



Test Report

Perceptra Inspectra CXR Version 3

โดย บริษัทเพอเซ็ปตรา จำกัด

รายงานผลการทดสอบ


โดยราชวิทยาลัยรังสีแพทย์แห่งประเทศไทย

ทดสอบใช้กับภาพรังสีทรวงอก ในกรณี

- คัดกรอง (screening) วัณโรคปอด
- อ่านผลซ้ำ (double reading) ให้กับรังสีแพทย์ เพื่อเพิ่มคุณภาพการวินิจฉัย
- เพิ่มความแม่นยำในการค้นหาพยาธิสภาพให้กับรังสีแพทย์
- ประมาณความยาก-ง่ายในการแปลผล
- จัดลำดับความเร่งด่วน (triage) ในการแปลผลให้แก่รังสีแพทย์

รายงานผลการทดสอบปัญญาประดิษฐ์ในภาพถ่ายรังสีทรวงอก

บริษัทที่ยื่นทดสอบ

หน่วยงาน	บริษัท เพอเซ็ปทรา จำกัด	 PERCEPTRA SEIZE THE FUTURE
ที่อยู่	อาคารศุภาคาร ห้องเลขที่ L2 ชั้น L เลขที่ 723 ถนนเจริญนคร แขวงคลองตันใหม่ เขตคลองสาน กรุงเทพมหานคร 10600	
ชื่อผู้ติดต่อ	ชัยวัฒน์ พุพิสุทธิ์ และสุพิชญา พุพิสุทธิ์	

บริษัทผู้พัฒนา

บริษัท	บริษัท เพอเซ็ปทรา จำกัด	
ที่อยู่	อาคารศุภาคาร ห้องเลขที่ L2 ชั้น L เลขที่ 723 ถนนเจริญนคร แขวงคลองตันใหม่ เขตคลองสาน กรุงเทพมหานคร 10600	
ประเทศ	ไทย	
เว็บไซต์	https://www.perceptra.tech	

