



ความไวและความจำเพาะของการแปลผลภาพถ่ายทางรังสีทรวงอกด้วยปัญญาประดิษฐ์ในผู้ป่วยมะเร็งปอด

วีรยา น้อยศิริ พ.บ.*, ชมพูนุท วิจิตรสงวน พ.บ.*, สายใจ เลิศโรจน์ปัญญา พ.บ.*, กิตติกา เจียมจิต พ.บ.*, จิราวรรณ ฉายจรรณ วท.บ.*, จาตุรงค์ ตันติบัณฑิต วศ.บ.**

* กลุ่มงานรังสีวิทยา สถาบันโรคทรวงอก ตำบลบางกระสอ อำเภอเมืองนนทบุรี

จังหวัดนนทบุรี 11000

** ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต อำเภอคลองหลวง จังหวัดปทุมธานี 12120

Abstract: Sensitivity and Specificity of Artificial Intelligence for Chest Diagnostic Radiology in Lung Cancer

Weeraya Noisiri M.D.*, Chomphunut Vijitrsaguan, M.D.*, Saijai Lertrojpanya, M.D.*, Kittika Jiamjit, M.D.*, Jirawan Chayjaroon, B.Sc.*, Charturong Tantibundhit, Ph.D.**

*Department Of Radiology, Central Chest Institute of Thailand, Bang Kraso, Mueang Nonthaburi, Nonthaburi, 11000

**Department of Electrical and Computer Engineering, Faculty of Engineering, Thammasat University Rangsit Campus, Khlong Luang, Pathumthani 12120 (E-mail:weeraya_tat@yahoo.com)

(Received: June 30, 2020; Revised: September 3, 2020; Accepted: September 15, 2020)

Background: AIChest4All is the model development for screening abnormalities in chest radiograph and classification as normal, suspected active TB, suspected lung malignancy, abnormal heart and great vessels, intrathoracic abnormal findings and extrathoracic abnormal findings. The purpose of this artificial intelligence (AI) was to aid radiologists and clinicians, especially in the rural areas. **Objectives:** To analyze the sensitivity and specificity of artificial intelligence for chest radiograph in diagnosis of lung malignancy. **Methods:** The pathological and cytological reports were retrospectively reviewed. 800 patients of malignancy and 716 patients of non malignancy were randomly selected. The chest radiographs in a 3-month period before the procedures were collected. The chest radiographs were reviewed by three radiologists and classified as lung malignancy and non lung malignancy groups. The same radiographs were evaluated by AI and reported as percent of probability. The cut point for detected lung cancer was analyzed. The sensitivity and specificity of lung cancer diagnosis by radiologist and AI were calculated. **Results:** The sensitivity and specificity in diagnosis of lung malignancy by radiologist was 67.5% and 83.1%, respectively. The sensitivity and specificity of AI to detected lung malignancy in chest radiographs was 50.0% and 84.8%, respectively. The cut point of probability percent which presented by AI for appropriated sensitivity and specificity was 52.5. **Conclusions:** The sensitivity of AI in diagnosis of lung malignancy from chest radiographs was slightly less than radiologist. The specificity of AI was comparable to the diagnosis by radiologist.

Keywords: Sensitivity, Specificity, Artificial intelligence, Chest radiograph, Lung cancer

บทคัดย่อ

ภูมิหลัง: AIChest4All คือ ปัญญาประดิษฐ์ที่พัฒนาขึ้นเพื่อช่วยตรวจคัดกรองโรคที่พบจากภาพเอกซเรย์ทรวงอก การรายงานผลจะแบ่งเป็นปกติ สงสัยวัณโรคระยะแพร่กระจายเชื้อ สงสัยมะเร็ง

ปอด ความผิดปกติของหัวใจและหลอดเลือด ความผิดปกติในทรวงอกและความผิดปกตินอกทรวงอก โดยมีจุดมุ่งหมายเพื่อช่วยแพทย์หรือรังสีแพทย์ในการวินิจฉัยโรคโดยเฉพาะตามต่างจังหวัด ที่อาจจะมีความรู้รังสีแพทย์ไม่เพียงพอ **วัตถุประสงค์:** เพื่อหาความ

ไวและความจำเพาะของปัญญาประดิษฐ์ในการวินิจฉัยมะเร็งปอด

วิธีการ: ผลพยาธิวิทยาและเซลล์วิทยาของผู้ป่วยถูกนำมาศึกษาย้อนหลัง ผลที่เป็นมะเร็งปอดจำนวน 800 ราย และที่ไม่ใช่มะเร็งปอดจำนวน 716 ราย ถูกคัดเลือกมาโดยวิธีสุ่มภาพรังสีทรวงอกของผู้ป่วยในช่วงระยะเวลาไม่เกิน 3 เดือนจากการทำหัตถการจะได้รับการอ่านและแปลผลโดยรังสีแพทย์ 3 ท่านว่าพบมะเร็งปอดหรือไม่ ภาพเอกซเรย์เดียวกันถูกนำมาวิเคราะห์โดยปัญญาประดิษฐ์และรายงานค่าความสงสัยมะเร็งปอดออกมาในรูปร้อยละ ผลที่ได้จะถูกนำมาเปรียบเทียบกับผลชิ้นเนื้อ เพื่อหาจุดตัดที่เหมาะสมหาความไวและความจำเพาะในการวินิจฉัยมะเร็งปอดของรังสีแพทย์และปัญญาประดิษฐ์ **ผล:** ความไวและความจำเพาะในการวินิจฉัยมะเร็งปอดของรังสีแพทย์คือร้อยละ 67.5 และ ร้อยละ 83.1 ตามลำดับ ส่วนปัญญาประดิษฐ์มีความไวและความจำเพาะเท่ากับร้อยละ 50.0 และร้อยละ 84.8 ตามลำดับ จุดตัดของความน่าจะเป็นในการวินิจฉัย คือ 52.5 **สรุป:** ความจำเพาะในการวินิจฉัยมะเร็งปอดของปัญญาประดิษฐ์มีค่าใกล้เคียงกับรังสีแพทย์ ส่วนความไวของปัญญาประดิษฐ์มีค่าน้อยกว่ารังสีแพทย์เล็กน้อย

