The Clean Paper

WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

एक clinical AI सुरक्षित दिखीऔर paperwork बेहतर हुआ — लेकिन मरीज़ोके outcomes बेहतर नहींहुए

Kenya की16 primary-care facilities में pragmatic cluster-randomized trial ने clinical officers के record system में जोड़े गए generative-AI decision-support tool कोपरखा। 9,691 मरीज़ोमें 14 दिन कासख्त expert-adjudicated composite treatment-failure outcome AI tool के साथ 2.2% और बिनाtool 2.0% था(adjusted odds ratio 0.77, 95% CI 0.55–1.08, P = 0.13) — महत्वपूर्ण अंतर नहीं। Safety signal नहींमिला; documentation सभीrated domains में बेहतर हुई; prescribing और satisfaction नहींबदले; antibiotic costs घटने की दिशामें trend था। यह "failed study" नहीं, बल्कि दुर्लभ ईमानदार अध्ययन है — benchmarks पर नहीं, मरीज़ो पर मापागया— और इसकासाधारण निष्कर्ष यहीहै कि tool सुरक्षित दिखाऔर प्रक्रियामें मदद की, पर दो हफ्तोंमें मरीज़ोकोdemonstrably बेहतर नहींकिया। इतनाछोटाclinical benefit पकड़ने के लिए बहुत बड़ाtrial चाहिए होगा।

Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

Paper: *Generative AI-enabled clinical decision support system in primary care: a pragmatic, cluster-randomized trial*, Ambrose Agweyu, Paul Mwaniki, Vaishnavi Menon, Robert Korom, Lynda Isaaka, Conrad Wanyama, Xiaoxuan Liu & Bilal A. Mateen (and colleagues), Nature Medicine (2026) · DOI: 10.1038/s41591-026-04503-6

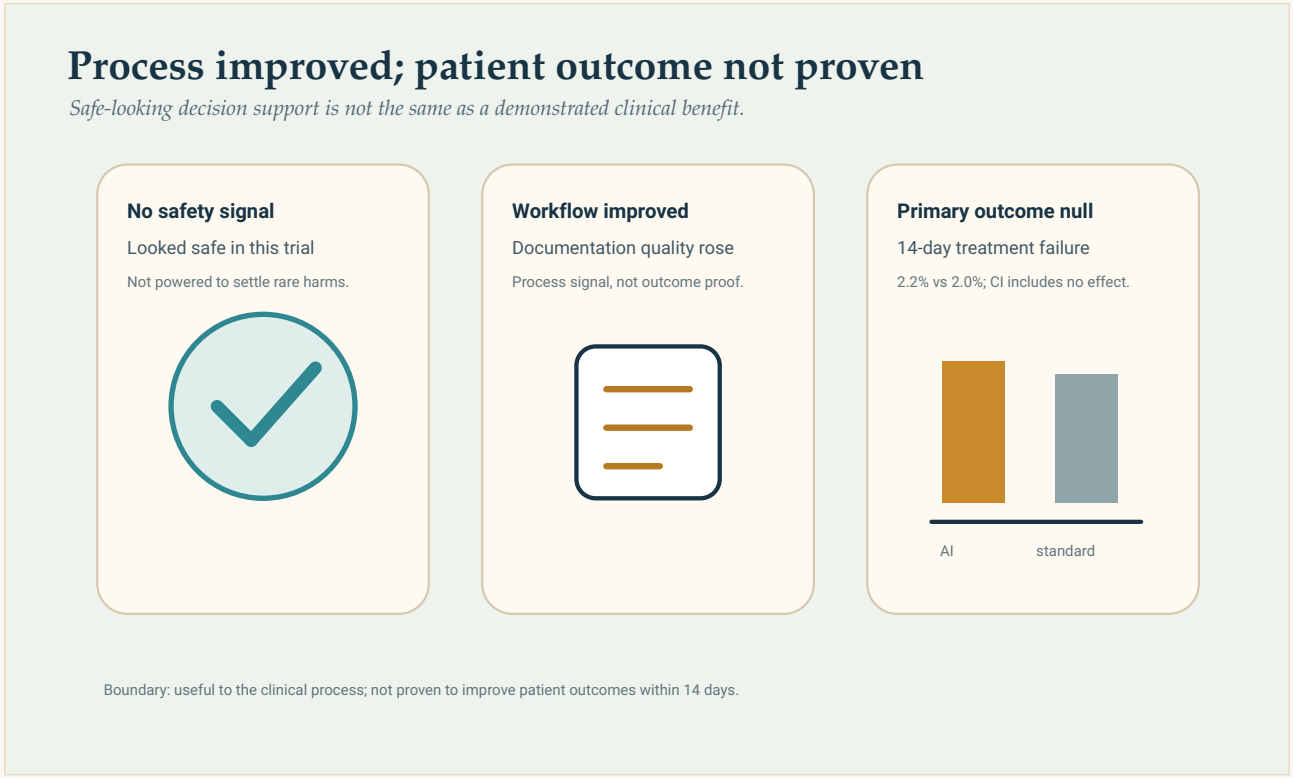
सुरक्षित और सहायक दिखना प्रभावी होने के बराबर नहीं है

