

การประเมินความแม่นยำของแพทย์ทั่วไปในการแปลผลฟิล์มรังสีทั่วไปในช่วงเวรตึก โรงพยาบาลพนัสนิคม

Accuracy of Plain Radiograph Interpretation by General Physicians during Night Shifts at Phanatnikhom Hospital

น้ำฝน ยิ่งยง, พ.บ.*

Numfon Yingyong, M.D.*

*กลุ่มงานรังสีวิทยา โรงพยาบาลพนัสนิคม จังหวัดชลบุรี ประเทศไทย 20140

*Department of Radiology, Phanatnikhom Hospital, Chonburi Province, Thailand, 20140

Corresponding author, E-mail address: x-raygirl@windowslive.com

Received: 30 Jun 2025. Revised: 04 Sep 2025. Accepted: 26 Sep 2025.

บทคัดย่อ

หลักการและเหตุผล : การแปลผลฟิล์มรังสีทั่วไปให้ถูกต้องเป็นสิ่งสำคัญในการวินิจฉัยโรคในท้องฉุกเฉิน โรงพยาบาลระดับ M1 มักไม่มีระบบบริการรังสีแพทย์ในช่วงเวรตึก การแปลผลจึงเป็นหน้าที่ของแพทย์เวรทั่วไป ซึ่งจำกัดด้านประสบการณ์อาจมีความผิดพลาดและส่งผลกระทบต่อดูแลผู้ป่วย ปัจจุบันยังไม่มีการศึกษาความแม่นยำของการแปลผลฟิล์มโดยแพทย์กลุ่มนี้ ณ ห้องฉุกเฉิน

วัตถุประสงค์ : เพื่อประเมินความแม่นยำของแพทย์เวรทั่วไป ในการแปลผลฟิล์มรังสีช่วงเวรตึก ณ ห้องฉุกเฉิน และวิเคราะห์ลักษณะความผิดพลาดที่พบบ่อย

วิธีการศึกษา : การศึกษาเชิงพรรณน่าย้อนหลัง ทบทวนข้อมูลผู้ป่วยที่ได้รับการตรวจฟิล์มรังสีทั่วไป ในช่วงเวลา 00.00-08.00 น. ที่โรงพยาบาลพนัสนิคม ระหว่างเมษายน พ.ศ.2567 ถึง มีนาคม พ.ศ.2568 รวม 1,146 ราย โดยใช้ผลแปลของรังสีแพทย์หนึ่งท่าน เป็นมาตรฐาน เปรียบเทียบกับผลการแปลของแพทย์เวรและจำแนกความผิดพลาดเป็น false positive, false negative และ partial interpretation

ผลการศึกษา : แพทย์เวรทั่วไปมีความแม่นยำในการแปลผล ร้อยละ 95.9 โดยมี sensitivity ร้อยละ 95.0 specificity ร้อยละ 96.4 PPV ร้อยละ 93.6 และ NPV ร้อยละ 97.2 พบความผิดพลาดรวม 32 ราย (ร้อยละ 4.1) ได้แก่ false negative 14 ราย (ร้อยละ 1.8) ส่วนใหญ่พลาดการวินิจฉัย infiltration false positive 12 ราย (ร้อยละ 1.5) ส่วนใหญ่เป็นการวินิจฉัย infiltration เกินจริงและpartial interpretation 6 ราย (ร้อยละ 0.8) เช่น รายงาน pneumothorax แต่พลาด rib fracture หรือรายงาน cephalization แต่พลาด pleural effusion

สรุปผล : แพทย์เวรทั่วไปในโรงพยาบาลระดับ M1 มีความแม่นยำในการแปลผลฟิล์มรังสีช่วงเวรตึกในระดับสูง แต่ยังมีข้อผิดพลาด โดยเฉพาะ false negative และ partial interpretation ในรอยโรค subtle จึงควรมีการอบรมเพิ่มเติมและพิจารณาใช้เทคโนโลยี AI เพื่อช่วยลดข้อผิดพลาดและเพิ่มคุณภาพการดูแลผู้ป่วย

