The Clean Paper

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

Կլինիկական AI, որը safety signal չցուցադրեց և փաստաթղթավորումը բարելավեց — բայց հիվանդային outcome-ները չբարելավեց

Քենիայի 16 առաջնային բուժօգնության հաստատություններում cluster-randomized trial-ը փորձարկել է clinical officers-ի էլեկտրոնային գրառումներում ներկառուցված գեներատիվ AI decision-support գործիք: 9,691 հիվանդի մոտ 14-օրյա treatment-failure composite-ը AI խմբում 2.2% էր, control-ում՝ 2.0% (adjusted odds ratio 0.77, 95% CI 0.55–1.08, P = 0.13)՝ նշանակալի տարբերություն չկար: Գործիքը safety signal չբարձրացրեց և documentation-ը բարելավեց, բայց երկու շաբաթում patient benefit չապացուցեց:

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

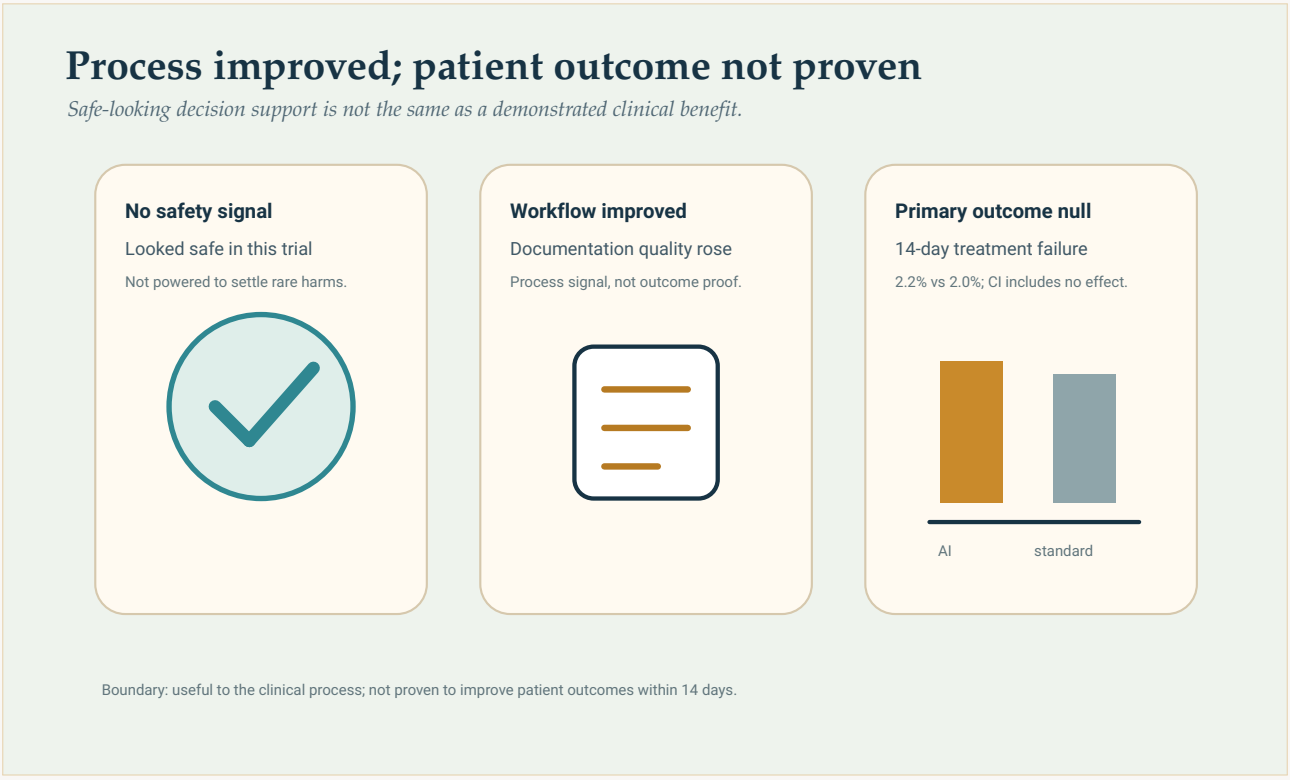
«Անվտանգ և օգտակար» նույնը չէ, ինչ «արդյունա