चिकित्सकीय AI पर अधिकांश सुर्खियाँ गलत तरह के अध्ययनों से बनती हैं। कोई मॉडल परीक्षा के प्रश्नों में अच्छा स्कोर करता है, या चुने हुए नैदानिक उदाहरणों पर डॉक्टरों से बेहतर दिखता है, और कहानी खुद लिखी हुई लगती है: मशीन क्लिनिक के लिए तैयार है।

इस परीक्षण ने कुछ अधिक कठिन और कहीं अधिक दुर्लभ किया। इसने जनरेटिव-AI आधारित निर्णय-सहायता उपकरण को वास्तविक प्राथमिक देखभाल केंद्रों में, वास्तविक मरीजों के साथ इस्तेमाल किया और वही सवाल पूछा जो वास्तव में मायने रखता है: क्या मरीजों के नतीजे बेहतर हुए?

ईमानदार उत्तर है—नहीं, कम से कम 14 दिनों में मापने योग्य रूप से नहीं। उपकरण से कोई सुरक्षा संकेत नहीं मिला। उसने चिकित्सकीय दस्तावेज़ीकरण की गुणवत्ता बेहतर की। संभव है कि उसने कुछ दवा-जनित भी घटाई हों। लेकिन उपचार-विफलता में सांख्यिकीय रूप से महत्वपूर्ण कमी नहीं आई, और लेखक सावधानी से कहते हैं कि यदि कोई लाभ है भी, तो शायद वह छोटा है।

यह अध्ययन की विफलता नहीं है। यही अध्ययन का काम है। चिकित्सा में AI पर जिम्मेदार साक्ष्य ऐसे दिखते हैं जब उन्हें बेंचमार्क की बजाय मरीजों पर मापा जाता है।



उपकरण से कोई सुरक्षा संकेत नहीं मिला और चिकित्सकीय दस्तावेज़ीकरण बेहतर हुआ, लेकिन पहले से तय 14-दिन का मरीज़-परिणाम सांख्यिकीय रूप से बेहतर नहीं हुआ। प्रक्रिया में मदद और सिद्ध मरीज़-लाभ एक ही बात नहीं हैं।

लेखकोंने क्याकिया

टीम ने केन्याके नैरोबीऔर कआिम्बू काउंटियोंमें नजिीस्वास्थ्य नेटवर्क Penda स्वास्थ्य द्वाराचलाए जाने वाले **16 प्राथमकि देखभाल केंद्रों**में एक व्यावहारकि, क्लस्टर-रैंडमाइज्ड परीक्षण कयिा। इन केंद्रोंमें देखभाल मुख्यत:**क्लनिकिल officers** देते हैं—तीन वर्षीय डप्लिोमावाले मध्यम-स्तरीय चकित्सिकीय पेशेवर—जनि्हें अक्सर वरषि्ठ वशिषज्ञ से तुरंत सलाह उपलब्ध नहींहोती।

