

KORELACE MEZI PROCENTUÁLNÍM ZACHYCEMÍ VOŠTINY NA HRCT PLIC VYHODNOCENÉ AUTOMATICKÝM PROGRAMEM CALIPER A VIZUÁLNÍ IDENTIFIKACÍ RADIOLOGY

CORRELATION BETWEEN PERCENTAGE OF HONEYCOMBING ON CHEST HRCT EVALUATED BY THE CALIPER AUTOMATED PROGRAM AND VISUAL IDENTIFICATION BY RADIOLOGIST

původní práce

Martin Hyršl¹
Eva Kočová¹
Pavel Eliáš¹
Jiří Vaňásek¹
Jiří Jandura¹
Vladimír Bartoš²
Jan Novosad³

¹Radiologická klinika LF UK a FN, Hradec Králové

²Plicní klinika LF UK a FN, Hradec Králové

³Ústav klinické imunologie a alergologie LF UK a FN, Hradec Králové

Přijato: 15. 8. 2018.

Korespondenční adresa:

MUDr. Martin Hyršl
 Radiologická klinika LF UK a FN
 Sokolská 581
 500 05 Hradec Králové
 e-mail: martin.hyrsl@fnhk.cz

Konflikt zájmů: žádný.

Hlavní stanovisko práce

Cílem této studie je určit procentuální hranici zastoupení voštiny vyhodnocenou programem CALIPER, která koreluje s vizuální identifikací přítomnosti voštiny skupinou radiologů.

SOUHRN

Hyršl M, Kočová E, Eliáš P, Vaňásek J, Jandura J, Bartoš V, Novosad J. Korelace mezi procentuálním zachycením voštiny na HRCT plic vyhodnocené automatickým programem CALIPER a vizuální identifikací radiology

Cíl: Jedná se o pilotní studii s cílem zjistit korelaci mezi procentuálním zachycením voštiny na HRCT plic vyhodnocené automatickým programem CALIPER a vizuální identifikací přítomnosti voštiny pěti zkušenými radiology u pacientů se známou intersticiální plicní fibrózou (IPF).

Metodika: Patnáct pacientů vybraných z abecedního seznamu pacientů z Fakultní nemocnice Hradec Králové (FNHK) se stanovenou diagnózou idiopatické plicní fibrózy (IPF). U 12 z nich byla provedena analýza programem CALIPER a následně i ohodnocení pěti radiology z FNHK.

Výsledky: Nejlepšího poměru mezi senzitivitou a specificitou bylo dosaženo při 3,5 procentuálním zastoupení voštiny v pravé plicí, kdy bylo dosaženo senzitivity 80 % a specificity 85,7 %. U dvou hodnocených pacientů došlo k výrazně rozdílné interpretaci přítomnosti voštiny.

Major statement

The main goal of this study is to specify the threshold of honeycombing calculated by program CALIPER that correlates with the visual identification of honeycombing presence by a group of radiologists.

SUMMARY

Hyršl M, Kočová E, Eliáš P, Vaňásek J, Jandura J, Bartoš V, Novosad J. Correlation between percentage of honeycombing on chest HRCT evaluated by the CALIPER automated program and visual identification by radiologist

Purpose: The pilot study to investigate the correlation between the percentage of HRCT honeycombing captured by the CALIPER automatic program and the visual identification of this morph presence by five radiological examinations in patients with known interstitial pulmonary fibrosis (IPF).

Methods: 15 patients selected from the alphabetical list of patients from the Faculty Hospital Hradec Králové (FNHK) with established diagnosis of idiopathic pulmonary fibrosis (IPF). For 12 of them, the CALIPER program was analysed, followed by the evaluation of five radiologists from FNHK.

Results: The best ratio between sensitivity and specificity was achieved at a 3.5 percent representation of honeycomb in right lung, when the sensitivity was 80% and the specificity was 85.7%. 13% of patients evaluated had a significantly different interpretation of honeycomb presence.