Բժշկական ԱԲի մասին վերնագրերի մեծ մասը կառուցվում է սխալ տեսակի ուսումնասիրությունների վրա: Մոդելը լավ է պատասխանում քննության հարցերին կամ գերազանցում է բժիշկներին ընտրված կլինիկական վինյետներում, և պատմությունն ինքն իրեն է գրում՝ մեքենան պատրաստ է կլինիկայի համար:

Այս փորձը շատ ավելի դժվար և հազվադեպ բան արեց: Գեներատիվ ԱԲի որոշումների աջակցման գործիքը տեղադրվեց իրական առաջնային բուժօգնության մեջ՝ իրական հաստատություններում, իրական հիվանդների հետ, և տրվեց այն հարցը, որն իրականում կարևոր է. հիվանդները ավելի լա՞վ արդյունք ունեցան:

Ազնիվ պատասխանը՝ ոչ: Գոնե ոչ չափելիորեն՝ 14 օրվա ընթացքում: Գործիքը անվտանգության ազդանշան չցուցադրեց: Այն բարելավեց կլինիկական փաստաթղթավորումը Հնարավոր է՝ նույնիսկ որոշ դեղերի ծախսը նվազեցրեց: Բայց բուժման ձախողումների թիվը վիճակագրորեն նշանակալիորեն չնվազեց, և հեղինակները զգուշորեն գրում են, որ եթե օգուտ կա, հավանաբար համեստ է:

Սա ուսումնասիրության ձախողում չէ: Սա հենց ուսումնասիրությունն է՝ աշխատող վիճակում: Այսպիսի տեսք ունի բժշկության մեջ ԱԲի պատասխանատու ապացույցը, երբ չափում են հիվանդներին, ոչ չափանիշային թեստները:



Գործիքը այս փորձի ընթացքում անվտանգության ազդանշան չցուցադրեց և բարելավեց կլինիկական փաստաթղթավորումը, բայց նախապես սահմանված 14-օրյա հիվանդային արդյունքը վիճակագրորեն նշանակալիորեն չբարելավվեց: Գործընթացային օգնությունը նույնը չէ, ինչ ապացուցված հիվանդային օգուտը:

Ի՞նչ արեցին հեղինակները

Թիմը իրականացրել է պրագմատիկ, կլաստերպատահականացված փորձ **16 առաջնային բուժօգնության հաստատություններում**, որոնք պատկանում են մասնավոր Penda Health ցանցին՝ Քենիայի Նայրոբի և Կիամբու շրջաններում: Այդ հաստատություններում բուժօգնությունը մեծ մասամբ տրամադրում են **կլինիկական officers**-ները՝ միջին մակարդակի բուժաշխատողներ, որոնք ունեն երեք տարվա դիպլոմ և հաճախ հեշտ հասանելիություն չունեն ավագ մասնագետի խորհրդատվությանը:

Պատահականացման միավորը բժիշկը/կլինիկական աշխատողն էր, ոչ հիվանդը: **103 կլինիկական officer** պատահականացվել է՝ 52-ը intervention խմբում, 51-ը վերահսկիչ խմբում: Երկու խմբերն էլ օգտագործում էին նույն ամպ-based էլեկտրոնային բժշկական գրառման համակարգը: Intervention խմբին լրացուցիչ հասանելի էր **ԱՔ խորհրդատվություն**-ը (տարբերակ 2.0)՝ **OpenAI GPT-4o** մեծ լեզվային մոդելի վրա կառուցված որոշումների աջակցման գործիք, որը ներկառուցված էր այդ գրառման համակարգում: Այն կարդում էր բժիշկի գրանցած տվյալները և կարող էր նշել հնարավոր խնդիրներ diagnosis-ի կամ բուժում plan-ի մեջ: Վերջնական որոշումը մնում էր բժիշկին. նա կարող էր ընդունել, փոխել կամ անտեսել առաջարկները:

Ո՞ր մոդելն էր և ինչո՞ւ են մանրամասները կարևոր

Գործիքը **ԱՔ խորհրդատվություն 2.0**-ն էր՝ **OpenAI GPT-4o**-ով (2025 թվականի մայիսյան թողարկում), որը հասանելի էր OpenAI-ի commercial API-ի միջոցով enterprise licence-ով և օգտագործվում էր ցածր պատահականություն պարամետրով (ջերմաստիճան 0.1): Այն աշխատում էր հատուկ կառուցված էլեկտրոնային գրառման համակարգում (EasyClinic’s EMR) և համակարգային հրահանգերով ուղղորդվում էր Քենիայի ազգային բուժման ուղեցույցներին համապատասխան: Հեղինակները հրապարակել են ամբողջ հրահանգ հրահանգը:
Ինչո՞ւ այսքան մանրամասն: Որովհետև արդյունքը վերաբերում է *մեկ կոնկրետ համակարգի*՝ մեկ մոդել տարբերակ, մեկ հրահանգ, մեկ գրանցում համակարգ և մեկ միջավայր: Այն «LLM-ները բժշկության մեջ» ընդհանուր դատավճիռ չէ: Հեղինակներն էլ նույնն են ասում՝ արդյունքը կոչելով *ժամանակային հենակետային փորձարկում, ոչ կարողության անփոփոխ գնահատական*: Ավելի նոր մոդելը, ուրիշ հրահանգը կամ ավելի քիչ թվայնացված կլինիկան կարող են փոխել արդյունքը:
Անկախության մասին. OpenAI-ն հետագայում տրամադրել է in-kind աջակցություն՝ ամպ-compute credits և API-ի տեխնիկական խորհրդատվություն, բայց հեղինակները նշում են, որ OpenAI-ն ընտրելու որոշումը կայացվել էր *մինչ այդ առաջարկը*, և ընկերությունը դեր չի

> ունեցել փորձարկման դիզայնի, տվյալներ collection-ի, վերլուծության կամ հրապարակում decision-ի մեջ:

2025 թվականի ապրիլի 22-ից հուլիսի 16-ը ընդգրկվել է **9,691 հիվանդ**: **հիմնական արդյունքը դիտավորյալ հիվանդակենտրոն և խիստ էր**՝ այցից հետո 14 օրվա ընթացքում expert-adjudicated *բուժում ձախողում composite*: Clinician-ների panel-ը, որը blinded էր ուսումնասիրություն թևի նկատմամբ, գնահատում էր՝ հիվանդն ունեցել է արդյոք վատ արդյունք, օրինակ՝ չլուծված կամ վատացած հիվանդություն: փորձարկումը նախապես գրանցվել էր Pan-African կլինիկական փորձարկումներ Registry-ում (202502499779176):

Դիզայնի հենց այս ընտրությունն է աշխատանքի կարևոր մասը: Հեշտ է ցույց տալ, որ ԱԲ գործիքը փոխում է այն, ինչ բժիշկը գրում է: Շատ ավելի դժվար և շատ ավելի նշանակալի է ցույց տալ, որ այն փոխում է այն, ինչ իրականում կատարվում է հիվանդի հետ:

Ի՞նչ գտան

հիմնական արդյունքը չբարելավվեց: Treatment ձախողում գրանցվել է ԱԲ խմբի 4,693 հիվանդից 102-ի մոտ (2.2%) և վերահսկիչ խմբի 4,654 հիվանդից 94-ի մոտ (2.0%): Հում տոկոսները փոքրինչ ավելի բարձր էին ԱԲի հետ, բայց adjustment-ից հետո կետային գնահատականը թեքվում էր օգուտի կողմը՝ **ճշգրտված շանսեր հարաբերակցություն 0.77 (95% վստահության միջակայք 0.55–1.08, P = 0.13)**: Սա վիճակագրորեն նշանակալի չէ: Հում և ճշգրտված թվերի ուղղության տարբերությունը հաշվարկային սխալ չէ. adjustment-ը հաշվի է առնում բժիշկ կլաստերների միջև տարբերությունները: Ամեն դեպքում վստահության միջակայքը հանգիստ ներառում է «ոչ մի ազդեցություն» տարբերակը, ուստի օգուտ պնդել չի կարելի, իսկ բացարձակ ազդեցությունը շատ փոքր էր:

շանսեր հարաբերակցությունները, վստահության միջակայքները և P արժեքները միասին կարդալու պարզ բացատրության համար տես կլինիկական արդյունքների ուղեցույցը:

Անվտանգության ազդանշան չկար՝ բայց սահմաններով: Ոչ մի լուրջ adverse դեպք գործիքի հետ պատճառահետևանքային կապով չգնահատվեց, իսկ անկախ գրախոսությունը անվտանգության ազդանշան չգտավ: Հեղինակները հստակ նշում են այդ հանգստացնող արդյունքի սահմանը. փորձարկումը նախատեսված չէր հազվադեպ, ծանր վնասները հայտնաբերելու համար, և չուներ նախապես սահմանված noninferiority կամ ֆորմալ անվտանգություն շրջանակ: Ուստի այն չի կարող *ապացուցել* հազվադեպ դեպքերի անվտանգությունը:

Փաստաթղթավորումը բարելավվեց: Blinded expert-ների կողմից վերանայված 2,000 encounter-ներում ԱԲ խորհրդատվություն օգտագործող բժիշկները բոլոր

գնահատված ոլորտներում ավելի լավ կլինիկական փաստաթղթավորում էին գրում՝ diagnosis-ի գրանցում, բուժում plan և ընդհանուր completeness։

նշանակումը գրեթե չփոխվեց։ նշանակումի մեջ, ներառյալ ճիշտ հակաբիոտիկ օգտագործելը, վիճակագրորեն նշանակալի տարբերություն չկար (ճշգրտված շանսեր հարաբերակցություն 0.86, 95% CI 0.48–1.55)։ Գործիքը հակաբիոտիկ նշանակում հաճախականությունը չփոխեց։

Հիվանդները տարբերություն չնկատեցին։ Satisfaction դիտարկում լրացրած 826 հիվանդների մոտ բավարարվածությունը երկու խմբերում գրեթե նույնն էր, և consultation ժամանակը նույնպես նման էր։

Ծախսերի ուղղությունը փոքրինչ նվազող էր։ ճշգրտված վերլուծությունում հակաբիոտիկկապված ծախսերը ԱԲ խմբում ավելի ցածր էր՝ հավանաբար ավելի էժան դեղերի ընտրության, ոչ ավելի քիչ նշանակումների պատճառով։ Մեկ հիվանդի հաշվով հակաբիոտիկ խնայողությունը թվում էր ավելի մեծ, քան գործիքի մեկ հիվանդի հաշվով գործարկման ծախսը։ Բայց հեղինակները սա ներկայացնում են որպես հուշող, ոչ հաստատված տնտեսական արդյունք. սեփականության ընդհանուր ծախսի ամբողջական հաշվարկ փորձարկման շրջանակից դուրս էր։

Հեղինակների մեկ նախադասությամբ ամփոփումն ամենամաքուրն է. այս սահմաններում LLM assistance-ը անվտանգության ազդանշան չի բարձրացրել, բայց 14 օրվա ընթացքում բուժում ձախողումը չի նվազեցրել, և եթե օգուտ կա, հավանաբար համեստ է։

Ինչո՞ւ «վիճակագրորեն նշանակալի տարբերությ չկա» չի նշանակում «չի աշխատում»

Null հիմնական արդյունքը հեշտ է չափազանց մեկնաբանել երկու ուղղությամբ էլ։ Երկու բան խանգարում է պարզ պատմությանը։

Առաջին՝ **փորձարկումը նախագծված էր ավելի մեծ ազդեցություն հայտնաբերելու համար, քան ստացվեց**։ Առաջնային բուժօգնությունում ծանր վատ արդյունքները հազվադեպ են՝ այստեղ մոտ 2%։ Փոքր իրական օգուտը աղմուկից տարբերելու համար շատ մեծ նմուշ է պետք։ Հեղինակների post-hoc վիճակագրական հզորություն հաշվարկը ցույց է տալիս, որ դիտարկվել է չափի ազդեցություն հայտնաբերելու համար անհրաժեշտ կլիներ **շատ ավելի մեծ փորձարկում՝ մոտավորապես 100,000-ից ավելի հիվանդ**։ Այս չափի ուսումնասիրության non-significant արդյունքը չի բացառում փոքր իրական օգուտը. այն նշանակում է, որ տվյալ նմուշը այն վստահորեն տարբերակել չի կարող։