रैंडमाइज़ेशन कीइकाई मरीज़ नहीं, चकित्सिक था। **103 क्लनिकिल officers** कोरैंडमाइज़ कयिागया। 52 हस्तक्षेप समूह में और 51 नयिंत्रण समूह में। दोनोंसमूहोंने एक हीक्लाउड-आधारति इलेक्ट्रॉनकि मेडकिल रकिॉर्ड इस्तेमाल कयिा। हस्तक्षेप समूह कोइसके अतरिक्ति **“AI Consult”** (संस्करण 2.0) मलिा, जो**OpenAI के GPT-4o** बड़े भाषामॉडल पर बनानरि्णय-सहायताउपकरण थाऔर उसीरकिॉर्ड में जुड़ाहुआ था। यह चकित्सिक द्वारादर्ज जानकारीपढ़कर नदिान या उपचार योजनामें संभावति समस्याओं कीओर ध्यान दलिासकताथा। अंतमि नरि्णय पूरीतरह चकित्सिक के पास रहा; वे सुझाव स्वीकार, बदल याअनदेखाकर सकते थे।

कौन-सामॉडल था, और यह ववरिण क्योंमहत्वरखताहै

उपकरण **AI Consult 2.0** था, जो**OpenAI के GPT-4o** (मई 2025 रलिीज़) पर चलताथा। इसे OpenAI के वाणज्यिकि API से एंटरप्राइज़ लाइसेंस के तहत चलायागयाऔर randomness कम रखीगई (तापमान 0.1)। यह EasyClinic के खास इलेक्ट्रॉनकि मेडकिल रकिॉर्ड (EMR) के भीतर थाऔर इसके प्रणालीप्रॉम्प्ट केन्याकीराष्ट्रीय उपचार-नरि्देशकिाओं के अनुरूप लखि गए थे; लेखकोंने पूराinstruction प्रॉम्प्ट प्रकाशति कयिा।
यह सब स्पष्ट करनाक्योंजरूरीहै? क्योंकि परणिाम *एक खास प्रणाली*के बारे में है—एक मॉडल संस्करण, एक प्रॉम्प्ट, एक रकिॉर्ड और एक सेटगि—न कि सामान्य रूप से “चकित्सिामें LLMs” के बारे में। लेखक भीयहीबात कहते हैं: वे अपने नष्किर्ष कोक्षमताकास्थरि अनुमान नहीं, बल्कि *समय-वशिषि्ट बेंचमार्क* कहते हैं। नया मॉडल, अलग प्रॉम्प्ट याकम डजिटिाइज्ड क्लनिकि परणिाम बदल सकते हैं।
स्वतंत्रताके बारे में: बात में OpenAI ने वस्तुरूप सहायतादी—cloud-compute credits और API उपयोग पर तकनीकी मार्गदर्शन—लेकनि लेखकोंके अनुसार OpenAI कोचुनने कानरि्णय उस प्रस्ताव से *पहले* लयिाजाचुकाथा, और OpenAI कीपरीक्षण-डजिाइन, डेटासंग्रह, वशि्लेषण याप्रकाशन के नरि्णय में कोई भूमकिानहींथी।

22 अप्रैल से 16 जुलाई 2025 के बीच **9,691 मरीज़** शामलि कए गए। **प्राथमकि परणिाम जानबूझकर मरीज़-केंद्रति और कठोर रखागयाथा**: वजि़टि के 14 दनिोंके भीतर *उपचार-वफिलताकासंयुक्त परणिाम*, जसिे वशिषज्ञोंने तय कयिा। चकित्सिकोंके एक पैनल ने अध्ययन-समूह जाने बनियाह आंकाकि मरीज़ कोबीमारीन सुधरने याबगिड़ने जैसे प्रतकिूल परणिाम हुए यानहीं। परीक्षण पहले से पंजीकृत था(Pan-African क्लनिकिल परीक्षण Registry 202502499779176)।

यहीडजिाइन-चयन अहम है। यह दखिानाआसान है कि AI उपकरण चकित्सिक के लखिने के तरीके कोबदलताहै। यह दखिानाकहींकठनि—और कहींअधकि सार्थक—है कि वह मरीज़ के साथ होने वालीचीज़ कोबदलताहै।