คำสำคัญ: ความไว ความจำเพาะ ปัญญาประดิษฐ์ ภาพรังสีทรวงอก มะเร็งปอด

บทนำ

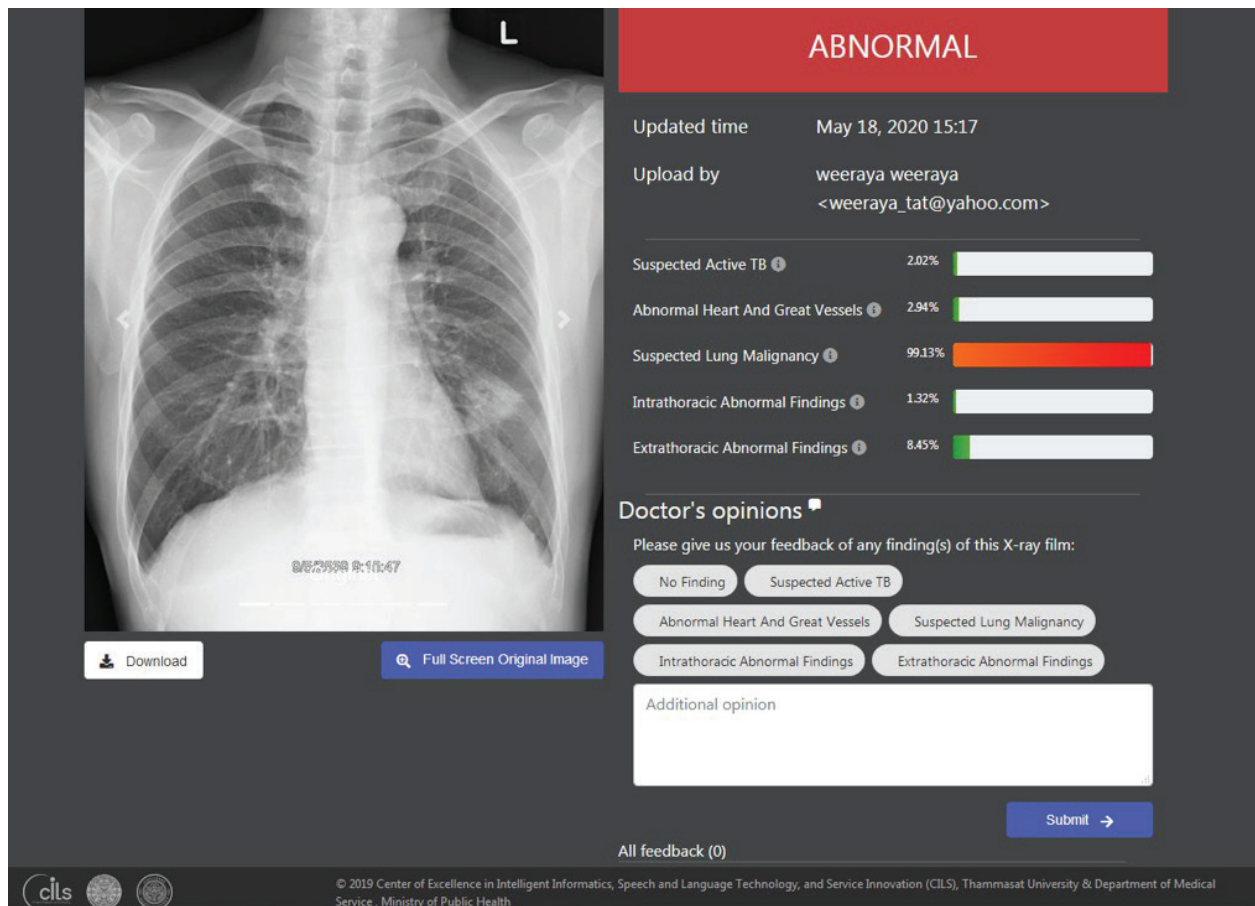
จากการศึกษาก่อนหน้านี้พบอุบัติการณ์ของมะเร็งประมาณ 18.1 ล้านรายและพบสาเหตุการตายจากมะเร็ง 9.6 ล้านรายทั่วโลก ในปี ค.ศ. 2018¹. มะเร็งปอดยังคงเป็นมะเร็งที่มีอุบัติการณ์และเป็นสาเหตุการตายอันดับต้นๆ โดยพบมะเร็งปอดรายใหม่ประมาณ 2.1 ล้านรายและมีอัตราการเสียชีวิตประมาณ 1.8 ล้านรายในปี 2018 คิดเป็นร้อยละ 18.4 ของสาเหตุการตายจากมะเร็งทั้งหมด

