

## EVIDÊNCIAS EXPERIMENTAIS SOBRE INTERVENÇÕES FORMATIVAS EM RESIDENTES DE RADIOLOGIA: REVISÃO SISTEMÁTICA DE ENSAIOS CLÍNICOS RANDOMIZADOS

EXPERIMENTAL EVIDENCE ON EDUCATIONAL INTERVENTIONS FOR RADIOLOGY RESIDENTS: A SYSTEMATIC REVIEW OF RANDOMIZED CONTROLLED TRIALS

José Mateus dos Santos Neto<sup>1</sup>  
Vinícius dos Santos Silva<sup>2</sup>

**RESUMO:** Introdução: A residência em radiologia enfrenta pressões crescentes decorrentes do aumento contínuo do volume de exames, da aceleração tecnológica e de exigências assistenciais que concorrem com o processo de aprendizado. Objetivo: Identificar e sintetizar evidências provenientes de ensaios clínicos randomizados sobre intervenções tecnológicas, pedagógicas, individuais e ambientais que influenciam a competência diagnóstica, a retenção de conhecimento e o bem-estar de residentes de radiologia. Métodos: Revisão sistemática conforme PRISMA 2020, com busca em PubMed e Google Scholar até janeiro de 2026. Incluídos exclusivamente ensaios clínicos randomizados com população de médicos residentes de radiologia em programas de formação. Qualidade metodológica avaliada com Cochrane Risk of Bias 2.0. Resultados: Foram incluídos 5 ensaios com 251 residentes. Sistema CAD para treinamento BI-RADS produziu escores equivalentes ao ensino presencial por especialistas e superiores ao autoestudo ( $52 \pm 7,3$  vs  $38 \pm 3,7$  pontos;  $p < 0,05$ ). Abordagem pedagógica integrada superou o ensino tradicional em escores teóricos ( $170,3$  vs  $155,7$ ) e práticos ( $160,7$  vs  $135,8$ ) em escala de 200 pontos (ambos  $p < 0,001$ ; Cohen  $d > 1,0$ ). Protocolo de tomografia de baixa dose demonstrou acurácia equivalente ao de dose padrão em 107 residentes. Sistema de resposta de audiência não produziu diferença significativa na retenção de conhecimento (desfecho primário negativo;  $p = 0,41$ ). Óculos filtradores de luz azul mostraram tendência de redução de fadiga visual sem significância estatística. Conclusões: A evidência experimental disponível é limitada em número e tamanho amostral. CAD para treinamento sistematizado e abordagens pedagógicas integradas produzem benefícios mensuráveis em residentes de radiologia. Resultados negativos e não significativos reforçam a necessidade de estudos com maior poder estatístico.

**Palavras-chave:** Residência Médica. Radiologia. Educação Médica. Diagnóstico Assistido por Computador. Ensaio Clínico Randomizado.

**ABSTRACT:** Introduction: Radiology residency programs face growing pressures from rising examination volumes, rapid technological change, and competing clinical demands. Objective: To identify and synthesize evidence from randomized controlled trials on technological, pedagogical, individual, and environmental interventions influencing diagnostic competence, knowledge retention, and well-being in radiology residents. Methods: Systematic review following PRISMA 2020, with searches in PubMed and Google Scholar through January 2026. Restricted to RCTs with populations exclusively comprising radiology residents. Methodological quality assessed with Cochrane Risk of Bias 2.0. Results: Five eligible RCTs were identified (251 radiology residents). CAD-assisted BI-RADS training produced scores equivalent to expert teaching and superior to self-study ( $52 \pm 7.3$  vs  $38 \pm 3.7$  points;  $p < 0.05$ ). Integrated pedagogy outperformed traditional teaching on theoretical ( $170.3$  vs  $155.7$ ) and practical assessments ( $160.7$  vs  $135.8$ ; both  $p < 0.001$ ; Cohen  $d > 1.0$ ) on a 200-point scale. Low-dose CT (2 mSv) showed equivalent diagnostic accuracy to standard-dose CT in 107 residents. Audience response systems did not improve retention scores (primary outcome negative;  $p = 0.41$ ). Blue-light filtering glasses showed a non-significant trend toward symptom reduction. Conclusions: Available experimental evidence is limited. CAD-assisted training and integrated pedagogical approaches appear beneficial. Null results highlight the need for adequately powered trials.