Závěr: Pilotní studie prokázala, že stanovení 3,5% hranice pro přítomnost voštiny má 80% senzitivitu a 85,7% specifitu v porovnání mezi programem CALIPER s vizuálním konsenzem radiologů. Pro limitace v této studii doporučujeme další výzkum.

Klíčová slova: CALIPER, intersticiální plicní fibróza, voština.

Conclusion: Pilot studies have shown that the 3.5% threshold for the presence of honeycombing has 80% sensitivity and 85.7% specificity when is program CALIPER compared to the visual consensus of radiologists. For limitations in this study, we recommend further research.

Key words: CALIPER, honeycombing, interstitial lung fibrosis.

ÚVOD

Idiopatická plicní fibróza (IPF) je definována jako specifická forma chronické progresivní fibrotizující intersticiální pneumonie neznámé příčiny (1). Jedná se o vzácné onemocnění, jehož incidence v čase narůstá, nyní je v Evropě udávána mezi 1,25 až 23,4 případy na 100 000 obyvatel (2). Postižení bývají nejčastěji muži mezi 60. a 70. rokem života (3). Typicky se projevuje nevysvětlitelnou dušností při zátěži, často spojenou s kašlem. Objektivně jsou přítomny inspirační chrupky při bázích plic a paličkovité prsty (4). Mezi faktory mající vliv na vývoj choroby patří kouření, znečištění prostředí, infekce, gastroezofageální reflux a v neposlední řadě genetická predispozice (5). Diagnostika se opírá o multidisciplinární konsilium, které se skládá z pneumologa, radiologa a patologa. Je to právě HRCT plic, které umožní zařadit pacienta s intersticiální plicní pneumonií do jedné ze tří kategorií, tzn. typická (typical) UIP, pravděpodobná (possible) UIP, nekonzistentní s UIP znaky (inconsistent with UIP Pattern) (6) (tab. 1). Jedna z nejvýznamnějších morf, která je specifická pro typickou UIP, je nález voštiny (honeycombing) (6).

Voština vzniká na podkladě destrukce a fibrotizace plicní tkáně. Fleischner Society's Glossary of Terms for Thoracic Imaging definuje voštinu jako vícečetné cystické prostory (shluky cyst) s dobře definovatelnou stěnou, vyskytující se nejčastěji subpleurálně (7).

Problém správné identifikace voštiny na výpočetní tomografii s vysokým rozlišením (high resolution computed tomography – HRCT) je již od roku 2013 známou skutečností (8). Identifikace voštiny je zcela zásadní pro správnou diagnostiku intersticiální plicní fibrózy jakožto podkladu pro diagnózu typické obvyklé plicní pneumonie (6). Nástup nových metod automatického vyhodnocování HRCT plic umožnil nový způsob hodnocení přítomnosti této morf. Cílem této pilotní studie je zjistit míru konsenzu mezi pěti radiology (z nichž dva mají více jak 15letou praxi v oboru plicní radiodiagnostiky) a programem – Computer Aided Lung Informatics for Pathology Evaluation and Rating (CALIPER). Program vznikl ve spolupráci s Mayo Clinic a FN Hradec Králové byl zapůjčen firmou Imbio.

METODIKA

Pro tuto pilotní studii bylo vybráno 15 pacientů z abecedního seznamu pacientů z Fakultní nemocnice Hradec Králové se stanovenou diagnózou idiopatické plicní fibrózy (IPF). Všichni tito pacienti jsou v současné době léčeni biologickou léčbou. Diagnostika proběhla standardní cestou na podkladě rozhodnutí multidisciplinárního týmu složeného z pneumologa, radiologa a patologa. Multidisciplinární tým měl přístup

Tab. 1. HRCT kritéria pro diagnostiku UIP
Table 1. HRCT criteria for diagnosing UIP