ในปัจจุบันถึงแม้ว่าการตรวจคัดกรองมะเร็งปอดด้วยเอกซเรย์คอมพิวเตอร์จะมีประสิทธิภาพสูงกว่าการตรวจด้วยเอกซเรย์ปอด แต่เอกซเรย์ปอดยังคงเป็นวิธีตรวจคัดกรองแรกเนื่องจากมีราคาถูก ใช้ปริมาณรังสีต่ำและสามารถเข้าถึงได้ง่ายกว่า²

ด้วยการร่วมมือกันของหลายสถาบันในประเทศ ได้แก่ ศูนย์แห่งความเป็นเลิศทางวิชาการด้านสารสนเทศอัจฉริยะ เทคโนโลยีสิ่งพูดและภาษาและนวัตกรรมด้านการบริการ (CILS) คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธรรมศาสตร์ โรงพยาบาลมะเร็งอุดรธานี สถาบันโรคทรวงอกและกรมการแพทย์ ทำให้เกิดการพัฒนปัญญาประดิษฐ์ที่ชื่อว่า AIChest4All (DMS TU) ขึ้นมาโดยมีวัตถุประสงค์เพื่อช่วยแพทย์และรังสีแพทย์ในแปลผลภาพรังสีทรวงอก โดยเฉพาะในโรงพยาบาลชุมชนที่อาจจะมีรังสีแพทย์ไม่เพียงพอ

ในการพัฒนปัญญาประดิษฐ์มีขั้นตอนที่สำคัญ 3 ขั้นตอน คือ 1. การประมวลผลภาพด้วยคอมพิวเตอร์เพื่อให้ได้ภาพที่ละเอียดคมชัดหรือสามารถนำข้อมูลที่อยู่ในภาพมาใช้งานในด้านอื่นๆ ต่อได้ (image processing) 2. การสอนให้ระบบคอมพิวเตอร์ใช้ข้อมูลในการเรียนรู้ด้วยตนเอง (machine learning) 3. อัลกอริทึมแบบระบบเรียนรู้เชิงลึกคล้ายวิธีการทำงานของระบบประสาทในสมองมนุษย์ (deep learning) โดยภาพรังสีทรวงอกที่จะนำมาให้ปัญญาประดิษฐ์เรียนรู้และจดจำ จะต้องผ่านการประเมินคุณภาพและจัดกลุ่มโดยทีมรังสีแพทย์จากโรงพยาบาลมะเร็งอุดรธานี และสถาบันโรคทรวงอก และเป็นภาพที่รังสีแพทย์ให้ความเห็นตรงกันอย่างน้อย 2 ใน 3 ท่าน พบว่ามีภาพจำนวน 89,733 ภาพ จากจำนวนภาพทั้งหมด 101,783 ภาพ ที่รังสีแพทย์วิเคราะห์ว่ามีคุณภาพและมีความเห็นตรงกัน จึงนำมาให้ปัญญาประดิษฐ์ใช้เรียนรู้

AIChest4All สามารถตรวจพบและรายงานว่าภาพเอกซเรย์นั้นมีความผิดปกติหรือไม่ โดยจะสรุปผลออกมาเป็นร้อยละของความน่าจะเป็นทั้งหมด 5 ประเภท คือ สงสัยวัณโรคระยะแพร่กระจายเชื้อ (suspected active TB), สงสัยมะเร็งปอด (suspected lung malignancy), ความผิดปกติของหัวใจและหลอดเลือด (abnormal heart and great vessels), ความผิดปกติในทรวงอก (intrathoracic abnormal findings) และความผิดปกติในนอกทรวงอก (extrathoracic abnormal findings) (ภาพที่ 1)



ภาพที่ 1 ตัวอย่างการรายงานผลของปัญญาประดิษฐ์

ในกลุ่มที่สงสัยมะเร็งปอดจะรวมทั้งมะเร็งที่เกิดในปอด (lung cancer) และมะเร็งที่แพร่กระจายมาที่ปอด (lung metastasis) ส่วนความผิดปกติใน mediastinum จะจัดอยู่ในกลุ่มความผิดปกติในทรวงอก มะเร็งที่กระจายมาที่เยื่อหุ้มปอดจะจัดในกลุ่มความผิดปกตินอกทรวงอก

การศึกษานี้จัดทำขึ้นเพื่อวิเคราะห์หาความไวและความจำเพาะของปัญญาประดิษฐ์ในการแยกโรคมะเร็งปอดโดยเปรียบเทียบกับผลวินิจฉัยของผู้ป่วยและหาจุดตัดที่เหมาะสมของร้อยละความน่าจะเป็นของมะเร็งปอดเพื่อใช้ในการวินิจฉัยมะเร็งปอด

วัตถุประสงค์และวิธีการ

การศึกษานี้ได้ผ่านการพิจารณาจากคณะกรรมการจริยธรรมเพื่อการวิจัยของสถาบันโรคทรวงอก เป็นการศึกษาแบบย้อนหลัง เก็บข้อมูลจากผลพยาธิวิทยาและผลเซลล์วิทยาของผู้ป่วยที่มีอายุตั้งแต่ 18 ปีขึ้นไปที่ได้รับการผ่าตัด การเจาะดูดชิ้นเนื้อหรือการตัดชิ้นเนื้อผ่านผนังทรวงอก โดยใช้เครื่องเอกซเรย์คอมพิวเตอร์หรืออัลตราซาวด์นำทาง หรือ โดยการส่องกล้องหลอดลม ระหว่างเดือนมกราคม พ.ศ. 2559 ถึงเดือนธันวาคม พ.ศ. 2562 ถ้ามีการทำหัตถการซ้ำในผู้ป่วยคนเดียวกัน จะเลือกใช้ผลชิ้นเนื้อและภาพรังสีทรวงอกเพียงครั้งเดียวโดยเลือกผลชิ้นเนื้อจากการผ่าตัดเป็นหลักเพราะมีความละเอียดและมีความน่าเชื่อถือมากกว่า ส่วนผล

