



Development and Implementation of a Criteria-Based Radiology Imaging Selection Platform (CRISP-AI)

Othria Ahmed, MD, Northwell Health; Nicholas Mynarski, MD; Ritesh D. Patel, MD

Introduction/Background

Utilization of radiologic imaging has increased exponentially over the past decades, with cross-sectional imaging volumes increasing by 1–5% annually (1). Recent studies have demonstrated a significant increase in radiologic imaging per patient encounter in the emergency department with decreased test positivity (2,3). Appropriate imaging selection is essential to promote high-value, evidence-based care while reducing unnecessary imaging. The American College of Radiology (ACR) Appropriateness Criteria provides evidence-based recommendations to guide imaging selection. Applying these guidelines consistently at the point of care is challenging for time-constrained clinicians.

Methods/Intervention

Criteria-based Radiology Imaging Selection Platform (CRISP-AI), a large language model (LLM)-based chatbot, was developed to provide evidence-based imaging recommendations using the ACR Appropriateness Criteria and our institution's protocols/policies. The platform incorporates these guidelines across seven radiologic subspecialties, creating a centralized clinical decision support resource for a broad range of imaging indications. Providers can enter patient-specific clinical information including age, sex, presenting symptoms, and relevant imaging history. When additional information is required to determine the most appropriate examination, CRISP-AI prompts targeted follow-up questions before generating an imaging recommendation. The chatbot also provides a link to the referenced guidelines for clinician review.

Results/Outcome

To evaluate performance, the platform was tested using 100 real-world clinical scenarios from our institution across multiple radiology subspecialties included in its knowledge base. Two radiology residents independently reviewed and compared CRISP-AI's recommendations against the current ACR Appropriateness Criteria and institutional protocols for each unique scenario. CRISP-AI demonstrated an overall concordance of 94% with the reviewers' assessments.

Conclusion

CRISP-AI demonstrates the potential of LLMs to serve as an accessible, evidence-based clinical decision support tool for radiologic imaging selection. This promotes enhanced diagnostic efficiency and optimal use of healthcare resources. Additionally, CRISP-AI may serve as a valuable resource for radiologists by providing guideline-based recommendations during report interpretation, particularly for clinical scenarios outside a radiologist's primary expertise. Future implementation within clinical workflows and electronic medical record systems will be necessary to evaluate its impact on imaging utilization, provider decision-making, and patient outcomes.

Statement of Impact

By translating the ACR Appropriateness Criteria into a real-time interactive platform, CRISP-AI aids physicians in enhancing diagnostic efficiency via selection of appropriate imaging at the point-of-care.

References:

1. McDonald RJ, Schwartz KM, Eckel LJ, et al. The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload. *Acad Radiol*. 2015;22(9):1191-1198. doi:10.1016/j.acra.2015.05.007
2. Prasad S, Hood CM, Young C, et al. Computed Tomographic Angiography and Yield for Gastrointestinal Bleeding in the Emergency Department. *JAMA Netw Open*. 2025;8(8):e2529746. doi:10.1001/jamanetworkopen.2025.29746
3. Dylla, L., Krothapalli, N., Tu, L., Payabvash, S., Burke, J. F., Sheth, K. N., Wira, C., & de Havenon, A. (2025). Trends in Head CT use in US emergency department patients from 2007 to 2022. *Neurology*, 105(12). <https://doi.org/10.1212/wnl.00000000000214347>

QRFBSUS-2479843-2-ANY.pdf

Figures 1-3 including workflow, sample responses from CRISP-AI, and reviewer concordance results

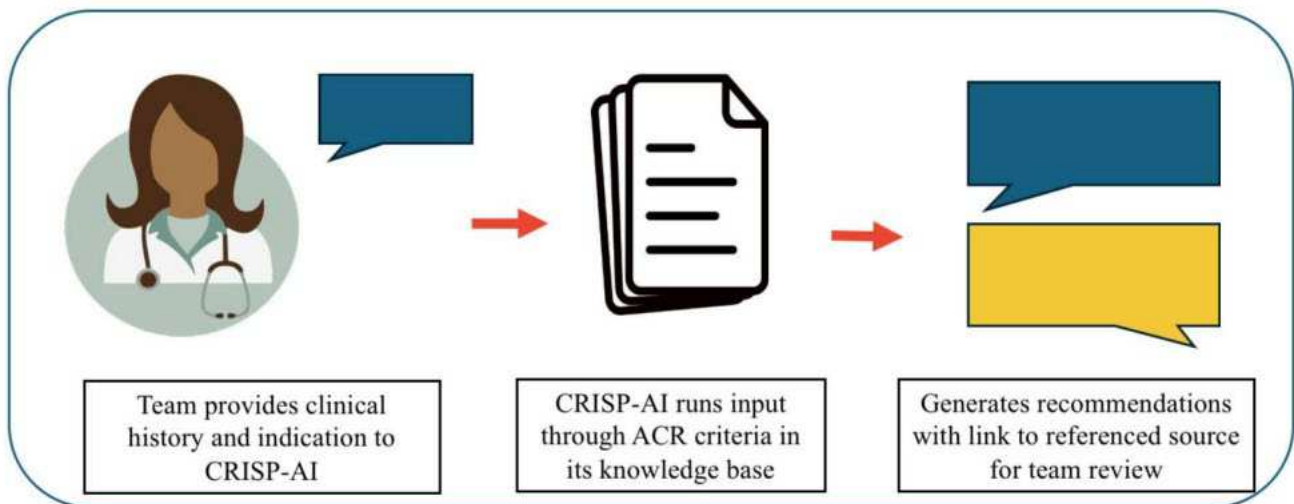


Fig 1. Workflow.