UIP Pattern (All Four Features)	Possible UIP Pattern (All Three Features)	Inconsistent with UIP Pattern (any of the Seven Features)
• Subpleural, basal predominance	• Subpleural, basal predominance	• Upper or mid-lung predominance
• Reticular abnormality	• Reticular abnormality	• Peribronchovascular predominance
• Honeycombing with or without traction bronchiectasis	• Absence of features listed as inconsistent with UIP pattern (see third column)	• Extensive ground glass abnormality (extent > reticular abnormality)
• Absence of features listed as inconsistent with UIP		• Profuse micronodules (bilateral, predominantly upper lobes)
		• Discrete cysts (multiple, bilateral, away from areas of honeycombing)
		• Diffuse mosaic attenuation/air-trapping (bilateral in three or more lobes)
		• Consolidation in bronchopulmonary segment(s)/lobe(s)

Tab. 2. **Míra jistoty radiologa ohledně přítomnosti voštiny na HRCT plic**
 Table 2. **Radiologist's certainty regarding the presence of honeycombing on HRCT lungs**

1	definitivně ano
2	pravděpodobná
3	nelze rozhodnout
4	nepravděpodobná
5	definitivně ne

k anamnéze, informacím o přítomnosti antigenní expozice, fyzikálnímu a funkčnímu plicnímu vyšetření, výsledkům rozpočtu tekutiny bronchoalveolární laváže, HRCT plic a v několika případech i k výsledkům plicní biopsie. Hodnocené skeny byly provedeny ve standardních rekonstrukcích. Ke studii byly použity skeny široké 0,6–1,5 mm, rekonstruované s ostrým kernelem 70 f, v plicním okně. Posouzení probíhalo metodou pěti bodového ohodnocení, které vyjadřovalo míru jistoty radiologa ohledně přítomnosti voštiny (tab. 2). Program CALIPER přiřazoval procentuální zastoupení voštiny v každé plíci a poté i souhrnné procentuální zastoupení v obou plících. Všichni hodnocení pacienti byli ve studii zaslepeni, bez informace o klinickém stavu pacienta.

Metodika statistického zkoumání

Deskriptivní statistiky jsou vyjádřeny jako průměry a směrodatné odchylky, resp. frekvence, průměry (mean) a interkvartilová rozpětí v případě kategoriálních dat či dat nesplňujících podmínky normality. Průměry byly porovnávány Mannovým-Whitneyovým testem pro nezávislé soubory. Vzájemné korelace mezi proměnnými byly kvantifikovány Spearmanovým koeficientem Rho, resp. Kendallovým Tau či koeficientem Eta. Vzájemné asociace mezi proměnnými byly hodnoceny metodou faktorové analýzy s využitím metody extrakce „analýzou z hlavních komponent“ (PCA – principal components analysis) a ortogonální rotací Varimax. Nejvhodnější cut-off hodnoty měření procentuální pří-

Tab. 3. **Průměrné hodnoty procentuálního zastoupení voštiny identifikované programem CALIPER**

Table 3. **The average values of the percentage of the honeycombing identified by the program CALIPER**

Deskriptiva	Statistika	
	průměr	SD
CALIPER celkově (%)	3,5	2,64575
CALIPER L plíce (%)	3,6667	2,99495
CALIPER P plíce (%)	3,5833	2,60971

Tab. 4. **Průměrné hodnoty míry jistoty přítomnosti voštiny na HRCT plic hodnoceno radiology**

Table 4. **The average values of the degree of assurance of honeycombing presence on HRCT lungs are evaluated by radiologists**

Deskriptiva	Statistika	
	průměr	SD
radiolog (1)	3,1667	1,80067
radiolog (2)	2,5833	1,50504
radiolog (3)	2,4167	1,44338
radiolog (4)	2,9167	1,37895
radiolog (5)	2,4167	1,62135

tomnosti voštiny programem CALIPER pro predikci klinicky významných změn v plicním parenchymu byly navrženy s využitím ROC (receiver operating characteristic) analýzy, výsledné hodnoty senzitivity a specifity jsou prezentovány. Jako statisticky významné byly označeny výsledky, pro něž $p < 0,05$. Statistická analýza byla provedena programem IBM SPSS Statistics pro Windows, Verze 22.0. Armonk, NY: IBM Corp.