Երկրորդ՝ **համեմատությունը մասամբ աղտոտվել էր**։ Configuration սխալի պատճառով որոշ վերահսկիչթև բժիշկներ կարճ ժամանակով հասանելիություն են ունեցել ԱԲ խորհրդատվությունին, իսկ նույն ցանցում աշխատող մարդիկ իրար հետ

խոսում և սովորություններ են փոխանցում խմբերի միջև: Երկուսն էլ intervention և վերահսկիչ խմբերը դարձնում են ավելի նման՝ ցանկացած իրական տարբերություն մղելով դեպի զրո: Բացի այդ, հյուրընկալող ցանցը արդեն համեմատաբար բարձր ստանդարտներով էր աշխատում, ինչը գործիքի համար բարելավման քիչ տեղ է թողնում:

Սրանցից ոչ մեկը չի փրկում «բեկում» վերնագիրը: Բայց ճիշտ ընթերցումը նաև «ԱԲն ձախողվեց» չէ: Ամենախիստ և ազնիվ վերջնակետում այս գործիքը երկու շաբաթվա ընթացքում հիվանդներին ցուցադրելի օգուտ չտվեց, մինչդեռ այն measurable կերպով բարելավեց փաստաթղթավորումը և անվտանգության ազդանշան չցուցադրեց:

Ի՞նչ սա *չի* ապացուցում

- Սա **չի ցույց տալիս**, որ ԱԲն բարելավել է հիվանդային արդյունքները: հիմնական 14-օր վերջնակետում նշանակալի օգուտ չկար:
- Սա **չի ցույց տալիս**, որ ԱԲն անօգուտ է: կետ գնահատելը հակված էր օգուտի կողմը, փաստաթղթավորումը բարելավվեց, դեղ ծախսը նվազման ուղղություն ուներ, իսկ null արդյունքը համատեղելի է փոքր իրական օգուտի հետ, որը փորձարկումը պարզապես բավական մեծ չէր հաստատելու համար:
- Սա **չի ապացուցում հազվադեպ վնասների անվտանգությունը**: անվտանգություն ազդանշան չկար, բայց փորձարկումը նախատեսված չէր հազվադեպ severe դեպքերի անվտանգություն certification-ի համար:
- Սա **չի ցույց տալիս**, որ «ԱԲն գերազանցում է բժիշկներին» կամ փոխարինում բժիշկին: Սա decision *support* է, և բժիշկը պահպանում էր ընդունելու կամ մերժելու ամբողջ իրավասությունը:
- Արդյունքը **ինքնաբերաբար չի generalize** այլ միջավայրերի վրա: փորձարկումը կատարվել է Քենիայի մեկ մասնավոր քաղաքային ցանցում: Rural, periurban կամ բարձր եկամուտ ունեցող միջավայրերում արդյունքը կարող է տարբեր լինել:
- Սա **չի հաստատում ծախսերի խնայողություն**: Ծախսերի նվազման ազդանշանը հուշող է, բայց ամբողջական տնտեսական գնահատում չէ:

Որքա՞ն ուժեղ է ապացույցը

Կենտրոնական պնդման՝ short-term հիվանդ harm-ի ապացուցված նվազեցում չգտնելու համար ապացույցները դիզայնի առումով ուժեղ է, իսկ եզրակացությունի առումով ճիշտ չափով համեստ: Prospective, pre-registered, կլաստերպատահականացված փորձարկումը blinded հիվանդմակարդակ composite արդյունքով մոտ է այսպիսի գործիքի համար հնարավոր լավագույն իրական աշխարհի ապացույցին: Այն շատ ավելի ինֆորմատիվ է, քան չափանիշային թեստ գնահատականը կամ vignette ուսումնասիրությունը:

Երկրորդային արդյունքները՝ ավելի լավ փաստաթղթավորում, դեղատոմսերի նշանակման էական փոփոխության բացակայություն, նույն գոհունակությունը և հակաբիոտիկների ավելի ցածր ծախսը, նույնպես օգտակար են, բայց պետք է հենց որպես երկրորդային արդյունքներ կարդալ։ Դրանք աջակցող ազդանշաններ են, ոչ գլխավոր եզրակացություն, և ենթարկվում են նույն խմբերի խառնման ու մեկ ցանցում կատարված փորձարկման սահմանափակումներին։

անվտանգության համար ապացույցը հանգստացնող է, բայց սահմանափակ. ազդանշան չի գտնվել, բայց սա հազվադեպ harm-եր գտնելու համար կառուցված փորձարկում չէր։

Ամենաօգտակար դիրքորոշումը ոչ «աշխատում է», ոչ «ձախողվեց» է։ Ավելի ճիշտ է ասել. զգույշ իրական աշխարհի փորձարկումը ցույց տվեց, որ այս ԱԲ գործիքը անվտանգության ազդանշան չի բարձրացրել և բարելավել է բուժօգնություն գործընթացը, բայց երկու շաբաթվա ընթացքում չի ապացուցել հիվանդարդյունք benefit, իսկ նման փոքր օգուտ հայտնաբերելու համար շատ ավելի մեծ ուսումնասիրություն է պետք։

Ինչո՞ւ է սա կարևոր

Բժշկական ԱԲի քննարկման մեջ հենց այսպիսի ապացույցները շատ քիչ է։ Կան հազարավոր աշխատանքներ, որտեղ մոդելները լավ են հանձնում քննություններ կամ համեմատվում բժիշկների հետ կոկիկ ընտրված դեպքերում։ Շատ քիչ են մեծ, պրագմատիկ, պատահականացված փորձարկումները, որոնք չափում են՝ իրական հիվանդները ավելի լա՞վ արդյունք ունեն։ Սա դրանցից մեկն է, և այն հանգում է ոչ փայլուն, բայց կարևոր ճշմարտությանը. թեստը լավ հանձնելը նույնը չէ, ինչ հիվանդին օգնելը։

Այս բացն ամբողջ պատմությունն է։ Գործիքը կարող է իրականում օգտակար լինել բժիշկին՝ ավելի պարզ գրառումներ, փոքր դեղ bills, երկրորդ «զույգ աչք», և միաժամանակ երկու շաբաթում hard հիվանդ արդյունք չփոխել։ Երկու փաստերն էլ կարող են միաժամանակ ճիշտ լինել, և հասուն առողջապահական համակարգը պետք է երկուսն էլ պահի, ոչ ընտրի միայն հարմար մեկը։

Այն նաև օգտակար ուղղությամբ փոխում է ապացույցի բեռը։ Եթե ընկերությունը ցանկանում է պնդել, որ իր կլինիկական ԱԲն բարելավում է բուժօգնությունը, համապատասխան ապացույցները leaderboard-ը չէ։ Պետք է այսպիսի փորձարկում՝ հիվանդների համար կարևոր արդյունքներով, և իդեալական դեպքում ավելի մեծ, որովհետև այստեղի ազնիվ դասը այն է, որ համեստ օգուտները տեսնելու համար մեծ ուսումնասիրություններ են պետք։