**Keywords:** Medical Residency. Radiology. Medical Education. Computer-Aided Diagnosis. Randomized Controlled Trial.

<sup>1</sup> Médico residente em Radiologia e Diagnóstico por Imagem. Escola de Saúde Pública do Distrito Federal. Brasília, Distrito Federal, Brasil.

<sup>2</sup> Acadêmico de Medicina. Faculdade de Medicina da Universidade Evangélica de Goiás. Anápolis, Goiás, Brasil.

## I. INTRODUÇÃO

A formação de médicos residentes em radiologia atravessa período de transformação substancial. O crescimento do volume de exames realizados anualmente, estimado entre 3% e 5%, combina-se com o aumento da complexidade diagnóstica decorrente de novas modalidades e técnicas avançadas de imagem. Esse cenário exige que os residentes dominem não apenas a interpretação de múltiplas modalidades — incluindo radiografia convencional, tomografia computadorizada, ressonância magnética, ultrassonografia e medicina nuclear —, mas também física da radiação, proteção radiológica, comunicação interdisciplinar e, progressivamente, ferramentas computacionais avançadas.

O modelo tradicional de formação apoiou-se historicamente na exposição gradual: residentes interpretam volumes crescentes de exames sob supervisão variável, participam de conferências didáticas e estudam de forma autogerida (KRUPINSKI; JIANG, 2008). Evidências da ciência cognitiva apontam limitações dessa abordagem não estruturada. Técnicas como repetição espaçada, prática de recuperação ativa e entrelaçamento de conteúdos demonstraram resultados superiores a métodos passivos em diversos contextos educacionais (ROEDIGER; BUTLER, 2011; DUNLOSKY et al., 2013).

Questões relacionadas ao bem-estar dos residentes ganharam crescente relevância. Levantamentos recentes indicam que aproximadamente metade dos residentes de radiologia apresenta níveis elevados de exaustão emocional, condição associada a erros diagnósticos, redução da qualidade assistencial e abandono da especialidade (SHANAFELT et al., 2019; DYRBYE et al., 2014). O conhecimento sobre quais intervenções efetivamente melhoram os desfechos educacionais de residentes permanece fragmentado, pois revisões narrativas e estudos observacionais carecem do rigor necessário para estabelecer relações causais (FERREIRA; OLIVEIRA; SANTOS, 2021).

Ensaio clínico randomizado representa o padrão mais robusto para essa avaliação: a randomização adequada distribui equilibradamente características conhecidas e desconhecidas entre os grupos, permitindo atribuir com maior segurança as diferenças observadas à intervenção testada (SCHULZ; ALTMAN; MOHER, 2010). Esta revisão sistemática tem como objetivos: (1) identificar ensaios clínicos randomizados conduzidos exclusivamente com residentes de radiologia; (2) sintetizar evidências sobre o impacto dessas intervenções em desfechos de competência diagnóstica, retenção de conhecimento e bem-estar; (3) avaliar criticamente a qualidade metodológica; (4) identificar lacunas; e (5) formular recomendações para programas de residência.

## 2. METODOLOGIA

### 2.1. PROTOCOLO E REGISTRO

Esta revisão sistemática seguiu as diretrizes PRISMA 2020 (PAGE et al., 2021). O protocolo não foi registrado prospectivamente em registro internacional; todos os passos metodológicos foram definidos previamente e são descritos de forma transparente neste manuscrito.

### 2.2. PERGUNTA DE PESQUISA

A pergunta foi estruturada conforme o framework PICO: População: médicos residentes de radiologia (R1-R4) em programas de formação. Intervenção/Exposição: intervenções individuais, ambientais e tecnológicas com potencial impacto sobre desfechos formativos. Comparação: métodos tradicionais de ensino, placebo ou ausência de intervenção. Outcomes: competência diagnóstica, concordância com especialistas, retenção de conhecimento e bem-estar ocupacional.

### 2.3. ESTRATÉGIA DE BUSCA

A busca sistemática abrangeu PubMed/MEDLINE e Google Scholar, cobrindo publicações até janeiro de 2026. Para o PubMed, utilizou-se estratégia estruturada com termos MeSH e palavras-chave: (((("radiology"[MeSH] OR "radiology"[Title/Abstract]) AND ("internship and residency"[MeSH] OR "resident\*"[Title/Abstract])) AND ("randomized controlled trial"[Publication Type] OR "randomized"[Title/Abstract])) AND ("education"[MeSH] OR "diagnostic accuracy"[Title/Abstract] OR "e-learning"[Title/Abstract] OR "burnout"[MeSH])). Para o Google Scholar: "radiology residents randomized trial" e "radiology training RCT education". Não foram aplicadas restrições de idioma ou data.