VÝSLEDKY

Do pilotní studie bylo zařazeno 15 pacientů s IPF. Hodnocení programem CALIPER bylo provedeno u 12 z nich. Pro technicky nedostatečné parametry provedených CT byli tři pacienti vyřazeni. Do studie bylo zapojeno deset mužů a dvě ženy, medián věku byl 74 let (věkové rozmezí pacientů bylo 64–86 let).

Průměrné hodnoty procentuálního zastoupení voštiny identifikované programem CALIPER, a to jak pro plíce jako celek, tak i izolovaně pro levou a pravou plíci, včetně směrodatné odchylky (SD), jsou uvedeny v tabulce 3. Tabulka 4 ukazuje průměrné hodnoty míry jistoty přítomnosti voštiny hodnotícími radiology na škále 1 až 5 (viz tab. 2). Všechny korelační koeficienty (až na jeden) mezi hodnotiteli a analýzami programu CALIPER (celkové hodnocení plic jako celku) byly nesignifikantní (tab. 5).

Metodou faktorové analýzy jednotlivých hodnotitelů a výsledků získaných z programu CALIPER využitím „analýzy hlavních komponent“ byly extrahovány dvě základní komponenty (virtuální proměnné), které dávají vznik dvěma odlišným faktorům. Jeden obsahoval pouze výsledky výpočtu z programu CALIPER, druhý všechny hodnotitele (tab. 6a,b, graf 1); Kaiserův-Meyerův-Olkinův KMO test = 0,065.

Jednotlivé HRCT obrazy byly dále rozděleny do dvou skupin dle průměrného hodnocení všech rentgenologů; skupina 1: voština přítomna s vysokou pravděpodobností (průměrné skóre v intervalu 1 až 2), skupina 2: přítomnost voštiny nehodnotitelná či nepravděpodobná (průměrné skóre bylo vyšší než 2). V obou skupinách byly následně porovnány naměřené hodnoty procentuálního zastoupení voštiny měřené programem CALIPER. Ty dosahovaly vyšších hodnot v případě vysoké pravděpodobnosti voštiny dle rentgenologů (tab. 7), rozdíl však nedosahovaly statistické významnosti (pravá plíce: $p = 0,068$, levá plíce: $p = 0,326$, celkově: $p = 0,097$), Mannův-Whitneyův test (graf 2, 3, 4).

Pro odhad kritické hodnoty procentuální přítomnosti voštiny měřené programem CALIPER, vhodné k odlišení výše definovaných dvou skupin HRCT nálezů, byla využita analýza

Tab. 5. **Všechny korelační koeficienty mezi hodnotiteli a analýzami programu CALIPER**

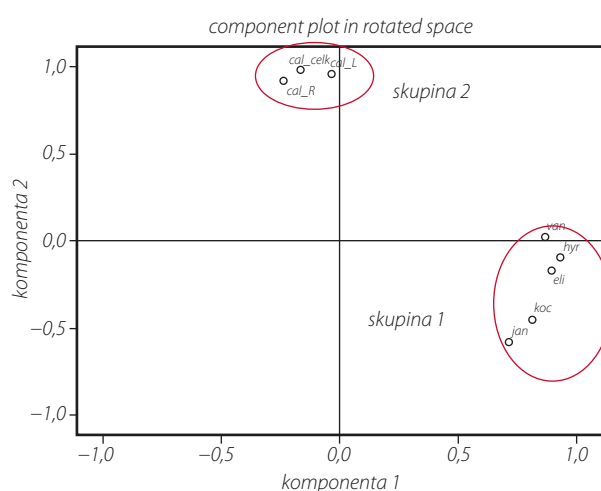
Table 5. **All correlation coefficients between evaluators and CALIPER analyses**

Korelace	CALIPER celkově (%)	
	radiolog 1 (15 let praxe)	553
	radiolog 2 (25 let praxe)	308
	radiolog 3 (10 let praxe)	155
	radiolog 4 (10 let praxe)	660
	radiolog 5 (2 roky praxe)	265