คำสำคัญ : ความแม่นยำ ฟิล์มรังสีทั่วไป แพทย์เวร เวรตึก โรงพยาบาลชุมชน

ABSTRACT

- Background** : Accurate interpretation of plain radiographs is crucial for diagnosing diseases in the emergency department. M1-level hospitals often lack a radiologist consultation system during night shifts, so interpretation falls to general physicians. Their limited experience may lead to errors that affect patient care. Currently, there is a lack of evidence on the diagnostic accuracy of this group of physicians in the emergency setting.
- Objective** : To assess the diagnostic accuracy of general physicians in interpreting plain radiographs during night shifts in the emergency department and to analyze the characteristics of common errors.
- Methods** : This retrospective descriptive study reviewed 1,146 plain radiographs obtained between April 2024 and March 2025 at Phanatnikhom Hospital during night duty hours. A radiologist's interpretation was used as the reference standard. Their findings were compared with the interpretations by general physicians, and errors were categorized as false positives, false negatives, or partial interpretations.
- Results** : General physicians demonstrated an overall interpretation accuracy of 95.9%, with a sensitivity of 95.0%, a specificity of 96.4%, a PPV of 93.6%, and an NPV of 97.2%. A total of 32 errors (4.1%) were found: 14 false negatives (1.8%), mostly missed infiltrations 12 false positives (1.5%), primarily over-diagnoses of infiltration; and 6 partial interpretations (0.8%), such as reporting pneumothorax but missing a rib fracture, or reporting cephalization but missing a pleural effusion.
- Conclusions** : General physicians at an M1-level hospital demonstrate high accuracy in interpreting plain radiographs during night shifts. However, errors, particularly false negatives and partial interpretations of subtle findings, remain a concern. Further training and the consideration of using AI-assisted tools could help reduce these errors and improve the quality of patient care.
- Keywords** : accuracy, plain radiographs, general physician, night shifts, community hospital.

หลักการและเหตุผล

ฟิล์มรังสีทั่วไป (plain radiograph) เป็นเครื่องมือพื้นฐานที่มีบทบาทสำคัญในการวินิจฉัยโรคในเวชปฏิบัติ โดยเฉพาะในแผนกฉุกเฉิน เนื่องจากสามารถให้ข้อมูลเบื้องต้นที่ช่วยประกอบการตัดสินใจทางคลินิกได้อย่างรวดเร็วและทันที่⁽¹⁾ อย่างไรก็ตาม ความถูกต้องของการแปลผลฟิล์มรังสีขึ้นอยู่กับทักษะและประสบการณ์ของผู้แปลผลเป็นสำคัญ⁽²⁾

จากการศึกษาก่อนหน้าในต่างประเทศ ความแม่นยำของแพทย์ในการแปลผลฟิล์มรังสีมีความหลากหลาย โดยมีอัตราความถูกต้องอยู่ในช่วงกว้างระหว่างร้อยละ 22.2-97 ขึ้นอยู่กับระดับประสบการณ์และชนิดของพยาธิสภาพที่ตรวจพบ เช่น Jimah และคณะ พบว่าเมื่อใช้ความเห็นของรังสีแพทย์เป็นมาตรฐานและไม่แจ้งประวัติทางคลินิกของผู้ป่วยแก่แพทย์

ผู้ประเมินมีเพียงร้อยละ 40.0 ของแพทย์ประจำบ้าน ร้อยละ 22.2 ของแพทย์ใช้ทุน ร้อยละ 24.2 ของแพทย์ฝึกหัด และร้อยละ 13.2 ของนักศึกษาแพทย์เท่านั้นที่สามารถแปลผลเอกซเรย์ปอดได้ถูกต้อง ในขณะที่เมื่อมีประวัติทางคลินิกประกอบทุกกลุ่มมีคะแนนความแม่นยำมากกว่าร้อยละ 75 แต่ผู้เข้าร่วมทุกคนยังคงมีปัญหาในการระบุพยาธิสภาพ ได้แก่ ภาวะลมในช่องเยื่อหุ้มปอด (pneumothorax) ภาวะลมในช่องกลางอก (pneumomediastinum) หรือกระดูกซี่โครงซี่ซ้ายหัก⁽²⁾ ในขณะที่ Halim และคณะ ทำการศึกษาความถูกต้องในการแปลผลภาพรังสีทั่วไปโดยแพทย์ประจำบ้านสาขารังสีวิทยา โดยใช้ความเห็นของรังสีแพทย์เป็นมาตรฐาน พบว่าอัตราความแม่นยำโดยรวมอยู่ที่ร้อยละ 69⁽³⁾ Bergus GR และคณะ ได้ทำการศึกษาการแปลผลภาพถ่ายรังสีของแพทย์เวชปฏิบัติครอบครัวเทียบกับรังสีแพทย์หนึ่งท่าน พบว่าแพทย์เวชปฏิบัติครอบครัวแปลผลภาพรังสีได้ถูกต้องคิดเป็นร้อยละ 92.4 ของทั้งหมด⁽⁴⁾ Petinaux B และคณะ ได้ทำการศึกษาการแปลผลภาพถ่ายรังสีของแพทย์เวชศาสตร์ฉุกเฉิน (EPs) เทียบกับรังสีแพทย์หนึ่งท่านพบว่า แพทย์เวชศาสตร์ฉุกเฉินแปลผลภาพรังสีได้ถูกต้องคิดเป็นร้อยละ 97 ของทั้งหมด ความผิดพลาดที่พบบ่อย ได้แก่ กระดูกหัก ข้อเคลื่อน ภาวะปอดอักเสบชนิด air-space disease และก้อนปอด⁽⁵⁾ แม้ว่าการศึกษาของ Petinaux B และคณะ⁽⁵⁾ จะศึกษาในกลุ่มแพทย์เวชศาสตร์ฉุกเฉิน แต่ก็แสดงให้เห็นถึงความแม่นยำที่สูงมากในบริบทของห้องฉุกเฉินซึ่งเป็นสภาพแวดล้อมที่คล้ายคลึงกับงานวิจัยนี้