ชิ้นเนื้อที่ไม่เพียงพอไม่สามารถให้การวินิจฉัยได้จะถูกคัดออก ก่อนบริเวณ mediastinum จะไม่รวมในการศึกษานี้ ผลชิ้นเนื้อที่ได้จะนำมาแบ่งผู้ป่วยเป็นสองกลุ่ม คือ กลุ่มที่มีมะเร็งปอดและกลุ่มที่ไม่มีมะเร็งปอด ถือเป็น การวินิจฉัยจากการตรวจมาตรฐาน

ภาพเอกซเรย์ทรวงอกแบบดิจิทัลในระยะเวลาไม่เกิน 3 เดือนก่อนการทำหัตถการจะถูกรวบรวมโดยรังสีแพทย์ 1 ท่าน ภาพเอกซเรย์ที่ไม่ได้มาตรฐาน มีน้ำในช่องเยื่อหุ้มปอดมาก หรือมีภาวะปอดแฟบมากจนบดบังก้อนหรือรอยโรคในปอดจะถูกตัดออกจากการศึกษานี้ ภาพเอกซเรย์จำนวน 800 ภาพได้ถูกสุ่มเลือกจากแต่ละกลุ่ม ในกลุ่มที่ไม่มีมะเร็งปอดและผลชิ้นเนื้อไม่ได้มาจากการผ่าตัดจะมีการติดตามภาพเอกซเรย์คอมพิวเตอร์หรือภาพเอกซเรย์ทรวงอกเพื่อความแน่ใจว่าผู้ป่วยไม่ได้เป็นมะเร็ง เช่น ก้อนหรือรอยโรคเล็กลงหรือหายไป รอยโรคไม่มีการเปลี่ยนแปลงเป็นระยะเวลามากกว่า 2 ปีขึ้นไป ภาพเอกซเรย์ของผู้ป่วยที่ไม่แน่ใจว่าเป็นมะเร็งปอดหรือไม่จะถูกคัดออก ภาพเอกซเรย์ที่เหลือทั้งหมดจะได้รับการอ่านและแปลผลแบบสุ่มโดยรังสีแพทย์ที่เหลืออีก 3 ท่านที่ทำงานทางด้านเอกซเรย์ทรวงอกโดยเฉพาะ ภาพเอกซเรย์ชุดเดียวกันนี้จะถูกป้อนเข้าไปและแปลผลโดยปัญญาประดิษฐ์ ทำให้สามารถหาความไวและความจำเพาะในการอ่านของรังสีแพทย์และปัญญาประดิษฐ์โดยเทียบกับผลชิ้นเนื้อซึ่งเป็นผลมาตรฐาน

นอกจากนี้ ยังได้ทำการวิเคราะห์ความไวและความจำเพาะ โดยแบ่งกลุ่มตามภาพเอกซเรย์ทรวงอกที่พบ คือ ก้อนเดี่ยว ก้อน จำนวนน้อยไม่เกิน 5 ก้อน ก้อนหลายตำแหน่งมากกว่า 5 ก้อน ก้อน กระจายเต็มปอดทั้ง 2 ข้าง, focal consolidation และ patchy infiltration

ความไวคำนวณจากจำนวนผู้ป่วยที่ผลอ่านเป็นมะเร็งและ ผลขึ้นเนื้อเป็นมะเร็งต่อจำนวนผู้ป่วยที่ผลขึ้นเนื้อเป็นมะเร็งทั้งหมด

ความจำเพาะคำนวณจากจำนวนผู้ป่วยที่ผลอ่านไม่เป็น มะเร็งและผลขึ้นเนื้อไม่เป็นมะเร็งต่อจำนวนผู้ป่วยที่ไม่เป็นมะเร็ง ทั้งหมด

ผล

ระหว่างเดือน มกราคม พ.ศ. 2559 จนถึงเดือนธันวาคม พ.ศ. 2562 ผู้ป่วยจำนวน 3,275 รายได้รับการผ่าตัด ตัดหรือเจาะ ดูดชิ้นเนื้อ ณ สถาบันโรคทรวงอก ผลขึ้นเนื้อจำนวน 1,075 จาก 4,350 รายถูกตัดออกเนื่องจากการเป็นกรทำหัตถการในผู้ป่วยราย เดิม อีก 470 รายได้รับการตัดออกเนื่องจากผลไม่สามารถให้การ วินิจฉัยได้ จากนั้น 800 ภาพรังสีทรวงอกถูกสุ่มเลือกมาจากแต่ละ กลุ่ม หลังจากรังสีแพทย์ประเมินภาพการติดตามเอกซเรย์ในกลุ่มที่

