2.....	27
Tab. 6: Výsledky diagnostické přesnosti – systém AI 3.....	28
Tab. 7: Komparace výkonnostních metrik testovaných AI systémů .....	28
Tab. 8: Klasifikace kazuistiky 1.....	29
Tab. 9: Klasifikace kazuistiky 2.....	30
Tab. 10: Klasifikace kazuistiky 3.....	31
Tab. 11: Klasifikace kazuistiky 4.....	32

## Seznam zkratek

AI	Artificial Intelligence (umělá inteligence)
AUC	Area Under the Curve (plocha pod ROC křivkou)
BA	Balanced Accuracy (vyvážená přesnost)
CE	Conformité Européenne (označení shody v EU)
CNN	Convolutional Neural Network (konvoluční neuronová síť)
DB	Database (databáze)
DICOM	Digital Imaging and Communications in Medicine (standard pro ukládání a přenos medicínských obrazových dat)
ESR	European Society of Radiology (Evropská radiologická společnost)
FDA	Food and Drug Administration (Úřad pro kontrolu potravin a léčiv USA)
FN	False Negative (falešně negativní nález)
FNOL	Fakultní nemocnice Olomouc
FNR	False Negative Rate (míra falešně negativních výsledků)
FP	False Positive (falešně pozitivní nález)
FPR	False Positive Rate (míra falešně pozitivních výsledků)
GDPR	General Data Protection Regulation (obecné nařízení o ochraně osobních údajů)
LLM	Large Language Model (velký jazykový model)
MDR	Medical Device Regulation (nařízení o zdravotnických prostředcích)
ML	Machine Learning (strojové učení)
NLR	Negative Likelihood Ratio (negativní věrohodnostní poměr)
PACS	Picture Archiving and Communication System (systém pro archivaci a přenos obrazové dokumentace)
PLR	Positive Likelihood Ratio (pozitivní věrohodnostní poměr)
RIS	Radiology Information System (radiologický informační systém)
ROC	Receiver Operating Characteristic (ROC křivka)
RTG	rentgenový
RWE	Real-World Evidence (důkazy z reálné klinické praxe)
SaMD	Software as a Medical Device (software jako zdravotnický prostředek)
TN	True Negative (pravý negativní nález)

TP	True Positive (pravý pozitivní nález)
XAI	Explainable Artificial Intelligence (vysvětlitelná umělá inteligence)

## Úvod

Rozvoj umělé inteligence v posledních letech zásadně ovlivňuje medicínskou diagnostiku, zejména radiologii. Systémy založené na strojovém učení a hlubokých neuronových sítích umožňují analyzovat rozsáhlé soubory obrazových dat a poskytují podporu při hodnocení rentgenových snímků. Současně se objevují otázky související se spolehlivostí modelů, úrovní vysvětlitelnosti a mírou důvěry mezi radiology.

Výzkum je zaměřen na komparativní analýzu nástrojů umělé inteligence určených pro hodnocení skeletálních rentgenových snímků a na zhodnocení postojů českých radiologů k jejich klinickému využití. Volba tématu vychází z rychlého rozvoje AI ve zdravotnictví a potřeby objektivně posoudit její přínos v diagnostické praxi.

Cílem práce je zhodnotit, jak čeští radiologové vnímají, přijímají a hodnotí nástroje umělé inteligence určené pro analýzu skeletálních rentgenových snímků. Práce má kombinovaný design, který propojuje kvantitativní a kvalitativní přístup. V kvantitativní části budou vybrané AI systémy otestovány na anonymizovaném souboru muskuloskeletálních snímků a hodnoceny budou jejich výkonnostní metriky – senzitivita, specificita, věrohodnostní poměry a vyvážená přesnost. Kvalitativní část bude vycházet z dotazníkového šetření mezi radiology. Jejím cílem je zmapovat postoje a důvěru odborníků k „black-box“ modelům, vnímání vysvětlitelnosti a bariéry implementace AI v klinickém prostředí.

Výsledná syntéza poskytne podklad pro doporučení zaměřená na efektivní a bezpečné zavádění nástrojů umělé inteligence do oblasti skeletální radiologie v českých zdravotnických zařízeních. Zjištění mohou přispět k dalšímu rozvoji výzkumu a podpořit praktickou aplikaci AI v diagnostice muskuloskeletálních onemocnění.

# 1 Teoretická část

Teoretická část bakalářské práce vymezuje základní pojmy, přístupy a souvislosti týkající se využití umělé inteligence v oblasti skeletální radiologie. Cílem kapitoly je vytvořit odborný rámec pro následnou komparativní analýzu vybraných nástrojů umělé inteligence určených pro hodnocení skeletálních rentgenových snímků. Pozornost je věnována principům umělé inteligence v radiologii, specifikům skeletálního zobrazování, typům dostupných AI nástrojů, přístupům k jejich hodnocení a faktorům ovlivňujícím jejich přijetí v klinické praxi.

V textu práce je používán pojem muskuloskeletální radiologie jako ustálené odborné označení oboru zabývajícího se zobrazováním kostí, kloubů a přilehlých struktur. Pro označení analyzovaných dat je používán popisný termín rentgenové snímky skeletu, který odpovídá běžné klinické praxi a odborné literatuře.

## 1.1 Umělá inteligence ve zdravotnictví a radiologii

Umělá inteligence představuje jeden z nejvýznamnějších technologických trendů současné medicíny. Zdravotnictví generuje rozsáhlé objemy heterogenních dat, zahrnujících obrazové, textové i numerické informace, které mohou být analyzovány algoritmy strojového učení s cílem podpořit diagnostická rozhodnutí, optimalizovat pracovní postupy nebo zvýšit bezpečnost pacientů. V posledních letech dochází k výraznému rozvoji výpočetních metod, jež umožňují efektivní zpracování velkých datových souborů a identifikaci vzorců, které mohou zůstat při manuálním hodnocení skryté. Zavádění umělé inteligence je patrné zejména v radiologii, protože oblast zobrazovacích metod disponuje velkými strukturovanými datovými soubory vhodnými pro trénování a validaci modelů hlubokého učení (Litjens et al., 2017).

V radiologii se uplatňují především modely hlubokého učení, které dokážou identifikovat patologické struktury, lokalizovat abnormality nebo klasifikovat snímky na základě diagnostických kritérií (Lundervold & Lundervold, 2019). Konvoluční neuronové sítě umožnily výrazný pokrok v automatizaci analýzy obrazových dat a vedly k rozvoji systémů schopných detekovat fraktury, degenerativní změny či jiné muskuloskeletální patologie na prostých rentgenových snímcích. Výsledky těchto modelů často dosahují úrovně srovnatelné s lidskými odborníky, přičemž významnou výhodou je konzistentní výkon nezávislý na únavě, časovém tlaku nebo individuální zkušenosti hodnotitele. To má zásadní význam zejména v prostředí urgentních příjmů a traumatologických pracovišť, kde je kladen důraz na rychlé a spolehlivé rozhodování.

Současný vývoj AI ve zdravotnictví je úzce spojen se změnou role radiologa i dalších lékařů podílejících se na hodnocení zobrazovacích vyšetření, zejména v prostředí urgentního příjmu. Namísto nahrazení lidského faktoru směřuje současný vývoj k modelu spolupráce, v němž algoritmy poskytují podpůrné informace a upozorňují na potenciálně patologické oblasti, zatímco finální klinické rozhodnutí zůstává na lékaři. Daný přístup umožňuje zvýšit efektivitu práce, snížit chybovost a zlepšit konzistenci hodnocení, aniž by byla narušena odborná odpovědnost zdravotnického pracovníka.

Nasazení umělé inteligence v klinické praxi však provází i otázky spojené s vysvětlitelností modelů, bezpečností rozhodovacích procesů a vlivem na profesní odpovědnost zdravotnických

pracovníků. Radiologové očekávají nejen vysokou přesnost algoritmů, ale také dostatečně transparentní výstupy, které umožní porozumět důvodům jednotlivých rozhodnutí a ověřit jejich klinickou relevanci (Muhammad & Bendechache, 2024; Nazir et al., 2023). Problematika tzv. „black-box“ modelů a rozvoj metod vysvětlitelné umělé inteligence jsou intenzivně diskutovány v kontextu medicínského zobrazování, přičemž jejich cílem je zvýšit interpretovatelnost algoritmických výstupů a posílit důvěru uživatelů v klinické praxi. Schopnost interpretace výstupů AI je proto považována za jeden z klíčových faktorů dlouhodobé udržitelnosti těchto technologií ve zdravotnictví.

Důležitým aspektem zavádění umělé inteligence do klinického prostředí je také regulace. Umělá inteligence využívaná pro medicínské účely je v Evropské unii klasifikována jako zdravotnický prostředek a podléhá legislativě Medical Device Regulation (MDR), která stanovuje požadavky na bezpečnost, klinickou validaci, řízení rizik a dohledatelnost rozhodovacích procesů (European Union, 2017). Podobný přístup uplatňuje i americký regulační úřad FDA, který eviduje a schvaluje systémy využívající algoritmy strojového učení pro klinické použití (FDA, 2022). Regulační rámce mají za cíl zajistit, aby nasazované systémy splňovaly požadavky na kvalitu, bezpečnost a odpovědnost, a zároveň umožňovaly jejich transparentní využití v každodenní klinické praxi. V posledních letech rychle narůstá počet AI aplikací schválených pro medicínské účely – FDA pravidelně aktualizuje veřejný seznam povolených AI/ML zdravotnických zařízení, který k roku 2025 zahrnuje již stovky položek, z nichž značná část spadá do oblasti radiologie. Také evropský trh zaznamenává příliv certifikovaných AI nástrojů nesoucích označení CE, přičemž všechny musí projít přísným posuzováním shody dle MDR. Striktní regulace a důkladná klinická validace jsou nezbytné pro zajištění, že zaváděné algoritmy budou fungovat spolehlivě a bezpečně v reálném prostředí u pacientů.

Úspěšná implementace umělé inteligence vyžaduje integraci do stávajících systémů nemocničního informačního prostředí, zejména PACS a RIS. Integrace musí zajistit technickou kompatibilitu, spolehlivý přenos dat a přehledné zobrazování výsledků přímo v prostředí, na které jsou uživatelé zvyklí. Výzkumy ukazují, že uživatelská zkušenost, ergonomie rozhraní a minimální zásah do zavedeného workflow patří mezi klíčové faktory ovlivňující přijetí umělé inteligence v klinické praxi (Kotter & Ranschaert, 2021; Farič et al., 2024). Pokud AI nástroj naruší plynulost práce (např. prodlevami v zobrazení výsledků nebo nutností přepínat do jiného softwaru), setká se pravděpodobně s odporem uživatelů. Naopak řešení, která jsou schopna generovat výstupy přímo ve formátu běžně používaném v klinické praxi (např. jako DICOM snímek s vyznačením nálezu či jako strukturovaná textová zpráva v rámci PACS), mají výrazně vyšší potenciál pro přijetí mezi zdravotníky. Radiologie tak představuje obor, kde se potenciál umělé inteligence již reálně projevuje, současně je však nutné věnovat zvýšenou pozornost bezpečnosti, etickým otázkám a odpovědnosti při interpretaci algoritmických výstupů (ESR, 2019).

Umělá inteligence přináší výrazné možnosti zlepšení diagnostické přesnosti a efektivity, zároveň však klade nové nároky na regulaci, klinickou validaci a praktickou implementaci. Radiologická praxe se postupně vyvíjí směrem k modelu, v němž člověk a algoritmus spolupracují, což umožňuje efektivnější a konzistentnější hodnocení zobrazovacích metod a vytváří předpoklady pro další rozvoj inteligentních podpurných systémů v klinickém prostředí. V neposlední řadě se otevírají i nové směry vývoje – například využití generativních modelů a velkých jazykových modelů (LLM) v radiologii. Uvedené nástroje mohou generovat syntetická data pro trénink

algoritmů, automaticky vytvářet předběžné radiologické zprávy nebo simulovat různé klinické scénáře. Do budoucna by tak mohly dále rozšířit možnosti, jak AI podpoří radiology v jejich práci, a přispět k ještě komplexnější integraci inteligentních systémů do zdravotnické praxe.

## 1.2 Skeletální zobrazování a typické diagnostické úlohy

Rentgenové vyšetření skeletu představuje základní zobrazovací metodu v diagnostice muskuloskeletálních poranění a onemocnění. Díky své široké dostupnosti, relativně nízkým nákladům a omezené radiační zátěži je rentgenové snímkování využíváno jako první linie vyšetření zejména v urgentní medicíně, traumatologii a ortopedii (Yu, 2015). Ve srovnání s výpočetní tomografií umožňuje rychlé zhodnocení stavu pacienta a rozhodnutí o dalším diagnostickém či terapeutickém postupu.

Navzdory těmto výhodám představuje hodnocení skeletálních rentgenových snímků významnou diagnostickou výzvu. Některé typy fraktur, zejména jemné fisury, únavové zlomeniny nebo léze v anatomicky složitých oblastech, mohou být na prostých snímcích obtížně rozpoznatelné. Například vlasové fraktury v oblasti kosti člunkové na zápěstí či nedislokované zlomeniny krčku femuru se na úvodních rentgenových snímcích často projeví jen minimálními známkami a mohou snadno uniknout pozornosti. Studie ukazují, že přehlédnuté fraktury patří mezi nejčastější příčiny diagnostických pochybení v urgentních příjmech, přičemž mohou vést k opožděné léčbě, zhoršení prognózy a v krajních případech i k právním sporům (Guly, 2001; Pinto et al., 2018). Chybná interpretace rentgenových snímků (zejména opomenutí patrné abnormality) přispívá k celkovému odhadu 3–5 % míry omylů v radiologii na denní bázi. V cíleně zaměřených studiích, které sledují výskyt chyb u specifických diagnóz nebo v určitých provozech, byly navíc zaznamenány ještě vyšší podíly opomenutých nálezů. Přes veškerá technická zlepšení zobrazovacích přístrojů tak zůstává lidský faktor limitujícím prvkem – radiolog může některé nálezy přehlédnout zejména tehdy, je-li pod časovým tlakem nebo čelí-li nadměrnému množství vyšetření.

Dalším faktorem ovlivňujícím kvalitu diagnostiky je rostoucí pracovní zátěž radiologů. Dlouhodobé trendy poukazují na nárůst počtu zobrazovacích vyšetření, který není vždy plně kompenzován odpovídajícím zvýšením počtu odborníků (Bhargavan et al., 2009). Zvyšující se objem práce na jednoho radiologa vyvolává obavy z efektu přetížení, protože nadměrná pracovní nálož může přispívat k únavě, chybovosti a profesnímu vyhoření. Experimentální studie zaměřené na vliv délky pracovní doby ukazují, že po dlouhém pracovním dni dochází ke statisticky významnému poklesu schopnosti detekovat patologické nálezy, včetně fraktur (Krupinski et al., 2010).

Otázka limitů pracovní zátěže a rychlosti interpretace vyšetření je proto předmětem odborné diskuse, přičemž je zdůrazňována potřeba vyvážení efektivity provozu a kvality diagnostiky (Alexander et al., 2022). Kombinace časového tlaku, vysokého objemu vyšetření a variability kvality snímků (ovlivněné např. přítomností artefaktů, implantátů či nestandardních projekcí) může zvyšovat riziko diagnostických pochybení a snižovat konzistenci hodnocení mezi jednotlivými hodnotiteli. Vysoká pracovní zátěž navíc omezuje prostor pro druhé čtení či konzultaci složitých případů, což může mít dopad na celkovou bezpečnost diagnostického procesu.

Specifickou výzvou skeletální radiologie je rovněž heterogenita vyšetřovaných anatomických oblastí a projekcí. Kostra zahrnuje desítky různých kostí a kloubů, z nichž každý má svou specifickou anatomii a vyžaduje odlišné projekce při rentgenovém zobrazení. Rozdílná kvalita snímků (např. vlivem tělesné konstituce pacienta či expozičních podmínek), přítomnost artefaktů (srůsty, šумы, kovové implantáty) nebo neobvyklých projekcí kladou vysoké nároky na zkušenost a pozornost hodnotitele. Například přítomnost ortopedických implantátů může částečně překrývat oblast zájmu a znesnadnit odhalení drobných linií lomů. Snímky drobných kostí zápěstí či hlezna v nestandardních incidencích zase vyžadují dobrou prostorovou představivost. V tomto kontextu se objevuje prostor pro využití umělé inteligence jako podpůrného nástroje, který může pomoci stabilizovat diagnostický výkon, snížit chybovost a zvýšit efektivitu práce v klinickém provozu. AI může například automaticky upozornit na podezřelou linii připomínající zlomeninu i v případech, kdy je nález nenápadný nebo překrytý artefaktem, a zajistit tak, že neunikne další pozornosti lékaře. V důsledku toho může být diagnostika skeletálních úrazů a onemocnění konzistentnější napříč různými pracovišti i různými časy během pracovního dne.

