The Clean Paper

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

AI agent ya gudanar da entire patient cases da kansa — a simulator, on past records

MIRA new kind of medical AI ne: maimakon answering single question, yana work entire case inside simulated hospital record — taking history, ordering da reading tests, reaching diagnosis, writing orders. On 574 retrospective cases from public database across eight pre-selected diagnoses, authors report it outperformed physicians on diagnostic accuracy and made largely guideline-concordant, medication-safe decisions. Every qualifier matters: sandbox on past records, text only; much edge on clear-cut test conditions; ordered about twice as many blood tests as doctors; several outcomes scored against original chart. Genuine advance agent that acts across whole workflow — not proof machine diagnoses better than doctor. Authors say generalization, safety and governance still need prospective real-world studies.

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

Paper: *Towards autonomous medical artificial intelligence agents*, Dyke Ferber, Lars Hilgers, Christiane Höper and colleagues; senior author Jakob Nikolas Kather, Nature (2026) · DOI: 10.1038/s41586-026-10675-5

Amsa tambaya ba daidai yake da tafiyar da dukan shari'ar mara lafiya ba

Yawancin labaran da ake yi game da basirar wucin-gadi a fannin lafiya suna magana ne kan samfurin da yake *amsa*. A ba shi taƙaitaccen bayanin mara lafiya ko tambayar jarrabawa, sai ya dawo da ganewar cuta ko sakin layi na shawara, kuma ya yi kyau a gwaji. Wannan yana da amfani, amma iyakarsa ta yi ƙunci: kamar ƙwararren mai ba da shawara ne wanda bai taɓa shiga kundin mara lafiya ya yi wani aiki ba.

MIRA an gina ta ne don yin wani abu dabam, kuma wannan bambancin shi ne ainihin labarin. Maimakon ta amsa tambaya guda, tana tafiyar da shari'ar mara lafiya gaba ɗaya: tana karanta kundin mara lafiya, tana tantance wane ƙarin tarihin lafiya take buƙata, tana umartar gwaje-gwajen dakin gwaje-gwaje, hoton jiki da ƙwayoyin cuta, tana karanta sakamakonsu, tana takaita jerin yiwuwar cututtuka, sannan tana rubuta umarnin da zai biyo baya — magunguna, kwantar da mara lafiya a asibiti, ko tura shi tiyata. Tana yin wannan ne ta hanyar aiwatar da matakai *a cikin* kundin lafiya na lantarki, kamar yadda likita yake yi, maimakon kawai ta samar da rubutu da mutum zai sake shigarwa.

Wannan babban mataki ne: daga tsarin da yake ba da shawara zuwa tsarin da yake aiwatar da matakai a dukkan aikin kula da mara lafiya. Amma irin wannan mataki ne kuma kan rikide a kanun labarai zuwa "AI ta fi likitoci." Saboda haka yana da muhimmanci a fayyace abin da aka gwada da kuma yanayin da aka gwada shi.

A taƙaice: MIRA ta tafiyar da shari'o'in mara lafiya gaba ɗaya da kanta kuma, a wannan gwajin, ta fi likitoci daidai wajen gano cuta — amma ta yi hakan ne a cikin muhalli na gwaji da aka killace, kan kundin marasa lafiya na baya, a cikin cututtuka takwas da aka zaɓa tun da farko, kuma sadarwarta rubutu kaɗai ce. Haka kuma, mafi yawan fifikon da ta samu ya fito ne daga cututtukan da ke da sakamakon gwaji mai bayyana sosai. Ci gaban gaskiya ne. Amma jadawalin maki ba asibiti ba ne.

MIRA showed a workflow capability, not a clinic verdict

A sandboxed EHR lets an agent act through a whole retrospective case. It is still not a real patient trial.

DEMONSTRATED

- end-to-end case work in a sandboxed EHR
- history, tests, differential diagnosis, orders
- 574 retrospective MIMIC-IV cases
- 8 pre-selected diagnoses
- higher diagnostic accuracy on this benchmark

NOT DEMONSTRATED

- real patients or live hospital deployment
- conditions beyond the eight-target set
- an equally informed head-to-head comparison
- freedom from public-data contamination
- safety and governance for clinical use

A capability shown in a sandbox -- not a diagnostician proven in a clinic.

The figure separates the workflow result from the deployment claim.

MIRA ta nuna cewa za ta iya tafiyar da aikin kula da mara lafiya daga farko zuwa ƙarshe a cikin EHR na gwaji. Ba ta nuna cewa an tura ta aiki a asibiti na gaske ba, ko cewa tana iya gano kowane irin cuta, ko cewa an yi mata gwaji da likitoci a yanayi da bayanai iri ɗaya gaba ɗaya.