โรงพยาบาลพนัสนิคมมีระบบการศึกษา รังสีแพทย์สำหรับการอ่านฟิล์มเฉพาะในช่วงเวลาราชการ ส่วนการให้บริการตรวจรักษาผู้ป่วยในห้องฉุกเฉินจะมีแพทย์เวชศาสตร์ฉุกเฉินและแพทย์ทั่วไปปฏิบัติงานร่วมกันในช่วงเวรเช้าและเวรบ่าย สำหรับเวรตีง (เวลา 00.00-08.00 น.) จะมีเพียงแพทย์ในช่วงปีแรกถึงปีที่สามหลังสำเร็จการศึกษา (early-career general physicians) ปฏิบัติงานโดยไม่มีระบบการศึกษาฟิล์มหรือการใช้ปัญญาประดิษฐ์มาช่วยเหลือ แพทย์กลุ่มนี้

อาจมีข้อจำกัดด้านประสบการณ์และทักษะเฉพาะทาง อีกทั้งยังมีภาวะเหนื่อยล้าจากการทำงานต่อเนื่อง ซึ่งอาจนำไปสู่การวินิจฉัยล่าช้า การรักษาที่ไม่เหมาะสม และเพิ่มความเสี่ยงต่อภาวะแทรกซ้อนในผู้ป่วย

ปัจจุบันยังไม่มีการศึกษาที่ประเมินความแม่นยำในการแปลผลฟิล์มรังสีโดยแพทย์กลุ่มนี้ที่ปฏิบัติงานในเวรตีงของห้องฉุกเฉินในบริบทของโรงพยาบาลระดับ M1 จึงมีความจำเป็นต้องดำเนินการศึกษานี้ เพื่อประเมินข้อมูลดังกล่าว และใช้เป็นแนวทางในการพัฒนาองค์ความรู้ การอบรม หรือระบบสนับสนุนการตัดสินใจทางรังสีวิทยาในอนาคต

วัตถุประสงค์

เพื่อศึกษาความแม่นยำของแพทย์เวรทั่วไปในช่วงปีแรกถึงปีที่สามหลังสำเร็จการศึกษาในการแปลผลฟิล์มรังสีทั่วไปในช่วงเวรตีงเปรียบเทียบกับผลการแปลผลของรังสีแพทย์หนึ่งท่านซึ่งในการศึกษานี้ถือเป็นมาตรฐาน (gold standard) และเพื่อวิเคราะห์ลักษณะของความผิดพลาดที่พบบ่อย เพื่อใช้เป็นข้อมูลพื้นฐานในการพัฒนาคุณภาพการดูแลผู้ป่วยในโรงพยาบาลระดับ M1 ต่อไป

จริยธรรมวิจัยและการพิทักษ์สิทธิของกลุ่มตัวอย่าง

งานวิจัยนี้ได้ผ่านการพิจารณาจากคณะกรรมการพิจารณาจริยธรรมการวิจัยเกี่ยวกับมนุษย์ สำนักงานสาธารณสุขจังหวัดชลบุรี ตามหนังสืออนุมัติเลขที่ 054-2568 ลงวันที่ 15 พฤษภาคม พ.ศ. 2568

ระเบียบวิธีการศึกษา

การวิจัยนี้เป็นการศึกษาเชิงพรรณนาย้อนหลัง (Retrospective Descriptive study)

คำจำกัดความ

1. แพทย์ทั่วไป หมายถึง แพทย์ที่อยู่ในช่วงปีแรกถึงปีที่สามหลังสำเร็จการศึกษา (early-career general physicians) ซึ่งเป็นแพทย์กลุ่มเดียวที่ปฏิบัติงานช่วงเวรตึก (เวลา 00.00-8.00 น.) ณ โรงพยาบาลพนสนิมคม ระดับ M1 ระหว่างเดือน เมษายน พ.ศ. 2567 ถึงมีนาคม พ.ศ. 2568

2. ความถูกต้องของการแปลผล แบ่งเป็น 2 เกณฑ์ว่า ตรงหรือไม่ตรงกับความเห็นรังสีแพทย์หนึ่งท่านของโรงพยาบาลพนสนิมคมที่มีประสบการณ์ 13 ปีในการทำงานด้านรังสีวิทยา ซึ่งถือเป็นมาตรฐาน (gold standard) ในการศึกษานี้แม้ไม่ได้มีการยืนยันด้วยเอกซเรย์คอมพิวเตอร์หรือผู้ประเมินมากกว่าหนึ่งท่าน อันเป็นข้อจำกัดของการศึกษา