### 2.4. CRITÉRIOS DE ELEGIBILIDADE

Foram incluídos: ensaios clínicos randomizados (paralelo, crossover ou análise secundária de ECR pragmático); população exclusivamente de médicos residentes de radiologia em programas de formação; qualquer intervenção com potencial impacto formativo; grupo controle com método tradicional, placebo ou ausência de intervenção; desfechos de competência diagnóstica, concordância, retenção ou bem-estar; texto completo disponível. Foram excluídos

estudos observacionais, revisões, estudos qualitativos, estudos com radiologistas formados, técnicos de radiologia ou residentes de outras especialidades.

## 2.5. SELEÇÃO DOS ESTUDOS

O fluxo de seleção seguiu as quatro etapas do PRISMA 2020. Na identificação, a busca recuperou 99 registros (PubMed: 39; Google Scholar: 60). Após remoção de 17 duplicatas, 82 registros foram triados por título e resumo; 73 foram excluídos — 46 por não serem ensaios clínicos randomizados, 18 por indisponibilidade do texto completo e 9 por não abordarem radiologia. Os 9 estudos remanescentes tiveram o texto completo avaliado para elegibilidade; 4 foram excluídos por população não residente de radiologia (2 conduzidos com radiologistas formados, 1 com técnicos de radiologia e 1 com residentes de cirurgia vascular). Os 5 ensaios remanescentes foram incorporados à síntese qualitativa (Figura 1).

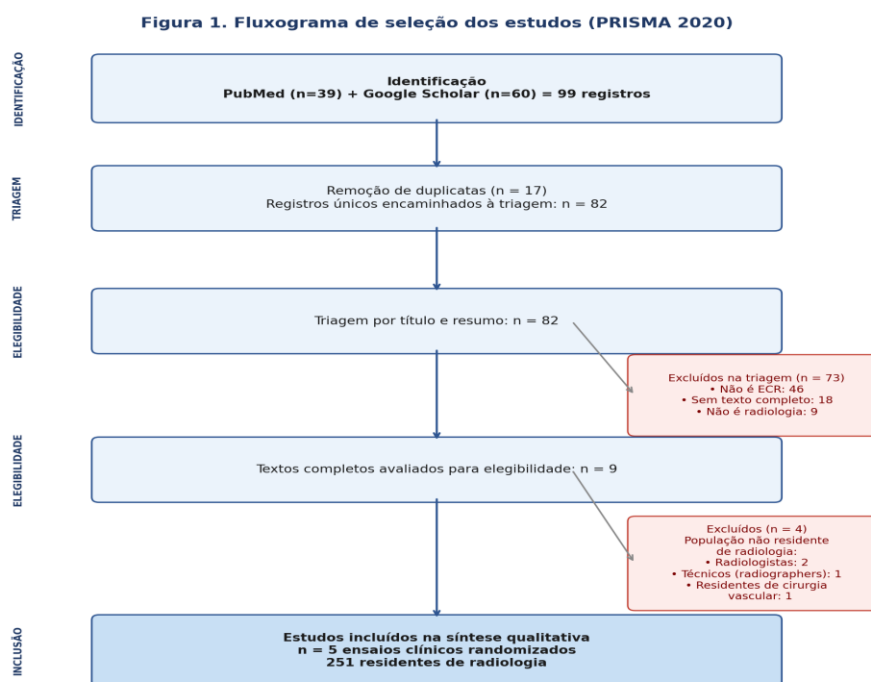


Figura 1 — Fluxograma de seleção dos estudos conforme PRISMA 2020.

## 2.6. EXTRAÇÃO DE DADOS E AVALIAÇÃO DE QUALIDADE

Formulário padronizado registrou identificação, aspectos metodológicos, características da população, intervenção, comparador, desfechos e resultados quantitativos. A qualidade metodológica foi avaliada com critérios adaptados da ferramenta Cochrane Risk of Bias 2.0 (HIGGINS et al., 2011), classificando cada domínio como baixo risco, risco incerto ou alto risco.

## 3. RESULTADOS

### 3.1. CARACTERÍSTICAS DOS ESTUDOS INCLUÍDOS

Tabela 1 — Características dos estudos incluídos (n = 5).