Original Aurora diagram — The Clean Paper CC BY 4.0

Abin da marubutan suka yi

MIRA — *Medical Intelligence for Reasoning and Action* — wata wakiliya ce mai cin gashin kanta da aka gina kan samfuran OpenAI. Sashen da yake tattaunawa da aiwatar da matakai yana amfani da **GPT-4o**, yayin da wani matakin tsara aiki dabam yake amfani da samfurin tunani na OpenAI **o1**. An saka waɗannan a cikin tsarin da aka ƙera musamman kuma ya bi ma'aunin HL7 FHIR — tsarin bayanai da asibitoci na gaske ke amfani da shi wajen musayar kundin lafiya — tare da kayan aiki fiye da dubu tamanin da MIRA za ta iya amfani da su. A cikin wannan muhalli na gwaji, MIRA za ta iya karɓo tarihin mara lafiya, umartar gwaje-gwaje da fassara sakamakonsu, ƙirƙirar jerin yiwuwar cututtuka, da tsara magani — ciki har da rubuta magunguna, tsara tiyata da shirya kwantar da mara lafiya a asibiti. Wani bayani da ya dace a lura da shi tun farko shi ne cewa masu tantance amsoshin MIRA ta atomatik su ma GPT-4o ne — daga dangin samfurin da ake gwadawa. Bayan wani mai bita ya nuna damuwa kan wannan, marubutan sun ƙara bitar wani likita mai cikakkiyar takardar ƙwarewa.

An samo gwajin daga **MIMIC-IV**, babban ma'ajin bayanan EHR na jama'a da aka cire bayanan da za su iya gane mutane. Daga kusan marasa lafiya 300,000 da aka kula da su a Beth Israel Deaconess Medical Center tsakanin 2008 da 2019, marubutan sun zaɓi **shari'o'i 574** da suka shafi **cututtuka takwas** — ciki har da appendicitis, pancreatitis, pneumonia da kamuwa da cutar mafitsara. A kowace shari'a, MIRA ta fara da bayanai kaɗan sannan ta yanke shawara mataki-mataki kan abin da za ta yi na gaba. Wannan yana kama da yadda ake binciken mara lafiya a zahiri, sai dai ana gudanar da shi ne kan kundin da aka riga aka san ƙarshen abin da ya faru.

Daga nan aka kwatanta aikinta da na likitoci a kan shari'o'in iri ɗaya.

Menene "wakili mai cin gashin kansa a cikin EHR na gwaji" a zahiri?

Kalmomi uku a nan suna ɗauke da ma'ana mai yawa, saboda haka ya dace a warware su.

Wakili yana nufin tsarin ba kawai chatbot ba ne da ke amsa umarni. Yana aiki a zagaye: ya duba halin da ake ciki, ya zaɓi mataki — misali ya umarci wani gwaji ko ya nemi ƙarin tarihin lafiya — ya karɓi sakamakon, sannan ya zaɓi mataki na gaba har sai ya kai ga ganewar cuta da tsarin magani. "Basira" tana fitowa daga samfurin harshe; ikon aiwatar da matakai yana fitowa daga tsarin da ke kewaye da shi, wanda yake juya rubutun samfurin zuwa ayyukan EHR da aka ba shi izini sannan ya dawo masa da sakamakon.

EHR na gwaji da aka killace yana nufin kwafin tsarin kundin lafiya da aka yi kwaikwayo, ba tsarin asibiti da ke aiki da marasa lafiya na gaske ba. MIRA za ta iya "umarci" gwaji ta karɓi sakamakon da mara lafiyar ya samu a zahiri, domin shari'ar ta tarihi ce kuma sakamakon yana cikin kundin. Babu wani abu da MIRA ke yi da ya shafi mara lafiya ko asibiti na gaske.

Mai cin gashin kanta yana nufin tana kammala shari'ar daga farko zuwa ƙarshe ba tare da mutum ya shiga tsakani ba — *a cikin wannan muhalli na gwaji*. Ba yana nufin an sake ta ta kula da marasa lafiya ba tare da sa ido ba. Waɗannan da'awoyi biyu ne masu bambanci sosai, kuma na farko kaɗai aka gwada.