Tab. 6a,b; graf 1. **Dvojice komponent pro radiology a program CALIPER, které vznikly metodou faktorové analýzy jednotlivých hodnotitelů a výsledků získaných z programu CALIPER využitím analýzy hlavních komponent**

Table 6a,b; graph 1. **The pair of components for radiologists and the CALIPER program, which were created by the method of factor analysis of the individual evaluators and the results obtained from the CALIPER program using the analysis of the main components**

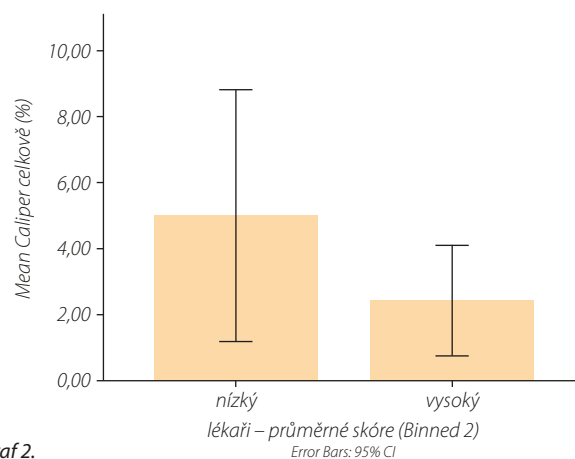
Skupina 1	Komponenta 1
radiolog 1	812
radiolog 2	894
radiolog 3	866
radiolog 4	713
radiolog 5	932
Skupina 2	Komponenta 2
CALIPER celkově (%)	0,98
CALIPER L plíce (%)	959
CALIPER P plíce (%)	921



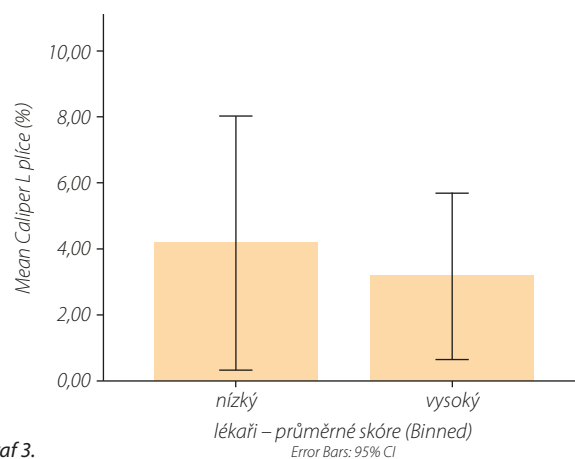
Tab. 7. **Jednotlivé HRCT obrazy byly rozděleny do dvou skupin. První skupina s vysokou pravděpodobností přítomnosti voštiny, druhá skupina s nehodnotitelnou či nepravděpodobnou voštinou. V obou skupinách byly následně porovnány naměřené hodnoty procentuálního zastoupení voštiny měřené programem CALIPER.**

Table 7. **Individual HRCT images were divided into two groups. The first group with high probability of honeycombing presence, the second group with an unpredictable or unlikely honeycombing. In both groups, the measured values of the percentage of the honeycombing measured by the CALIPER program were then compared.**

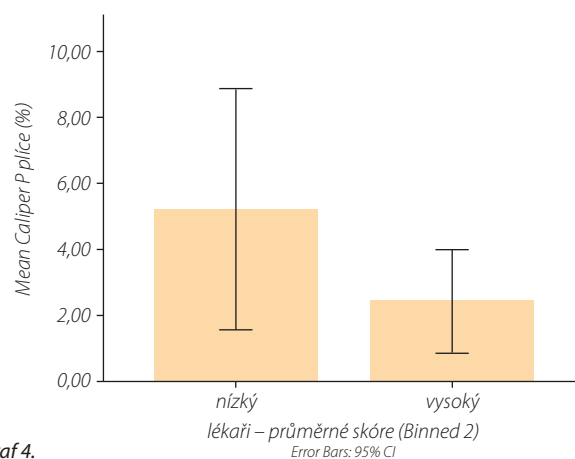
	Lékaři – průměrné skóre	N	Průměr	SD
CALIPER celkově (%)	nízké (1–2)	5	5,0000	3,08221
	vysoké (> 2)	7	2,4286	1,81265
CALIPER L plíce (%)	nízké (1–2)	5	4,8000	3,70135
	vysoké (> 2)	7	2,8571	2,34013
CALIPER P plíce (%)	nízké (1–2)	5	5,2000	2,94958
	vysoké (> 2)	7	2,4286	1,71825



Graf 2.