### 1.3 Nástroje umělé inteligence pro hodnocení skeletálních rentgenových snímků

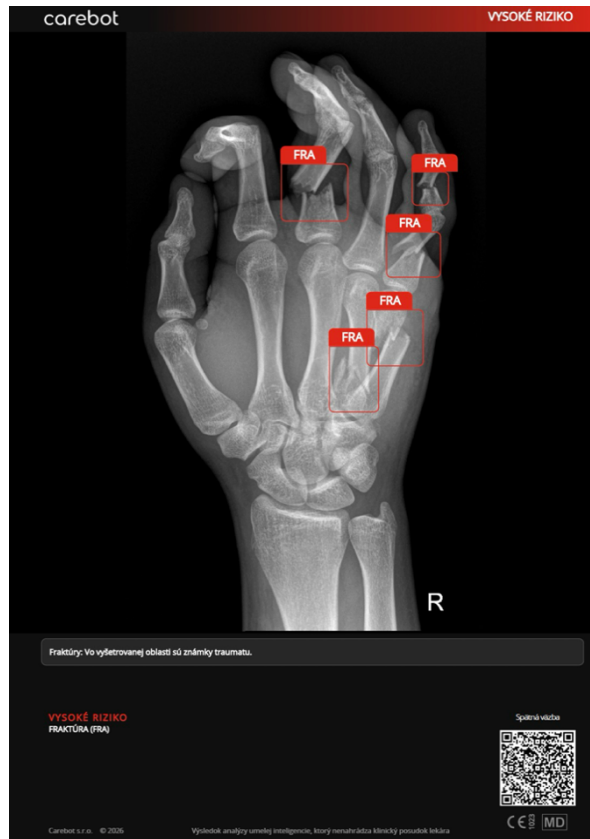
Nástroje umělé inteligence určené pro analýzu skeletálních rentgenových snímků tvoří heterogenní skupinu systémů, které se liší algoritmickým přístupem, rozsahem podporovaných indikací i způsobem zapojení do klinického workflow. Převážná část současných řešení je založena na metodách hlubokého učení, zejména na konvolučních neuronových sítích, které umožňují automatickou extrakci obrazových rysů a jejich následnou klasifikaci (Kutbi, 2024). Jednotlivé AI systémy se liší nejen architekturou modelů (např. čistě konvoluční sítě vs. hybridní modely kombinující více přístupů), ale také rozsahem anatomických oblastí, které jsou schopny analyzovat, a typu výstupů, jež uživatelům poskytují. Některé nástroje se zaměřují úzce na konkrétní regiony skeletu – například detekci zlomenin končetinových kostí – zatímco jiné se snaží pokrýt širší spektrum skeletálních struktur včetně osového skeletu (páteře) či pánevního pletence. Odlišnosti se projevují rovněž ve formě výstupů: některé algoritmy poskytují pouze informaci o přítomnosti či nepřítomnosti suspektního nálezu, jiné generují lokalizační mapy či zvýrazněné oblasti v obraze (tzv. heatmapy) indikující místo možného nálezu a další mohou nabízet doplňkové funkce usnadňující interpretaci nálezu v klinické praxi.

Rozsah podporovaných klinických úloh se postupně rozšiřuje. Zpočátku se komerčně dostupné AI systémy soustředily převážně na detekci akutních patologických nálezů, zejména zlomenin. V poslední době však přibývají i nástroje zaměřené na chronické či degenerativní stavy a kvantitativní hodnocení. Jako příklad lze uvést aplikace pro automatické měření zakřivení páteře u skoliózy nebo hodnocení deformit končetin pomocí úhlových parametrů – uvedené funkce směřují k podpoře ortopedické diagnostiky nad rámec pouhého záchytu akutních úrazů. Jiným případem je využití AI pro odhad kostního věku z rentgenů zápěstí, či predikce rizika osteoporózy na základě analýzy trabekulární struktury kosti na běžném snímku hrudníku. Rozšiřující se funkcionality ukazují, že AI nástroje pro skeletální radiologii mohou pokrýt jak urgentní diagnostiku (např. detekce fraktur a dislokací), tak plánovanou péči (např. měření progresu skoliózy nebo vyhledávání známek chronických onemocnění kostí).

Komerčně dostupné AI systémy se liší nejen architekturou modelů, ale také rozsahem anatomických oblastí, které jsou schopny analyzovat. Některé nástroje se zaměřují na konkrétní regiony, například horní či dolní končetiny, jiné se snaží pokrýt širší spektrum skeletálních struktur. Rozdíly se projevují rovněž ve formě výstupů poskytovaných uživateli, které mohou zahrnovat pravděpodobnostní skóre, lokalizační mapy nebo strukturované reporty integrované přímo do obrazové dokumentace.

Klíčovým aspektem klinického využití umělé inteligence je způsob, jakým je nástroj zapojen do stávající infrastruktury PACS/RIS. Efektivní nasazení vyžaduje kompatibilitu se standardy DICOM, podporu bezpečné komunikace mezi systémy a minimalizaci zásahů do zavedených pracovních postupů. Řešení, která umožňují automatické načtení snímků z PACS, jejich analýzu na pozadí a následné navrácení výsledků zpět do PACS ve formě, na jakou jsou radiologové zvyklí (např. jako další série obrázků se zvýrazněnými nálezy či jako připojený strukturovaný report), jsou zpravidla uživatelsky nejpřívětivější. Naproti tomu nástroje vyžadující manuální export dat do externí aplikace nebo čekání na výsledky mimo běžné workflow narážejí na menší ochotu k užívání. Z pohledu koncového uživatele – radiologa – by ideální AI systém měl fungovat víceméně neviditelně na pozadí: snímky by byly analyzovány ihned po pořízení a při otevření v prohlížeči by již byly k dispozici výsledky (např. graficky vyznačená podezřelá místa a pravděpodobnostní hodnocení nálezu). Takový způsob integrace maximalizuje přínos AI, neboť poskytuje informace bez prodlení a bez dodatečné námahy ze strany lékaře.

Jako ilustrativní příklad technické implementace umělé inteligence do klinického workflow lze uvést systém Carebot AI Bones, který demonstruje přístup založený na plné integraci do DICOM infrastruktury a generování strukturovaných výstupů kompatibilních s PACS (Carebot, 2022). Systém zpracovává skeletální rentgenové snímky přímo v rámci nemocniční sítě – obrazy jsou po získání automaticky odeslány ke zpracování a algoritmus následně vrací výsledek jako další DICOM objekt. Výstupem může být například kopie původního snímku s barevně vyznačenými oblastmi, kde byla detekována potenciální patologie, doplněná o textovou zprávu shrnující nálezy.



**Obr. 1: Výstup AI systému Carebot Bones při analýze rentgenového snímku ruky. Detekce fraktur je označena červenými rámečky s popiskem „FRA“; systém zároveň vyhodnotil celkové riziko jako vysoké.**

*Zdroj: Vlastní úprava, 2026, s laskavým svolením Carebot s.r.o.*

Radiolog si tak může v PACS prohlédnout výsledky AI podobně jako jakýkoli jiný snímek či sérii. Uvedený příklad slouží k ilustraci jednoho z možných technických přístupů k implementaci AI nástrojů v radiologii a nepředstavuje hodnocení výkonu ani srovnání s jinými dostupnými systémy. Na trhu existuje řada dalších řešení, jako jsou například algoritmy BoneView od společnosti Gleamer či Rayvolve od AZmed, zaměřené na detekci fraktur dlouhých kostí, nebo systém Radiobotics cílící na hodnocení změn kolenního kloubu u osteoartrózy. Každý z těchto nástrojů má odlišný rozsah indikací a lišit se může i přístup k prezentaci výsledků (od prostého upozornění na „suspektní“ snímek až po detailní popis nálezu). Všechny však spojuje nutnost důkladné klinické validace a snaha o co nejlepší souhru s lidským uživatelem.

## 1.4 Výkonové metriky AI (senzitivita, specificita, věrohodnostní poměry)

Hodnocení výkonnosti nástrojů umělé inteligence v radiologii vyžaduje použití standardizovaných metrik, které umožňují objektivní posouzení diagnostické přesnosti a následné porovnání různých systémů. Mezi nejčastěji používané ukazatele patří senzitivita, specificita, pozitivní a negativní prediktivní hodnota a plocha pod ROC křivkou (AUC), které poskytují informace o schopnosti algoritmu správně identifikovat patologické i normální nálezy (Saito & Rehmsmeier, 2015).

Tabulka 1 uvádí definice základních metrik diagnostické přesnosti používaných při hodnocení modelů umělé inteligence v medicínském zobrazování. Senzitivita (citlivost) vyjadřuje schopnost testu odhalit nemocné – tedy podíl případů s danou patologií, které algoritmus správně označí jako pozitivní nález. Specificita (specifičnost) naopak popisuje schopnost správně identifikovat zdravé – podíl případů bez patologie, které algoritmus správně vyhodnotí jako negativní nález. Základní ukazatele doplňují pozitivní a negativní věrohodnostní poměr (PLR a NLR), které jsou nezávislé na prevalenci onemocnění a poskytují robustní odhad klinické spolehlivosti algoritmu. Pro celkové zhodnocení výkonnosti se následně využívá vyvážená přesnost (BA), jež představuje průměr senzitivity a specifity a zabraňuje zkreslení u nevyvážených datových souborů. Vzorce pro výpočet vycházejí z kombinace hodnot TP (true positives), FP (false positives), FN (false negatives) a TN (true negatives), které tvoří základní prvky konfuzní matice.

Tab. 1: Seznam výkonových metrik

<b>Metrika</b>	<b>Definice (výpočet)</b>	<b>Interpretace (klinický význam)</b>
<b>Senzitivita</b> (citlivost)	$\frac{TP}{TP + FN}$	Schopnost testu správně odhalit jedince s onemocněním. Vysoká senzitivita znamená, že algoritmus zachytí většinu skutečných pozitivních případů (málo falešně negativních); je žádoucí u screeningových a urgentních vyšetření, kde opomenutí patologie může mít vážné následky.
<b>Specificita</b> (specifičnost)	$\frac{TN}{TN + FP}$	Schopnost testu správně identifikovat jedince bez onemocnění. Vysoká specificita znamená, že algoritmus většinou správně označí zdravé jako negativní (málo falešně pozitivních); je důležitá pro omezení falešných poplachů a zbytečných navazujících vyšetření.
<b>Pozitivní věrohodnostní poměr (PLR)</b>	$\frac{\text{Senzitivita}}{1 - \text{Specificita}}$	Udává, kolikrát je pravděpodobnější pozitivní nález AI u pacienta se skutečnou frakturou než u zdravého. Hodnoty nad 10 značí excelentní test a znamenají výrazné potvrzení diagnózy.
<b>Negativní věrohodnostní poměr (NLR)</b>	$\frac{1 - \text{Senzitivita}}{\text{Specificita}}$	Vyjadřuje spolehlivost negativního výsledku. Čím blíže je hodnota nule (ideálně pod 0,1), tím spolehlivěji dokáže AI vyloučit patologii a lékař může pacienta s jistotou propustit.
<b>Vyvážená přesnost (BA)</b>	$\frac{\text{Senzitivita} + \text{Specificita}}{2}$	Průměr senzitivity a specifity. Poskytuje objektivní zhodnocení celkové přesnosti algoritmu, a to i na nevyvážených datových souborech (např. při obrovské převaze zdravých snímků nad patologiemi).

Zdroj: Vlastní zpracování, 2026

Ve výsledcích hodnocení AI systémů se často uvádí také FPR (False Positive Rate, neboli  $1 - \text{Specificita}$ ) a FNR (False Negative Rate, neboli  $1 - \text{Senzitivita}$ ), případně další odvozené metriky, avšak uvedené ukazatele lze snadno odvodit z výše zmíněných základních hodnot. Pro komplexní posouzení modelu se používá i celková přesnost (podíl všech správně klasifikovaných případů z celku) a křivka ROC s plochou pod křivkou (AUC), která znázorňuje kompromis mezi senzitivitou a specificitou napříč různými možnými prahy rozhodnutí algoritmu.

Interpretace získaných metrik musí být vždy zasazena do klinického kontextu. Vysoká senzitivita může být klíčová v prostředí urgentní medicíny či screeningu, kde je prioritou minimalizace přehlédnutých patologických nálezů (tj. minimalizace falešně negativních výsledků). Naopak vysoká specifita může snižovat počet falešně pozitivních nálezů a tím omezovat zbytečná další vyšetření – to je důležité například u diagnostiky, kde každý falešný poplach může vést k invazivnímu doplňujícímu zákroku či zbytečné psychické zátěži pacienta. Zatímco tradiční prediktivní hodnoty bývají silně závislé na prevalenci hodnocené patologie v populaci a mohou u menších či nevyvážených datových souborů poskytovat zkreslený obraz, věrohodnostní poměry (PLR a NLR) daný nedostatek efektivně eliminují. Jejich nezávislost na četnosti výskytu patologie z nich činí vysoce spolehlivé ukazatele skutečného klinického přínosu AI nástrojů nezávisle na charakteru sledované kohorty. Vyvážená přesnost navíc zajišťuje, že celkový odhad výkonnosti nebude uměle nadhodnocen v situacích, kdy v testovacím vzorku drtivě převažují zdravé snímky nad těmi s patologií.

Pro zvýšení validity hodnocení se uplatňují multicentrické a multi-reader studie, které umožňují zohlednit variabilitu dat, rozdílné zobrazovací protokoly i individuální přístup hodnotitelů. Takový přístup přispívá k vyšší generalizovatelnosti výsledků a poskytuje realističtější obraz chování algoritmu v klinické praxi. Nedílnou součástí hodnocení je také definice referenčního standardu, podle něhož se výkonnost AI posuzuje. Referenční neboli zlatý standard bývá stanovován konsenzem zkušených odborníků (radiologů), případně je kombinován s dlouhodobým klinickým sledováním či konfrontací s peroperačními nebo histopatologickými nálezy. Důsledná definice referenční pravdy je nezbytná pro objektivní zjištění senzitivity a specifity – pokud by byl například chybný samotný referenční nález (např. radiologové v konziliu přehlédli drobnou frakturu), mohl by algoritmus paradoxně obdržet „trestné body“ za to, že danou patologii správně identifikoval. Kvalitní studie výkonnosti AI proto dbají na robustní nastavení referenčního standardu a externí validaci modelu na nezávislých datech, což omezuje riziko systematické chyby a lépe odráží skutečnou klinickou použitelnost algoritmu (Salehinejad et al., 2021). Zároveň je vhodné výsledky hodnotit nejen celkovými metrikami, ale i pomocí detailnějších analýz (např. lokalizační přesnost nálezů, hodnocení podle obtížnosti případů apod.), které mohou odhalit silné a slabé stránky posuzovaných systémů.

## 1.5 Vnímání a důvěra radiologů v technologie AI – vysvětlitelnost a etika

Úspěšné zavádění umělé inteligence do radiologické praxe je podmíněno nejen technickými parametry, ale také mírou důvěry uživatelů. Radiologové vnímají umělou inteligenci především jako podpůrný nástroj, jehož výstupy musí být klinicky interpretovatelné a transparentní. Schopnost porozumět algoritmičtým rozhodnutím hraje zásadní roli při jejich začlenění do každodenní praxe. Jak již bylo zmíněno, vysvětlitelnost algoritmů představuje jeden z klíčových faktorů ovlivňujících přijetí umělé inteligence. Metody vizualizace rozhodovacích procesů – například zvýraznění relevantních oblastí obrazu (tzv. saliency maps) – mohou přispět ke zvýšení důvěry a usnadnit kontrolu výstupů ze strany lékaře. Pokud AI systém označí na snímku konkrétní místo jako podezřelé a zároveň tuto oblast graficky zvýrazní, umožní tím radiologovi rychle posoudit, zda zachytil skutečnou patologii, nebo se „spletl“ vlivem artefaktu či běžné struktury. Tím se spolupráce mezi člověkem a AI stává přehlednější a lékař se může

na algoritmus lépe spolehnout. Nedostatečná transparentnost může naopak vést k odmítání systémů i v případech, kdy dosahují vysoké diagnostické přesnosti – uživatelé se zdráhají důvěřovat tzv. „černé skříňce“, z níž vypadává odpověď bez jasného zdůvodnění. Empirické studie ukazují, že míra důvěry radiologů v umělou inteligenci významně souvisí s transparentností modelu, způsobem prezentace výsledků a srozumitelností jeho výstupů (Ewals et al., 2024). Přehledové práce zároveň identifikují mezi hlavní bariéry adopce AI nedostatečné porozumění principům fungování algoritmů, obavy z profesní odpovědnosti a nejistotu ohledně klinické validace systémů (Eltawil et al., 2023). Zde se opět ukazuje význam zapojení radiologů do vývoje a testování AI – uživatelská zpětná vazba může pomoci navrhnout systém tak, aby poskytoval pro kliniku smysluplné informace (např. „algoritmus detekoval zlomeninu distálního radia, protože identifikoval přerušeni kortikální linie na předozadní projekci“ místo pouhého „nález: zlomenina ano/ne“).