Clinical teams and radiologists alike can provide relevant clinical history, demographics, and indication into CRISP-AI platform. Gemini 2.5 will utilize the ACR appropriateness criteria guidelines from 7 radiology subspecialties and our institution's protocols/policies within its knowledge base to generate an evidence-based recommendation.

A. Please input primary clinical concern and patient history. Answer the resulting questions, if asked, to obtain recommendations for radiological imaging based upon ACR's Appropriateness Criteria

B. What imaging to order for congenital hip dislocation?

To provide the most accurate recommendation, could you please specify the age of the child? Imaging recommendations for congenital hip dislocation (Developmental Dysplasia of the Hip) vary significantly based on the patient's age.

The patient is 2 months old

According to ACR Appropriateness Criteria for Developmental Dysplasia of the Hip-Child, for a 2-month-old patient with congenital hip dislocation (which falls under "Child, younger than 4 months of age. Physical findings of DDH. Initial imaging," or "Child, between 4 weeks to 4 months of age. Equivocal physical examination or risk factors for DDH. Initial imaging."), the following imaging studies are recommended:

US Hips
This study is usually appropriate for initial imaging in a 2-month-old with suspected developmental dysplasia of the hip. Ultrasound is highly effective in evaluating the cartilaginous structures of the infant hip, which are not well visualized on radiographs at this age. It allows for dynamic assessment of hip stability and morphology.

Radiography Pelvis
This study is usually not appropriate for initial imaging in a 2-month-old. Pelvic radiographs are limited in evaluating developmental dysplasia of the hip in infants younger than 3 months because the ossific nucleus of the femoral head typically appears later (between 4 to 6 months), and the acetabular margin is largely cartilaginous, making assessment of morphology and alignment difficult.

ACR Appropriateness Criteria document used: Developmental Dysplasia of the Hip-Child.pdf

Today 14:25 • Gemini 2.0 Flash • 2 sources

What imaging do I order for a patient 8 months old with history of congenital hip dysplasia?

According to ACR Appropriateness Criteria for a child older than 6 months of age with concern for Developmental Dysplasia of the Hip (DDH):

Radiography Pelvis
This procedure is usually appropriate. At 8 months of age, the femoral head ossific nucleus is typically present, making pelvic radiography the preferred imaging modality. It allows for the assessment of the femoral head ossific nucleus, the development of the proximal femur, and bony acetabular morphology. Clinical assessment at this age is often limited, leading to a greater reliance on imaging for confirmation and monitoring of DDH.

US Hips
This procedure is usually not appropriate. For children older than 6 months, there is insufficient evidence to recommend the use of ultrasound, as the evaluation may be inadequate due to suboptimal visualization of the hip joint anatomy from decreased acoustic penetration.

ACR Appropriateness Criteria document used: Developmental Dysplasia of the Hip-Child.pdf

Today 14:25 • Gemini 2.0 Flash • 2 sources

Fig 2. CRISP-AI Platform.

A. Initial prompt for users to provide their clinical scenario. **B.** CRISP-AI will interact with users, asking relevant clinical information including past medical/surgical history, demographics, patient setting, and prior imaging. CRISP-AI's responses will vary based on these demographics, such as age underlined in the figure. It will also provide a link to the imaging guidelines it utilized for the recommendation for providers to review.

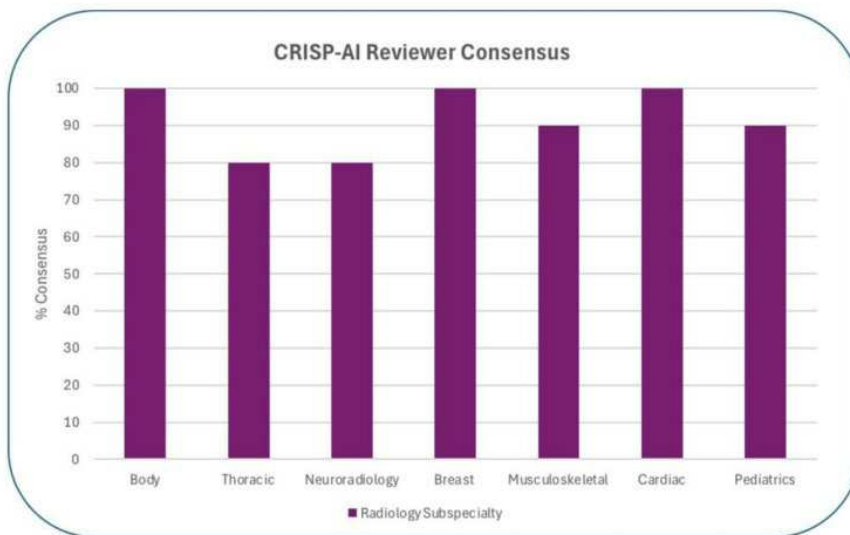


Fig 3. CRISP-AI Reviewer Consensus.

100 unique clinical scenarios taken from our institution were provided to the chatbot. The generated responses were reviewed by 2 independent radiology residents for accuracy in adhering to the ACR Appropriateness Criteria. CRISP-AI had 94% consensus, with slight variations in accuracy per subspecialty as depicted above.

Keywords

Artificial intelligence; Large Language Model; Radiology; ACR Appropriateness Criteria; Overutilization; Patient Safety