उन्हें क्यामलिा

प्राथमकि परणिाम बेहतर नहींहुआ। AI समूह के 4,693 मरीज़ोंमें 102 (2.2%) और नयिंत्रण समूह के 4,654 मरीज़ोंमें 94 (2.0%) में उपचार-वफिलताहुई। कच्चे प्रतशित AI के साथ थोड़ाअधकि थे, लेकनि समायोजन के बाद बदि्-अनुमान लाभ कीदशिामें गया**समायोजति odds अनुपात 0.77 था(95% वशि्वास-अंतराल 0.55 से 1.08, P = 0.13)**—जोसांख्यकिीय रूप से महत्वपूर्ण नहींथा। कच्चे और समायोजति परणिामोंकादशिाबदलनागणनाकीगलतीनहींहै; समायोजन चकित्सिक-क्लस्टरों

के बीच अंतर को ध्यान में रखता है। कसी भी तरह विश्वास-अंतराल आरंभ से “कोई प्रभाव नहीं” को शामिल करता है, इसलिए लाभ का दावा नहीं किया जा सकता, और पूरा अंतर बहुत छोटा था।

Odds अनुपात, विश्वास-अंतराल और P मान को साथ पढ़ने का आसान तरीका क्लिनिकल परिणाम पढ़ने की मार्गदर्शिका में है।

सीमाओं के भीतर कोई सुरक्षा संकेत नहीं मिला। कसी गंभीर प्रतिकूल घटना को उपकरण से संबंधित नहीं माना गया, और स्वतंत्र समीक्षा में भी सुरक्षा संकेत नहीं मिला। लेखक इस भरोसे की सीमा साफ़ बताते हैं: अध्ययन दुर्लभ गंभीर हानियों को पकड़ने के लिए पर्याप्त शक्ति वाला नहीं था, और इसमें पहले से तय noninferiority या औपचारिक सुरक्षा ढाँचा नहीं था। इसलिए यह असामान्य घटनाओं के लिए सुरक्षा को *सिद्ध* नहीं कर सकता।

दस्तावेज़ीकरण बेहतर हुआ। 2,000 मुलाक़ातों की blinded विशेषज्ञ समीक्षा में AI Consult इस्तेमाल करने वाले चिकित्सकों के रिकॉर्ड हर आंके गए क्षेत्र में बेहतर थे—दर्ज निदान, उपचार योजना और कुल पूर्णता।

दवा लिखने के तरीके में बहुत कम बदलाव आया। Prescribing में कोई महत्वपूर्ण अंतर नहीं था, जिसमें सही antibiotic उपयोग भी शामिल था (समायोजित odds अनुपात 0.86, 95% CI 0.48 से 1.55)। उपकरण ने antibiotic prescribing दर नहीं बदली।

मरीज़ों ने कोई अंतर महसूस नहीं किया। संतुष्टि सर्वे पूरा करने वाले 826 मरीज़ों में दोनों समूहों की संतुष्टि लगभग समान थी और consultation समय भी लगभग समान रहा।

लागत में हल्की गिरावट का संकेत मिला। समायोजित विश्लेषण में AI समूह में antibiotics से जुड़ी लागत कम थी—संभवतः कम prescriptions की बजाय सस्ते विकल्प चुने जाने से—और प्रति मरीज़ antibiotic बचत उपकरण चलाने की प्रति-मरीज़ लागत से अधिक दिखी। लेखक इसे संकेत मानते हैं, स्थापित निष्कर्ष नहीं। कुल ownership लागत का पूरा आर्थिक लेखा परीक्षण के दायरे में नहीं था।

लेखकों का अपना एक-पंक्ति सार सबसे साफ़ है: इन सीमाओं के भीतर LLM सहायता सुरक्षित दिखी, लेकिन उसने 14 दिनों में उपचार-विफलता कम नहीं की, यदि लाभ है, तो शायद वह छोटा है।

“कोई महत्वपूर्ण अंतर नहीं” का अर्थ “यह काम नहीं करता” नहीं है

नकारात्मक प्राथमिक परिणाम को किसी भी दिशा में ज्यादा पढ़ना आसान है। दो बातें सरल कहानी को रोकती हैं।