ระบบที่ยื่นทดสอบ

ชื่อระบบ	Inspectra CXR
เวอร์ชัน	V.3
ข้อมูลระบบ	

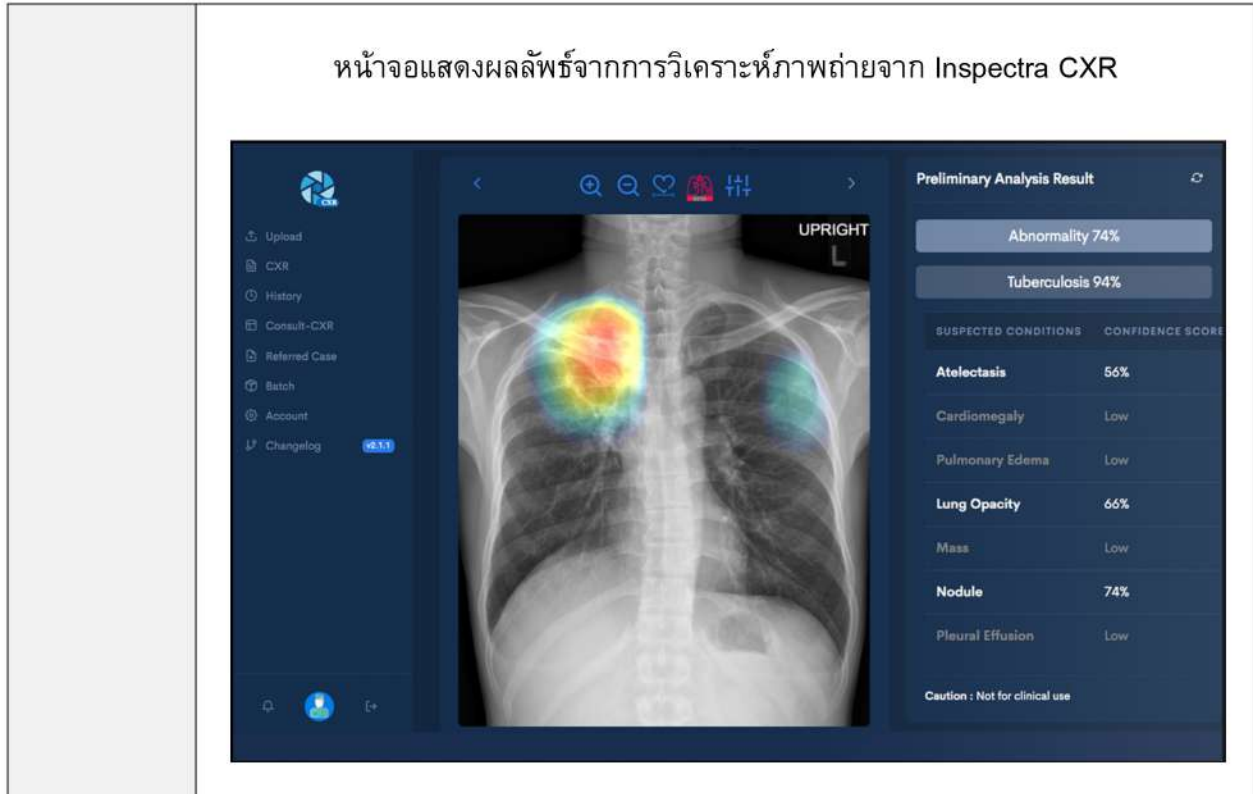


Inspectra CXR is an AI-based software that analyzes adult frontal chest x-rays for signs of abnormalities in order to prioritize them for clinical review. The device aims to assist in prioritizing and triaging chest x-ray images. Inspectra CXR identifies features suggestive of 8 abnormality conditions:

1. Pleural effusion
2. Tuberculosis
3. Lung opacity
4. Pulmonary edema
5. Mass
6. Nodule
7. Cardiomegaly
8. Atelectasis

The standalone software consists of both on-premise and on-cloud modules. The on-premise module interacts with the hospital Picture Archiving and Communication System (PACS) in order to retrieve input DICOM images and de-identify data before sending it to the on-cloud module to process. The on-cloud deep learning algorithms then analyze chest x-rays and output structured reports detailing probabilities of abnormalities in the image. The software is capable of producing a secondary capture DICOM image with AI analysis results to send back to the PACS system or display on MyInspectra web application. The software flags cases with suspected findings, showing passive notification on the work list to notify qualified physicians to urgently review the suspected cases. X-rays without an identified anomaly are placed on the worklist for a routine review, which is the current standard of care. Inspectra CXR does not mark, highlight, or direct user attention to a specific location on the original chest x-ray image. The device generates a secondary image from the original image with an overlay that has to be turned on by qualified physicians when additional information is needed in the clinical review.

The system is compatible with adult chest radiographs starting from 15 years old. The system sorts out pediatric chest radiographs, Posterior-Anterior (PA) Upright radiographs, irrelevant radiographs, and non-digital chest radiographs before processing. The system provides preliminary screening which allows only images indicated as the lung and with the patient's age stated. Images without the part of the patient's body identified or without the age stated will not be processed. The device intends to only assist radiologists /trained clinicians to facilitate clinical review of chest x-rays. It does not include clinical or patient historical data or recommends treatments or diagnoses. The final decision has to be provided by attending physicians or radiologists.



ชุดข้อมูลที่ใช้ทดสอบ

รหัสชุดข้อมูล	1A
จำนวนภาพถ่าย	300 ภาพ

คุณลักษณะของข้อมูลที่ใช้ทดสอบ

จำนวนภาพถ่าย 300 ภาพได้ถูกสุ่มมาจากชุดข้อมูลทดสอบจำนวน 1,500 ภาพซึ่งได้รับการสนับสนุนจาก โรงพยาบาลสงขลานครินทร์ จังหวัดสงขลา, โรงพยาบาลเชียงรายประชารักษ์ จังหวัดเชียงราย, โรงพยาบาลอุดรธานี จังหวัดอุดรธานี, โรงพยาบาลสุทธาเวช จังหวัดมหาสารคาม, และกองวัณโรค กรมควบคุมโรค กระทรวงสาธารณสุข โดยทุกภาพจะถูกอ่านโดยรังสีแพทย์ B Reader จำนวน 3 ท่าน

คุณลักษณะของภาพถ่ายรังสีในโครงการ:

- ภาพถ่ายรังสีทรวงอกของผู้ป่วยที่มีอายุตั้งแต่ 15 ปี ด้วยเครื่องเอกซเรย์คอมพิวเตอร์
- ไม่มีภาพถ่ายจากผู้ป่วยที่มีสถานะ HIV Serology เป็นบวก
- ไม่มีภาพถ่ายจากผู้ป่วยที่มีการติดเชื้อที่ปอดแบบฉวยโอกาสอื่น ๆ หรือการติดเชื้อร่วม เช่น การติดเชื้อไมโครแบคทีเรีย อีสโตพลาสโมซิส ไครปโตคอคโคซิส เมลิออยโดซิส และแอคติโนมัยโคซิส

การวัดความตรงภายในในชุดข้อมูล ใช้ตัววัดดังต่อไปนี้

- Pairwise Agreement ค่าเฉลี่ยของความเหมือนกันกันระหว่างแต่ละคู่ของรังสีแพทย์
- Intraclass Agreement (ICC) ค่าเฉลี่ยของ Pearson's correlation แบบ ICC(2,3) เมื่อมีรังสีแพทย์ 3 ท่าน อ่านภาพถ่ายรังสีแบบสุ่ม
- การวิเคราะห์ความตรงระหว่างผู้ประเมินโดยใช้สถิติ Pairwise Cohen's Kappa และ Fless' Kappa

จำนวนรอยโรคในชุดข้อมูล

ตารางที่ 1 จำนวนรอยโรคที่ระบุโดยรังสีแพทย์ ในภาพถ่ายรังสีที่ได้รับการวินิจฉัยไม่เป็นวัณโรคปอด ($N_{\text{Non-TB}}$) และในภาพถ่ายรังสีได้รับยืนยันผลวัณโรคปอดโดยผลตรวจเสมหะย้อมเชื้อหรือผลตรวจเพาะเชื้อ (N_{TB}) รวมถึงจำนวนรอยโรคที่ได้รับการยืนยันว่ามีรอยโรคจากรังสีแพทย์ส่วนใหญ่ ($N_{\text{Consensus}}$) ในชุดข้อมูล 1A จำนวน 300 ภาพ แต่ละภาพมีการอ่านโดยรังสีแพทย์ “B” Reader จำนวน 3 รายแบบสุ่ม

Finding	$N_{\text{Non-TB}}$	N_{TB}	$N_{\text{Consensus}}$
Abnormality	53	448	158
Small opacity	33	374	135
Primary nodular	1	274	98
Primary reticular	32	87	21
Secondary nodular	9	224	76
Secondary reticular	24	125	30
Large opacity	2	349	119
Mass/nodule	2	114	29
Cavity	0	242	81
Fibrosis	7	196	58
Calcification	5	61	12
Pleural effusion	0	99	34
Pleural thickening	7	122	39
Pneumothorax	0	1	0
Hilar adenopathy	2	87	18
Mediastinal adenopathy	1	30	7
Consistent with tuberculosis	0	422	144
Active Tuberculosis	0	387	136
Patchy infiltration	0	265	99
Cavity with surrounding consolidation	0	210	73
Unilateral hilar/paratracheal lymph node enlargement	0	39	5
Pleural effusion	0	52	15
Miliary nodules	0	74	20
Indeterminate tuberculosis	0	35	6
Reticulonodular infiltration	0	22	4
Destroyed lung or bronchiectasis	0	3	0
Inconsistent with tuberculosis	450	28	156

ความตรงภายในของชุดข้อมูล

ตารางที่ 2 ความตรงภายในระหว่างผู้ประเมิน (Inter-rater Reliability) ของแต่ละรอยโรคในชุดข้อมูล 1A จำนวน 300 ภาพ ของรังสีแพทย์ “B” Reader จำนวน 3 ราย ซึ่งวัดโดยค่าสถิติ Pairwise Agreement, ICC(2,3), Cohen’s kappa และ Fleiss’ kappa