Akwai ƙarin muhimmin bayani game da wannan muhalli. MIRA za ta iya *umarci* duk gwajin da take so,

> amma tsarin zai ba ta sakamako ne kawai idan an yi wannan gwajin ga mara lafiyar a rayuwa ta gaske. Idan ta nemi gwajin jini da aka yi wa mara lafiyar, za ta samu ainihin ƙimomin. Idan ta nemi hoton da ba a taɓa yi ba, za ta samu "N/A — ba a samu damar yin gwajin ba," ba wani adadi da aka ƙirƙira ba. Tarihin lafiya ma haka yake: idan kundin bai ƙunshi abin da MIRA ta tambaya ba, mara lafiyar kwaikwayo zai ce bai sani ba. Don haka MIRA ba za ta iya kirawo gwaji mai yanke hukunci wanda ba a taɓa yi ba; tana aiki ne da abin da binciken mara lafiyar na gaske ya ƙunsa.
> Wannan bambancin yana da muhimmanci domin kalmar da ta fi jan hankali — "mai cin gashin kanta" — ita ce mafi sauƙin fahimta fiye da abin da aka gwada.

Abin da suka gano

A wannan gwajin, **MIRA ta fi likitoci daidai wajen gano cuta**. Amma adadin da ake yawan ɗauka a matsayin kanun labari ya ƙunshi lambobi uku dabam, don haka ya dace a raba su.

Da kanta, a duk **shari'o'i 574**, MIRA ta faɗi ganewar cuta daidai a **88.9%** na lokuta, idan aka kwatanta da ganewar cutar lokacin sallama da aka rubuta a MIMIC-IV. A gwajin kai-tsaye da mutane, likitocin ba su yi duk shari'o'i 574 ba; sun yi wani rukuni na **shari'o'i 311** da MIRA ma ta yi, da kundin bayanai da kayan aiki iri ɗaya. A waɗannan shari'o'i 311, MIRA ta samu **87.8%**. An kwatanta ta da ƙungiyoyin likitoci biyu. Na farko, **manyan likitoci** huɗu masu cikakkiyar takardar ƙwarewa da shekaru 7 zuwa 11 na gogewa, sun samu **78.1%**. Na biyu, **ƙungiya mai matakan ƙwarewa daban-daban**, galibinsu likitocin zama tare da ƙwararru biyu, sun samu **71.1%**. Shari'o'i iri ɗaya, kayan aiki iri ɗaya: MIRA ta zo ta farko, manyan likitoci na biyu, ƙungiyar da ta fi cike da ƙananan likitoci ta uku. Takardar ta yi amfani da lafazi mai taka-tsantsan: MIRA ta kasance "daidai da likitoci akai-akai, kuma sau da yawa ta zarce su" a cututtukan takwas.

Shawarwarin da suka biyo bayan ganewar cuta ma sun yi kyau: galibinsu sun dace da ka'idojin jinya, magungunan sun kasance masu aminci, kuma shawarar kwantar da mara lafiya ta dace. Marubutan ba su ga kuskuren magani mai tsananin haɗari a rukuni biyar na aminci ba — hulɗar magunguna, daidaita kashi ga aikin koda, rashin jituwa, haɗarin QT da rubuta opioids — kuma magunguna 467 cikin 468 sun dace. Duk da haka, su kansu sun jaddada cewa tsarin "bai kai aminci 100% ba." Idan aka ɗauki sakamakon yadda yake, abin lura ne: wakili ya tafiyar da shari'ar gaba ɗaya sannan ya fi likitoci a daidaiton ganewar cuta.

Amma *inda* MIRA ta samu wannan fifiko yana da muhimmanci kamar fifikon kansa. Masana masu zaman kansu da suka karanta takardar, ciki har da waɗanda suka yi sharhi a Science Media Centre, sun lura cewa mafi yawan bambancin ya fito ne daga cututtukan da ke da **gwajin da yake yanke hukunci sosai**, kamar appendicitis da pancreatitis, inda hoton jiki ko wani gwajin dakin gwaje-gwaje zai iya warware tambayar. A pneumonia da kamuwa da cutar mafitsara — biyu daga cikin dalilan da mutane ke yawan zuwa sashen gaggawa — MIRA da likitoci duk sun fi yin rauni, kuma tazarar da ke tsakaninsu ta fi ƙanƙanta.

Haka kuma bangarorin biyu ba su yi amfani da bayanai daidai gwargwado ba. MIRA ta fi dogaro da gwajin dakin gwaje-gwaje: ta yi amfani da kusan **51%** na ma'aunan jini da ke samuwa a kulawar yau da kullum, yayin da manyan likitoci suka yi amfani da kusan **28%** — matsakaicin ƙarin ma'aunan jini bakwai a kowace shari'a. Marubutan sun nuna cewa wannan har yanzu ya yi ƙasa da abin da ke cikin kundin kulawar yau da kullum, kuma ba su ga ƙaruwa a hotunan jiki masu tsada ba. Duk da haka, ƙarin bayanai kansu na iya ƙara daidaiton ganewar cuta. Saboda haka

ba cikakken gwaji ne tsakanin bangarori biyu da suka samu bayanai iri ɗaya ba. Likita da aka ba shi damar umartar kowane gwajin jini mai arha ba tare da tunanin kuɗi, rashin jin daɗin mara lafiya ko jinkiri ba shi ma zai iya yin kyau a takarda.