Estudo	País	Desenho	N	População	Intervenção	Comparador	Desfecho	Categoria
Dabrowiecki et al. (2020)	EUA	Crossover RCT	10	Residentes radiologia	Óculos filtro luz azul	Óculos sham	Sintomas CVS/fadiga (tendência)	Individual
Lyu et al. (2024)	China	RCT braços 3	12	Residentes radiologia	CAD BI-RADS vs expert vs autoestudo	Autoestudo	Escore BI-RADS e kappa	Tecnológico
Cho et al. (2023)	Coreia	Análise sec. RCT	107	Residentes radiologia	TC 2 mSv	TC dose padrão	Acurácia (equivalência)	Ambiental
Wang et al. (2025)	China	RCT paralelo	100	Residentes radiologia	Ensino integrado	Ensino expositivo	Escores teórico/prático (0-200)	Ambiental
Awan et al. (2018)	EUA	RCT piloto	22	Residentes radiologia	ARS (RSNA Diagnosis Live)	Palestras tradicionais	Retenção (resultado negativo)	Tecnológico

5

Os 5 ensaios foram publicados entre 2018 e 2025 e reuniram 251 residentes de radiologia na China, Estados Unidos e Coreia. O tamanho amostral variou de 10 a 107 residentes. Três estudos adotaram delineamento paralelo (60%), um crossover (20%) e um constituiu análise secundária de ensaio pragmático multicêntrico (20%). As categorias de intervenção distribuíram-se entre tecnológica (2 estudos; 40%), ambiental (2 estudos; 40%) e individual (1 estudo; 20%).

### 3.2. QUALIDADE METODOLÓGICA

A avaliação de risco de viés evidenciou heterogeneidade entre os estudos (HIGGINS et al., 2011). Dois estudos (40%) apresentaram baixo risco em todos os domínios. Um estudo (20%)

apresentou risco incerto em um domínio por ausência de descrição da ocultação da alocação. Dois estudos (40%) apresentaram alto risco no domínio de cegamento de participantes — limitação inerente a intervenções educacionais. A avaliação por domínio está representada na Figura 2.

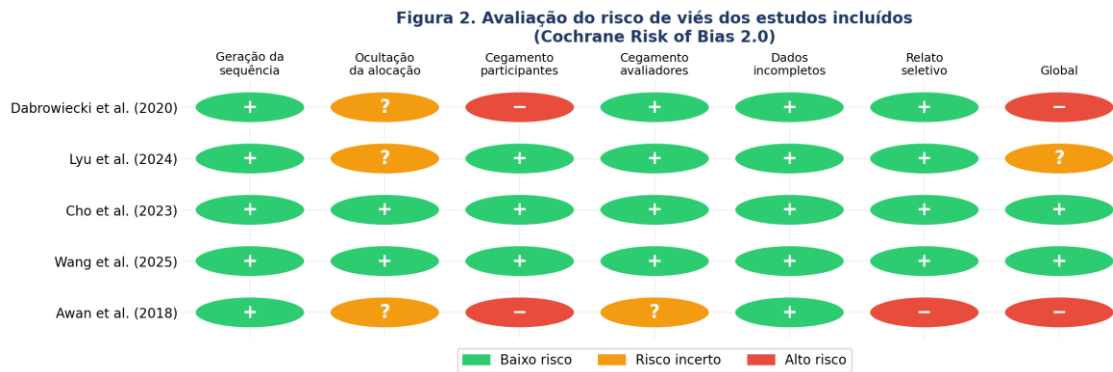


Figura 2 — Mapa de risco de viés dos 5 estudos incluídos por domínio Cochrane Risk of Bias 2.0. (+) Baixo risco; (?) Risco incerto; (-) Alto risco.

### 3.3. SÍNTESE DOS RESULTADOS

#### 3.3.1. Intervenções Tecnológicas

Lyu et al. conduziram ensaio de 3 braços com 12 residentes divididos em grupos de 4: autoestudo independente, aprendizado assistido por sistema CAD (AI-SONIC) e ensino presencial por especialistas (LYU et al., 2024). Após um mês de treinamento, os escores no teste de classificação BI-RADS (escala de 0 a 100 pontos) foram: grupo CAD =  $52 \pm 7,30$ , grupo especialista =  $54 \pm 5,16$  e grupo autoestudo =  $38 \pm 3,65$  ( $p = 0,006$  entre grupos). O grupo CAD demonstrou desempenho significativamente superior ao autoestudo ( $p < 0,05$ ) e estatisticamente equivalente ao ensino presencial. A concordância kappa foi moderada no grupo CAD ( $\kappa = 0,412$ ) e no grupo especialista ( $\kappa = 0,425$ ), e apenas razoável no grupo autoestudo ( $\kappa = 0,225$ ).