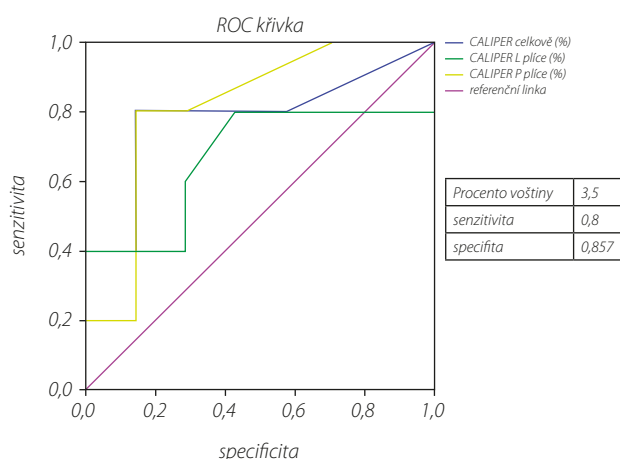
Graf 3.



Graf 4.

Graf 2, 3, 4. **Mannův-Whitneyův test; pravá plíce: $p = 0,068$, levá plíce: $p = 0,326$, celkově: $p = 0,097$**

Graph 2, 3, 4. **Mann-Whitney's test; right lung: $p = 0.068$, left lung: $p = 0.326$, total: $p = 0.097$**



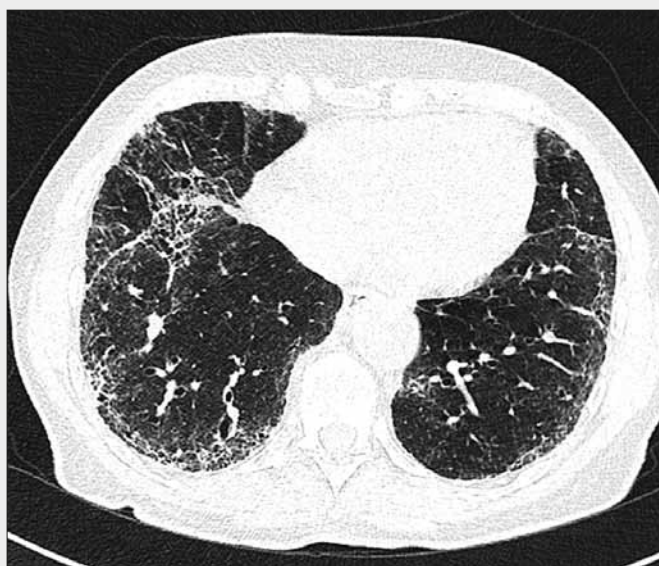
Graf 5. Odhad kritické procentuální přítomnosti voštiny měřené programem CALIPER za pomoci ROC analýzy
Graph 5. Estimation of critical percentage of honeycombing measured by CALIPER with ROC analysis

ROC. Jako optimální se ukázala být hodnota 3,5 % (senzitivita 0,8, specifita 0,857) měřená v pravé plicí (graf 5). Tento výsledek však nedosahoval statistické významnosti ($p = 0,074$).