पहली, **परीक्षण उस प्रभाव से बड़ा प्रभाव पकड़ने के लिए बनाया गया था जो वास्तव में दिखा।** प्राथमिक देखभाल में गंभीर खराब परिणाम दुर्लभ हैं—यहाँ लगभग 2%—इसलिए छोटे वास्तविक लाभ को शोर से अलग करने के लिए बहुत बड़े नमूने चाहिए। लेखकों की बताई शक्ति गणनाएँ बताती हैं कि उनके देखे गए आकार का प्रभाव पकड़ने के लिए **कहीं बड़ा परीक्षण, लगभग 100,000 से अधिक मरीज़ों के स्तर का** चाहिए होगा। इस आकार के अध्ययन में non-significant परिणाम छोटे, वास्तविक लाभ को खारिज नहीं करता। इसका अर्थ है कि यह अध्ययन उसे स्पष्ट नहीं कर सका।

दूसरी, **तुलना कुछ हद तक धुंधली हो गई थी।** configuration त्रुटि के कारण कुछ समय के लिए नियंत्रण समूह के कुछ चिकित्सकों को AI Consult मिल गया, और एक ही नेटवर्क के चिकित्सक आपस में बात करते हैं तथा आदतें समूह-सीमा पार ले जा सकते हैं। दोनों बातें समूहों को अधिक समान दिखाती हैं और वास्तविक अंतर को शून्य की ओर धकेल सकती हैं। इसके अलावा, मेज़बान नेटवर्क पहले से अपेक्षाकृत उच्च मानकों पर चलता था, इसलिए सुधार के लिए जगह कम थी।

इनमें से कोई बात “breakthrough” वाली सुर्खी को बचाती नहीं है। लेकिन सही पढ़ाई निराशावादी भी नहीं है: सबसे कठिन और सबसे ईमानदार एंडपॉइंट पर यह उपकरण दो हफ़्तों में मरीज़ों को मापने योग्य लाभ नहीं दे सका—जबकि उसने रिकॉर्ड-रखरखाव को

स्पष्ट रूप से बेहतर कयिाऔर सुरक्षति दखिा।

यह क्यासद्धि नहीकरता

- यह **नहीं**दखिताकि AI ने मरीज़ोंके परणिम बेहतर कए। प्राथमकि 14-दनि एंडपॉइंट पर महत्वपूर्ण लाभ नहींथा।
- यह **नहीं**दखिताकि AI बेकार है। बद्धि-अनुमान उसके पक्ष में था, दस्तावेज़ीकरण हर क्षेत्र में सुधराऔर दवा-लागत नीचे जाने कासंकेत मलिा; null परणिम छोटे वास्तवकि लाभ के साथ संगत है जसिे यह परीक्षण पुष्ट करने के लए बहुत छोटाथा।
- यह दुर्लभ हानयिोंके लए उपकरण कीसुरक्षासद्धि **नहीं**करता। कोई सुरक्षासंकेत नहींमलिा, लेकनि परीक्षण दुर्लभ गंभीर घटनाओं कीसुरक्षाप्रमाणति करने के लए डज़िाइन याpowered नहींथा।
- यह **नहीं**दखिताकि "AI डॉक्टरोंकोहरादेताहै" याचकित्सिक कोबदल देताहै। यह decision *समर्थन* था, सुझाव स्वीकार याअस्वीकार करने कापूराअधकिार चकित्सिक के पास था।
- यह अपने-आप समान्यीकृत **नहीं**होता। परीक्षण केन्याके एक नजिीशहरीनेटवर्क में हुआ; ग्रामीण, periurban याअधकि आय वाले परविश अलग होसकते हैं।
- यह लागत-बचत **स्थापति नहीं**करता। लागत कासंकेत आशाजनक है, पूर्ण आर्थकि मूल्यांकन नहीं।

साक्ष्य कतिनामजबूत है?