ไม่ใช่มะเร็ง ผู้ป่วยอีก 84 รายถูกตัดออกเนื่องจากไม่มีความแน่นอน ในการวินิจฉัยว่าไม่เป็นมะเร็ง ภาพเอกซเรย์ทั้งหมด 1,516 ภาพได้ ถูกรวบรวมโดยรังสีแพทย์ ผู้ป่วย 907 ราย (ร้อยละ 59.8) เป็นผู้ชาย 609 ราย (ร้อยละ 40.2) เป็นผู้หญิง ผู้ป่วยทั้งหมดมีอายุอยู่ระหว่าง 18 ปีถึง 88 ปี อายุเฉลี่ยของผู้ป่วยในกลุ่มที่เป็นมะเร็ง 52 ± 16 ปี (51-53 ปี) น้อยกว่าผู้ป่วยในกลุ่มที่ไม่เป็นมะเร็ง 64 ± 11 ปี (63-65 ปี) ($p < 0.001$) ขนาดเฉลี่ยของก้อนในกลุ่มที่ไม่เป็นมะเร็งปอด 1.3 ± 2.0 ซม. (1.2-1.5 ซม.) จะเล็กกว่าในกลุ่มมะเร็งปอดคือ 3.7 ± 2.2 ซม. (3.6-3.9 ซม.) ($p < 0.001$) (ตารางที่ 1)

ในกลุ่มที่เป็นมะเร็งพบผลขึ้นเนื้อเป็น adenocarcinoma มากที่สุด คือ 412 ราย (ร้อยละ 51.5) รองลงมาเป็น non-small cell carcinoma 253 ราย (ร้อยละ 31.6), squamous cell carcinoma 71 ราย (ร้อยละ 8.9) ส่วนอีกร้อยละ 8.0 เป็นมะเร็ง ชนิดอื่นๆ

ส่วนกลุ่มที่ไม่เป็นมะเร็งพบผลขึ้นเนื้อเป็น การอักเสบของ ปอด 151 ราย (ร้อยละ 21.0) วัณโรคปอด 146 ราย (ร้อยละ 20.4) ถุงลมโป่งพองรวมถึง blebs และ bullae 118 ราย (ร้อยละ 16.4) และปอดอักเสบจากเชื้อรา (aspergillosis) 60 ราย (ร้อยละ 8.3)

ตารางที่ 1 ลักษณะพื้นฐานของผู้ป่วย

ลักษณะพื้นฐาน	กลุ่มที่เป็นมะเร็ง (n = 800)	กลุ่มที่ไม่เป็นมะเร็ง (n = 716)	p-value
ผู้ชาย	469 (58.6)	438 (61.1)	0.319
อายุ (ปี)	63.77 ± 11.50	52.16 ± 16.13	<0.001
ขนาดก้อน* (ซม.)	3.74 ± 2.27	1.32 ± 1.97	<0.001

*ขนาดก้อน = ขนาดของก้อนที่พบโดยวัดจากระยะที่มากที่สุดของก้อนในภาพรังสีทรวงอก

ข้อมูลแสดงเป็นจำนวน (ร้อยละ) หรือ ค่าเฉลี่ย $\pm$ ส่วนเบี่ยงเบนมาตรฐาน ค่า $p < 0.05$ มีนัยสำคัญทางสถิติ

ความไวและความจำเพาะของรังสีแพทย์เท่ากับ 67.5 % (95% CI = 64.2, 70.7) และ 83.1% (95% CI = 80.2, 85.7) ตาม ลำดับ (ตารางที่ 2)

ตารางที่ 2 การเปรียบเทียบผลการวินิจฉัยโดยรังสีแพทย์กับผลขึ้นเนื้อในโรคมะเร็งปอด

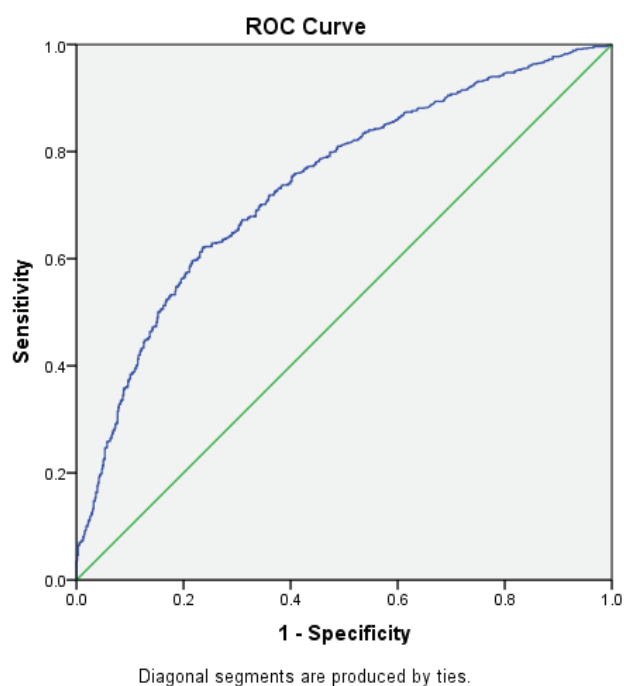
		ผลขึ้นเนื้อ		
		เป็นมะเร็งปอด (ราย)	ไม่เป็นมะเร็งปอด (ราย)	รวม
การวินิจฉัยโดยรังสีแพทย์	เป็นมะเร็งปอด (ราย)	540	121	661
	ไม่เป็นมะเร็งปอด (ราย)	260	595	855
	รวม	800	716	1,516

จุดตัดของความน่าจะเป็นในการวินิจฉัยมะเร็งปอด มีค่า เท่ากับ 52.5 ทำให้ได้ความไวและความจำเพาะของปัญญาประดิษฐ์ เท่ากับ 50.0% (95% CI = 46.5, 53.5) และ 84.8% (95% CI

= 82.0, 87.2) ตามลำดับ สามารถคำนวณได้จากตารางที่ 3 และ แสดงเป็น receiver operating characteristic (ROC) curves (ภาพที่ 2)