Finding	Agreement	ICC	Cohen’s	Fleiss’
Abnormality	0.8956	0.9181	0.7884	0.7884
Small opacity	0.8356	0.8584	0.6681	0.6681
Primary nodular	0.8422	0.8357	0.6281	0.6282
Primary reticular	0.7911	0.2301	0.0892	0.0897
Secondary nodular	0.7222	0.5352	0.2751	0.2761
Secondary reticular	0.7578	0.3006	0.1260	0.1233
Large opacity	0.8800	0.8993	0.7477	0.7478
Mass/nodule	0.8378	0.5365	0.2777	0.2776
Cavity	0.8600	0.8449	0.6446	0.6439
Fibrosis	0.7978	0.6866	0.4212	0.4212
Calcification	0.8889	0.4027	0.1829	0.1825
Pleural effusion	0.9578	0.9163	0.7838	0.7844
Pleural thickening	0.8711	0.7319	0.4762	0.4752
Pneumothorax	0.9978	0.0000	0.0000	0.0000
Hilar adenopathy	0.8644	0.4867	0.2395	0.2394
Mediastinal adenopathy	0.9556	0.6008	0.3130	0.3318
Consistent with tuberculosis	0.9644	0.9751	0.9286	0.9286
Active Tuberculosis	0.9178	0.9373	0.8323	0.8323
Patchy infiltration	0.8178	0.7940	0.5613	0.5614
Cavity with surrounding consolidation	0.8333	0.7755	0.5333	0.5342
Unilateral hilar/paratracheal lymph node enlargement	0.9244	0.2289	0.0846	0.0887
Pleural effusion	0.9444	0.7431	0.4941	0.4897
Miliary nodules	0.8978	0.5905	0.3212	0.3227
Indeterminate tuberculosis	0.9444	0.5114	0.2588	0.2568
Reticulonodular infiltration	0.9600	0.3705	0.1947	0.1613
Destroyed lung or bronchiectasis	0.9933	0.0067	0.0015	0.0033
Inconsistent with tuberculosis	0.9644	0.9751	0.9286	0.9286

ตารางที่ 3 การตีความค่า ICC และ Kappa ตาม Landis and Koch (1977)¹

ICC/Kappa Statistic	Strength of Agreement
<0.00	Poor
0.00 – 0.20	Slight
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 – 1.00	Almost Perfect

¹Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. In *Biometrics* (Vol. 33, Issue 1, p. 159). JSTOR. <https://doi.org/10.2307/2529310>

ผลการทดสอบ

การวัดความตรงภายนอกนั้นวัดโดยใช้ Pairwise Agreement หรือค่าเฉลี่ยของความเหมือนกันระหว่างคู่ของแต่ละรังสีแพทย์และปัญญาประดิษฐ์และ Pairwise Cohen's Kappa หรือค่าเฉลี่ยของค่าสถิติ Cohen's Kappa ระหว่างคู่ของแต่ละรังสีแพทย์และปัญญาประดิษฐ์ เพื่อเปรียบเทียบระหว่างรังสีแพทย์กับปัญญาประดิษฐ์ ("B" vs AI) และรังสีแพทย์ด้วยกัน ("B" vs "B")

ตารางที่ 4 การวิเคราะห์ความตรงกันภายนอกระหว่างรังสีแพทย์และปัญญาประดิษฐ์

Finding	N	Threshold	Pairwise Agreement		Cohen's Kappa	
			"B" vs "B"	"B" vs AI	"B" vs "B"	"B" vs AI
Abnormality	501	*	0.9007	0.8811	0.7993	0.7578
Tuberculosis	422	0.43	0.9635	0.9411	0.9261	0.8817
Opacity [#]	464	0.38	0.9087	0.9244	0.8190	0.8488
Pleural Effusion	99	0.17	0.9578	0.9544	0.7838	0.7704

[#]Opacity combines small and large opacities.