Etické aspekty využití umělé inteligence zahrnují otázky odpovědnosti za diagnostické rozhodnutí, ochrany osobních údajů a bezpečnosti dat. Zdravotnické AI systémy podléhají regulaci zdravotnických prostředků a musí splňovat požadavky na klinickou validaci, řízení rizik a dohledatelnost rozhodovacích procesů (European Union, 2017). Jasné vymezení role algoritmu v diagnostickém procesu je základním předpokladem pro dlouhodobé a odpovědné využívání umělé inteligence v radiologii. Z právního hlediska vyvstává otázka „Kdo nese odpovědnost za diagnózu, pokud se umělá inteligence mylí?“. Současná praxe staví finální odpovědnost na lékaře, který výstup AI využívá jako poradní stanovisko. Pokud by však došlo k diagnostické chybě na základě doporučení AI, je obtížné jednoznačně určit míru zavinění – nabízí se úvaha, zda odpovědnost částečně nesou i tvůrci a distributoři daného softwaru. Problematika je zatím spíše teoretická, neboť v praxi je AI vnímána jen jako podpůrný nástroj, nikoli autonomní diagnostik. Přesto do budoucna nelze vyloučit, že s rostoucí autonomií algoritmů budou muset legislativní rámce upřesnit, jak nakládat s chybami umělé inteligence a kdo za ně ponese právní důsledky.

Další významnou oblastí je ochrana dat a soukromí. Vývoj a provoz AI systémů ve zdravotnictví vyžaduje přístup k velkým objemům klinických dat, často včetně obrazové dokumentace pacientů. Zajištění anonymizace těchto dat a respektování platné legislativy (např. GDPR v Evropě) je nezbytné pro udržení důvěry pacientů i zdravotníků. Je nutné zavést taková opatření, aby data využívaná k tréninku algoritmů nemohla být zpětně spojena s konkrétními pacienty. Kromě technických řešení (odstranění identifikačních údajů z DICOM souborů, rozostření obličejů na snímcích apod.) sem patří i etický požadavek informovaného souhlasu – pacient by měl být ideálně srozuměn s tím, že jeho data mohou být použita pro vývoj a testování AI nástrojů, a dát k tomu souhlas. V praxi se toto často ošetřuje rámcovým souhlasem při vstupu do studie či léčby v daném zařízení.

Za zmínku stojí i problematika možných biasů (zkreslení) v algoritmech a jejich dopad na etiku péče. Pokud je AI natrénována na nereprezentativních datech, může poskytovat systematicky horší výkony pro určité skupiny pacientů (např. odlišné výsledky pro různé etnické či věkové skupiny, pokud nebyly dostatečně zastoupeny v tréninku). To by mohlo vést k nerovnému přístupu v péči. Proto je doporučováno, aby dataset pro trénink i validaci AI byl co nejpestřejší a zahrnoval různorodé případy z více pracovišť – jedině tak lze ověřit, že algoritmus funguje spolehlivě napříč populací a za různých podmínek. Zde opět hraje roli regulátor, který při schvalování AI jako zdravotnického prostředku posuzuje i uvedené aspekty (FDA, 2022).

Celkově platí, že míra důvěry radiologů v AI technologie bude do značné míry ovlivněna tím, nakolik se dané nástroje osvědčí jako přínosné a bezpečné v praxi. Časné pilotní studie často ukázaly potenciál AI zlepšit detekci nálezů či urychlit práci, avšak pro širší přijetí bude třeba dalšího vývoje zaměřeného na uživatelskou přívětivost a spolehlivost. Transparentní implementace AI v radiologii – zahrnující jasně definované místo algoritmu v diagnostickém postupu, možnosti kontroly jeho výstupů a zajištění jeho správné funkce – představuje základní předpoklad pro dlouhodobé a odpovědné využívání těchto technologií. Radiologové by měli být nejen koncovými uživateli, ale i aktivními účastníky procesu zavádění AI, aby mohli formulovat své požadavky a obavy. Otevřený dialog mezi vývojáři AI a zdravotnickou obcí je nezbytný k tomu, aby nové nástroje skutečně reflektovaly potřeby klinické praxe a byly přijímány s důvěrou. Jak uvádí European Society of Radiology, radiologové musí v éře nástupu umělé inteligence převzít aktivní roli – vzdělávat se v oblasti AI, podílet se na vývoji standardů a usilovat o to, aby nasazení AI vždy probíhalo v souladu s etikou a se zaměřením na prospěch pacienta. Jedině tak bude AI v radiologii chápána jako vítaný pomocník, a ne jako hrozba či nekontrolovatelný černý box. V blízké budoucnosti se dá očekávat stále těsnější propojení práce radiologa a umělé inteligence – klíčové však je, aby toto propojení bylo založeno na důvěře, jasných pravidlech a průběžném vyhodnocování přínosů i rizik pro kvalitu péče o pacienty.

## 1.6 Přehled certifikovaných AI nástrojů v radiologii

Klinické nasazení nástrojů umělé inteligence v radiologii je v současnosti podmíněno jejich oficiální certifikací. V rámci Evropské unie musí být uvedené systémy posouzeny podle nařízení Medical Device Regulation (MDR) a označeny značkou CE, která potvrzuje jejich bezpečnost, účinnost a soulad s požadavky zdravotnického práva. Ve Spojených státech je za certifikaci odpovědný úřad Food and Drug Administration (FDA), jenž schvaluje konkrétní algoritmy jako tzv. Software as a Medical Device (SaMD).

Certifikované nástroje umělé inteligence se v radiologii využívají především k detekci a klasifikaci patologických struktur. V muskuloskeletální oblasti jde nejčastěji o identifikaci zlomenin, degenerativních změn a dalších skeletálních abnormalit na prostých rentgenových snímcích. Využívány jsou rovněž nástroje zaměřené na snímky plic, prsní žlázy nebo mozku. Rozsah funkcí se liší podle výrobce – některé systémy generují pravděpodobnostní skóre přítomnosti nálezu, jiné poskytují lokalizační mapy nebo strukturované zprávy přímo v rámci integrovaného prostředí PACS.

Většina certifikovaných systémů je postavena na algoritmech hlubokého učení, zejména konvolučních neuronových sítích. Klinická validace probíhá prostřednictvím retrospektivních nebo prospektivních studií, často s využitím referenčního standardu vytvořeného skupinou expertů. Certifikace bývá udělena na základě důkazů o bezpečnosti, výkonnosti a přínosu systému pro diagnostický proces.

Dostupnost certifikovaných nástrojů se v jednotlivých zemích liší v závislosti na regulačních požadavcích, komerční politice výrobce a míře integrace s existujícími nemocničními systémy. I přes rostoucí počet dostupných řešení zůstává výběr vhodného systému závislý na konkrétních klinických potřebách a technických možnostech pracoviště.

## 2 Metodika výzkumu

Kapitola detailně vymezuje metodologický rámec, který byl navržen a aplikován pro naplnění stanovených cílů práce. Výzkumný proces je logicky rozčleněn do několika fází, jež pokrývají jak exaktní měření diagnostické výkonnosti vybraných systémů, tak zkoumání postojů odborné veřejnosti. V následujících podkapitolách je podrobně popsán výzkumný design, způsob selekce a anonymizace testovaných nástrojů umělé inteligence, protokol pro získávání a přípravu klinických dat, definice referenčního standardu a v neposlední řadě také struktura nástrojů pro sběr kvalitativních dat.

### 2.1 Výzkumný design a teoretická východiska zvoleného přístupu

Pro komplexní uchopení problematiky integrace umělé inteligence do radiologické praxe byl zvolen smíšený výzkumný design (mixed-methods research). Zvolený metodologický přístup umožňuje efektivně propojit kvantitativní a kvalitativní rovinu zkoumání, čímž je dosaženo vyšší validity a komplexnosti výsledků. Samotná evaluace prediktivních schopností algoritmů poskytuje pouze dílčí pohled na jejich klinickou využitelnost. Jak upozorňují aktuální studie, vysoká diagnostická přesnost (accuracy) je podmínkou nutnou, nikoliv však postačující pro úspěšnou implementaci v reálném zdravotnickém prostředí (Tejani et al., 2024).

Kvantitativní část výzkumu je koncipována jako retrospektivní observační komparativní analýza. Její podstatou je objektivní zhodnocení výkonnosti tří nezávislých softwarových řešení na sdíleném anonymizovaném datovém souboru rentgenových snímků. Kvalitativní fáze je koncipována jako dotazníkové šetření zaměřené na postoje radiologů k využití AI nástrojů v klinické praxi. Zatímco kvantitativní data odpovídají na otázku technické spolehlivosti systémů, kvalitativní poznatky identifikují sociotechnické faktory ovlivňující uživatelskou akceptaci, vnímaná rizika a požadavky na transparentnost rozhodovacích procesů (van der Velden et al., 2022). Obě výzkumné linie se následně syntetizují a umožňují formulovat ucelená doporučení pro praxi.

### 2.2 Výběr, charakteristika a zaslepení testovaných AI nástrojů

Do srovnávací analýzy byly zařazeny tři komerčně dostupné nástroje umělé inteligence, které jsou primárně dedikovány pro automatizovanou detekci skeletálních patologií, zejména traumatických změn na prostých rentgenových snímcích. Základním a mandatorním inkluzním kritériem pro zařazení softwaru do testování byla existence platné certifikace pro klinické použití na území Evropské unie. Všechny hodnocené systémy tedy plně odpovídají požadavkům nařízení o zdravotnických prostředcích MDR a jsou opatřeny certifikační značkou CE (European Union, 2017). Uplatněním definovaného kritéria bylo zajištěno, že komparace neprobíhá na experimentálních modelech, nýbrž na produktech, které jsou v současnosti reálně nabízeny zdravotnickým zařízením.

Všechny tři systémy využívají k analýze obrazu architektury hlubokých neuronových sítí (Deep Learning), konkrétně konvoluční neuronové sítě (CNN), jež představují současný standard v oblasti medicínského počítačového vidění. Výstupem analyzovaných systémů je jednak binární

klasifikace (přítomnost či nepřítomnost traumatické změny), jednak přesná vizuální lokalizace detekovaných fraktur v rámci obrazu.

Z důvodu striktního dodržení legislativních podmínek, ochrany obchodního tajemství dodavatelů a licenčních ujednání bylo přistoupeno k úplné systematické anonymizaci všech hodnocených nástrojů. Všechny tři analyzované systémy jsou proto v celé práci označovány výhradně kódy AI 1, AI 2 a AI 3, bez uvedení jejich obchodních názvů, výrobce či jiných identifikačních znaků. Zvolené označení slouží pouze k rozlišení jednotlivých řešení a nepředstavuje žádné pořadí, preferenci ani hodnotící hierarchii. Uplatněná anonymizace přispívá k omezení rizika konfirmačního zkreslení (bias) při interpretaci výsledků a k posílení objektivity při komparaci výkonnostních metrik.

### 2.3 Protokol pro extrakci dat, anonymizace a strukturace datového souboru

Obrazová data nezbytná pro kvantitativní testování byla získána z databáze Fakultní nemocnice Olomouc. Zdrojem dat byl přímo nemocniční systém PACS. Vzhledem k objemu požadovaných datových položek a nutnosti optimalizace časového harmonogramu výzkumu byla provedena datová extrakce formou hromadného databázového exportu (DB dump). Zvolený způsob hromadného exportu je technicky zřetelně efektivnější než manuální či sekvenční stahování jednotlivých studií prostřednictvím standardního grafického rozhraní PACS a umožňuje zpracování rozsáhlejších obrazových datasetů.

Klíčovým a nezbytným krokem metodiky byla důsledná deidentifikace extrahovaných dat. Obrazové soubory v medicínském standardu DICOM (Digital Imaging and Communications in Medicine) obsahují rozsáhlá metadata. Veškeré identifikační údaje pacientů (jméno, rodné číslo, datum narození, identifikační číslo pacienta) a informace identifikující ošetřujícího lékaře či konkrétní oddělení byly z hlaviček DICOM souborů trvale a nevratně odstraněny (World Health Organization, 2021). Aplikovaná deidentifikace dat zajistila bezvýhradný soulad s nařízením GDPR (Obecné nařízení o ochraně osobních údajů) a s etickými standardy pro nakládání se zdravotnickou dokumentací pro výzkumné účely.

Pro sestavení validního a reprezentativního datového souboru byla definována striktní exkluzní kritéria. Cílem bylo vytvořit homogenní vzorek, na kterém lze adekvátně porovnat čistou prediktivní sílu modelů. Ze zdrojového datasetu byly proto systematicky vyřazeny:

- kontrolní rentgenové snímky pořízené přes sádrovou fixaci,
- snímky pacientů s nasazenými vnějšími fixatery nebo dlahami,
- rentgenogramy obsahující masivní kovové osteosyntetické materiály (šrouby, dlahy, hřeby).

Důvodem pro aplikaci vymezených exkluzních kritérií je skutečnost, že přítomnost cizích materiálů a artefaktů v obraze podstatně modifikuje distribuci pixelových intenzit. Vzniklé anomální hrany a vysoce kontrastní oblasti u modelů hlubokého učení často vedou k falešně pozitivním detekcím nebo naopak maskují subtilní linie fraktur, což zkresluje celkové hodnocení výkonnosti algoritmu (Kuo et al., 2022). Datový soubor je tedy tvořen výhradně nativními snímky skeletu. Pro kvantitativní část výzkumu byl využit soubor vycházející z vyšetření realizovaných

v období od 5. do 11. března 2025. Jeho strukturovaná podoba je k dispozici v elektronické podobě v Příloze B.1.

## 2.4 Kvantitativní vyhodnocení výkonnosti a definice referenčního standardu

Pro možnost výpočtu objektivních klasifikačních metrik bylo nezbytné stanovit absolutní referenční standard (tzv. ground truth), vůči kterému jsou výstupy umělé inteligence porovnávány. V rámci navrženého výzkumu byly jako referenční standard zvoleny původní definitivní radiologické popisy (reporty), které byly vytvořeny atestovanými lékaři Fakultní nemocnice Olomouc v čase vzniku snímku. Popsaný evaluační model vychází z konceptu RWE (Real-World Evidence) a konfrontuje schopnosti umělé inteligence přímo se standardem běžné zdravotnické péče.

Proces hodnocení probíhal porovnáním predikce systému (AI 1, AI 2, AI 3) s referenčním standardem. Pro účely statistické analýzy byly následně kvantifikovány základní ukazatele diagnostické přesnosti, jejichž detailní teoretické vymezení a matematická definice jsou podrobně rozepsány v podkapitole 1.4 práce. Konkrétně byly vyhodnocovány parametry senzitivity, specifity, pozitivního a negativního věrohodnostního poměru (PLR a NLR) a vyvážené přesnosti (Balanced Accuracy).

U všech tří testovaných systémů byla data statisticky analyzována na bázi matice záměn (confusion matrix) s cílem odhalit průkazné diference v chybovosti, a to především v distribuci falešně pozitivních (FP) a falešně negativních (FN) výsledků.

## 2.5 Nástroje a postupy pro sběr kvalitativních dat

Kvalitativní šetření, mapující interakci lidského faktoru se systémy AI, bylo založeno na strukturovaném elektronickém dotazníku, jehož tvorba a následná distribuce probíhala prostřednictvím platformy Google Forms. Dotazníkové šetření bylo zacíleno na českou radiologickou komunitu, přičemž zkoumanou populaci tvořili lékaři v předatestační přípravě i atestovaní odborníci napříč spektrem zdravotnických zařízení.

Konstrukce dotazníku byla logicky rozdělena do tří klíčových sekcí, jejichž přehled a obsahové zaměření detailně ilustruje Tabulka 2. Kompletní znění použitého dotazníkového formuláře je k dispozici v Příloze A.1. Datový soubor výsledků dotazníkového šetření mezi radiology je k dispozici v elektronické podobě v Příloze B.2.

Tab. 2: Struktura dotazníkového šetření

<b>Sekce dotazníku</b>	<b>Zkoumaná oblast</b>	<b>Zaměření a typ položek</b>
<b>Demografický a expertní profil</b>	Profesní zázemí a zkušenosti respondentů	Délka praxe, typ zdravotnického pracoviště, úroveň dosavadních praktických zkušeností s AI technologiemi.
<b>Vnímání spolehlivosti a vysvětlitelnosti (XAI)</b>	Důvěra k systému typu „black-box“	Měření ochoty spoléhat na algoritmická rozhodnutí. Vyhodnocení nezbytnosti vizuálních zdůvodnění pro zvýšení diagnostické jistoty (hodnoceno primárně prostřednictvím standardizovaných Likertových škál).
<b>Procesní a legislativní bariéry</b>	Úskalí implementace do klinické praxe	Obavy ze sdílené profesní odpovědnosti za diagnostickou chybu, ergonomické požadavky na integraci AI výstupů přímo do pracovního prostředí (RIS/PACS) (Kanakaraj et al., 2022).