ตารางที่ 3 การเปรียบเทียบผลการวินิจฉัยโดยปัญญาประดิษฐ์กับผลชิ้นเนื้อในโรคมะเร็งปอด

		ผลชิ้นเนื้อ		รวม
		เป็นมะเร็งปอด (ราย)	ไม่เป็นมะเร็งปอด (ราย)	
การวินิจฉัยโดยปัญญาประดิษฐ์	เป็นมะเร็งปอด (ราย)	400	109	509
	ไม่เป็นมะเร็งปอด (ราย)	400	607	1,007
	รวม	800	716	1,516



ภาพที่ 2 Receiver operating characteristic (ROC) curves

ความไวและความจำเพาะของปัญญาประดิษฐ์แบ่งตามลักษณะที่พบในภาพเอกซเรย์ (ตารางที่ 4)

ตารางที่ 4 ความไวและความจำเพาะของปัญญาประดิษฐ์แบ่งตามลักษณะที่พบในภาพเอกซเรย์

ลักษณะที่พบ	ความไว (95% CI)	ความจำเพาะ (95% CI)
Nodules/ mass (n = 996)	53.2 (49.6, 56.8)	68.8 (63.0, 74.1)
One nodule/ mass (n = 761)	53.5 (49.4, 57.5)	64.7 (57.6, 71.2)
Size: ≤ 3 cm (n = 325)	36.5 (30.4, 43.0)	67.7 (67.7, 83.8)
>3-5 cm (n = 262)	56.2 (49.3, 62.9)	59.0 (46.5, 70.5)
>5-7 cm (n = 110)	75.8 (66.3, 83.3)	40.0 (19.8, 64.3)
>7 cm (n = 64)	73.2 (60.4, 83.0)	0 (0.0, 0.0)
Few nodules/ masses (n = 163)	54.4 (45.3, 63.2)	71.4 (57.6, 82.2)
Several nodules/ masses (n = 53)	45.2 (29.1, 62.2)	90.9 (72.2, 97.5)
Diffuse nodules/ masses (n = 19)	50 (26.8, 73.2)	100 (56.6, 100)
Focal consolidation (n=82)	27.6 (14.7, 45.7)	67.7 (70.8, 90.8)
Patchy infiltration (n=201)	6.9 (1.9, 22.0)	85.9 (91.8, 98.0)

ความน่าจะเป็นที่ผู้ป่วยจะเป็นโรคมะเร็งจริงเมื่อการตรวจให้ผลบวก คือ ปัญหาประดิษฐ์บอกว่า เป็นมะเร็ง (positive predictive value) เท่ากับ 78.6% (95% CI = 74.8, 82.0) และความน่าจะเป็นที่ผู้ป่วยไม่ได้เป็นโรคมะเร็งเมื่อการตรวจให้ผลเป็นลบ คือ ปัญหาประดิษฐ์บอกว่าไม่เป็นมะเร็ง (negative predictive value) เท่ากับ 60.3% (95% CI = 57.2, 63.3) ความถูกต้องแม่นยำ (accuracy) มีค่าเท่ากับ 66.4% (95% CI = 64.0, 68.8)

วิจารณ์

จากการศึกษาก่อนหน้า³⁻⁶ พบว่าความไวในการตรวจพบมะเร็งปอดจากภาพรังสีทรวงอกโดยการอ่านและแปลผลของแพทย์อยู่ระหว่างร้อยละ 36 ถึงร้อยละ 84 ขึ้นกับประชากรที่ศึกษา ส่วนความจำเพาะมีค่าอยู่ในช่วงร้อยละ 76.1 ถึง ร้อยละ 99

ด้วยความก้าวหน้าของเทคโนโลยี ได้มีการคิดค้นและพัฒนาาระบบคอมพิวเตอร์เพื่อมาช่วยในการตรวจค้นหาความผิดปกติ (computer-aided detection) และได้นำมาใช้กับภาพรังสีทรวงอกตั้งแต่ปี ค.ศ. 1970⁷ และในปี ค.ศ. 1992 Matsumoto⁸ เริ่มใช้เทคนิคนี้ในการตรวจหามะเร็งปอดจากภาพเอกซเรย์ในผู้ป่วยที่ไม่มีอาการพบว่ามีค่าความไวถึงร้อยละ 62 แต่มีค่าผลบวกหลงเฉลี่ยเท่ากับ 15 จุดต่อหนึ่งภาพ (false positive per image) ข้อมูลจากอีกหลายการศึกษาพบความไวของ computer-aided detection ในการตรวจพบก้อนในปอดจากภาพรังสีทรวงอกอยู่ระหว่างร้อยละ 51 ถึงร้อยละ 80^{6,9-15} และสามารถช่วยแพทย์และรังสีแพทย์ตรวจหามะเร็งปอดได้ดีขึ้น