Me ya sa "ta doke likitoci" ba yana nufin "ta fi likita kyau" ba

Samun maki mafi girma a gwaji yana da sauƙin wuce gona da iri wajen fassara shi. Akwai abubuwa uku tsakanin "MIRA ta samu maki mafi girma" da "MIRA ta fi likita."

Na farko, **abin da aka yi amfani da shi a matsayin ma'aunin gaskiya**. Farfesa Julie Jacko ta lura cewa wasu muhimman sakamakon MIRA an ayyana su ne bisa abin da aka rubuta a ma'ajin bayanan asali. Wato tsarin yana samun lada idan ya **maimaita abin da aka rubuta cewa likitoci sun yi**, ba lallai idan ya nuna mafi kyawun kulawa da za a iya yi ba. Kundin tarihi ya zama kamar amsar jarrabawa. Wannan hanya ce mai ma'ana ta gina gwaji, amma tana auna daidaituwa da abin da aka yi ne, ba cikakkiyar ingancin kulawa ba. Wannan matsalar ta fi shafar ma'aunan daidaiton magani — ko umarnin MIRA ya dace da abin da aka rubuta a kundin — yayin da daidaiton ganewar cuta aka auna shi da ganewar lokacin sallama, wanda ya fi kusa da sakamako na gaske.

Na biyu, **bambancin yawan bayanan da aka samu**: idan MIRA ta yi amfani da gwajin jini da yawa — 51% maimakon 28% — tana da ƙarin shaida. Wani ɓangare na tazarar da ta samu na iya fitowa daga samun ƙarin bayanai, ba daga tunani mafi kyau kaɗai ba.

Na uku, **inda aka gudanar da gwajin**. Muhalli ne na kwaikwayo, kan shari'o'in baya, a cututtuka takwas da aka riga aka zaɓa, kuma ta rubutu kaɗai. A asibiti, ganewar mara lafiya ba ta dogara da abin da ya faɗa kaɗai ba; yadda ya faɗa, gwajin jiki, halayen da likita ya gani da sauran bayanai ma suna da muhimmanci. Wakilin da yake karanta rubutu ba ya ganin waɗannan. Kuma mara lafiya da matsalarsa ba ta cikin cututtukan takwas gaba ɗaya bai shiga wannan binciken ba.

Akwai kuma wata damuwa mai sauƙin mantawa: **yiwuwar bayanan horo sun taɓa haɗuwa da bayanan gwajin**. MIMIC-IV ma'ajin jama'a ne kuma an yi rubuce-rubuce da yawa a kansa. Samfurin harshe da aka horar da shi da bayanan intanet na iya taɓa ganin takardu, tattaunawar shari'o'i ko wasu bayanai da ke da alaƙa da shi. Idan wani ɓangare na amsoshin ya shiga bayanan horo, wani ɓangare na aikin na iya zama tunawa maimakon tunanin likitanci. Bincike mai zaman kansa ne kaɗai zai iya ware waɗannan. Marubutan da kansu sun ce sakamakon "za a iya fassara shi cikin taka-tsantsan a matsayin wata iyaka ta sama" kuma yana iya "yin ƙarin hasashen ikon tsarin fiye da yadda zai yi kan wasu shari'o'in jama'a." Wato takardar kanta ta sanya iyaka ga kanun labarinta.

Babu ɗayan waɗannan da ya sa MIRA ta zama marar ban sha'awa. Abin da ya dace shi ne a auna da'awar: a wannan gwajin shari'o'in baya, wakilin da ya tafiyar da kundin gaba ɗaya ya fi likitoci a wasu ganewar cuta, musamman inda gwaje-gwaje suka bayar da amsa mai tsabta — kuma ya yi amfani da gwaje-gwaje da yawa. Wannan hujja ce ta ainihin ƙwarewa, ba hukunci cewa na'urori sun riga sun fi likitoci gano cuta ba.