*Zdroj: Vlastní zpracování, 2026*

Původně zvažované doplnění dotazníkového šetření o polostrukturované rozhovory s experty nebylo z organizačních důvodů realizováno, a proto nebylo zahrnuto do analytické části práce.

### 3 Výsledky a analýza dat

Kapitola prezentuje zjištění získaná na základě definovaného metodologického postupu. První část analyzuje reálnou diagnostickou výkonnost tří vybraných softwarových řešení na retrospektivním souboru rentgenových snímků. Následující sekce předkládá výsledky dotazníkového šetření mapujícího postoje a vnímání radiologické komunity k integraci algoritmů do klinického prostředí.

#### 3.1 Kvantitativní část – výsledky testování AI nástrojů

Základem pro objektivní komparativní analýzu byl retrospektivní datový soubor získaný z obrazového archivu PACS Fakultní nemocnice Olomouc. Po aplikaci exkluzních kritérií vznikl finální testovací vzorek o celkovém rozsahu 300 nativních rentgenových snímků skeletu. Z demografického hlediska byl soubor genderově absolutně vyvážený (150 mužů a 150 žen) a pokrýval široké věkové rozpětí od 3 do 99 let. Popsané věkové rozložení zajišťuje vysokou variabilitu kostní struktury, od vyvíjejícího se dětského skeletu až po osteoporotické změny u seniorů, což představuje klíčový předpoklad pro robustní otestování algoritmů.

Zkoumaný obrazový materiál zahrnoval široké spektrum anatomických lokalizací. Nejpočetněji byla zastoupena oblast kolene a nohy, následovaná snímky ruky a zápěstí. Detailní distribuci vyšetřovaných částí těla ilustruje Tabulka 3.

**Tab. 3: Distribuce vyšetřovaných anatomických oblastí**

<b>Vyšetřovaná oblast (Body Part)</b>	<b>Absolutní četnost (n)</b>	<b>Relativní četnost (%)</b>
<b>Koleno / Noha (Knee/Leg)</b>	191	63,67
<b>Ruka / Zápěstí (Hand/Wrist)</b>	44	14,67
<b>Rameno / Klíční kost (Shoulder/Clavicle)</b>	25	8,33
<b>Noha / Kotník (Foot/Ankle)</b>	21	7,00
<b>Loket / Paže (Elbow/Arm)</b>	10	3,33
<b>Pánev / Kyčelní kloub (Pelvis/Hip Joint)</b>	9	3,00
<b>Celkem</b>	300	100,00

*Zdroj: Vlastní zpracování, 2026*

Komparací s referenčním standardem atestovaných radiologů byla stanovena přesná distribuce nálezů. Z celkového počtu 300 vyšetření obsahovalo 117 snímků reálnou patologii (frakturu), zatímco 183 snímků představovalo fyziologický nález bez strukturálních změn.

Při analýze diagnostické úspěšnosti vykazoval systém AI 1 mimořádnou spolehlivost. Z celkového počtu pozitivních případů nedokázal algoritmus rozpoznat pouze pět patologií. Zároveň generoval naprosté minimum falešných poplachů. Kompletní přehled absolutních hodnot z matice záměn a z nich odvozených výkonnostních metrik pro systém AI 1 poskytuje Tabulka 4.

Tab. 4: Výsledky diagnostické přesnosti – systém AI 1

<b>Sledovaný parametr</b>	<b>Dosažená hodnota</b>
<b>Pravdivě pozitivní nálezy (TP)</b>	112
<b>Falešně negativní nálezy (FN)</b>	5
<b>Pravdivě negativní nálezy (TN)</b>	179
<b>Falešně pozitivní nálezy (FP)</b>	4
<b>Senzitivita</b>	95,73 %
<b>Specifická</b>	97,81 %
<b>Pozitivní věrohodnostní poměr (PLR)</b>	43,79
<b>Negativní věrohodnostní poměr (NLR)</b>	0,04
<b>Vyvážená přesnost (Balanced Accuracy)</b>	96,77 %

Zdroj: Vlastní zpracování, 2026

Systém AI 2 dosáhl v rámci testování rovněž vysoce nadprůměrných výsledků, nicméně v porovnání s prvním hodnoceným modelem u něj byla zaznamenána mírně vyšší chybovost v záchytu reálných patologií (osm falešně negativních případů). Detailní statistické rozložení predikcí systému AI 2 sumarizuje Tabulka 5.

Tab. 5: Výsledky diagnostické přesnosti – systém AI 2

<b>Sledovaný parametr</b>	<b>Dosažená hodnota</b>
<b>Pravdivě pozitivní nálezy (TP)</b>	109
<b>Falešně negativní nálezy (FN)</b>	8
<b>Pravdivě negativní nálezy (TN)</b>	178
<b>Falešně pozitivní nálezy (FP)</b>	5
<b>Senzitivita</b>	93,16 %
<b>Specifická</b>	97,27 %
<b>Pozitivní věrohodnostní poměr (PLR)</b>	34,10
<b>Negativní věrohodnostní poměr (NLR)</b>	0,07
<b>Vyvážená přesnost (Balanced Accuracy)</b>	95,22 %

Zdroj: Vlastní zpracování, 2026

U třetího hodnoceného nástroje, kódovaného jako AI 3, proběhl výrazný propad diagnostické výkonnosti ve všech měřených parametrech. Algoritmus selhal při identifikaci 31 reálných fraktur a současně vygeneroval trojnásobné množství falešných poplachů oproti předchozím řešením. Přehled naměřených hodnot ukazuje Tabulka 6.

Tab. 6: Výsledky diagnostické přesnosti – systém AI 3

<b>Sledovaný parametr</b>	<b>Dosažená hodnota</b>
<b>Pravdivě pozitivní nálezy (TP)</b>	86
<b>Falešně negativní nálezy (FN)</b>	31
<b>Pravdivě negativní nálezy (TN)</b>	168
<b>Falešně pozitivní nálezy (FP)</b>	15
<b>Senzitivita</b>	73,50 %
<b>Specifická</b>	91,80 %
<b>Pozitivní věrohodnostní poměr (PLR)</b>	8,97
<b>Negativní věrohodnostní poměr (NLR)</b>	0,29
<b>Vyvážená přesnost (Balanced Accuracy)</b>	82,65 %

Zdroj: Vlastní zpracování, 2026

Závěrečné souhrnné srovnání klíčových diagnostických metrik prezentuje Tabulka 7.

Tab. 7: Komparace výkonnostních metrik testovaných AI systémů

<b>Systém</b>	<b>Senzitivita</b>	<b>Specifická</b>	<b>PLR</b>	<b>NLR</b>	<b>Vyvážená přesnost</b>
<b>AI 1</b>	95,73 %	97,81 %	43,79	0,04	96,77 %
<b>AI 2</b>	93,16 %	97,27 %	34,10	0,07	95,22 %
<b>AI 3</b>	73,50 %	91,80 %	8,97	0,29	82,65 %

Zdroj: Vlastní zpracování, 2026

Při přímém srovnání testovaných systémů vystupují do popředí dvě zcela odlišné úrovně vyspělosti certifikovaných algoritmů. Systémy AI 1 a AI 2 dosáhly v analyzovaném souboru nejvyšších hodnot sledovaných metrik. Hodnota PLR nad 10 je v medicíně považována za znak vysoce přesného testu – výsledek prvních dvou modelů přesahující hodnotu 34 znamená, že při upozornění umělé inteligence může být pozitivní výstup těchto systémů spojen s vysokou pravděpodobností přítomnosti patologie.

Naopak systém AI 3 za konkurencí markantně zaostává. Vysoká hodnota NLR (0,29) prokazuje riziko – ani pokud algoritmus označí snímek jako čistý, nemůže se na výsledek lékaře s jistotou spolehnout a hrozí opomenutí skryté fraktury. Naměřená výkonnostní mezera empiricky dokazuje fakt, že samotná existence certifikace CE u softwaru automaticky negarantuje vysokou diagnostickou robustnost v rutinní klinické praxi.

### 3.1.1 Analýza diskrepantních nálezů

Pro hlubší pochopení limitů jednotlivých modelů hlubokého učení byly z testovacího datasetu vybrány reprezentativní kazuistiky, u nichž došlo k neshodě mezi algoritmičnou predikcí a reálným nálezem.

**Kazuistika 1: Falešně negativní nález při detekci deprese laterálního tibiálního plata**

Referenční hodnocení rentgenového snímku pravého kolenního kloubu potvrdilo depresi laterálního tibiálního plata bez výraznější dislokace, současně se známkami gonartrózy. Anamnesticky byl přítomen pád na pravé koleno a následné přetrvávání obtíží, které vedlo k dalšímu terapeutickému řešení. Na snímku se tak současně uplatňovaly akutní posttraumatické změny i chronická degenerativní přestavba kloubu.

Při porovnání výstupů jednotlivých systémů bylo zjištěno, že AI 1 a AI 3 označily snímek správně jako patologický, a odpovídaly tedy klasifikaci true positive. Systém AI 2 naopak nález nevyhodnotil jako suspektní, přestože fraktura byla referenčním standardem potvrzena. Výsledek AI 2 proto odpovídal falešně negativní klasifikaci.

Z hlediska interpretace je podstatné, že traumatický nález vznikl v terénu již přítomných degenerativních změn. Právě kombinace artrózy, sklerotizace a nepravidelnosti kloubních kontur může snižovat přehlednost oblasti laterálního plata tibie a ztěžovat rozpoznání impresivní složky fraktury. Negativní algoritmický výstup by v podobné situaci mohl vést k oslabení podezření na akutní poranění, a potvrzuje tak nutnost odborného hodnocení snímku v souvislosti s klinickým stavem pacienta.



**Obr. 2: Rentgenový snímek pravého kolenního kloubu s depresí laterálního tibiálního plata**

*Zdroj: anonymizovaný RTG snímek z PACS archivu FNOL, vlastní zpracování, 2026*

**Tab. 8: Klasifikace kazuistiky 1**

<b>Hodnotitel</b>	<b>Výsledek hodnocení</b>
<b>Referenční standard (Ground Truth)</b>	Pozitivní (potvrzeno lékařem)
<b>AI 1</b>	Detekováno (true positive)
<b>AI 2</b>	Nedetekováno (false negative)
<b>AI 3</b>	Detekováno (true positive)

*Zdroj: Vlastní zpracování, 2026*

### Kazuistika 2: Falešně negativní nález při detekci fraktur přednoží

Na rentgenovém snímku pravé nohy byly referenčním hodnocením potvrzeny fraktury baze proximálního článku II. a IV. prstu a současně fraktury hlaviček II.–V. metatarzu. Klinicky byly přítomny bolesti po došlapu a lokální citlivost, což odpovídalo posttraumatickému poškození přednoží. Jednalo se tedy o jednoznačně patologický snímek s vícečetným kostním poraněním.

Největší rozdíl mezi hodnocenými systémy se projevil v rozsahu zachycení nálezu. AI 2 a AI 3 vyhodnotily snímek správně jako pozitivní a odpovídaly klasifikaci true positive. Naproti tomu AI 1 patologii nezachytil, ačkoli přítomnost fraktur byla potvrzena referenčním standardem. U systému AI 1 se proto jednalo o falešně negativní výsledek.

Přednoží představuje anatomicky členitou oblast s vysokou koncentrací drobných kostních struktur, jejich vzájemným překryvem a proměnlivou přehledností v jednotlivých projekcích. Vícečetné poranění proto nemusí automaticky znamenat snadnější detekci, protože patologické změny mohou být rozloženy v několika sousedních lokalitách a nemusí vytvářet jeden dominantní obraz. Přehlédnutí podobného nálezu by mohlo vést k podcenění rozsahu poranění a k nepřesnému zhodnocení funkční závažnosti poškození.



**Obr. 3: Rentgenový snímek pravé nohy s frakturami přednoží**

*Zdroj: anonymizovaný RTG snímek z PACS archivu FNOL, vlastní zpracování, 2026*

**Tab. 9: Klasifikace kazuistiky 2**

<b>Hodnotitel</b>	<b>Výsledek hodnocení</b>
<b>Referenční standard (Ground Truth)</b>	Pozitivní (potvrzeno lékařem)
<b>AI 1</b>	Nedetkováno (false negative)
<b>AI 2</b>	Detkováno (true positive)
<b>AI 3</b>	Detkováno (true positive)

*Zdroj: Vlastní zpracování, 2026*

### Kazuistika 3: Falešně negativní nález při detekci stresové fraktury II. metatarzu

U třetího analyzovaného snímku levé nohy byla referenčním hodnocením potvrzena starší fraktura II. metatarzu v suprabazální oblasti vlevo. Doplnující vyšetření zároveň prokázalo edém v okolí baze II. metatarzu a obraz odpovídající hojení stresové zlomeniny. Patologický nález tedy nebyl spojen s výrazným obrazem akutního traumatu, ale spíše s méně nápadnými známkami již probíhající reparační reakce.

Rozdílnost výsledků jednotlivých AI systémů byla v tomto případě výrazná. AI 1 vyhodnotil snímek správně jako patologický a odpovídal klasifikaci true positive. Oproti tomu AI 2 i AI 3 nález neidentifikovaly, přestože byl potvrzen referenčním standardem. V obou případech tak šlo o falešně negativní výsledek.

Podstata obtížnosti spočívala především v charakteru samotného nálezu. Hojící se nebo stresová fraktura obvykle nevykazuje tak výrazné morfologické změny jako čerstvé dislokované poranění, a na prostém rentgenovém snímku proto může působit méně přesvědčivě. Negativní algoritmičtý výstup je v takové situaci rizikový zejména tím, že může snižovat pozornost věnovanou subtilním, ale klinicky relevantním změnám. Případ tím zdůrazňuje význam propojení obrazového hodnocení s anamnézou, klinickými obtížemi a výsledky doplňujících vyšetření.



**Obr. 4: Rentgenový snímek levé nohy se stresovou frakturou II. metatarzu**  
Zdroj: anonymizovaný RTG snímek z PACS archivu FNOL, vlastní zpracování, 2026

**Tab. 10: Klasifikace kazuistiky 3**

<b>Hodnotitel</b>	<b>Výsledek hodnocení</b>
<b>Referenční standard (Ground Truth)</b>	Pozitivní (potvrzeno lékařem)
<b>AI 1</b>	Detekováno (true positive)
<b>AI 2</b>	Nedetekováno (false negative)
<b>AI 3</b>	Nedetekováno (false negative)

Zdroj: Vlastní zpracování, 2026

#### Kazuistika 4: Falešně negativní nález při detekci fraktury hlavičky radia

Rentgenový snímek levého loketního kloubu byl referenčním hodnocením uzavřen jako fraktura hlavičky radia bez dislokace. V klinickém obrazu dominovala bolestivost a omezení hybnosti po úrazovém mechanismu pádu. Šlo tedy o nález, který byl v souladu jak s anamnézou, tak s odborným radiologickým závěrem.

V porovnání s ostatními případy byla zde chyba soustředěna pouze u jednoho systému. AI 1 a AI 2 klasifikovaly snímek správně jako pozitivní a odpovídaly klasifikaci true positive. AI 3 však patologii nedetekoval, ačkoli byla fraktura referenčním standardem potvrzena. Výstup systému AI 3 byl proto hodnocen jako falešně negativní.

Nedislokované fraktury v oblasti loketního kloubu mohou být na prostém rentgenovém snímku méně nápadné, zejména pokud lomná linie není výrazná a není přítomna zjevná změna vzájemného postavení kostních struktur. V případě hlavičky radia se navíc uplatňuje i složitější kloubní anatomie a částečný překryv okolních struktur. Přehlédnutí podobného poranění může mít dopad na další režimová a terapeutická opatření, a ukazuje proto, že automatizované hodnocení nemůže být odděleno od klinického kontextu a odborného posouzení.