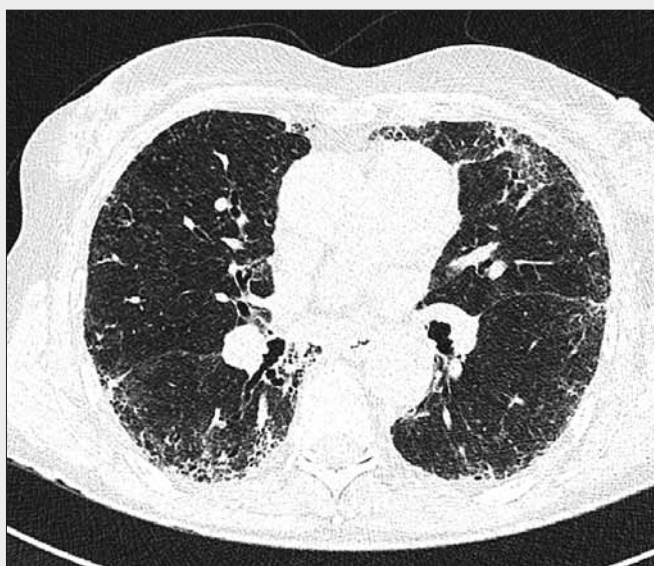
DISKUSE

Cílem této pilotní studie je zjistit, jaké procento voštiny zachycené programem CALIPER nejlépe koreluje s vizuálním zhodnocením voštiny radiology. Na základě výsledků statistické analýzy lze konstatovat, že nejlepší senzitivity a specifity na přítomnost voštiny dosahoval program CALIPER při hranici procentuálního zastoupení voštiny 3,5 % a více. Tato hranice platila jak pro přítomnost v jedné plicí, tak i pro celkové procento v obou plicích. Senzitivita při této hranici dosahuje 80 % a specifita 85,7 %.

Všichni pacienti v této skupině měli podle multidisciplinárního týmu diagnostikováno IPF. Přítomnost voštiny jakožto určujícího faktoru pro diagnostiku typické UIP byla



▲ Obr. 1



▲ Obr. 2



▲ Obr. 3

Obr. 1, 2, 3. HRCT plic, axiální a koronární rovina. Radiology byly identifikované vícečetné retikulace subpleurálně, trakční bronchiectázie a denzity mléčného skla. Voštiny radiology identifikována nebyla. Program CALIPER identifikoval 7% zastoupení v levé plicí a 6% v pravé plicí.

Fig. 1, 2, 3. HRCT lung, axial and coronal lane. Radiologists have identified multiple subpleural reticulation, traction bronchiectasis and ground glass opacity. Honeycombing was not identified. CALIPER identified presence of honeycombing in 7% of the left lung and 6% of the right lung.



▲ Obr. 4



▲ Obr. 5

Obr. 4, 5. **HRCT plic, axiální, koronální rovina.** Radiology identifikována pokročilá plicní fibróza v dorzobazální části plic, byla rozpoznána přítomnost voštiny. Program CALIPER přítomnost voštiny nepotvrdil.

Fig. 4, 5. **HRCT lung, axial, coronal plane.** Radiologist identified advanced pulmonary fibrosis in the dorsal part of the lungs has been recognized the presence of honeycombing. The CALIPER program did not confirm the presence of the honeycombing.

původními hodnotícími radiology přiznána celkem u sedmi z těchto 12 pacientů.

Výsledky studie ukázaly, že rozpoznání voštiny na HRCT plic je obtížným úkolem, a to i pro odborníky s dlouholetou praxí. Problém identifikovat voštinu radiologem vzniká z řady důvodů. Nejčastěji je to správná interpretace toho, co pojem voština znamená. Právě rigidní dodržování této definice může vést radiologa k posouzení subpleurálních bronchiolektázií či paraseptálního emfyzému jako přítomnosti voštiny. Proto je vždy nutné přistupovat k hodnocení plicního parenchymu opatrně a uvažovat o plicích komplexně, nikoliv pouze izolovaně identifikovat přítomnost morf.

Program CALIPER byl vyvinut institutem Biomedical Imaging Resource Laboratory na Mayo Clinic, Rochester, MN, USA. Tento program je za pomoci histogramových map schopen analyzovat patologické morfy na HRCT plic a následně i identifikovat jak celkové procento postižení v jedné plicí, tak i celkové procento postižení v obou plicích, včetně lokalizace této patologie (9).