केंद्रीय दावा—कम समय में मरीज़-हानि घटने कासद्धि प्रमाण नहींमलिा—के लए अध्ययन-डज़िाइन मजबूत है और नष्किर्ष उचति रूप से संयमति है। भारी pre-registered, समूहति होनायादृच्छकिीकृत परीक्षण, जसिमें blinded और मरीज़-स्तरीय संयुक्त परणिम है, ऐसे उपकरण के लए वास्तवकि दुनयिाके सर्वोत्तम प्रकार के साक्ष्योंमें आताहै। यह बेंचमार्क स्कोर याकेस-ववरिण अध्ययन से कहींअधकि जानकारीदेताहै।

द्वतिीयक नष्किर्ष—बेहतर दस्तावेज़ीकरण, अपरविर्तति prescribing, समान संतुष्ट, कम antibiotic लागत—अच्छे साक्ष्य हैं, लेकनि उन्हें द्वतिीयक हीपढ़नाचाहए: सहायक संकेत, मुख्य सुर्खीनहीं, और वहीदूषण तथाsingle-network सीमाएँ इन पर भीलागू होतीहैं।

सुरक्षाके लए साक्ष्य आश्वस्त करने वालालेकनि सीमति है: कोई संकेत नहींमलिा, पर अध्ययन दुर्लभ हानयिाँखोजने के लए बनानहींथा।

सबसे उपयोगीरुख न "यह काम करताहै" है, न "यह वफिल हुआ"। सहीरुख है: सावधानीपूर्वक वास्तवकि दुनयिाके परीक्षण में इस AI उपकरण से कोई सुरक्षासंकेत नहींमलिाऔर देखभाल प्रक्रयिाबेहतर हुआ, लेकनि दोहफ्तोंमें मरीज़-परणिम कालाभ सद्धि नहींहुआ—और ऐसालाभ पकड़ने के लए कहींबड़ाअध्ययन चाहए।

यह क्योंमायने रखताहै

चकित्सिकीय AI कीबहस कोठीक इसीतरह के साक्ष्य कीकमीहै। हजारोंशोधपत्र हैं जनिमें मॉडल परीक्षामें अव्वल आते हैं यासुथरे क्लनिकिल मामले पर चकित्सिक से मेल खाते हैं। वास्तवकि मरीज़ोंकोफायदाहुआ यानहीं, यह मापने वाले बड़े pragmatic रैंडमाइज्ड परीक्षण बहुत कम हैं। यह उनमें से एक है, और इसकासावधाण-सासत्य महत्वपूर्ण है: परीक्षापास

करना मरीज़ की मदद करने के बराबर नहीं है।

यही अंतर पूरी कहानी है। कोई उपकरण चिकित्सक के लिए सच में उपयोगी हो सकता है—साफ़ notes, कम दवा-खर्च, एक अतिरिक्त जाँच—और फिर भी पंद्रह दिनों से भी कम समय में किसी कठोर मरीज़ परिणाम को न बदल पाए। दोनों बातें एक साथ सच हो सकती हैं, और परिपक्व स्वास्थ्य-व्यवस्थाओं को सुविधाजनक कहानी चुनने की बजाय दोनों को साथ रखना होगा।

यह प्रमाण का भार भी सही दिशा में रीसेट करता है। यदि कोई कंपनी दावा करती है कि उसकी क्लिनिकल AI देखभाल बेहतर करती है, तो प्रासंगिक साक्ष्य लीडरबोर्ड नहीं है। वह ऐसा परीक्षण है, उन परिणामों पर जो मरीज़ों के लिए मायने रखते हैं—और आदर्श रूप से इससे बड़ा क्योंकि यहाँ ईमानदार सबक यही है कि छोटे लाभ देखने के लिए बड़े अध्ययन चाहिए।

साफ़ सारांश

केन्या के 16 प्राथमिक देखभाल केंद्रों में एक pragmatic, समूह-यादृच्छिकीकृत परीक्षण ने क्लिनिकल officers द्वारा इस्तेमाल किए जाने वाले इलेक्ट्रॉनिक रिकॉर्ड में जोड़े गए generative-AI निर्णय-सहायता उपकरण (“AI Consult”) को परखा। 9,691 मरीज़ों में 14 दिनों के भीतर expert-adjudicated treatment-failure composite उपकरण के साथ 2.2% और बिना उपकरण 2.0% था (समायोजित odds अनुपात 0.77, 95% CI 0.55–1.08, P = 0.13)—कोई महत्वपूर्ण अंतर नहीं। उपकरण से कोई सुरक्षा संकेत नहीं मिला, हर आंके गए क्षेत्र में क्लिनिकल दस्तावेज़ीकरण बेहतर हुआ, prescribing नहीं बदली, मरीज़-संतुष्टि समान रही और antibiotic लगात कुछ कम होने से जुड़ी थी। लेखकों का निष्कर्ष है कि इन सीमाओं के भीतर उपकरण सुरक्षित दिखा लेकिन उपचार failure कम नहीं हुआ; यदि लाभ है तो शायद वह छोटा है, और देखे गए आकार का प्रभाव पकड़ने के लिए लगभग 100,000 मरीज़ों के स्तर का कहीं बड़ा परीक्षण चाहिए। परिणाम न यह दिखाता है कि क्लिनिकल AI मरीज़ों के नतीजे बेहतर करती है, न यह कि वह बेकार है—यह दिखाता है कि सुरक्षित दिखने और प्रक्रिया में मदद करने वाला उपकरण एक शहरी नैरोबी नेटवर्क में दो हफ़्तों में मरीज़ों को मापने योग्य लाभ नहीं दे सका।