AIChest4All เป็นอัลกอริทึม (algorithm) ที่ใช้ในการตรวจคัดกรองภาพรังสีทรวงอกแบบอัตโนมัติและแยกความผิดปกติที่พบเป็น 5 แบบ การศึกษานี้มีวัตถุประสงค์เพื่อหาจุดตัดที่เหมาะสมของผลการอ่านจากปัญญาประดิษฐ์ในการวินิจฉัยมะเร็งปอด ความไวและความจำเพาะในการวินิจฉัยมะเร็งปอดเมื่อเทียบกับผลชิ้นเนื้อซึ่งถือเป็นมาตรฐานในการวินิจฉัยโรคมะเร็ง สำหรับจุดตัดของความน่าจะเป็นในการวินิจฉัยมะเร็งปอด เนื่องจากการแจกแจงของข้อมูลเป็นแบบไม่ปกติ มีค่าสังเกตบางค่าผิดปกติ คือ สูงหรือต่ำกว่าค่าสังเกตค่าอื่นมากๆ จึงใช้ค่ามัธยฐานของกลุ่มที่เป็นมะเร็งปอดเป็นจุดตัด ซึ่งมีค่าเท่ากับ 52.5 ความจำเพาะของปัญญาประดิษฐ์นั้นนับว่ามีความใกล้เคียงกับรังสีแพทย์ ส่วนความไวในการวินิจฉัยเมื่อเปรียบเทียบกับรังสีแพทย์จะมีค่าน้อยกว่าเล็กน้อย แต่ยังคงมีค่าใกล้เคียงกับการศึกษาก่อนหน้า แต่การศึกษานี้มีความแตกต่างจากหลายการศึกษาก่อนหน้าที่เป็นการศึกษา Lung nodule เน้นที่การตรวจพบก้อน ในขณะที่การศึกษานี้เป็นการตรวจภาพรังสีทรวงอกโดยรวมเหมือนที่พบในชีวิตประจำวันจึงมีลักษณะภาพที่พบหลายแบบ การจะแปลผลเป็นบวกจะมีความซับซ้อนกว่าเล็กน้อยคือ ต้องอาศัยการตรวจพบรอยโรค ซึ่งอาจจะมีลักษณะเป็นก้อนหรือไม่ก็ได้และการให้การวินิจฉัยว่ารอยโรคนั้นเหมือนมะเร็ง

หรือไม่ ถ้าเทียบความไวในการตรวจพบก้อนมะเร็งปอดของการศึกษานี้คือร้อยละ 53.2 ก็มีความใกล้เคียงกับการศึกษาก่อนหน้านี้ที่มีความไวในการตรวจพบก้อน¹¹⁻¹⁵ อยู่ในช่วงร้อยละ 51-74 โดยมีผลลบลงในช่วง 1-4.2 จุดต่อภาพ ส่วนความจำเพาะนั้นไม่มีระบุไว้ชัดเจนในการศึกษาก่อนหน้า

จากการแบ่งกลุ่มตามลักษณะที่พบในภาพเอกซเรย์พบว่าความไวในการวินิจฉัยจะเพิ่มขึ้นในกลุ่มที่ก้อนมีขนาดใหญ่ขึ้น จากน้อยกว่า 3 ซม.จนถึงกลุ่มที่มีขนาดมากกว่า 5 ซม.จนถึง 7 ซม. เนื่องจากตรวจพบก้อนได้ง่ายขึ้นและมีโอกาสเป็นมะเร็งมากขึ้น ความไวในกลุ่มที่มีหลายก้อนจะต่ำกว่าก่อนเดียว ส่วนความไวในกลุ่มของ consolidation และ patchy infiltration มีค่าค่อนข้างต่ำ เนื่องจากเป็นลักษณะภาพที่พบในกลุ่มที่ไม่ใช่มะเร็งปอด เช่น ปอดอักเสบ หรือ วัณโรคปอดมากกว่า

วัณโรคปอดเป็นโรคที่พบบ่อยในประเทศไทย ลักษณะของภาพเอกซเรย์ทรวงอกบางประเภทอาจมีความคล้ายคลึงกับโรคมะเร็ง เช่น miliary pattern หรือรอยโรคบริเวณด้านบนของปอด จึงทำให้การวินิจฉัยมีความยากขึ้นกว่าการศึกษาในประเทศที่พบวัณโรคน้อย พบว่าปัญญาประดิษฐ์ให้การวินิจฉัยว่าเป็นวัณโรค 93 ราย คิดเป็นร้อยละ 23.3 ของผลลบลงทั้งหมด

ข้อจำกัดในการวินิจฉัยมะเร็งปอดจากภาพรังสีทรวงอกส่วนหนึ่งมาจากข้อจำกัดในการตรวจหาก้อนบริเวณที่เป็นจุดซ่อนเร้น ได้แก่ ยอดปอดด้านบนที่ซ้อนทับกระดูกซี่โครง บริเวณซั้วปอดทั้งสองข้าง ด้านหลังเงาหัวใจ และฐานปอดทั้งสองข้างที่ซ้อนทับเงาของกระดูกบังลม โดยพบว่าปัญญาประดิษฐ์มีผลลบลงในบริเวณต่างๆ เหล่านี้มากกว่ารังสีแพทย์

ส่วนข้อจำกัดในการวินิจฉัยมะเร็งปอดเมื่อเทียบกับการตรวจคัดกรองด้วยภาพเอกซเรย์คอมพิวเตอร์ คือ subsolid nodule เนื่องจากมีความเข้มของก้อนน้อยและขอบเขตเห็นไม่ชัดเจน ทำให้มีความไวต่ำกว่า อีกประการคือเอกซเรย์คอมพิวเตอร์สามารถพิจารณาก่อนได้แบบสามมิติ เอกซเรย์เห็นภาพแบบสองมิติ ไม่สามารถเห็น density ที่อยู่ในก้อนและขอบหน้าหลังของก้อนหรือกรณีก้อนมะเร็งปอดที่พบในภาพเอกซเรย์ปอดของผู้ป่วยที่มีผลเป็นจากวัณโรคหรือมีรอยโรคอื่นบดบัง ทำให้วินิจฉัยได้ยากเมื่อเทียบกับเอกซเรย์คอมพิวเตอร์ที่เห็นรอยโรคแยกออกจากกันได้ชัดเจนกว่า