3. ประเภทของความผิดพลาดในการแปลผล แบ่งเป็น 3 กลุ่ม ได้แก่

3.1 การวินิจฉัยผิด (false positive) การแปลผลว่าฟิล์มของผู้ป่วยมีภาวะผิดปกติ แต่ในความเป็นจริงจากเกณฑ์มาตรฐานไม่มีความผิดปกติ

3.2 การวินิจฉัยพลาด (false negative) การแปลผลว่าฟิล์มของผู้ป่วยปกติ แต่ในความเป็นจริงจากเกณฑ์มาตรฐานมีภาวะผิดปกติ

3.3 การแปลผลไม่ครบถ้วน (partial interpretation) การแปลผลว่าฟิล์มของผู้ป่วยมีภาวะผิดปกติแต่แปลผลความผิดปกติไม่ได้ไม่ครบทุกภาวะ

4. โรงพยาบาลระดับ M1 หมายถึง โรงพยาบาลชุมชนขนาดกลางที่มีจำนวนเตียงระหว่าง 120-239 เตียง ให้บริการรักษาพยาบาลทั่วไปและเฉพาะทางบางสาขา ทำหน้าที่เป็นโรงพยาบาลแม่ข่ายของโรงพยาบาลชุมชนในเครือข่าย และโดยทั่วไปจะไม่มีการรับผู้ป่วยประจำตลอด 24 ชั่วโมง ซึ่งโรงพยาบาลพนสนิมคมอยู่ในกลุ่มนี้

ประชากรและกลุ่มตัวอย่าง

ประชากรที่ศึกษา ได้แก่ ผู้ป่วยที่ได้รับการตรวจภาพถ่ายเอกซเรย์รังสีธรรมดาในช่วงเวรตึก ณ โรงพยาบาลพนสนิมคม ระหว่างเดือนเมษายน พ.ศ. 2567 ถึงเดือนมีนาคม พ.ศ. 2568 โดยการแปลผลเบื้องต้นดำเนินการโดยแพทย์เวรทั่วไป

เกณฑ์คัดเข้า (Inclusion criteria):

1. ภาพรังสีทั่วไปที่ถ่ายในช่วงเวลา 00.00-08.00 น. ระหว่างเดือนเมษายน พ.ศ. 2567 ถึงมีนาคม พ.ศ. 2568 ณ โรงพยาบาลพนสนิมคม
2. มีบันทึกผลการแปลภาพรังสีโดยแพทย์เวรทั่วไปในระบบเวชระเบียนคอมพิวเตอร์ของโรงพยาบาล
3. ภาพรังสีมีคุณภาพเพียงพอสำหรับการวินิจฉัย (diagnostic quality image)

เกณฑ์คัดออก (Exclusion criteria):

1. ข้อมูลเวชระเบียนไม่สมบูรณ์ เช่น ขาดข้อมูลประชากร ขาดผลการวินิจฉัยจากแพทย์เวรหรือไม่สามารถเข้าถึงภาพรังสีได้
2. ภาพรังสีที่ไม่มีการบันทึกผลแปลโดยแพทย์เวร หรือไม่สมารถพบทวนผลได้
3. ภาพรังสีที่มีคุณภาพไม่เพียงพอต่อการวินิจฉัย เช่น ภาพเบลอ มี artifact รบกวนอย่างชัดเจน หรือมองไม่เห็นโครงสร้างสำคัญ

การคำนวณขนาดกลุ่มตัวอย่าง: sample size calculation ใช้สูตรการคำนวณขนาดตัวอย่างสำหรับการศึกษาความแม่นยำ (proportion with specified precision)

$$n = \frac{Z_{1-\alpha/2}^2 \times P \times (1 - P)}{d^2}$$

N = จำนวนกลุ่มตัวอย่าง

$Z_{1-\alpha/2}$ = ค่าคงที่ทางสถิติที่ระดับความเชื่อมั่น 95% = 1.96

P = ความแม่นยำที่คาดการณ์ไว้ (estimated accuracy): คาดว่าแพทย์เวรทั่วไปมีความแม่นยำ P = 0.80

d = ค่าความคลาดเคลื่อนที่ยอมรับได้ (margin of error): ยอมให้ค่าคลาดเคลื่อน d = 0.03

ขนาดกลุ่มตัวอย่างขั้นต่ำที่ต้องการ = อย่างน้อย 683 ราย

เพื่อให้สามารถวิเคราะห์ข้อมูลได้อย่างเหมาะสม ควรเผื่อกรณีข้อมูลไม่ครบหรือสูญเสียประมาณร้อยละ 10 ดังนั้น ควรเก็บข้อมูลอย่างน้อย 752 ราย

การดำเนินการวิจัย

รังสีแพทย์เป็นผู้รวบรวมข้อมูลย้อนหลังจากเวชระเบียนและระบบภาพรังสีของโรงพยาบาลพนสนิม โดยดึงข้อมูลผู้ป่วยที่ได้รับการตรวจภาพถ่ายเอกซเรย์รังสีธรรมดาในช่วงเวรตึก ระหว่างเดือนเมษายน พ.ศ. 2567 ถึงเดือนมีนาคม พ.ศ. 2568 โดยมีการแปลผลภาพรังสีโดยแพทย์ทั่วไปบันทึกไว้ในระบบเวชระเบียนคอมพิวเตอร์