बिना बढ़ा-चढ़ाकर जाँच

पेपर क्या दिखाता है: वास्तविक दुनिया के रैंडमाइज़्ड परीक्षण में प्राथमिक देखभाल रिकॉर्ड में LLM आधारित निर्णय-सहायता उपकरण जोड़ने से कोई सुरक्षा संकेत नहीं मिला, क्लिनिकल दस्तावेज़ीकरण बेहतर हुआ, prescribing नहीं बदली और 14-दिन के कठोर treatment-failure composite में महत्वपूर्ण कमी नहीं आई।

क्या संभव है लेकिन सिद्ध नहीं: उपकरण उपचार failure में इतना छोटा वास्तविक लाभ दे कि यह परीक्षण उसे पकड़ न सका, कुल लागत जोड़ने पर यह पैसा बचाए; बेहतर दस्तावेज़ीकरण आगे चलकर बेहतर देखभाल में बदले।

यह क्या नहीं दिखाता कि क्लिनिकल AI मरीज़ों के परिणाम बेहतर करती है; कि दुर्लभ घटनाओं के लिए यह असुरक्षित है; कि यह चिकित्सक को बदलती या उनसे बेहतर है; कि यही परिणाम ग्रामीण या अधिक आय वाले परिवेश में भी होंगे; कि लागत-बचत स्थापित है।

मुख्य सीमाएँ: अध्ययन उस प्रभाव से बड़ा प्रभाव पकड़ने के लिए powered था जो देखा गया (दुर्लभ परिणाम के लिए बहुत बड़े नमूने चाहिए); configuration त्रुटि से नियंत्रण arm में दूषण हुआ और सहायक नेटवर्क की आदतें तुलना को धुंधला करती हैं—दोनों अंतर को शून्य की ओर झुका सकते हैं; पहले से ऊँचे मानकों वाला एक हीन-जीशहरी नेटवर्क; पहले से तय noninferiority या सुरक्षा फ़्रेमवर्क नहीं, केवल 14-दिन का horizon।

सामान्य पाठक को कितना भरोसा रखना चाहिए? इस परीक्षण में उपकरण से कोई सुरक्षा संकेत नहीं मिला और दस्तावेज़ीकरण बेहतर हुआ—इस पर उच्च भरोसा। यहाँ 14 दिनों में मरीज़-परिणाम मापने योग्य रूप से बेहतर नहीं हुए—इस पर भी उच्च

भरोसा। लेकिन “यह काम करता है” या “यह विफल है” जैसे व्यापक patient-benefit दावों पर कम भरोसा—वह प्रश्न वास्तव में खुला है और कहीं बड़े अध्ययन की जरूरत है। उचित रुख: ऐसा सावधानीपूर्वक वास्तविक-दुनिया परिणाम जिसमें उपकरण सुरक्षित दिखा और देखभाल प्रक्रिया बेहतर हुआ, न कि यह फैसला कि AI प्रायः कि देखभाल को बदल देती है या बर्बाद कर देती है।

स्रोत

Agweyu et al., Nature Medicine (2026)
https://doi.org/10.1038/s41591-026-04503-6

Pan-African Clinical Trials Registry, registration 202502499779176
https://pactr.samrc.ac.za/

EurekAlert / PATH press release on the trial (2026)
https://www.eurekalert.org/news-releases/1133583