ข้อจำกัดของการศึกษานี้คือเป็นการศึกษาแบบย้อนหลัง ซึ่งอาจมีการสูญหายของข้อมูลหรือขาดความครบถ้วนในบางจุด การวินิจฉัยมะเร็งปอดในการศึกษานี้เป็นการแปลผลจากภาพรังสีทรวงอกเพียงภาพเดียว แตกต่างจากการวินิจฉัยตามปกติที่สามารถใช้ภาพเอกซเรย์เก่ามาประกอบการพิจารณาว่าก้อนที่พบมีการโตขึ้นจากเดิมหรือไม่ ซึ่งเป็นอีกลักษณะที่สำคัญในการวินิจฉัยมะเร็งปอด และอาจทำให้ค่าความไวและความจำเพาะเพิ่มขึ้น

สรุป

ปัญญาประดิษฐ์ที่ได้รับการพัฒนาและใช้ข้อมูลจากผู้ป่วยในประเทศไทยมีความจำเพาะในการวินิจฉัยมะเร็งปอดจากภาพรังสีทรวงอกใกล้เคียงกับรังสีแพทย์ แต่ยังมีควมมีความไวที่ต่ำ ควรจะต้องมีการศึกษาและพัฒนาเพิ่มเติม หรือใช้ร่วมกับการคัดกรองและการตรวจวินิจฉัยอื่นๆ

แนวทางที่จะพัฒนาปัญญาประดิษฐ์ให้ดียิ่งขึ้น คือ อาจเพิ่มข้อมูลให้ปัญญาประดิษฐ์ได้เรียนรู้เพิ่มขึ้นในจุดที่ยังมีข้อผิดพลาด เช่น ก้อนที่พบบริเวณจุดซ่อนเร้นต่างๆ นอกจากนี้

AIChest4All ยังมีระบบให้รังสีแพทย์ช่วยให้ความเห็นและส่งข้อมูลกลับไปเพื่อปรับปรุงและพัฒนาปัญญาประดิษฐ์ต่อไป

กิตติกรรมประกาศ

ขอขอบคุณ นายแพทย์เอนก กนกศิลป์ ผู้อำนวยการสถาบันโรคทรวงอก ทีมงานจากสถาบันโรคทรวงอกและภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธรรมศาสตร์ ที่ช่วยสนับสนุนให้การศึกษาในครั้งนี้สำเร็จด้วยดี

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; 68:394-424.
2. van Beek EJ, Mirsadraee S, Murchison JT. Lung cancer screening: Computed tomography or chest radiographs? *World J Radiol* 2015; 7:189-93.
3. Toyoda Y, Nakayama T, Kusunoki Y, Iso H, Suzuki T. Sensitivity and specificity of lung cancer screening using chest low-dose computed tomography. *Br J Cancer* 2008; 98:1602-7.
4. Gavelli G, Giampalma E. Sensitivity and specificity of chest X-ray screening for lung cancer: review article. *Cancer* 2000; 89(11 Suppl):2453-6.
5. Quekel LG, Kessels AG, Goei R, van Engelshoven JM. Detection of lung cancer on the chest radiograph: a study on observer performance. *Eur J Radiol* 2001; 39:111-6.
6. Xu Y, Ma D, He W. Assessing the use of digital radiography and a real-time interactive pulmonary nodule analysis system for large population lung cancer screening. *Eur J Radiol* 2012; 81:e451-6.
7. Toriwaki J, Suenaga Y, Negoro T, Fukumura T. Pattern recognition of chest X-ray images. *Comput Graph Image Process*. 1973; 2: 252-71.
8. Matsumoto T, Yoshimura H, Giger ML, Doi K, Macmahon H, Montner SM, et al. Potential usefulness of computerized nodule detection in screening programs for lung cancer. *Invest Radiol* 1992; 27:471-5.
9. Doi K. Current status and future potential of computer-aided diagnosis in medical imaging. *Br J Radiol* 2005; 78:S3-19.
10. Schilham AM, van Ginneken B, Loog M. A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database. *Med Image Anal* 2006; 10:247-58.
11. Shiraishi J, Abe H, Engelmann R, Doi K. Effect of high sensitivity in a computerized scheme for detecting extremely subtle solitary pulmonary nodules in chest radiographs: observer performance study. *Acad Radiol* 2003;10: 1302-11.
12. Schalekamp S, van Ginneken B, Koedam E, Snoeren MM, Tiehuis AM, Wittenberg R, et al. Computer-aided detection improves detection of pulmonary nodules in chest radiographs beyond the support by bone-suppressed images. *Radiology* 2014; 272:252-61.
13. Coppini G, Diciotti S, Falchini M, Villari N, Valli G. Neural networks for computer-aided diagnosis: detection of lung nodules in chest radiograms. *IEEE Trans Inf Technol Biomed* 2003; 7:344-57.
14. Sakai S, Soeda H, Takahashi N, Okafuji T, Yoshitake T, Yabuuchi H, et al. Computer-aided nodule detection on digital chest radiography: validation test on consecutive T1 cases of resectable lung cancer. *J Digit Imaging* 2006; 19:376-82.
15. Shiraishi J, Li Q, Suzuki K, Engelmann R, Doi K. Computer-aided diagnostic scheme for the detection of lung nodules on chest radiographs: localized search method based on anatomical classification. *Med Phys* 2006; 33:2642-53.