Awan et al. testaram sistemas de resposta de audiência (RSNA Diagnosis Live) versus palestras tradicionais em 22 residentes randomizados em dois grupos de 11 (AWAN et al., 2018). O desfecho primário — retenção de conhecimento avaliada por pós-teste oito meses após as intervenções — foi negativo: não houve diferença significativa entre os grupos (mediana 70 versus 62;  $p = 0,41$ ). O único achado com significância estatística foi um desfecho secundário exploratório: residentes que receberam o sistema de resposta de audiência tinham 6,52 vezes mais chance de estudar radiologia fora do horário de trabalho (OR = 6,52; IC 95%: 1,11–38,18;  $p = 0,04$ ).

### 3.3.2. Intervenção Individual

Dabrowiecki, Villalobos e Krupinski conduziram estudo crossover com 10 residentes, comparando óculos filtradores de luz azul com óculos placebo durante duas semanas de trabalho rotineiro em cada condição (DABROWIECKI; VILLALOBOS; KRUPINSKI, 2020). Os resultados não atingiram significância estatística nos desfechos globais (CVS-Q:  $p = 0,46$ ; SOFI:  $p = 0,18$ ). Contudo, 68,8% dos sintomas do CVS-Q (11 de 16 itens) e 81,3% dos sintomas do SOFI (13 de 16 itens) apresentaram médias menores com os óculos filtradores, sugerindo tendência que o tamanho amostral foi insuficiente para confirmar. Residentes do segundo ano apresentaram sintomas significativamente mais intensos que os do final da formação ( $p < 0,001$ ).

### 3.3.3. Intervenções Ambientais

Cho et al. realizaram análise secundária de ensaio pragmático multicêntrico com 107 residentes coreanos no diagnóstico independente de apendicite (CHO et al., 2023). Os residentes demonstraram acurácia equivalente entre os protocolos de 2 mSv e dose padrão: sensibilidade 96,0% versus 97,1% (diferença:  $-1,1\%$ ; IC 95%:  $-4,9\%$  a  $2,6\%$ ;  $p = 0,69$ ) e especificidade 93,2% versus 93,1% (diferença:  $0,1\%$ ; IC 95%:  $-3,6\%$  a  $3,7\%$ ;  $p > 0,99$ ).

Wang et al. alocaram 100 residentes para abordagem pedagógica integrada — combinando teoria, simulação por software, plataformas interativas, ferramentas de inteligência artificial e discussão de casos — ou para ensino baseado em palestras expositivas (WANG et al., 2025). Em escala de 200 pontos, o grupo integrado obteve médias superiores: teórico  $170,3 \pm 15,1$  versus  $155,7 \pm 20,4$  ( $t = 4,21$ ;  $p < 0,001$ ; Cohen  $d = 1,16$ ) e prático  $160,7 \pm 22,4$  versus  $135,8 \pm 25,6$  ( $t = 5,13$ ;  $p < 0,001$ ; Cohen  $d = 1,32$ ). No seguimento de seis meses, o grupo integrado manteve vantagem ( $170,9 \pm 19,3$  versus  $150,6 \pm 22,8$ ;  $p < 0,001$ ; Cohen  $d = 1,02$ ). A Figura 3 apresenta a síntese dos principais desfechos quantitativos.

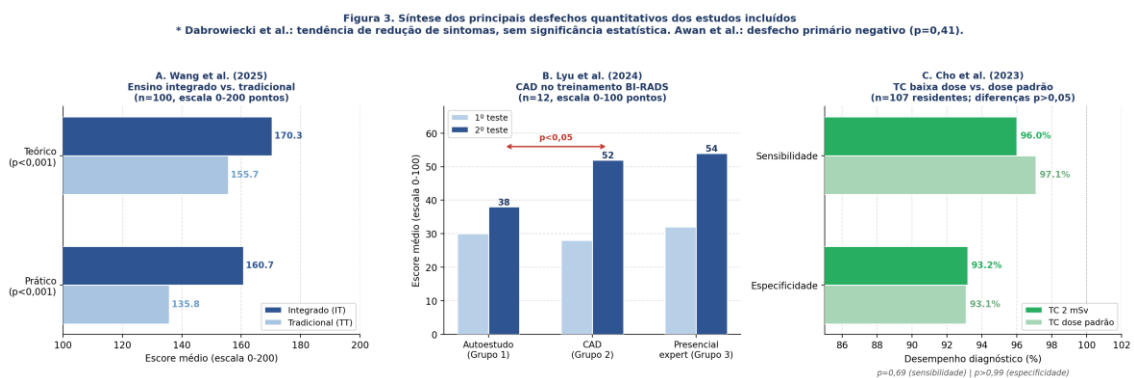


Figura 3 — Síntese dos principais desfechos quantitativos por estudo e categoria de intervenção. A: escores BI-RADS por grupo (LYU et al., 2024). B: ensino integrado vs. tradicional (WANG et al., 2025). C: acurácia