Obr. 5: Rentgenový snímek levého loketního kloubu s frakturou hlavičky radia  
Zdroj: anonymizovaný RTG snímek z PACS archivu FNOL, vlastní zpracování, 2026

Tab. 11: Klasifikace kazuistiky 4

<i>Hodnotitel</i>	<i>Výsledek hodnocení</i>
<b>Referenční standard (Ground Truth)</b>	Pozitivní (potvrzeno lékařem)
<b>AI 1</b>	Detekováno (true positive)
<b>AI 2</b>	Detekováno (true positive)
<b>AI 3</b>	Nedetkováno (false negative)

Zdroj: Vlastní zpracování, 2026

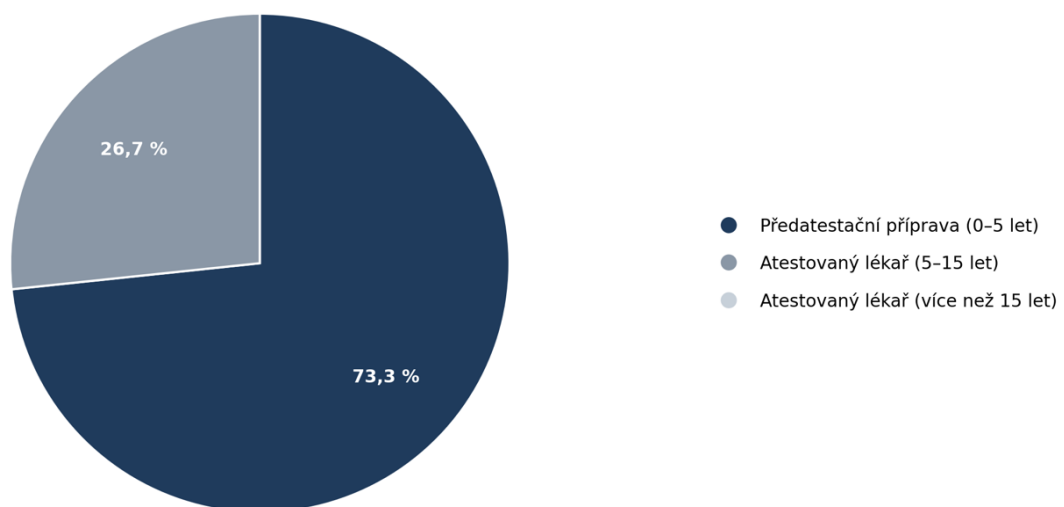
Diskrepanční kazuistiky ukázaly, že k chybným klasifikacím docházelo především v případech méně nápadných fraktur, při současné přítomnosti degenerativních změn nebo v anatomicky složitějších oblastech s horší přehledností. Opakovaně se potvrdilo, že samotná přítomnost patologického nálezu nezaručuje jeho spolehlivou detekci všemi hodnocenými systémy. Rozdíly mezi jednotlivými modely naznačují, že jejich výkonnost není dána pouze celkovými souhrnnými metrikami, ale také citlivostí vůči konkrétním typům nálezů. Analýza kazuistik proto doplňuje kvantitativní výsledky o detailnější pohled na situace, v nichž mohou být výstupy AI klinicky méně spolehlivé.

## 3.2 Kvalitativní část – postoje radiologů a vnímání AI nástrojů

Kvalitativní šetření bylo realizováno prostřednictvím strukturovaného elektronického dotazníku a bylo zacíleno na odbornou radiologickou komunitu v České republice. Z celkového počtu 41 cíleně oslovených lékařů bylo získáno 15 validních odpovědí, což představuje návratnost (response rate) 36,6 %. Dosažený vzorek poskytuje velmi cenný vhled do aktuálního vnímání nástrojů umělé inteligence, úrovně důvěry v algoritmičká rozhodnutí a umožňuje exaktně identifikovat hlavní bariéry bránící plošné implementaci analyzovaných technologií do nemocniční praxe.

### 3.2.1 Demografický a expertní profil respondentů

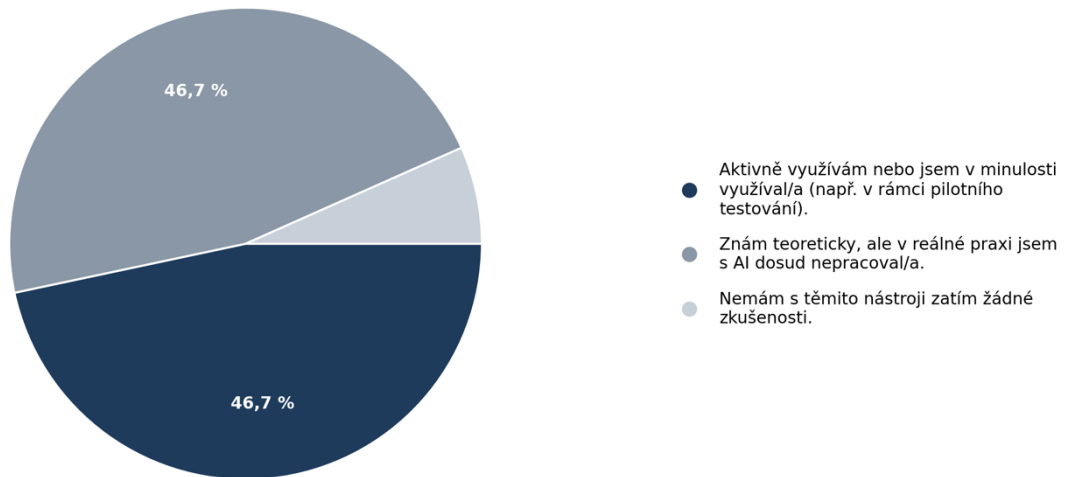
Z hlediska délky klinické praxe byl zkoumaný vzorek dominantně tvořen mladší generací lékařů v předatestační přípravě (s praxí 0–5 let), kteří představovali 73,3 % všech dotazovaných (11 respondentů). Zbytek vzorku (26,7 %, 4 respondenti) tvořili atestovaní lékaři s praxí 5 až 15 let. Z pohledu primárního působiště převažovali respondenti z fakultních nemocnic (53,3 %), následovaní lékaři z krajských a městských zdravotnických zařízení (40 %). Nejnížší zastoupení měly soukromé ambulance a polikliniky (6,7 %). Zastoupení mladších lékařů je pro výzkum vysoce relevantní, jelikož právě nastupující generace radiologů bude s nástroji umělé inteligence v budoucnu pracovat nejčastěji a bude formovat standardy budoucí diagnostiky.



Obr. 6: Délka klinické praxe respondentů v oboru radiologie

Zdroj: Vlastní zpracování, 2026

Klíčovým parametrem pro kontextualizaci odpovědí byla dosavadní praktická zkušenost s využíváním nástrojů AI. Téměř polovina dotazovaných (46,7 %) deklarovala aktivní zkušenost (ať už v rutinním provozu, nebo v rámci pilotního testování). Zcela identický podíl vzorku (46,7 %) znal systémy umělé inteligence primárně teoreticky z odborné literatury či kongresů, aniž by s nimi dosud měl přímou uživatelskou interakci. Lékaři zcela bez povědomí o popisovaných technologiích tvořili pouze marginální menšinu (6,6 %). Získaná data jasně prokazují, že umělá inteligence již přestává být v české radiologické obci pouhým teoretickým konceptem, jelikož významná část nastupující generace lékařů již s těmito nástroji měla možnost klinicky pracovat.



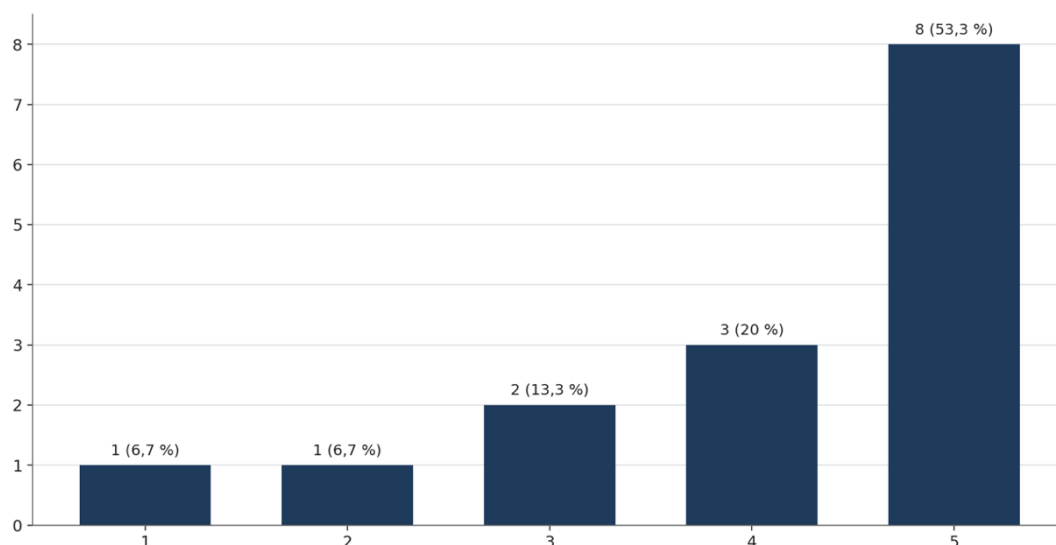
**Obr. 7: Dosavadní praktická zkušenost respondentů s AI nástroji**

*Zdroj: Vlastní zpracování, 2026*

### 3.2.2 Vnímání spolehlivosti a požadavky na vysvětlitelnost (XAI)

Druhý blok dotazníku se soustředil na problematiku důvěry (trust) a na fenomén tzv. „black-box“ (černé skříňky) systémů. Výsledky prokazují, že česká radiologická obec je vůči netransparentním modelům značně skeptická a vyžaduje přesné vizuální zdůvodnění algoritmických kroků. Výrazná většina respondentů (73,3 % při součtu hodnot 4 a 5 na pětibodové Likertově škále) se přiklonila k názoru, že modelům bez vizualizace nálezu v obraze nelze plně důvěřovat.

Prezentovaný postoj naprosto koresponduje s rostoucím tlakem na vývoj vysvětlitelné umělé inteligence (Explainable AI, XAI). Vizuální zdůvodnění formou ohraničujících boxů (bounding boxes) či teplotních map (heatmaps) představuje pro 73,3 % dotazovaných kritický faktor, který významně zvyšuje jejich diagnostickou jistotu. Samotná textová informace o přítomnosti fraktury je pro lékaře nedostačující; expert potřebuje vidět, na základě jakých obrazových bodů algoritmus ke svému závěru dospěl.

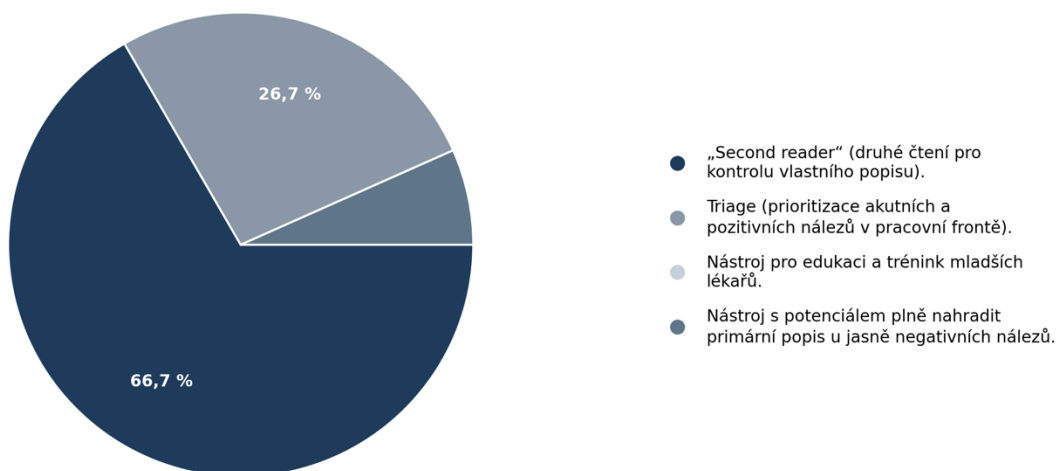


**Obr. 8: Vliv vizualizace (XAI) na důvěru v diagnostický výsledek**

*Zdroj: Vlastní zpracování, 2026*

Vysokou míru kritického uvažování a rezistenci vůči algoritmické autoritě potvrdila otázka simulující přímý konflikt názorů. V situaci, kdy by systém vyhodnotil snímek jako negativní, avšak subjektivní klinické podezření na patologii by u lékaře na základě anamnézy přetrvávalo, by 80 % respondentů (12 lékařů) algoritmické doporučení s vysokou pravděpodobností ignorovalo a řídilo se vlastním úsudkem. Zjištěný postoj demonstruje, že radiologové nepřístupují k AI jako k neomylnému orákulu, ale zachovávají si zdravou míru klinické skepse.

Pokud jde o ideální zařazení softwaru do klinického workflow, dominantní část vzorku (66,7 %, 10 respondentů) vnímá systémy umělé inteligence primárně jako „second reader“ (nástroj pro kontrolu a potvrzení vlastního popisu). Významná menšina (26,7 %, 4 respondenti) preferuje využití pro prioritizaci akutních případů (triage). Model plné automatizace (nasazení AI s potenciálem plně nahradit primární popis u jasně negativních nálezů) podpořil pouze jediný respondent (6,7 %). Popsané rozložení potvrzuje převažující konzervativní přístup komunity k přenechání diagnostické autonomie strojům.

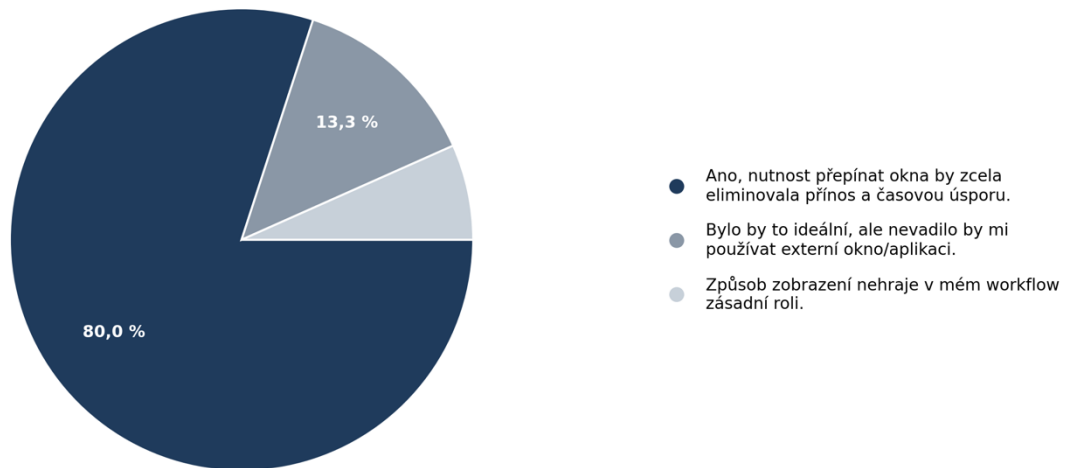


**Obr. 9: Preferovaný způsob zařazení AI do klinického workflow**

*Zdroj: Vlastní zpracování, 2026*

### 3.2.3 Procesní a legislativní bariéry

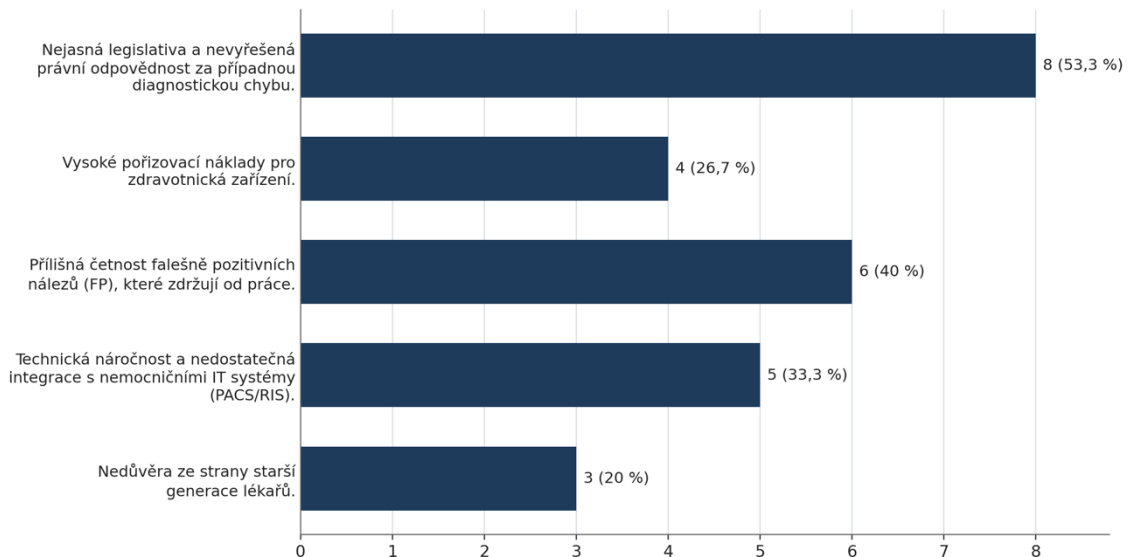
Závěrečná doména šetření analyzovala praktické, technické a systémové překážky bránící adopci inovací do rutinního nemocničního provozu. Z dat jednoznačně vyplývá stěžejní role ergonomie pracovního prostředí. Pro absolutní většinu dotazovaných (80 %, 12 respondentů) je naprosto nezbytné, aby byly výstupy AI integrovány přímo do nativního prohlížeče PACS. Nutnost otevírání externích aplikací či webových prohlížečů je vnímána jako rušivý element, který zcela eliminuje slibovanou časovou úsporu plynoucí z automatizace. Menšina respondentů (13,3 %) by akceptovala využití externí aplikace jako kompromis, zatímco pro okrajovou část (6,7 %) nehraje způsob zobrazení ve workflow zásadní roli.