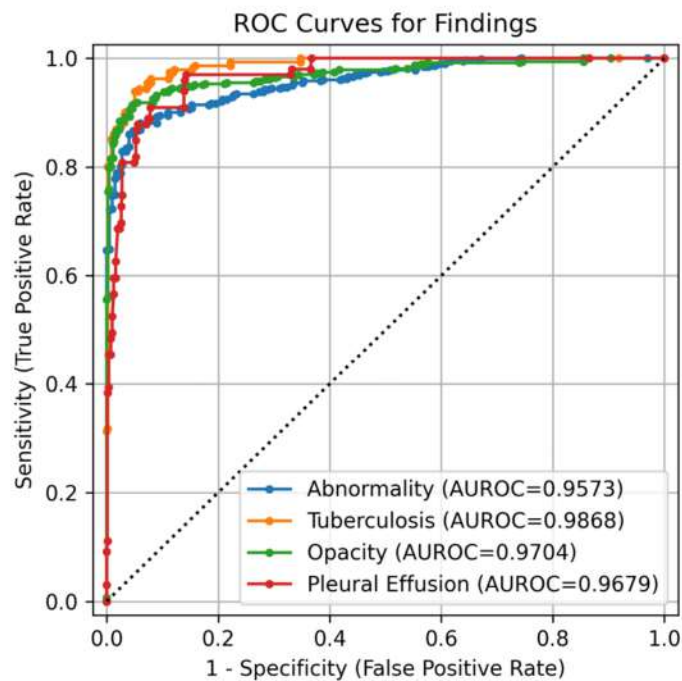
ในการวัดประสิทธิภาพในการวินิจฉัยของแต่ละรอยโรคนั้นใช้เกณฑ์ ความไว (Sensitivity), ความจำเพาะ (Specificity), อัตราส่วนการทำนายผู้ป่วยที่เป็นโรคที่ถูกต้อง (Positive Prediction Rate, PPR) และอัตราส่วนการทำนายผู้ป่วยที่ไม่เป็นโรคที่ถูกต้อง (Negative Prediction Rate, NPR) ด้วยเกณฑ์การวินิจฉัย (Diagnostic Threshold) ที่ระบุโดยผู้ผลิต รวมถึงพื้นที่ใต้กราฟ ROC เช่นกัน

ตารางที่ 5 ประสิทธิภาพในการวินิจฉัยของแต่ละรอยโรคของปัญญาประดิษฐ์เปรียบเทียบกับรังสีแพทย์

Finding	N	Threshold	Specificity	Sensitivity	PPV	NPV	AUROC
Abnormality	501	*	0.9248	0.8802	0.9363	0.8601	0.9573
Tuberculosis	422	0.43	0.9498	0.9313	0.9424	0.9400	0.9868
Opacity [#]	464	0.38	0.9312	0.9181	0.9342	0.9144	0.9704
Pleural Effusion	99	0.17	0.9725	0.8081	0.7843	0.9762	0.9679

[#]Opacity combines small and large opacities.

รูปภาพที่ 1 เส้นโค้ง ROC ของแต่ละรอยโรค



บทวิจารณ์

สำหรับการทดสอบ Spectra CXR V.3 มีรอยโรคจำนวน 3 รายการที่ตรงกับรายการรอยโรคในแบบทดสอบ คือ Tuberculosis, Lung Opacity และ Pleural Effusion

ตารางที่ 6 ความแตกต่างระหว่าง Pairwise Agreement และ Cohen's Kappa

Finding	Pairwise Agreement			Cohen's Kappa		
	B vs "B"	"B" vs AI	Diff	"B" vs "B"	"B" vs AI	Diff
Abnormality	0.9007	0.8811	-1.96%	0.7993	0.7578	-4.15%
Tuberculosis	0.9635	0.9411	-2.24%	0.9261	0.8817	-4.44%
Opacity [#]	0.9087	0.9244	+1.57%	0.8190	0.8488	+2.98%
Pleural Effusion	0.9578	0.9544	-0.34%	0.7838	0.7704	-1.34%

[#]Opacity combines small and large opacities.