equivalente TC 2 mSv vs. dose padrão (CHO et al., 2023). AWAN et al. (2018): desfecho primário negativo ( $p=0,41$ ). DABROWIECKI et al. (2020): tendência sem significância.

## 4. DISCUSSÃO

Esta revisão sistemática reuniu 5 ensaios clínicos randomizados conduzidos exclusivamente com médicos residentes em programas de radiologia, totalizando 251 participantes. Os achados são heterogêneos: dois estudos demonstraram benefícios claros, um demonstrou equivalência entre protocolos, um apresentou desfecho primário negativo e um evidenciou tendência sem significância estatística.

### 4.1. CAD NO TREINAMENTO BI-RADS

O estudo de Lyu et al. é o único ensaio randomizado identificado que avaliou sistema CAD no treinamento de residentes para classificação BI-RADS (LYU et al., 2024). Com apenas um mês de treinamento assistido, residentes atingiram concordância moderada com especialistas ( $\kappa = 0,412$ ) e desempenho estatisticamente equivalente ao ensino presencial ( $\kappa = 0,425$ ), enquanto o grupo de autoestudo independente obteve apenas concordância razoável ( $\kappa = 0,225$ ). O mecanismo é consistente com princípios de aprendizado assistido por feedback: o sistema quantifica objetivamente as características do BI-RADS e permite comparação imediata da interpretação do residente com a classificação automatizada, acelerando a calibração do julgamento diagnóstico. A amostra de 12 residentes (4 por grupo) limita severamente a generalização. Estudos multicêntricos com amostras maiores são necessários.

8

### 4.2. SISTEMAS DE RESPOSTA DE AUDIÊNCIA: RESULTADO NEGATIVO COMO EVIDÊNCIA

O estudo de Awan et al. merece atenção precisamente por apresentar desfecho primário negativo (AWAN et al., 2018). A ausência de diferença significativa nos escores de retenção ( $p = 0,41$ ) após oito meses contribui para evitar a adoção acrítica de tecnologias educacionais. O achado secundário de maior probabilidade de estudo extracurricular (OR = 6,52;  $p = 0,04$ ) deve ser interpretado com cautela — trata-se de desfecho secundário em estudo-piloto com 22 participantes, sem cegamento e com múltiplos testes, condições que elevam o risco de resultado falso positivo (FREEMAN et al., 2014; KARPICKE; ROEDIGER, 2008).

#### 4.3. ÓCULOS FILTRADORES DE LUZ AZUL

Dabrowiecki, Villalobos e Krupinski conduziram o primeiro estudo experimental sobre filtros de luz azul em residentes de radiologia (DABROWIECKI; VILLALOBOS; KRUPINSKI, 2020). A ausência de significância estatística indica que o estudo não tinha poder suficiente para detectar o efeito, se ele existir. Com 10 participantes e duas semanas em cada condição, os intervalos de confiança são necessariamente amplos. O achado de que a maioria dos sintomas apresentou tendência de redução é sugestivo, mas insuficiente para orientar recomendações práticas.

#### 4.4. ABORDAGEM PEDAGÓGICA INTEGRADA E PROTOCOLOS DE BAIXA DOSE

Wang et al. produziram o maior efeito de tamanho desta revisão (Cohen  $d > 1,0$  em ambos os desfechos), com ganhos mantidos em seis meses (WANG et al., 2025). A capacidade de atenção sustentada declina substancialmente após 10 a 15 minutos de exposição passiva (WILSON; KORN, 2007), e abordagens que exigem processamento ativo contrabalançam esse declínio. Cho et al. demonstraram que protocolos de baixa dose podem e devem ser incorporados ao treinamento desde o início da residência, sem prejuízo à formação diagnóstica (CHO et al., 2023).