Կարճ ամփոփում

Քենիայի 16 առաջնային բուժօգնության հաստատություններում պրագմատիկ, կլաստերպատահականացված փորձարկումը փորձարկել է գեներատիվ ԱԲ որոշումների աջակցման գործիք՝ «ԱԲ խորհրդատվություն», որը ավելացվել էր կլինիկական officers-ի օգտագործած էլեկտրոնային գրառման համակարգին: 9,691 հիվանդների մեջ 14 օրվա ընթացքում expert-adjudicated բուժումնախողում composite-ը ԱԲ խմբում 2.2% էր, վերահսկիչում՝ 2.0% (ճշգրտված շանսեր հարաբերակցություն 0.77, 95% CI 0.55–1.08, P = 0.13)՝ վիճակագրորեն նշանակալի տարբերություն չկար: Գործիքը անվտանգության ազդանշան չցուցադրեց, բարելավեց կլինիկական փաստաթղթավորումը, նշանակումը և հիվանդ satisfaction-ը գրեթե չփոխեց և կապված էր որոշակի ավելի ցածր հակաբիոտիկ ծախսի հետ: Հեղինակների եզրակացությունը՝ այս սահմաններում գործիքը անվտանգության ազդանշան չի բարձրացրել, բայց բուժում ձախողումը չի նվազեցրել, իսկ ցանկացած օգուտ հավանաբար համեստ է: Observed ազդեցությունը վստահորեն տեսնելու համար կարող էր պահանջվել մոտ 100,000-ից ավելի հիվանդի փորձարկում: Արդյունքը չի ցույց տալիս, որ կլինիկական ԱԲն հիվանդային արդյունքներ է բարելավում, բայց նաև չի ցույց տալիս, որ այն անօգուտ է. այն ցույց է տալիս, որ բուժօգնություն գործընթացին օգնող և այս փորձարկման ընթացքում անվտանգության ազդանշան չտված գործիքը երկու շաբաթվա ընթացքում ցուցադրելի հիվանդ benefit չի տվել:

Առանց չափազանցության

Ի՞նչ է ցույց տալիս աշխատանքը: Իրական պատահականացված փորձարկման ընթացքում LLM-based որոշումների աջակցման գործիքը անվտանգության ազդանշան չի բարձրացրել, բարելավել է փաստաթղթավորումի որակը, նշանակումը չի փոխել և 14-օրյա strict հիվանդ բուժումնախողում composite-ը վիճակագրորեն նշանակալիորեն չի նվազեցրել:

Ի՞նչն է հավանական, բայց ապացուցված չէ: Որ գործիքը տալիս է փոքր իրական բուժումնախողում նվազում, որը փորձարկումը չափազանց փոքր էր հայտնաբերելու համար, որ ամբողջական ծախս accounting-ի դեպքում գումար է խնայում, կամ որ լավ փաստաթղթավորումը ժամանակի ընթացքում վերածվում է ավելի լավ բուժօգնությունի:

Ի՞նչ չի ցույց տալիս: Որ կլինիկական ԱԲն բարելավում է հիվանդ արդյունքներ, որ հազվադեպ դեպքների համար այն ապացուցված անվտանգ է, որ փոխարինում կամ գերազանցում է բժիշկներին, որ արդյունքը նույնն է rural կամ բարձր եկամուտ ունեցող միջավայրերում, կամ որ ծախս խնայողությունը հաստատված է:

Գլխավոր սահմանափակումները: փորձարկումը powered էր դիտարկվել էից ավելի մեծ ազդեցության համար, իսկ հազվադեպ արդյունքները մեծ նմուշ են պահանջում: Configuration սխալը որոշ վերահսկիչ բժիշկների ԱԲ հասանելիություն

տվեց, իսկ shared-ցանց սովորությունները նույնպես կարող էր խմբերը մոտեցնել: Ուսումնասիրությունը մեկ մասնավոր քաղաքային ցանցում էր՝ արդեն բարձր ստանդարտներով, չուներ նախապես սահմանված noninferiority կամ ֆորմալ անվտանգություն շրջանակ և հետևել է միայն 14 օր:

Որքա՞ն վստահություն պետք է ունենա ընդհանուր ընթերցողը: Բարձր՝ որ այս փորձարկման ընթացքում գործիքը անվտանգության ազդանշան չի բարձրացրել և փաստաթղթավորումը բարելավել է: Բարձր՝ որ այստեղ 14-օրյա հիվանդ արդյունքը ցուցադրելիորեն չի բարելավել: Ցածր՝ «աշխատում է» կամ «ձախողվել է» վերջնական դատավճռի նկատմամբ: Այդ հարցը դեռ բաց է և պահանջում է շատ ավելի մեծ փորձարկում: Ճիշտ դիրքորոշումը՝ զգույշ իրական աշխարհի արդյունք գործիքի մասին, որը բուժօգնություն գործընթացին օգնել է, ոչ verdict այն մասին, թե ԱԲն առաջնային բուժօգնությունը փրկում կամ փչացնում է:

Աղբյուրներ

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