**Obr. 10: Požadavky na integraci AI do nativního prohlížeče PACS**

*Zdroj: Vlastní zpracování, 2026*

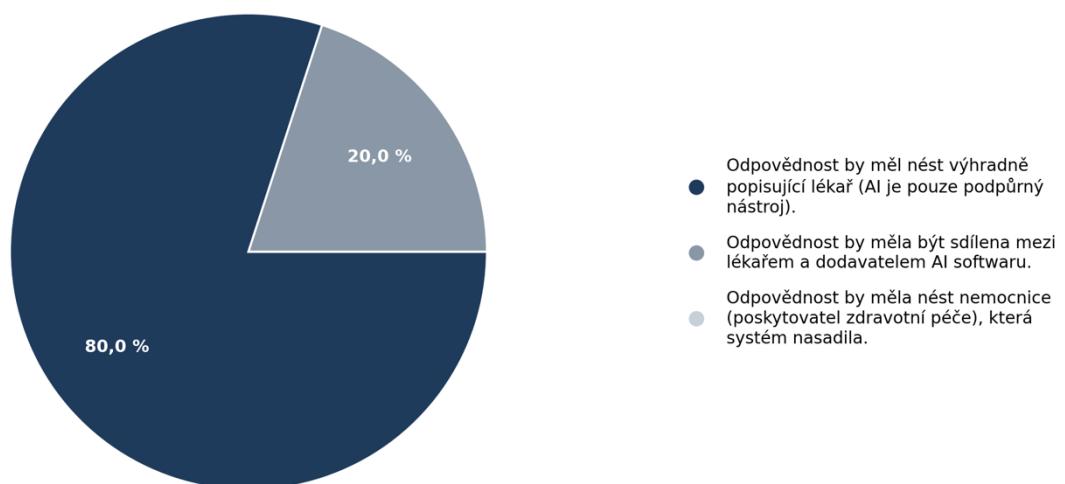
Analýza nejvýznamnějších aktuálních bariér odhalila roztříštěnost klíčových problémů, přičemž respondenti měli možnost označit dvě hlavní překážky. Jako vůbec nejvýznamnější bariéra byla identifikována nejasná legislativa a nevyřešená právní odpovědnost (53,3 %). Druhým nejčastěji zmiňovaným problémem se stala přílišná četnost falešně pozitivních nálezů (40 %), jež v hodnocení překonala i technickou náročnost integrace s IT systémy (33,3 %) či vysoké pořizovací náklady (26,7 %). Obava z falešných poplachů, které lékaře zdržují a nutí je přehodnocovat jasně zdravé snímky, přímo koreluje s výsledky kvantitativní části práce; konkrétně systém AI 3 prokázal vysokou chybovost generováním artefaktových detekcí. Pětina dotazovaných (20 %) spatřuje překážku rovněž v nedůvěře ze strany starší generace lékařů.



**Obr. 11: Největší aktuální bariéry pro plošné nasazení AI v ČR**

*Zdroj: Vlastní zpracování, 2026*

Z hlediska budoucí implementace přinesla zcela zásadní zjištění analýza vnímání profesní odpovědnosti. Navzdory volání po lepším a jasnějším legislativním rámci se výrazná většina respondentů (80 %, 12 lékařů) domnívá, že i v případě převzetí chybného (falešně negativního) výsledku vygenerovaného umělou inteligencí by odpovědnost za přehlédnutí fraktury měl nést výhradně popisující lékař. Pětina dotazovaných (20 %, 3 respondenti) naproti tomu zastává názor, že odpovědnost by měla být sdílena mezi lékařem a dodavatelem softwaru. Získaná data demonstrují obrovské uvědomění si osobní profesní zodpovědnosti. Komunita přijímá AI výlučně jako poradní a podpůrný hlas, nikoliv jako autonomní entitu schopnou plně převzít právní či medicínskou garanci za pacienta.



**Obr. 12: Názor respondentů na nesení právní odpovědnosti za chybu AI**

*Zdroj: Vlastní zpracování, 2026*

Hlubší kvalitativní analýza textových komentářů v otevřených otázkách navíc odhalila významné obavy z psychologického fenoménu zvaného automation bias (nekritické spoléhání se na technologii). Identifikované obavy jsou o to relevantnější, že zkoumaný vzorek tvořili dominantně mladí lékaři v předatestační přípravě. Jak explicitně uvedli někteří respondenti

na základě zkušeností z nemocničního pilotního testování, existuje reálné riziko, že u začínajících lékařů by vinou neustálé softwarové podpory mohlo dojít k degradaci jejich vlastních diagnostických schopností. Lékaři vnímají hrozbu ztráty sebevědomí a jistoty při samostatném hodnocení složitějších patologií ve chvílích, kdy by systém z technických důvodů měl výpadek nebo by nebyl pro danou anatomickou oblast certifikován. Zkvalitnění detekce a redukce opomenutých fraktur prostřednictvím algoritmů tak musí být vždy vyvažováno systematickým a kontinuálním vzděláváním personálu; umělá inteligence smí sloužit pouze jako bezpečnostní pojistka, nikoliv jako náhrada lidské erudice.

Výsledky dotazníkového šetření ukázaly, že postoje radiologů k využití umělé inteligence nejsou určovány pouze obecnou otevřeností vůči novým technologiím, ale především vnímanou spolehlivostí výstupů, požadavkem na jejich srozumitelnost, kvalitou integrace do systémů PACS/RIS a jasným vymezením profesní a právní odpovědnosti. Současně bylo potvrzeno, že úspěšná implementace AI v radiologii není podmíněna jen technickou přesností algoritmu, ale také schopností systému zapadnout do reálného klinického provozu a získat důvěru uživatelů. Kvalitativní část tak ukázala, že přijetí AI nástrojů v praxi je výsledkem souběhu technologických, organizačních i profesních faktorů.

### 3.3 Komparace a syntéza výsledků

Zvolený smíšený výzkumný design umožnil propojit dvě komplementární roviny hodnocení zkoumané problematiky. V kvantitativní části byla posouzena reálná diagnostická výkonnost vybraných AI nástrojů na souboru anonymizovaných skeletálních rentgenových snímků, zatímco kvalitativní část zachytila postoje radiologů k jejich využití v klinické praxi. Teprve vzájemná konfrontace obou datových linií umožnila formulovat ucelenější závěry o tom, za jakých podmínek mohou být nástroje umělé inteligence v oblasti skeletální radiologie považovány za klinicky využitelné, důvěryhodné a bezpečné.

Syntéza výsledků ukázala, že přijetí AI nástrojů radiology není určováno pouze deklarovanou inovativností technologie ani přítomností formální certifikace, ale především reálně vnímanou spolehlivostí výstupů, typem generovaných chyb, způsobem prezentace výsledků a mírou integrace do pracovního prostředí. Současně bylo potvrzeno, že ani vysoká souhrnná výkonnost systému sama o sobě nezaručuje jeho bezproblémové využití v každodenním provozu. Klinická akceptace AI nástroje se ukázala být výsledkem souběhu technologických, organizačních a profesních faktorů, které musí být splněny současně.

#### 3.3.1 Rozpor mezi certifikací a reálnou klinickou spolehlivostí

Kvantitativní část prokázala výrazné rozdíly v diagnostické výkonnosti mezi hodnocenými systémy. Zatímco AI 1 a AI 2 dosahovaly ve sledovaných metrikách stabilně vysokých hodnot, systém AI 3 vykazoval podstatně nižší přesnost a vyšší chybovost. Uvedené zjištění je významné zejména proto, že všechny hodnocené nástroje byly formálně určeny ke klinickému využití. Bylo tak potvrzeno, že samotná přítomnost regulačního nebo certifikačního rámce nemůže být interpretována jako automatická záruka srovnatelné diagnostické robustnosti v reálných podmínkách.

V kvalitativní části se obdobná nejistota promítla do zdrženlivého postoje radiologů vůči plné autonomii algoritmických systémů. Respondenti opakovaně zdůrazňovali, že finální důvěra v AI nemůže být postavena pouze na tom, že nástroj splňuje formální podmínky pro nasazení, ale musí vycházet z jeho ověřené spolehlivosti při práci s konkrétním typem obrazových dat a v konkrétním klinickém kontextu. Výsledky obou částí se tedy v tomto bodě vzájemně potvrzují. Bylo ukázáno, že nedůvěra části radiologů nepředstavuje odpor vůči technologii jako takové, ale racionální reakci na možnost, že rozdíly mezi jednotlivými systémy mohou být v praxi klinicky významné.

Význam uvedeného zjištění spočívá v tom, že proces implementace AI nástrojů nemůže být redukován na otázku regulatorního schválení. Nezbytnou součástí bezpečného nasazení musí být také lokální validace výkonu systému, průběžné sledování jeho chybovosti a kritické hodnocení toho, v jakých typech nálezů může docházet k selhání. Teprve kombinace formální způsobilosti a empiricky ověřené spolehlivosti může vytvořit dostatečný základ pro důvěru klinických uživatelů.

### 3.3.2 Falešné poplachy jako kritický faktor narušující klinické workflow

Kvalitativní část ukázala, že jednou z nejvýznamnějších bariér plošného nasazení AI v radiologii je obava z nadměrného množství falešně pozitivních výstupů. Respondenti vnímali opakované upozorňování na neexistující patologii jako faktor, který může snižovat plynulost práce, zvyšovat kognitivní zátěž a postupně oslabovat ochotu výstupy systému respektovat. Význam uvedené bariéry byl potvrzen i kvantitativními výsledky, protože mezi hodnocenými systémy byly zjištěny výrazné rozdíly právě v četnosti chyb tohoto typu.

V klinickém provozu nemá falešně pozitivní výstup pouze statistický význam, ale přímý dopad na organizaci práce. Každý suspektní signál musí být lékařem znovu zhodnocen, což prodlužuje čas potřebný k interpretaci, odvádí pozornost od skutečně patologických nálezů a může vytvářet stav postupného otupění vůči upozorněním systému. V literatuře bývá podobný jev popisován jako alert fatigue, v prostředí radiologie však nabývá zvláštní závažnosti, protože práce s obrazovou dokumentací je již sama o sobě spojena s vysokými nároky na koncentraci a přesnost.

Syntéza obou částí proto ukazuje, že specifická algoritmu nepředstavuje pouze technickou charakteristiku, ale také důležitý předpoklad jeho provozní akceptovatelnosti. Nástroj, který generuje nadměrné množství falešných poplachů, nemusí být klinicky přínosný ani tehdy, pokud dosahuje relativně dobré senzitivity. Pro úspěšnou implementaci je nezbytné, aby systém nejen rozpoznával patologické nálezy, ale současně nepřispíval k narušení workflow a nezvyšoval pracovní zátěž radiologa. Uvedené zjištění podporuje požadavek respondentů na takové AI řešení, které bude integrováno přímo do PACS a současně bude poskytovat výstupy v rozsahu, jenž nebude rušivý ani nadměrně zahlcující.

### 3.3.3 Reziduální chybovost, automation bias a zachování role radiologa

Ani nejlépe hodnocené systémy nevykazovaly absolutní bezchybnost. Analýza diskrepantních případů ukázala, že k selháním docházelo zejména u méně nápadných fraktur, v anatomicky složitějších oblastech nebo v situacích, kdy byl traumatický nález překryt jinými strukturálními změnami. Uvedené zjištění je pro interpretaci kvalitativních dat zásadní, protože vysvětluje, proč

radiologové nevnímají AI jako náhradu lidského rozhodování, ale spíše jako podpůrný nástroj, jehož výstupy musí být průběžně ověřovány.

V otevřených komentářích respondentů se opakovaně objevovala obava z fenoménu automation bias, tedy z nekritického přebírání algoritmických závěrů bez dostatečné vlastní verifikace. Zvláštní pozornost byla věnována riziku, že u lékařů v předatestační přípravě by dlouhodobé spoléhání na softwarovou podporu mohlo oslabovat schopnost samostatného hodnocení obrazové dokumentace. Kvantitativní zjištění přitom uvedenou obavu nepřímo potvrzují. Jestliže i výkonný systém může v jednotlivých případech patologii přehlédnout, pak by nekritické převzetí jeho negativního závěru mohlo vést k reálnému diagnostickému pochybení.

Z hlediska klinické bezpečnosti tedy bylo potvrzeno, že AI nemůže být v současné fázi vývoje chápána jako autonomní diagnostická autorita. Její role je vhodněji vymezena jako podpora druhého čtení, nástroj pro zvýšení pozornosti nebo mechanismus doplňkové kontroly. Takové pojetí je v souladu i s postoji respondentů k právní odpovědnosti, protože většina z nich přisuzovala finální odpovědnost za interpretaci nálezu popisujícímu lékaři. Zachování centrální role radiologa se tak neukazuje jako projev profesní rezistence, ale jako racionální důsledek reziduální chybovosti současných systémů a potřeby kritického klinického úsudku.

Synteticky vzato bylo prokázáno, že bezpečné a udržitelné využití AI ve skeletální radiologii předpokládá zachování lidské expertízy jako rozhodujícího prvku diagnostického procesu. Výstup algoritmu musí být interpretován jako podpůrná informace, nikoli jako definitivní závěr. Dlouhodobá implementace proto musí být spojena nejen s technickým nasazením systému, ale také s průběžným vzděláváním uživatelů, kultivací kritického přístupu k výstupům AI a jasným vymezením situací, ve kterých má lidské hodnocení přednost.

Na základě komparace kvantitativních a kvalitativních výsledků lze uzavřít, že klinicky přijatelný model využití AI v oblasti skeletálních rentgenových snímků musí splňovat několik souběžných podmínek. Požadována je dostatečně vysoká diagnostická výkonnost, nízká četnost falešně pozitivních i falešně negativních výstupů, transparentní prezentace suspektních oblastí, plynulá integrace do systému PACS/RIS a zachování radiologa jako finálního interpreta nálezu. Úspěch implementace tedy není určován pouze tím, zda algoritmus „funguje“, ale také tím, zda jeho fungování odpovídá požadavkům klinické praxe, profesní odpovědnosti a důvěry uživatelů.

Pokud je AI využívána jako nástroj druhého čtení s možností vizuální verifikace nálezu, může představovat významný přínos pro zvýšení bezpečnosti a standardizace diagnostického procesu. Pokud by však byla chápána jako samostatný rozhodovací mechanismus bez kritického lidského dohledu, nebyly by pro takové využití na základě získaných výsledků splněny dostatečné podmínky. Syntéza výsledků proto podporuje model, v němž je umělá inteligence integrována jako podpůrná technologie rozšiřující schopnosti radiologa, nikoli jako nástroj nahrazující jeho odborný úsudek.

## 4 Diskuse

Na provedenou syntézu výsledků navazuje diskuse, jejímž účelem je interpretovat získaná zjištění v širším odborném a klinickém rámci. Pozornost již není soustředěna na samotný popis rozdílů mezi hodnocenými systémy a postoji respondentů, ale na význam zjištěných skutečností pro skeletální radiologii, na jejich vztah k dosavadnímu poznání a na podmínky bezpečné implementace umělé inteligence do klinické praxe.

Předložený výzkum propojil technickou evaluaci vybraných AI nástrojů s analýzou postojů českých radiologů. Právě propojení obou rovin umožnilo posoudit nejen diagnostickou výkonnost algoritmů, ale také míru jejich klinické přijatelnosti, důvěryhodnosti a provozní využitelnosti. Získaná zjištění ukázala, že rozhodování o zavádění AI do skeletální radiologie nemůže vycházet pouze z výkonových metrik ani pouze z deklarované technologické vyspělosti systému. Rozhodující význam mají rovněž typ generovaných chyb, způsob prezentace výstupu, kvalita integrace do pracovního prostředí a zachování role radiologa jako finálního interpreta nálezu.

### 4.1 Interpretace výsledků ve vztahu k odborné literatuře

Výsledky kvantitativní části prokázaly výrazné rozdíly v diagnostické výkonnosti mezi hodnocenými systémy. Uvedený závěr odpovídá poznatkům odborné literatury, podle níž nelze na AI nástroje pro detekci fraktur nahlížet jako na jednotnou skupinu řešení se srovnatelnou klinickou hodnotou. Přehledové práce i metaanalýzy ukazují, že AI v oblasti detekce fraktur dosahuje často velmi dobrých výsledků, avšak výkonnost zůstává závislá na konkrétním modelu, typu trénovacích dat, anatomické lokalitě i způsobu klinického nasazení (Husarek et al., 2024; Nowroozi et al., 2024; Lo Mastro et al., 2025).

Vlastní výsledky uvedené poznání potvrzují. U dvou hodnocených systémů byla zaznamenána relativně stabilní výkonnost, zatímco třetí systém vykazoval podstatně vyšší chybovost a nižší diagnostickou spolehlivost. Význam získaného zjištění spočívá mimo jiné v tom, že všechny hodnocené nástroje byly formálně určeny ke klinickému využití. Z odborného hlediska proto byla potvrzena skutečnost, že regulační nebo certifikační rámec nemůže být považován za dostatečný důkaz srovnatelné robustnosti v lokálních podmínkách. V literatuře bývá v obdobné souvislosti zdůrazňována potřeba lokální validace algoritmů, průběžného sledování výkonnosti a kritického hodnocení chyb v reálném provozu (Mello-Thoms & Mello, 2023).