จากนั้นรังสีแพทย์จะทำการทบทวนภาพรังสีโดยเปรียบเทียบผลกับที่แพทย์เวรได้แปลไว้ และจัดหมวดผลออกเป็น 2 กลุ่ม คือ แปลผลตรงกันกับรังสีแพทย์ หรือ แปลผลไม่ตรงกัน

ในกรณีที่ตรงกัน จะวิเคราะห์ต่อไปว่าอยู่ในกลุ่ม false positive, false negative หรือ partial interpretation ตามคำจำกัดความข้างต้น เพื่อใช้ในการวิเคราะห์ทางสถิติต่อไป

การวิเคราะห์ข้อมูล

ข้อมูลพื้นฐานเกี่ยวกับผู้ป่วย รายงานโดยใช้สถิติเชิงพรรณนา ได้แก่ ประเภทของฟิล์มรังสี ผลการแปลผลฟิล์มรังสี (ปกติ/ผิดปกติ) ลักษณะของความผิดพลาดในการแปลผล นำเสนอผลในรูปของจำนวน (n) และร้อยละ (%) เพื่อแสดงความถี่ของข้อมูลแต่ละกลุ่ม

การประเมินความแม่นยำในการแปลผลเปรียบเทียบผลการแปลฟิล์มรังสีของแพทย์เวรทั่วไปกับผลการทบทวนของรังสีแพทย์ซึ่งถือเป็นมาตรฐาน (gold standard) โดยสร้างตาราง 2x2 และคำนวณค่าทางสถิติ ดังนี้ Sensitivity (ความไว) Specificity (ความจำเพาะ) Positive Predictive Value: PPV (ค่าทำนายทางบวก) Negative Predictive Value: NPV (ค่าทำนายทางลบ) Accuracy (ความแม่นยำ)

การวิเคราะห์ลักษณะของความผิดพลาดจำแนกและรายงานประเภทของความผิดพลาดในการแปลผล ได้แก่ การวินิจฉัยผิด (false positive) การวินิจฉัยพลาด (false negative) และการแปลผลไม่ครบถ้วน (partial interpretation) จะใช้สถิติเชิงพรรณนา โดยรายงานความถี่สัมพัทธ์ (ร้อยละ)

ผลการศึกษา

จากการรวบรวมข้อมูลฟิล์มรังสีทั่วไปในช่วงเวรตึก (เวลา 00.00-08.00 น.) ระหว่างเดือนเมษายน พ.ศ. 2567 ถึงเดือนมีนาคม พ.ศ. 2568 พบว่า มีฟิล์มทั้งหมด 1,146 ชิ้น ผู้ป่วยมีอายุระหว่าง 3 เดือน ถึง 97 ปี อายุเฉลี่ย 45 ปี เพศชาย 688 ราย (ร้อยละ 60.0) และเพศหญิง 458 ราย (ร้อยละ 40.0) โดยชนิดของฟิล์มที่ส่งตรวจมากที่สุด ได้แก่ ภาพถ่ายทรวงอก (CXR) จำนวน 737 ราย (ร้อยละ 64.3) รองลงมาคือ Acute Abdominal Series (ร้อยละ 10.6) (ตารางที่ 1)

ตารางที่ 1 ลักษณะพื้นฐานของผู้ป่วย (Baseline characteristics) (n=1,146)

ตัวแปร	จำนวน	ร้อยละ
อายุ (ปี)	3 เดือน - 97 ปี, เฉลี่ย 45 ปี	-
เพศ		
- ชาย	688	60.0%
- หญิง	458	40.0%
ประเภทฟิล์มรังสี		
- Chest radiograph (CXR)	737	64.3%
- Skull	27	2.4%
- Cervical spine	18	1.6%
- Thoracolumbar spine	4	0.3%
- Lumbosacral spine	14	1.2%
- Abdomen	30	2.6%
- Abdominal series (AAS)	121	10.6%
- Pelvis	52	4.5%
- Upper extremities	77	6.7%
- Lower extremities	66	5.8%

จากฟิล์มทั้งหมด 1,146 ชิ้น พบว่ามี เมื่อเปรียบเทียบกับรังสีแพทย์ (gold standard) การแปลผลโดยแพทย์เวรจำนวน 778 ราย (ร้อยละ 67.9) พบว่ามีความถูกต้องตรงกัน 746 ภาพ คิดเป็น โดยในจำนวนนี้พบผลปกติ 496 ราย (ร้อยละ 43.3) และ ร้อยละ 95.9 (ตารางที่ 2) ผลผิดปกติ 282 ราย (ร้อยละ 24.6) ขณะที่ 368 ราย (ร้อยละ 32.1) ไม่มีข้อมูลผลแปลผลจากแพทย์เวร