Dva pacienti byli výrazně jinak zhodnoceni radiology a programem CALIPER. U pacienta 6 jsou zřejmé četné subpleurální retikulace s trakčními bronchiektáziemi a ojedinělými denzitami mléčného skla. Přítomnost voštiny nebyla pěti radiology identifikována, pouze jeden radiolog rozpoznal přítomnost voštiny dorzobazálně vpravo. Program CALIPER identifikoval až 7% zastoupení voštiny v levé plicí a 6% v pravé plicí (obr. 1, 2, 3).

Odlišným případem byl pacient 9. Na HRCT plic (obr. 4, 5) byla identifikována pokročilá plicní fibróza, v dorzobazální

části plic (zejména vlevo) byla všemi radiology rozpoznána přítomnost voštiny. Program CALIPER však v této lokalitě identifikoval pouze přítomnost retikulací, nikoliv přítomnost voštiny. Jednoznačné vyhodnocení přítomnosti voštiny by mohla přinést pouze kryobiopsie, která však zatím nebyla provedena.

Hlavní limitací této pilotní retrospektivní studie je relativně malý vzorek pacientů, který byl ještě snížen z důvodu náročnosti programu CALIPER na kvalitu pořízeného vyšetření HRCT plic. Retrospektivní povaha této studie navíc neumožňovala jednotný protokol vyšetření u všech plic, což mohlo vést k částečnému ovlivnění výsledného procenta zachycené voštiny.

Dalším limitujícím faktorem byla samotná identifikace voštiny, která stála na konsenzu pěti radiologů, kteří se mezi sebou lišili jak délkou praxe, tak i zkušenostmi v hodnocení plicních patologií. Významnou limitací je dále neexistence bi-optického průkazu přítomnosti voštiny, která je jako jediná s jistotou schopna potvrdit či vyvrátit její existenci.

ZÁVĚR

Rozpoznání voštiny u pacientů s intersticiálním postižením hraje klíčovou roli pro identifikaci UIP. Prezentovaná pilotní studie prokázala, že stanovení 3,5% hranice pro přítomnost voštiny má 80% senzitivitu a 85,7% specifitu v porovnání s vizuálním konsenzem radiologů. Pilotní studie prokázala potřebnost a účelnost dalšího výzkumu v této oblasti.

LITERATURA

1. American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). *Am J Respir Crit Care Med* 2000; 161(2 Pt 1): 646–664.
2. Nalysnyk L, Cid-Ruzafa J, Rotella P, Esser D. Incidence and prevalence of idiopathic pulmonary fibrosis: review of the literature. *European Respiratory* 2012; 21(126): 355–361.
3. Raghu G, Chen SY, Hou Q, Yeh WS, Collard HR. Incidence and prevalence of idiopathic pulmonary fibrosis in US adults 18–64 years old. *European Respiratory Journal* 2016; 48(1): 179–186.
4. Nakamura Y, Suda T. Idiopathic pulmonary fibrosis: diagnosis and clinical manifestations. *Clinical Medicine Insights Circulatory, Respiratory and Pulmonary Medicine* 2015; 9(Suppl 1): 163–171.
5. Raghu G, Collard HR, Egan JJ, et al. An Official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *American Journal of Respiratory and Critical Care Medicine* 2011; 183(6): 788–824.
6. Lynch DA, Sverzellati N, Travis WD. Diagnostic criteria for idiopathic pulmonary fibrosis: a fleischner society white paper. *The lancet Respiratory Medicine* 2018; 6(2): 138–153.
7. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J. Fleischner society: glossary of terms for thoracic imaging. *Radiology* 2008; 246(3): 697–722.
8. Watadani T, Sakai F, Johkoh T, et al. Interobserver variability in the CT assessment of honeycombing in the lungs. *Radiology* 2012; 266(3): 936–944.
9. Bartholmai BJ, Raghunath S, Karwowski RA, et al. Quantitative CT imaging of interstitial lung diseases. *Journal of Thoracic Imaging* 2013; 28(5): 10.