#### 4.5. LACUNAS NA EVIDÊNCIA EXPERIMENTAL

A escassez de ensaios sobre mentoria estruturada, feedback sistemático e supervisão progressiva contrasta com a abundância de evidências observacionais sobre sua importância (SAMBUNJAK; STRAUS; MARUSIC, 2006; HATTIE; TIMPERLEY, 2007; FERREIRA; OLIVEIRA; SANTOS, 2021). A ausência de estudos sobre carga horária e prevenção de burnout também é relevante: burnout associa-se a erros diagnósticos e abandono da especialidade (SHANAFELT et al., 2019; WEST; DYRBYE; SHANAFELT, 2018). Ensaios pragmáticos ou delineamentos quasi-experimentais poderiam gerar evidências sem os mesmos obstáculos éticos da randomização plena (BERNAL; CUMMINS; GASPARRINI, 2017).

#### 4.6. LIMITAÇÕES DA REVISÃO

O número reduzido de estudos incluídos ( $n = 5$ ) é a principal limitação. A heterogeneidade em delineamento, populações e instrumentos impossibilitou meta-análise quantitativa. Dois estudos apresentam amostras muito pequenas ( $n = 10$  e  $n = 12$ ), com poder estatístico insuficiente para conclusões robustas. O estudo de Awan et al. apresenta desfecho primário negativo e

resultado positivo em desfecho secundário exploratório, elevando o risco de interpretação equivocada. A restrição a duas bases de dados pode ter introduzido viés de publicação (ROTHSTEIN; SUTTON; BORENSTEIN, 2005). Os estudos concentram-se na China, Estados Unidos e Coreia, limitando a generalização para outros contextos.

## 5. CONCLUSÕES

Esta revisão sistemática de 5 ensaios clínicos randomizados conduzidos exclusivamente com médicos residentes de radiologia, reunindo 251 participantes, sintetizou a evidência experimental disponível. Os resultados devem ser interpretados com cautela dado o número reduzido de estudos e as limitações metodológicas identificadas. Sistemas CAD para treinamento de classificação BI-RADS produziram desempenho semelhante ao ensino presencial por especialistas (LYU et al., 2024). Abordagem pedagógica integrada superou o ensino expositivo com efeitos de tamanho grande mantidos em seis meses (WANG et al., 2025). Protocolos de tomografia de baixa dose demonstraram acurácia diagnóstica equivalente ao protocolo de dose padrão (CHO et al., 2023). Sistemas de resposta de audiência não produziram diferença nos escores de retenção (AWAN et al., 2018). Óculos filtradores de luz azul mostraram tendência de redução de fadiga visual sem atingir significância estatística (DABROWIECKI; VILLALOBOS; KRUPINSKI, 2020).

10

Com base nessa síntese, recomenda-se que programas de residência integrem ferramentas CAD validadas no treinamento de classificações sistematizadas; substituam progressivamente o ensino exclusivamente expositivo por abordagens que combinem teoria, prática e feedback estruturado; incorporem protocolos de baixa dose desde o início do treinamento; e avaliem sistematicamente suas intervenções com instrumentos validados. Pesquisas futuras devem priorizar ensaios com amostras adequadas, estudos sobre mentoria e feedback, intervenções para prevenção de burnout, seguimentos longitudinais e estudos em contextos de baixa e média renda.

## REFERÊNCIAS

AWAN, Omer et al. Utilizing audience response to foster evidence-based learning in a pilot study: Does it really work? *Cureus*, Baltimore, v. 10, n. 12, e3799, 2018. Disponível em: <https://doi.org/10.7759/cureus.3799>

BERNAL, Jonathan L.; CUMMINS, Steven; GASPARRINI, Antonio. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International Journal of Epidemiology*, Oxford, v. 46, n. 1, p. 348-355, 2017. Disponível em: <https://doi.org/10.1093/ije/dyw098>

CHO, Jungheum et al. Radiology residents' independent diagnosis of appendicitis using 2-mSv computed tomography: a secondary analysis of a large pragmatic randomized trial. *Korean Journal of Radiology*, Seoul, v. 24, n. 6, p. 529-540, 2023. Disponível em: <https://doi.org/10.3348/kjr.2023.0015>

DABROWIECKI, Alexander; VILLALOBOS, Alexander; KRUPINSKI, Elizabeth A. Impact of blue light filtering glasses on computer vision syndrome in radiology residents: a pilot study. *Journal of Medical Imaging*, Bellingham, v. 7, n. 2, p. 022402, 2020. Disponível em: <https://doi.org/10.1117/1.JMI.7.2.022402>