Zvláštní interpretační význam měla analýza diskrepantních kazuistik. Bylo zjištěno, že vyšší riziko selhání se objevovalo zejména u diagnosticky méně zřejmých fraktur, u nálezů v anatomicky složitějších oblastech a v případech, kdy byl traumatický obraz překryt jinou strukturální odchylkou, například degenerativní přestavbou. Odborná literatura upozorňuje na obdobný problém opakovaně, protože souhrnné výkonové metriky sice informují o celkovém výkonu systému, avšak nevystihují plně jeho chování v klinicky obtížných nebo netypických situacích (Husarek et al., 2024; Bachmann et al., 2024; Lo Mastro et al., 2025). Z uvedeného důvodu byla kvalitativní analýza chybových případů vyhodnocena jako nezbytný doplněk kvantitativní evaluace.

Kvalitativní část práce ukázala, že postoje radiologů k AI odpovídají dlouhodobému trendu popisovanému v radiologické literatuře. Umělá inteligence nebývá vnímána jako náhrada lidského rozhodování, ale spíše jako podpůrný mechanismus, jehož význam je podmíněn důvěryhodností, transparentností a praktickou použitelností (Mello-Thoms & Mello, 2023). V etických a odborných diskusích bývá opakovaně zdůrazňováno, že AI má v radiologii sloužit k posílení lidské expertízy, nikoli k eliminaci lidské odpovědnosti. Současně bývá zdůrazňována potřeba transparency, dependability a zachování lidské accountability za finální klinické rozhodnutí (Geis et al., 2019).

Podobně byla v práci potvrzena významná role vysvětlitelnosti modelu. Respondenti preferovali řešení, která umožňují vizuální lokalizaci suspektní oblasti a poskytují radiologovi možnost kriticky ověřit, z jakého důvodu algoritmus dospěl k určitému závěru. Uvedený požadavek je v souladu s literaturou věnovanou explainable AI, podle níž představuje srozumitelnost výstupu důležitou podmínku klinické akceptace, zejména v prostředí medicínského zobrazování, kde je třeba propojit algoritmičké doporučení s odpovědným odborným úsudkem (de Vries et al., 2023; Borys et al., 2023b).

Souhrnně lze uvést, že výsledky práce jsou s dostupnou odbornou literaturou ve shodě. Potvrzen byl vysoký potenciál AI pro podporu detekce skeletálních patologií, avšak zároveň byla zdůrazněna nutnost opatrné interpretace výkonových metrik, lokální validace, kritického posuzování chybových situací a zachování role radiologa jako finálního interpreta nálezu. Přínos předloženého výzkumu spočívá v tom, že uvedené obecné závěry byly konkretizovány pro prostředí české skeletální radiologie.

## 4.2 Hlavní bariéry a přínosy implementace AI v české skeletální radiologii

Za hlavní přínos hodnocených AI nástrojů lze považovat možnost podpory radiologa při detekci skeletálních patologií a snížení rizika přehlédnutí nálezu. Takový přínos nabývá zvláštního významu v prostředí, kde je popis rentgenových snímků zatížen časovým tlakem, vysokým počtem vyšetření a proměnlivou složitostí nálezů. V případě dostatečně přesného a provozně dobře integrovaného systému může algoritmičká podpora sloužit jako druhé čtení, zvýšit konzistenci hodnocení a přispět ke standardizaci diagnostického procesu. Obdobný potenciál bývá připisován AI nástrojům i v přehledových pracích zaměřených na detekci fraktur a klinickou implementaci AI v radiologii (Nowroozi et al., 2024; Lo Mastro et al., 2025).

Přínos AI však nebyl vyhodnocen jako bezpodmínečný. Z výsledků práce vyplynulo, že klinická využitelnost závisí nejen na schopnosti algoritmu patologii rozpoznat, ale také na charakteru chyb, které systém generuje. Významnou bariéru představují falešně pozitivní výstupy. Každé neopodstatněné upozornění zvyšuje čas potřebný k interpretaci, odvádí pozornost od skutečně patologických nálezů a může oslabovat ochotu radiologa výstupy systému respektovat. V literatuře bývá obdobný efekt popisován v souvislosti s workflow disruption a alert fatigue, tedy se stavem postupného otupění vůči upozorňovacím mechanismům, které jsou příliš časté nebo málo spolehlivé (Dikici et al., 2020; Mello-Thoms & Mello, 2023).

Vedle falešně pozitivních výstupů zůstává zásadním problémem i přítomnost falešně negativních chyb. Klinická závažnost podobných omylů spočívá v riziku přehlédnutí fraktury nebo jiné

patologické změny a v následném opoždění správného diagnostického či terapeutického postupu. Analýza kazuistik ukázala, že ani výkonnější systémy nejsou vůči podobným selháním plně imunní, zejména u méně nápadných nálezů a v anatomicky komplikovanějších oblastech. Z praktického hlediska proto bylo potvrzeno, že AI nemůže být využívána izolovaně, ale pouze jako podpůrná informace, která zvyšuje pozornost radiologa a doplňuje jeho vlastní interpretaci.

Významná bariéra byla identifikována také v oblasti důvěry. Rezervovaný postoj radiologů nebyl spojen s obecnou neochotou přijímat nové technologie, ale s racionálními pochybnostmi o spolehlivosti, transparentnosti a odpovědnosti za algoritmický výstup. Důvěra v AI tedy nebyla založena na deklarované modernosti nástroje, ale na možnosti výstup smysluplně ověřit a bezpečně využít v konkrétní klinické situaci. Právě proto byla zdůrazněna potřeba vizuální lokalizace suspektní oblasti, srozumitelné prezentace výsledku a možnosti kritické verifikace ze strany radiologa. Odborná literatura řadí uvedené faktory mezi klíčové determinanty klinické akceptace AI již delší dobu (de Vries et al., 2023; Borys et al., 2023a).

Další podstatná bariéra se týká právní a profesní odpovědnosti. Většina respondentů přisuzovala finální odpovědnost za interpretaci nálezu nadále popisujícímu lékaři, a to i při využití AI jako součásti diagnostického procesu. Uvedený postoj lze považovat za logický důsledek reziduální chybovosti současných systémů a za výraz profesní odpovědnosti, nikoli za projev technologické rezistence. Etické práce věnované AI v radiologii docházejí k obdobnému závěru, když uvádějí, že odpovědnost za finální klinické rozhodnutí zůstává v dohledné době na straně lidských tvůrců a operátorů systému, v klinické praxi tedy především na straně zdravotnického pracovníka využívajícího algoritmický výstup (Geis et al., 2019).

V českém prostředí byla jako zvlášť významná identifikována také ergonomie pracovního prostředí. Akceptace AI byla spojena s požadavkem, aby výstupy byly dostupné přímo v nástrojích běžně používaných radiologem a aby nenarušovaly plynulost klinického workflow. Praktická použitelnost systému proto nebyla dána pouze jeho přesností, ale rovněž tím, jak nenápadně a efektivně se dokáže začlenit do každodenního provozu. Systém vyžadující práci v externím prostředí nebo generující nadměrné množství rušivých upozornění by byl pravděpodobně vnímán spíše jako další provozní zátěž než jako přínosný nástroj.

Souhrnně lze uzavřít, že přínosy a bariéry implementace AI tvoří vzájemně propojený celek. Rychlost detekce může znamenat časovou úsporu, avšak při nadměrné chybovosti může naopak pracovní proces zpomalovat. Automatická lokalizace nálezu může podporovat orientaci v obraze, avšak při nedostatečné transparentnosti může zvyšovat nedůvěru. Praktická hodnota AI proto nebyla spojena s představou automatizace za každou cenu, ale s takovým způsobem nasazení, který posiluje bezpečnost, standardizaci a efektivitu bez oslabení kritické role lékaře.

### 4.3 Možnosti integrace AI do klinického workflow

Výsledky práce ukázaly, že otázka integrace AI do klinického workflow představuje jeden z rozhodujících faktorů jejího přijetí. Nešlo přitom pouze o technickou možnost propojit algoritmus se stávající nemocniční infrastrukturou, ale především o způsob provedení, který nebude narušovat zavedené pracovní návyky radiologů. Požadavek na přímou integraci do PACS byl vnímán jako praktická nutnost, nikoli jako doplňková výhoda. V odborné literatuře bývá

obdobně zdůrazňováno, že AI má být zapojena do již existujícího workflow co nejplynuleji a s co nejmenším množstvím rušivých zásahů do běžného provozu (Dikici et al., 2020).

Za nejvhodnější model integrace lze považovat uspořádání, v němž je snímek po pořízení automaticky analyzován na pozadí, výsledek je vrácen přímo do prostředí PACS a radiolog má při popisu současně k dispozici původní obraz i algoritmickou anotaci. Takový model zachovává plynulost workflow a současně umožňuje okamžitou kritickou verifikaci výstupu. Z provozního hlediska se jako vhodné jeví, aby algoritmus neposkytoval pouze binární informaci o přítomnosti patologie, ale rovněž vizuální označení suspektní oblasti a přiměřeně srozumitelnou informaci o míře jistoty výsledku. Právě transparentní prezentace nálezu zvyšuje pravděpodobnost, že bude výstup využit jako podpůrný prvek rozhodování, nikoli jako obtížně interpretovatelný black box (de Vries et al., 2023; Borys et al., 2023b).

V návaznosti na kvalitativní zjištění lze dále konstatovat, že AI by měla být v klinickém procesu postavena do role nenápadné, ale okamžitě dostupné vrstvy podpory. Vhodné je zejména využití pro druhé čtení, doplňkovou kontrolu nebo případnou prioritizaci snímků se suspektním nálezem. Jako problematické by naopak bylo vnímáno nasazení, které by radiologa nutilo mechanicky reagovat na každý algoritmický výstup bez možnosti selektivního a kritického zhodnocení. Úspěšná integrace tedy neznamená maximální viditelnost systému, ale funkční a ergonomické začlenění do procesu popisu.

Z praktického hlediska by měla být součástí implementace i možnost zpětné vazby a průběžného auditu. Bylo-li prokázáno, že jednotlivé systémy vykazují odlišnou chybovost a že část jejich selhání vzniká v konkrétních typech nálezů, je žádoucí, aby nasazený systém nebyl vnímán jako statický produkt, ale jako technologie, jejíž výkon je třeba průběžně sledovat. Transparentní evidence algoritmických výstupů, jejich následné korekce radiologem a možnost analyzovat opakující se typy chyb mohou významně přispět jak k bezpečnosti provozu, tak k budování důvěry uživatelů. Potřeba takového feedback loop je popsána rovněž v literatuře věnované workflow integraci AI v radiologii (Dikici et al., 2020).

Do úvahy musí být zahrnut i dopad AI na vzdělávání a profesní rozvoj lékařů. V práci byla identifikována obava, že nadměrné spoléhání na softwarovou podporu může oslabovat samostatné diagnostické uvažování, zejména pokud uživatel nekriticky přejímá algoritmická doporučení. Implementace AI proto nemá být chápána pouze jako technické nasazení nového nástroje, ale také jako změna pracovního a vzdělávacího prostředí. Aby nedocházelo k posilování fenoménu automation bias, musí být uživatelé vedeni k průběžnému ověřování algoritmických výstupů, k reflexi jejich limitů a k udržení schopnosti samostatného posouzení snímku i bez algoritmické asistence (Dratsch et al., 2023; Geis et al., 2019).

Interpretace získaných výsledků musí být současně vztahována k limitům studie. Kvantitativní část byla založena na omezeném počtu vybraných komerčních nástrojů a na konkrétním retrospektivním souboru skeletálních rentgenových snímků. Kvalitativní část zachytila postoje vybraného souboru českých radiologů a nehodnotila dlouhodobý dopad AI přímo v živém klinickém provozu. Získaná zjištění proto nelze bez dalšího zobecnit na všechny dostupné systémy ani na všechny klinické kontexty. Přesto však bylo možné formulovat několik podstatných implikací. Přijetí AI ve skeletální radiologii je podmíněno souběhem dostatečné diagnostické výkonnosti, nízké rušivosti, transparentního výstupu, plynulé integrace do pracovního prostředí a zachování finální role radiologa.

Diskuse podporuje model implementace, v němž je umělá inteligence využívána jako podpůrná technologie rozšiřující schopnosti radiologa, nikoli jako nástroj nahrazující jeho odborný úsudek. Uvedené pojetí je v souladu jak s kvantitativními výsledky, které prokázaly rozdílnou výkonnost a přetrvávající reziduální chybovost systémů, tak s kvalitativními zjištěními, podle nichž jsou důvěra, srozumitelnost a provozní nenáročnost klíčovými předpoklady praktické akceptace. Hlavní odpověď na výzkumnou otázku tedy spočívá v závěru, že přínos AI pro diagnostiku skeletálních patologií je reálný, avšak plně se uplatní pouze za podmínky odpovědné implementace, kritického využívání výstupů a úzké vazby na lidskou expertízu.

## Závěr

Bakalářská práce byla zaměřena na komparativní analýzu nástrojů umělé inteligence pro hodnocení skeletálních rentgenových snímků a na posouzení jejich vnímání českými radiology v kontextu klinické praxe. Teoretická část vymezila základní principy využití umělé inteligence v radiologii, specifika skeletálního zobrazování, výkonové metriky diagnostických systémů a faktory ovlivňující důvěru odborníků v algoritmická rozhodnutí. Metodická část vycházela z kombinovaného výzkumného přístupu, v němž byla propojena kvantitativní evaluace vybraných AI nástrojů na anonymizovaném souboru skeletálních rentgenových snímků s kvalitativním šetřením zaměřeným na postoje radiologů.

Výsledky kvantitativní části ukázaly, že mezi hodnocenými systémy existují výrazné rozdíly v diagnostické výkonnosti. Bylo potvrzeno, že ani formální určení nástroje ke klinickému využití samo o sobě nezaručuje srovnatelnou robustnost při práci s reálnými daty. Analýza diskrepantních případů současně ukázala, že k selháním docházelo zejména u méně nápadných fraktur, v anatomicky složitějších oblastech a v situacích, kdy byl traumatický nález překryt dalšími strukturálními změnami.

Kvalitativní část prokázala, že radiologové nepřístupují k využití umělé inteligence apriorně odmítavě, její přijetí je však podmíněno několika zásadními požadavky. Za klíčové byly označeny spolehlivost výstupů, nízká četnost falešně pozitivních nálezů, možnost vizuální verifikace suspektní oblasti, přímá integrace do prostředí PACS a jasné vymezení profesní a právní odpovědnosti. Umělá inteligence byla vnímána především jako podpůrný nástroj, nikoli jako autonomní náhrada radiologa.

Propojení obou výzkumných linií umožnilo formulovat hlavní závěr práce: přínos umělé inteligence pro diagnostiku skeletálních patologií je reálný, avšak výrazně podmíněný. Klinická využitelnost AI nástrojů nezávisí pouze na dosažených výkonových metrikách, ale rovněž na typu generovaných chyb, způsobu prezentace výsledků a míře začlenění do klinického workflow. Za bezpečný a udržitelný model využití lze považovat takové uspořádání, v němž AI funguje jako nástroj druhého čtení nebo doplňkové kontroly, zatímco finální odpovědnost za interpretaci nálezu zůstává na straně radiologa.

Přínos práce spočívá v propojení technického hodnocení vybraných AI systémů s analýzou postojů odborné veřejnosti, čímž byly vytvořeny podklady pro odpovědnější posuzování jejich zavádění do klinické praxe.

Za jedno z omezení předložené práce je třeba považovat skutečnost, že kvalitativní část nebyla v konečné podobě rozšířena o původně zamýšlené polostrukturované rozhovory s lékaři. Jejich realizace se ukázala jako obtížně proveditelná zejména s ohledem na vysokou pracovní vytíženost oslovených odborníků a omezené možnosti osobního setkání v rámci klinického provozu. Kvalitativní zjištění proto vycházejí výhradně z dotazníkového šetření. Absence polostrukturovaných rozhovorů však nesnižuje relevanci získaných dat pro základní orientaci v postojích radiologů k implementaci umělé inteligence, pouze vymezuje prostor pro jejich budoucí hlubší rozpracování.

Další výzkum by měl směřovat k rozšíření datového souboru, k multicentrickému ověření získaných výsledků a k prospektivnímu sledování dopadu AI na diagnostickou přesnost, pracovní zátěž radiologů a organizaci klinického workflow.