ตารางที่ 2 ผลการเปรียบเทียบการแปลผล (2x2 table)

	รังสีแพทย์		รวม
	ผิดปกติ	ปกติ	
แพทย์เวร	264	18	282
	14	482	496
รวม	278	500	778

พบว่าแพทย์เวรทั่วไปไม่มีอัตรา false negative (ร้อยละ 1.5) และ partial interpretation (ร้อยละ 0.8) สูงที่สุด (ร้อยละ 1.8) รองลงมาคือ false positive (ตารางที่ 3)

ตารางที่ 3 ค่าความแม่นยำในการแปลผล

ค่าความแม่นยำในการแปลผล	ร้อยละ	(95% CI)
Sensitivity	95.0%	(91.7-97.2)
Specificity	96.4%	(94.4-97.9)
Positive Predictive Value	93.6%	(90.3-95.9)
Negative Predictive Value	97.2%	(95.4-98.3)
Accuracy	95.9%	(94.2-97.2)

จากความผิดพลาดรวม 32 ราย พบว่าเป็น false positive 12 ราย (ร้อยละ 1.5) ส่วนใหญ่จำนวน 8 รายเกิดจากการวินิจฉัยผิดว่ามี infiltration ส่วนที่เหลืออย่างละ 1 วินิจฉัยผิดว่ามีกระดูกสันหลังหัก นิ้วในข้อต่อก่อนในปอด และน้ำในช่องเยื่อหุ้มปอด

False negative 14 ราย (ร้อยละ 1.8) มักพลาดการวินิจฉัย infiltration 9 ราย ส่วน cardiomegaly,

lung cavity, fracture rib, periosteal reaction, lung nodule พลาดการวินิจฉัยไปอย่างละ 1 ราย

Partial interpretation 6 ราย (ร้อยละ 0.8) เช่น รายงาน pneumothorax แต่พลาด fracture rib รายงาน cephalization แต่พลาด pleural effusion, พบ fracture clavicle แต่พลาด fracture rib เป็นต้น (ตารางที่ 4)

ตารางที่ 4 ประเภทของความผิดพลาดและตัวอย่างพยาธิสภาพ (n=32)

Type of error	Number	Percentage	พยาธิสภาพที่ผิดพลาดมากที่สุด
False positive	12	1.5%	Infiltration 8 ราย
False negative	14	1.8%	Infiltration 9 ราย
Partial interpretation	6	0.8%	Missed fracture rib

อภิปรายผล

การศึกษานี้เป็นการประเมินความแม่นยำในการแปลผลฟิล์มรังสีทั่วไปโดยแพทย์เวรทั่วไปในช่วงเวรตึกของโรงพยาบาลระดับ M1 ซึ่งก่อนหน้านี้ยังไม่มีการศึกษามาก่อน ผลการศึกษาพบว่าแพทย์เวรทั่วไปมีความแม่นยำในการแปลผลสูง (accuracy ร้อยละ 95.9) โดยมีค่า sensitivity ร้อยละ 95.0 และ specificity ร้อยละ 96.4 ซึ่งสะท้อนให้เห็นถึงศักยภาพของแพทย์กลุ่มนี้ในการดูแลผู้ป่วยในภาวะฉุกเฉิน แม้จะมีข้อจำกัดด้านประสบการณ์และการขาดระบบสนับสนุนด้านรังสีแพทย์ในช่วงเวลาดังกล่าว

เมื่อเปรียบเทียบกับการศึกษาก่อนหน้านี้ พบว่าผลการศึกษานี้ให้ค่า accuracy สูงกว่าบางงานวิจัย เช่น Jimah และคณะ รายงานความแม่นยำของแพทย์ในหลายระดับเพียงร้อยละ 22.2-40 หากไม่มีข้อมูลคลินิกประกอบ และเพิ่มขึ้นมากกว่าร้อยละ 75 เมื่อมีข้อมูลคลินิก⁽²⁾ รวมทั้งสูงกว่าผลการศึกษาของ Halim และคณะ ที่รายงานความแม่นยำของแพทย์ประจำบ้านรังสีวิทยาอยู่ที่ร้อยละ 69⁽³⁾ อย่างไรก็ตาม ผลการศึกษานี้ใกล้เคียงกับรายงานของ Bergus GR และคณะ ในแพทย์เวชปฏิบัติครอบครัวร้อยละ 92.4⁽⁴⁾ และสอดคล้องกับ Petinaux B และคณะ ที่รายงานว่าแพทย์เวชศาสตร์ฉุกเฉินมี accuracy สูงถึงร้อยละ 97 เมื่อเทียบกับรังสีแพทย์หนึ่งท่าน⁽⁵⁾ โดยสะท้อนว่าการแปลผล