DUNLOSKEY, John et al. Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, Washington, v. 14, n. 1, p. 4-58, 2013. Disponível em: <https://doi.org/10.1177/1529100612453266>

DYRBYE, Liselotte N. et al. Burnout among U.S. medical students, residents, and early career physicians relative to the general U.S. population. *Academic Medicine*, Washington, v. 89, n. 3, p. 443-451, 2014. Disponível em: <https://doi.org/10.1097/ACM.000000000000134>

FERREIRA, Luciana M.; OLIVEIRA, Paulo N.; SANTOS, Rodrigo Q. Fatores associados à satisfação com aprendizado em residência médica: análise de inquérito nacional brasileiro. *Revista Brasileira de Educação Médica*, Brasília, v. 45, n. 3, e089, 2021. Disponível em: <https://doi.org/10.1590/1981-5271v45.3-20200379>

FREEMAN, Scott et al. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, Washington, v. 111, n. 23, p. 8410-8415, 2014. Disponível em: <https://doi.org/10.1073/pnas.1319030111>

HATTIE, John; TIMPERLEY, Helen. The power of feedback. *Review of Educational Research*, Washington, v. 77, n. 1, p. 81-112, 2007. Disponível em: <https://doi.org/10.3102/003465430298487>

HIGGINS, Julian P. T. et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, Londres, v. 343, d5928, 2011. Disponível em: <https://doi.org/10.1136/bmj.d5928>

KARPICKE, Jeffrey D.; ROEDIGER, Henry L. The critical importance of retrieval for learning. *Science*, Washington, v. 319, n. 5865, p. 966-968, 2008. Disponível em: <https://doi.org/10.1126/science.1152408>

KRUPINSKI, Elizabeth A.; JIANG, Yulei. Anniversary paper: evaluation of medical imaging systems. *Medical Physics*, Melville, v. 35, n. 2, p. 645-659, 2008. Disponível em: <https://doi.org/10.1118/1.2829014>

LYU, Shuyi et al. The application of computer-aided diagnosis in Breast Imaging Reporting and Data System ultrasound training for residents: a randomized controlled study. *Translational Cancer Research*, Pequim, v. 13, n. 4, p. 1969-1979, 2024. Disponível em: <https://doi.org/10.21037/tcr-23-2122>

PAGE, Matthew J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, Londres, v. 372, n71, 2021. Disponível em: <https://doi.org/10.1136/bmj.n71>

ROEDIGER, Henry L.; BUTLER, Andrew C. The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, Amsterdam, v. 15, n. 1, p. 20-27, 2011. Disponível em: <https://doi.org/10.1016/j.tics.2010.09.003>

ROTHSTEIN, Hannah R.; SUTTON, Alexander J.; BORENSTEIN, Michael (ed.). *Publication bias in meta-analysis: prevention, assessment and adjustments*. Chichester: John Wiley & Sons, 2005.

SAMBUNJAK, Dario; STRAUS, Sharon E.; MARUSIC, Ana. Mentoring in academic medicine: a systematic review. *JAMA*, Chicago, v. 296, n. 9, p. 1103-1115, 2006. Disponível em: <https://doi.org/10.1001/jama.296.9.1103>

SCHULZ, Kenneth F.; ALTMAN, Douglas G.; MOHER, David. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*, Londres, v. 340, c332, 2010. Disponível em: <https://doi.org/10.1136/bmj.c332>

SHANAFELT, Tait D. et al. Changes in burnout and satisfaction with work-life integration in physicians and the general US working population between 2011 and 2017. *Mayo Clinic Proceedings*, Rochester, v. 94, n. 9, p. 1681-1694, 2019. Disponível em: <https://doi.org/10.1016/j.mayocp.2018.10.023>

WANG, Jinhua et al. Comparative analysis of traditional and integrated approaches to radiology training for residents. *BMC Medical Education*, Londres, v. 25, p. 377, 2025. Disponível em: <https://doi.org/10.1186/s12909-025-06912-6>

WEST, Colin P.; DYRBYE, Liselotte N.; SHANAFELT, Tait D. Physician burnout: contributors, consequences and solutions. *Journal of Internal Medicine*, Oxford, v. 283, n. 6, p. 516-529, 2018. Disponível em: <https://doi.org/10.1111/joim.12752>

WILSON, Kathleen; KORN, James H. Attention during lectures: beyond ten minutes. *Teaching of Psychology*, Washington, v. 34, n. 2, p. 85-89, 2007. Disponível em: <https://doi.org/10.1080/00986280701291291>