## Seznam použité literatury

- ALEXANDER, R. G., WAITE, S., BRUNO, M. A., KRUPINSKI, E. A., BERLIN, L., MACKNIK, S. L. a MARTINEZ-CONDE, S. Mandating limits on workload, duty, and speed in radiology. *Radiology* [online]. 2022, 304(2), 274–282 [cit. 2026-03-24]. Dostupné z: doi:10.1148/radiol.212631
- BACHMANN, R., GUNES, G., HANGAARD, S., BLANKE, F., JENSEN, K. K., LARSEN, L. P., NYBING, J. D. a BRORSON, S. Improving traumatic fracture detection on radiographs with artificial intelligence support: a multi-reader study. *BJR Open* [online]. 2024, 6(1), tzae011 [cit. 2026-03-24]. Dostupné z: doi:10.1093/bjro/tzae011
- BHARGAVAN, M., KAYE, A. H., FORMAN, H. P. a SUNSHINE, J. H. Workload of radiologists in United States in 2006–2007 and trends since 1991–1992. *Radiology* [online]. 2009, 252(2), 458–467 [cit. 2026-03-24]. Dostupné z: doi:10.1148/radiol.2522081895
- BORYS, K., SCHMITT, Y. A., NAUTA, M., SEIFERT, C., KRÄMER, N., FRIEDRICH, C. M. a NENSA, F. Explainable AI in medical imaging: An overview for clinical practitioners – Beyond saliency-based XAI approaches. *European Journal of Radiology* [online]. 2023a, 162, 110786 [cit. 2026-03-24]. Dostupné z: doi:10.1016/j.ejrad.2023.110786
- BORYS, K., SCHMITT, Y. A., NAUTA, M., SEIFERT, C., KRÄMER, N., FRIEDRICH, C. M. a NENSA, F. Explainable AI in medical imaging: An overview for clinical practitioners – Saliency-based XAI approaches. *European Journal of Radiology* [online]. 2023b, 162, 110787 [cit. 2026-03-24]. Dostupné z: doi:10.1016/j.ejrad.2023.110787
- CAREBOT. DICOM Conformance Statement: Carebot AI Bones. Praha: Carebot s.r.o., 2022.
- DE VRIES, B. M., ZWEZERIJNEN, G. J. C., BURCHELL, G. L., VAN VELDEN, F. H. P., MENKE-VAN DER HOUVEN VAN OORDT, C. W. a BOELLAARD, R. Explainable artificial intelligence (XAI) in radiology and nuclear medicine: a literature review. *Frontiers in Medicine* [online]. 2023, 10, 1180773 [cit. 2026-03-24]. Dostupné z: doi:10.3389/fmed.2023.1180773
- DIKICI, E., BIGELOW, M., PREVEDELLO, L. M., WHITE, R. D. a ERDAL, B. S. Integrating AI into radiology workflow: levels of research, production, and feedback maturity. *Journal of Medical Imaging* [online]. 2020, 7(1), 016502 [cit. 2026-03-24]. Dostupné z: doi:10.1117/1.JMI.7.1.016502
- DRATSCH, T., CHEN, X., REZAZADE MEHRIZI, M., KLOECKNER, R., MÄHRINGER-KUNZ, A., PÜSKEN, M., BAEßLER, B., SAUER, S., MAINTZ, D. a DOS SANTOS, D. P. Automation Bias in Mammography: The Impact of Artificial Intelligence BI-RADS Suggestions on Reader Performance. *Radiology* [online]. 2023, 307(4), e222176 [cit. 2026-03-24]. Dostupné z: doi:10.1148/radiol.222176
- ELTAWIL, F. A., ATALLA, M., BOULOS, E., AMIRABADI, A. a TYRRELL, P. N. Analyzing barriers and enablers for the acceptance of artificial intelligence innovations into radiology practice: A scoping review. *Tomography* [online]. 2023, 9(4), 1443–1455 [cit. 2026-03-24]. Dostupné z: doi:10.3390/tomography9040115
- EUROPEAN SOCIETY OF RADIOLOGY. What the radiologist should know about artificial intelligence – an ESR white paper. *Insights into Imaging* [online]. 2019, 10(1), 44 [cit. 2026-03-24]. Dostupné z: doi:10.1186/s13244-019-0738-2

- EUROPEAN UNION. Regulation (EU) 2017/745 of the European Parliament and of the Council on medical devices (MDR). Brussels: Official Journal of the European Union, 2017.
- EWALS, L. J. S., HEESTERBEEK, L. J. J., YU, B., VAN DER WULP, K., MAVROEIDIS, D., FUNK, M., SNIJDERS, C. C. P., JACOBS, I., NEDEREND, J. a PLUYTER, J. R. The impact of expectation management and model transparency on radiologists' trust and utilization of AI recommendations. *JMIR AI* [online]. 2024, 3, e52211 [cit. 2026-03-24]. Dostupné z: doi:10.2196/52211
- FARIČ, N., HINDER, S., WILLIAMS, R., RAMAESH, R., BERNABEU, M. O., VAN BEEK, E. a CRESSWELL, K. Early experiences of integrating an artificial intelligence-based diagnostic decision support system into radiology settings: a qualitative study. *Journal of the American Medical Informatics Association* [online]. 2024, 31(1), 24–34 [cit. 2026-03-24]. Dostupné z: doi:10.1093/jamia/ocad191
- FOOD AND DRUG ADMINISTRATION. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. Silver Spring: U.S. Food and Drug Administration, 2022.
- GEIS, J. R., BRADY, A. P., WU, C. C., SPENCER, J., RANSCHAERT, E., JAREMKO, J. L., LANGER, S. G., BORONDY KITTS, A., BIRCH, J., SHIELDS, W. F. a TANG, A. Ethics of Artificial Intelligence in Radiology. *Radiology* [online]. 2019, 293(2), 436–440 [cit. 2026-03-24]. Dostupné z: doi:10.1148/radiol.2019191586
- GULY, H. R. Diagnostic errors in an accident and emergency department. *Emergency Medicine Journal* [online]. 2001, 18(4), 263–269 [cit. 2026-03-24]. Dostupné z: doi:10.1136/emj.18.4.263
- HUSAREK, J., HESS, S., RAZAEIAN, S., RUDER, T. D., SEHMISCH, S., MÜLLER, M. a LIODAKIS, E. Artificial intelligence in commercial fracture detection products: a systematic review and meta-analysis. *Scientific Reports* [online]. 2024, 14, 23053 [cit. 2026-03-24]. Dostupné z: doi:10.1038/s41598-024-73058-8
- KANAKARAJ, P., RAMADASS, K., BAO, S., BASFORD, M., JONES, L. M., LEE, H. H., XU, K., SCHILLING, K. G., CARR, J. J., TERRY, J. G., HUO, Y., SANDLER, K. L., NEWTON, A. T. a LANDMAN, B. A. Workflow integration of research AI tools into a hospital radiology rapid prototyping environment. *Journal of Digital Imaging* [online]. 2022, 35(4), 1023–1033 [cit. 2026-03-24]. Dostupné z: doi:10.1007/s10278-022-00601-2
- KOTTER, E. a RANSCHAERT, E. Challenges and solutions for introducing artificial intelligence in daily clinical workflow. *European Radiology* [online]. 2021, 31(1), 5–7 [cit. 2026-03-24]. Dostupné z: doi:10.1007/s00330-020-07148-2
- KRUPINSKI, E. A., BERBAUM, K. S., CALDWELL, R. T., SCHATZ, K. M. a KIM, J. Long radiology workdays reduce detection and accommodation accuracy. *Journal of the American College of Radiology* [online]. 2010, 7(9), 698–704 [cit. 2026-03-24]. Dostupné z: doi:10.1016/j.jacr.2010.03.004
- KUO, R. Y. L., HARRISON, C., CURRAN, T. A., JONES, B., FREETHY, A., CUSSONS, D., STEWART, M., COLLINS, G. S. a FURNISS, D. Artificial intelligence in fracture detection: a systematic review and meta-analysis. *Radiology* [online]. 2022, 304(1), 50–62 [cit. 2026-03-24]. Dostupné z: doi:10.1148/radiol.211785

- KUTBI, M. Artificial intelligence-based applications for bone fracture detection using medical images: a systematic review. *Diagnostics* [online]. 2024, 14(17), 1879 [cit. 2026-03-24]. Dostupné z: doi:10.3390/diagnostics14171879
- LITJENS, G., KOOI, T., EHTESHAMI BEJNORDI, B., SETIO, A. A. A., CIOMPI, F., GHAFORIAN, M., VAN DER LAAK, J. A. W. M., VAN GINNEKEN, B. a SÁNCHEZ, C. I. A survey on deep learning in medical image analysis. *Medical Image Analysis* [online]. 2017, 42, 60–88 [cit. 2026-03-24]. Dostupné z: doi:10.1016/j.media.2017.07.005
- LO MASTRO, A., GRASSI, E., BERRITTO, D., IACOBELLIS, F., BRUNESE, L. a PINTO, A. Artificial intelligence in fracture detection on radiographs: a literature review. *Japanese Journal of Radiology* [online]. 2025, 43(4), 551–585 [cit. 2026-03-24]. Dostupné z: doi:10.1007/s11604-024-01702-4
- LUNDERVOLD, A. S. a LUNDERVOLD, A. An overview of deep learning in medical imaging focusing on MRI. *Computer Vision and Image Understanding* [online]. 2019, 189, 102–126 [cit. 2026-03-24]. Dostupné z: doi:10.1016/j.zemedi.2018.11.002
- MELLO-THOMS, C. a MELLO, C. A. B. Clinical applications of artificial intelligence in radiology. *The British Journal of Radiology* [online]. 2023, 96(1150), 20221031 [cit. 2026-03-24]. Dostupné z: doi:10.1259/bjr.20221031
- MUHAMMAD, D. a BENDECHACHE, M. Unveiling the black box: A systematic review of explainable artificial intelligence in medical image analysis. *Computational and Structural Biotechnology Journal* [online]. 2024, 24, 542–560 [cit. 2026-03-24]. Dostupné z: doi:10.1016/j.csbj.2024.08.005
- NAZIR, S., DICKSON, D. M. a AKRAM, M. U. Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks. *Computers in Biology and Medicine* [online]. 2023, 156, 106668 [cit. 2026-03-24]. Dostupné z: doi:10.1016/j.compbimed.2023.106668
- NOWROOZI, A., SALEHI, M. A., SHOBEIRI, P., VAHDAT, S., HASHEMI, S. M., KHOSRAVI, B., KHATAMI, F. a REZAEI, N. Artificial intelligence diagnostic accuracy in fracture detection from plain radiographs and comparing it with clinicians. *Clinical Radiology* [online]. 2024, 79(8), 579–588 [cit. 2026-03-24]. Dostupné z: doi:10.1016/j.crad.2024.04.009
- SAITO, T. a REHMSMEIER, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE* [online]. 2015, 10(3), e0118432 [cit. 2026-03-24]. Dostupné z: doi:10.1371/journal.pone.0118432
- SALEHINEJAD, H., KITAMURA, F. C., KANDEMIR, M., DITKOFKY, N., LIN, A., BHARATHA, A., SUTHIPHOSUWAN, S., LIN, H. M., WILSON, J. R., MAMDANI, M. a COLAK, E. A real-world demonstration of machine learning generalizability in the detection of intracranial hemorrhage. *Scientific Reports* [online]. 2021, 11, 17051 [cit. 2026-03-24]. Dostupné z: doi:10.1038/s41598-021-95533-2
- TEJANI, A. S., NG, Y. S., XI, Y. a RAYAN, J. C. Understanding and mitigating bias in imaging artificial intelligence. *Radiographics* [online]. 2024, 44(5), e230067 [cit. 2026-03-24]. Dostupné z: doi:10.1148/rg.230067

- VAN DER VELDEN, B. H. M., KUIJF, H. J., GILHUIJS, K. G. A. a VIERGEVER, M. A. Explainable artificial intelligence in deep learning-based medical image analysis. *Medical Image Analysis* [online]. 2022, 79, 102470 [cit. 2026-03-24]. Dostupné z: doi:10.1016/j.media.2022.102470
- WORLD HEALTH ORGANIZATION. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization, 2021.
- YU, J. S. Easily missed fractures in the lower extremity. *Radiologic Clinics of North America* [online]. 2015, 53(4), 737–755 [cit. 2026-03-24]. Dostupné z: doi:10.1016/j.rcl.2015.02.003

## Přílohy

## **Přílohy A Formuláře**

### **Příloha A.1 Formulář dotazníkového šetření pro radiology**

Postoje radiologů k integraci AI do diagnostiky skeletálních RTG snímků

Vážená paní doktorko, vážený pane doktore, předkládaný dotazník slouží ke sběru primárních dat pro výzkumnou část bakalářské práce zabývající se integrací nástrojů umělé inteligence (AI) do diagnostiky skeletálních rentgenových snímků. Cílem je zmapovat reálné postoje, úroveň důvěry a vnímané bariéry adopce AI ze strany odborné radiologické komunity. Dotazník je plně anonymní, jeho vyplnění zabere přibližně 5 minut a získaná data budou využita výhradně pro akademické účely.

Blok 1: Demografický a expertní profil

1. Délka klinické praxe v oboru radiologie:

- Předatestační příprava (0–5 let)
- Atestovaný lékař (5–15 let)
- Atestovaný lékař (více než 15 let)

2. Typ primárního pracoviště:

- Fakultní nemocnice
- Krajská / Městská nemocnice
- Soukromá ambulance / Poliklinika

3. Dosavadní praktická zkušenost s využíváním nástrojů umělé inteligence pro analýzu obrazu v klinické praxi:

- Aktivní využívání v současnosti nebo v minulosti (např. v rámci pilotního testování).
- Teoretická znalost bez dosavadní reálné praxe s AI.
- Zcela bez předchozích zkušeností.

Blok 2: Vnímání spolehlivosti a vysvětlitelnosti (XAI)

*U následujících tvrzení prosím vyznačte míru Vašeho souhlasu na škále 1 až 5 (1 = Zcela nesouhlasím, 5 = Zcela souhlasím).*

4. Prediktivním modelům umělé inteligence (tzv. black-box), které poskytnou pouze textový výsledek bez přesné vizualizace nálezu v obraze, nelze v klinické praxi plně důvěřovat.

- 1 (Zcela nesouhlasím) – 2 – 3 – 4 – 5 (Zcela souhlasím)

5. Vizualní zdůvodnění rozhodnutí algoritmu (např. ohraničující boxy nebo teplotní mapy označující frakturu) významně zvyšuje diagnostickou jistotu při přebírání výsledku.

- 1 (Zcela nesouhlasím) – 2 – 3 – 4 – 5 (Zcela souhlasím)

6. Pokud by systém AI vyhodnotil rentgenový snímek jako negativní (bez fraktury), ale subjektivně by přetrvávalo podezření na patologii, s vysokou pravděpodobností by bylo algoritmické rozhodnutí ignorováno.

- 1 (Zcela nesouhlasím) – 2 – 3 – 4 – 5 (Zcela souhlasím)

7. Nástroje umělé inteligence jsou vnímány primárně jako nástroj pro:

- „Second reader“ (druhé čtení pro kontrolu vlastního popisu).
- Triage (prioritizace akutních a pozitivních nálezů v pracovní frontě).
- Nástroj pro edukaci a trénink mladších lékařů.
- Nástroj s potenciálem plně nahradit primární popis u jasně negativních nálezů.

Blok 3: Procesní a legislativní bariéry

8. Je považováno za absolutně nezbytné, aby byly výstupy AI integrovány přímo do nativního prohlížeče PACS (bez nutnosti otevírat externí aplikaci nebo webový prohlížeč)?

- Ano, nutnost přepínat okna by zcela eliminovala přínos a časovou úsporu.
- Ideální stav, nicméně využití externího okna/aplikace by nepředstavovalo překážku.
- Způsob zobrazení nehraje ve workflow zásadní roli.

9. Co je vnímáno jako největší aktuální bariéra pro plošné nasazení AI do české radiologické praxe? (Lze vybrat maximálně 2 možnosti)

- Nejasná legislativa a nevyřešená právní odpovědnost za případnou diagnostickou chybu.
- Vysoké pořizovací náklady pro zdravotnická zařízení.
- Přílišná četnost falešně pozitivních nálezů (FP), které zdržují od práce.
- Technická náročnost a nedostatečná integrace s nemocničními IT systémy (PACS/RIS).
- Nedůvěra ze strany starší generace lékařů.

10. Jak by měla být v současnosti řešena právní odpovědnost v případě, že je převzat falešně negativní výsledek od certifikovaného AI nástroje a je tak přehlédnuta fraktura?

- Odpovědnost by měl nést výhradně popisující lékař (AI je pouze podpůrný nástroj).
- Odpovědnost by měla být sdílena mezi lékařem a dodavatelem AI softwaru.
- Odpovědnost by měla nést nemocnice (poskytovatel zdravotní péče), která systém nasadila.

11. (Nepovinné) Další komentář k úskalím či přínosům zavádění AI do hodnocení skeletálních rentgenových snímků:

- [Textové pole pro volnou odpověď]

**Přílohy B Datové soubory**

**Příloha B.1 Datový soubor pro vyhodnocení výzkumné části práce**

Příloha je přiložena v elektronické podobě.

Formát souboru: .xlsx

**Příloha B.2 Datový soubor výsledků dotazníkového šetření mezi radiology**

Příloha je přiložena v elektronické podobě.

Formát souboru: .csv