จากตารางที่ 6 เมื่อเทียบ Pairwise Agreement และ Cohen's Kappa ระหว่างรังสีแพทย์กับปัญญาประดิษฐ์ ("B" vs AI) และรังสีแพทย์ด้วยกัน ("B" vs "B") แล้ว Spectra CXR V.3 มีประสิทธิภาพใกล้เคียงกับรังสีแพทย์ (ความแตกต่างน้อยกว่า 5%) ในส่วนของการวินิจฉัยวัณโรค (Tuberculosis) ซึ่งปัญญาประดิษฐ์มีคะแนนน้อยกว่ารังสีแพทย์อยู่ 2.24% สำหรับ Pairwise Agreement และ คะแนนน้อยกว่ารังสีแพทย์อยู่ 4.44% สำหรับ Cohen's Kappa (N = 422), การตรวจพบ Lung Opacity ซึ่งปัญญาประดิษฐ์มีคะแนนมากกว่ารังสีแพทย์อยู่ 1.57% สำหรับ Pairwise Agreement และ คะแนนมากกว่ารังสีแพทย์อยู่ 2.98% สำหรับ Cohen's Kappa (N = 464), และการตรวจพบ Pleural Effusion ซึ่งปัญญาประดิษฐ์มีคะแนน น้อยกว่ารังสีแพทย์อยู่ 0.34% สำหรับ Pairwise Agreement และ คะแนนน้อยกว่ารังสีแพทย์อยู่ 1.34% สำหรับ Cohen's Kappa (N = 99)

ในส่วนของการคัดกรองวัณโรคปอด Spectra CXR V.3 เมื่อวิเคราะห์บนชุดข้อมูล 1A จำนวน 300 ภาพ พบว่าประสิทธิภาพในการวินิจฉัยวัณโรคใกล้เคียงกับรังสีแพทย์ และมีค่าพื้นที่ใต้กราฟตัวรับ (AUROC) อยู่ที่ 0.9868 ค่าความไว (Sensitivity) อยู่ที่ 0.9313 และค่าความจำเพาะ (Specificity) อยู่ที่ 0.9498 ที่ Threshold เป็น 0.43

อ้างอิงถึงเกณฑ์ [The Target Product Profiles \(TPPs\) for a rapid non-sputum-based biomarker test for tuberculosis detection](#) ของ World Health Organisation (WHO) ตามตารางที่ 7 จะเห็นได้ว่าแต่ละเกณฑ์การวัดมีค่า Sensitivity และ Specificity ที่แตกต่างกัน

ตารางที่ 7 TPP for a rapid non-sputum-based biomarker test for tuberculosis detection

	Minimal Requirements		Optimal Requirements	
	Sensitivity	Specificity	Sensitivity	Specificity
Smear-replacement test	Overall >80%	98%	Overall >95%	98%
	Positive >99%		Positive >99%	
	Negative >60%		Negative >68%	
Non-sputum based biomarker test	Overall >65%	98%	Positive >98%	98%
	Positive >98%		Negative >68%	
Triage test	90%	70%	95%	80%

เอกสารอ้างอิง: https://academic.oup.com/ijid/article/211/suppl_2/S29/2490781

ผลลัพธ์ของการคัดกรองวัณโรคปอด Spectra CXR V.3 ใน Threshold ต่าง ๆ เมื่อเทียบกับเกณฑ์ WHO TPP โดยกำหนดให้เป็นค่า Threshold ที่สูงที่สุดที่จะทำให้ค่า Specificity นั้นมีความใกล้เคียงกับ WHO TPP มากที่สุด ผลลัพธ์ที่ได้นั้นเป็นดังตารางที่ 8

ตารางที่ 8 ค่า Sensitivity และ Specificity ที่ Threshold ต่าง ๆ ตามเกณฑ์ WHO TPP

Threshold	Sensitivity	Specificity
0.6198	0.8697	0.9832
0.4300	0.9313	0.9498
0.1761	0.9858	0.8410
0.1453	0.9929	0.7782

ทั้งนี้เมื่อนำผลลัพธ์ที่ได้เปรียบเทียบกับเกณฑ์ WHO TPP พบว่า Inspectra CXR V.3 นั้นผ่านเกณฑ์ Triage test ทั้งใน Minimal Requirements และ Optimal Requirements และผ่านเกณฑ์ Smear-replacement test และ Non-sputum based biomarker test สำหรับเกณฑ์ Minimal Requirements ซึ่งผลการทดสอบ ได้ถูกสรุปในตารางที่ 9

ตารางที่ 9 ผลการทดสอบการคัดกรองวัณโรคปอด Inspectra CXR V.3 ตามเกณฑ์ WHO TPP

	Minimal Requirements	Optimal Requirements
Smear-replacement test	ผ่าน	ไม่ผ่าน
Non-sputum based biomarker test	ผ่าน	ไม่ผ่าน
Triage test	ผ่าน	ผ่าน

/8 เมษายน 2566