ฟิล์มรังสีโดย non-radiologists สามารถมีความแม่นยำสูงได้ หากอยู่ในบริบทของการปฏิบัติงานจริงที่มีข้อมูลทางคลินิกประกอบ

สาเหตุที่โรงพยาบาลพหุศาสตร์มี accuracy สูงอาจอธิบายได้จากหลายปัจจัย ได้แก่ (1) การแปลผลในบริบทจริงที่แพทย์เวรได้รับข้อมูลทางคลินิกจากผู้ป่วย การตรวจร่างกายและผลตรวจทางห้องปฏิบัติการร่วมด้วย ทำให้สามารถประเมินภาพรังสีได้แม่นยำกว่าการศึกษาที่เป็น simulation (2) ลักษณะผู้ป่วยที่เข้ารับบริการส่วนใหญ่เป็นโรคที่พบบ่อยในเวชปฏิบัติ เช่น ปอดอักเสบ กระดูกหัก ทำให้แพทย์คุ้นเคยกับลักษณะรังสีวิทยา และ (3) แพทย์เวรทั่วไปแม้จะอยู่ในช่วง early-career แต่ได้รับการเรียนการสอนรังสีวิทยาในหลักสูตรแพทยศาสตรบัณฑิตที่ครอบคลุมเบื้องต้น รวมทั้งมีการเรียนรู้ต่อเนื่องระหว่างการปฏิบัติงานจริง

เมื่อวิเคราะห์ลักษณะของความผิดพลาด พบว่า false negative มีสัดส่วนสูงที่สุดร้อยละ 1.8 โดยส่วนใหญ่เป็นการพลาดการวินิจฉัย infiltration และรอยโรค subtle อื่น ๆ เช่น cardiomegaly, lung cavity และ rib fracture ซึ่งอาจทำให้เกิดการวินิจฉัยล่าช้าและมีผลต่อการรักษาผู้ป่วย ขณะที่ partial interpretation พบร้อยละ 0.8 เป็นการแปลผลพบพยาธิสภาพบางส่วนแต่ไม่ครบถ้วน เช่น รายงาน pneumothorax แต่พลาด

rib fracture หรือรายงาน cephalization แต่พลาต pleural effusion ซึ่งแม้จะไม่พลาตการวินิจฉัยหลัก แต่ยังมีความเสี่ยงต่อการจัดการรักษาที่ไม่สมบูรณ์ การแยก false negative ออกจาก partial interpretation จึงมีความสำคัญ เนื่องจาก false negative อาจทำให้ผู้ป่วยไม่ได้รับการวินิจฉัยเลย ขณะที่ partial interpretation แม้จะพบรอยโรคหลักแต่ยังขาดรายละเอียดที่อาจมีผลต่อการรักษา

สำหรับ false positive ร้อยละ 1.5 ส่วนใหญ่ เกิดจากการวินิจฉัย infiltration เกินจริง ซึ่งอาจนำไปสู่ การรักษาที่มากเกินความจำเป็น (over-treatment) เช่น การให้ยาปฏิชีวนะโดยไม่จำเป็น

การวินิจฉัยผิดพลาดในการศึกษานี้ เป็นไปในทิศทางเดียวกันกับการศึกษาก่อนหน้าที่พบว่า รอยโรคที่วินิจฉัยผิดพลาดมากที่สุดคือ รอยโรคขนาดเล็กหรือภาวะกระดูกหัก⁽⁶⁾

ในปัจจุบันมีการพัฒนาเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI) สำหรับการแปลผลภาพรังสีทรวงอก โดยมีงานวิจัยยืนยันว่า AI สามารถช่วยในการตรวจหาพยาธิสภาพที่ subtle เช่น pneumothorax ขนาดเล็ก หรือรอยโรคของกระดูกที่ตรวจยาก ได้อย่างมีนัยสำคัญ⁽⁷⁻¹⁰⁾ การนำ AI มาใช้ร่วมกับการแปลผลของแพทย์เวร จึงอาจช่วยลดความผิดพลาดในทุกกลุ่ม และเป็นแนวทางที่ควรพิจารณาในการพัฒนาคุณภาพการดูแลผู้ป่วยในอนาคต

จุดแข็งของการศึกษานี้คือ เป็นงานวิจัยที่ทำในบริบทของโรงพยาบาลระดับ M1 ซึ่งสะท้อนสถานการณ์จริงของประเทศไทยและเป็นข้อมูลเชิงประจักษ์ชุดแรกๆ ที่แสดงให้เห็นว่าแพทย์ทั่วไปในช่วงปีแรกถึงปีที่สาม สามารถมี accuracy สูงในการแปลผลฟิล์มรังสีทรวงอก ข้อมูลนี้สามารถใช้สนับสนุนการพัฒนากระบวนการดูแลผู้ป่วยและการจัดอบรม targeted training สำหรับโรคที่มักพลาต เช่น infiltration หรือ subtle fracture

อย่างไรก็ตาม การศึกษานี้มีข้อจำกัด ได้แก่ การใช้รังสีแพทย์เพียง 1 ท่านเป็น gold standard โดยไม่มีการยืนยันด้วยเอกซเรย์คอมพิวเตอร์ การใช้ปัญญาประดิษฐ์หรือผู้ประเมินมากกว่าหนึ่งคน ซึ่งอาจมี bias

บางประการ นอกจากนี้ ผู้ป่วยบางส่วนไม่มีการบันทึกผลแปลฟิล์มโดยแพทย์เวรร้อยละ 32.1 ทำให้ข้อมูลไม่ครบถ้วนทั้งหมด รวมถึงเป็นการศึกษาเชิง retrospective ซึ่งอาจมีข้อจำกัดด้านคุณภาพข้อมูล ข้อจำกัดเหล่านี้สามารถนำไปพัฒนาการศึกษาในอนาคตได้

สรุป

แพทย์เวรทั่วไปในโรงพยาบาลระดับ M1 มีความแม่นยำในการแปลผลฟิล์มรังสีทรวงอกอยู่ในระดับสูง (accuracy ร้อยละ 95.9) แต่ยังคงพบความผิดพลาด โดยเฉพาะ false negative และ partial interpretation ในรอยโรค subtle ดังนั้น ควรมีการพัฒนาแนวทางการฝึกอบรมและพิจารณานำเทคโนโลยี AI มาใช้เป็นเครื่องมือช่วยแปลผล เพื่อลดความผิดพลาดและเพิ่มคุณภาพการดูแลผู้ป่วยในอนาคต

กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณนางสาวจุฬารภรณ์ พูลเยี่ยม นักสถิติ หน่วยการจัดการข้อมูลเพื่อการวิจัย ฝ้ายวิจิตร คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล ที่ได้ให้คำปรึกษาและความช่วยเหลือในการวิเคราะห์ข้อมูลทางสถิติอย่างละเอียดและครบถ้วน การแนะนำเกี่ยวกับการเลือกใช้วิธีการทางสถิติที่เหมาะสม

เอกสารอ้างอิง

1. Geftter WB, Post BA, Hatabu H. Commonly Missed Findings on Chest Radiographs: Causes and Consequences. *Chest* 2023;163(3): 650-61. doi: 10.1016/j.chest.2022.10.039.
2. Jimah BB, Baffour Appiah A, Sarkodie BD, Anim D. Competency in Chest Radiography Interpretation by Junior Doctors and Final Year Medical Students at a Teaching Hospital. *Radiol Res Pract* 2020;2020:8861206. doi: 10.1155/2020/8861206.

3. A Halim NI, Mohd Zaki F, Manan HA, Mohamed Z. An Evaluation of the Quality of Plain Radiograph Interpretations by Radiology Trainees: A Single Institution Experience. *Diagnostics (Basel)* 2022;12(8):1954. doi: 10.3390/diagnostics12081954.
4. Bergus GR, Franken EA Jr, Koch TJ, Smith WL, Evans ER, Berbaum KS. Radiologic interpretation by family physicians in an office practice setting. *J Fam Pract* 1995;41(4): 352-6. PMID: 7561708
5. Petinaux B, Bhat R, Boniface K, Aristizabal J. Accuracy of radiographic readings in the emergency department. *Am J Emerg Med* 2011;29(1):18-25. doi: 10.1016/j.ajem.2009.07.011.
6. Pinto A, Reginelli A, Pinto F, Lo Re G, Midiri F, Muzj C, et al. Errors in imaging patients in the emergency setting. *Br J Radiol* 2016;89(1061):20150914. doi: 10.1259/bjr.20150914.
7. Huang J, Neill L, Wittbrodt M, Melnick D, Klug M, Thompson M, et al. Generative Artificial Intelligence for Chest Radiograph Interpretation in the Emergency Department. *JAMA Netw Open* 2023;6(10):e2336100. doi: 10.1001/jamanetworkopen.2023.36100.
8. Ahn JS, Ebrahimian S, McDermott S, Lee S, Naccarato L, Di Capua JF, et al. Association of Artificial Intelligence-Aided Chest Radiograph Interpretation With Reader Performance and Efficiency. *JAMA Netw Open* 2022;5(8):e2229289. doi:10.1001/jamanetworkopen.2022.29289.
9. Bennani S, Regnard NE, Ventre J, Lassalle L, Nguyen T, Ducarouge A, et al. Using AI to Improve Radiologist Performance in Detection of Abnormalities on Chest Radiographs. *Radiology* 2023;309(3):e230860. doi: 10.1148/radiol.230860.
10. รุ่งนภา อุนพวงศ์พิพัฒน์. การเปรียบเทียบการแปลผลภาพถ่ายรังสีทรวงอก: กรณีการแปลผลโดยรังสีแพทย์และเทคโนโลยีปัญญาประดิษฐ์ โรงพยาบาลกาฬสินธุ์. *วารสารโรคมะเร็ง* 2567;44(3):173-83.