Abin da wannan bai tabbatar ba

- **Bai nuna cewa MIRA tana aiki kan marasa lafiya na gaske ba.** Dukkan shari'o'in na baya ne kuma an yi kwaikwayonsu; babu mara lafiya da MIRA ta kula da shi kai tsaye.
- **Bai nuna cewa tana aiki a wajen cututtuka takwas ba.** Matsalolin da ba su cikin rukunin da aka zaɓa — waɗanda su ne yawancin rikice-rikicen asibiti — ba a gwada su ba.
- **Bai nuna cewa an shirya tura ta aiki cikin aminci ba.** Marubutan sun ce har yanzu ana buƙatar binciken gaba-gaba na ainihin marasa lafiya kan faɗaɗa aiki, aminci da tsarin kula da amfani.
- **Bai tabbatar da cikakken gwaji mai adalci da likitoci ba.** MIRA ta yi amfani da gwajin jini da yawa sosai, kuma an auna wasu sakamakon magani da abin da aka rubuta a kundin tarihi.
- **Bai tabbatar cewa babu tasirin tunawa da bayanan horo ba.** MIMIC-IV ma'ajin jama'a ne; ana buƙatar maimaitawa mai zaman kanta.
- **Ba yana nufin "AI za ta maye gurbin likitoci" ba.** Wannan kayan taimakon yanke shawara ne da zai iya aiwatar da matakai a kundin; amfani na gaske zai fi yiwuwa ya kasance tare da likitocin da ke riƙe da ikon yanke hukunci kuma suke kawo bayanan da rubutu ba ya ɗauke da su.

Yaya ƙarfin shaidar yake?

Ya kamata a raba da'awar gida biyu.

A matsayin nuna ƙwarewa, wato cewa wakilin da ke amfani da samfurin harshe zai iya tafiyar da shari'ar mara lafiya gaba ɗaya a cikin EHR na gwaji — tarihin lafiya, gwaje-gwaje, ganewar cuta da umarni — sakamakon sabo ne kuma mai gamsarwa sosai. Wannan shi ne sabon ɓangaren da ya fi cancantar kulawa.

A matsayin da'awar cewa MIRA ta fi likitoci, shaidar tana da iyaka. Ta yi aiki ne a wani takamaiman gwajin shari'o'in baya, a cututtuka takwas, tare da bambancin yawan bayanan da bangarorin suka samu, wasu ma'aunai da suka dogara da kundin tarihi, da yiwuwar haɗuwar bayanan horo. Ya isa a ce "wakilin ya yi matuƙar kyau a wannan gwajin." Bai isa a ce "ya fi likita iya gano cuta" gaba ɗaya ba, kuma marubutan ba su yi wannan da'awar ta biyu ba.

Matsayin da ya fi amfani ba "AI ta doke likitoci" ba ne, kuma ba "abin wasa ne kawai" ba. Shi ne: wani sabon nau'in AI na lafiya — wanda yake aiki a cikin kundin maimakon amsa tambayoyi kaɗai — ya yi kyau a gwaji mai wahala na shari'o'in baya. Mataki na gaba shi ne ya tabbatar da kansa a kan marasa lafiya na gaske da matsalolin da ba a zaɓa tun da farko ba, a bincike na gaba-gaba kuma ƙarƙashin tsarin sa ido.

Me ya sa yake da muhimmanci

Shekaru da dama, tambayar da ta fi jan hankali a AI na lafiya ta sauya a hankali. Da farko tambayar ita ce: "samfuri zai iya gano cuta daidai?" Samfura da yawa sun rika amsa eh a kan gwaje-gwaje masu tsabta. MIRA tana nuna matsawa zuwa tambaya mafi wahala: "samfuri zai iya *yin aikin* — tattara bayanai, umartar gwaji, fassara sakamako, yanke shawara da aiwatar da mataki — a dukkan tsarin aiki?" Wannan tambaya ce mafi amfani kuma mafi gaskiya,

domin a aiwatarwa ne wahala da haɗari suke bayyana.

Shi ya sa yadda ake gabatar da labarin yake da muhimmanci. Kanun "AI ta fi likitoci" yana jawo hankali zuwa ɓangaren da ba shi ne sabon abu mafi girma ba kuma shaidarsa ta fi iyaka. Abin da ya fi sabo ya fi ƙanƙanta a lafazi amma ya fi tasiri: wakili da zai iya bi ta cikin kundin mara lafiya daga farko zuwa ƙarshe. Idan wannan ƙwarewar ta tsaya a binciken gaba, za ta fara sauya **yadda ake gudanar da aiki** tun kafin ta sauya wanda ke da ikon yanke shawara. Ba likita ne ake cirewa ba; umartar gwaji, bin sakamako da fassara su — tsarin aikin ne zai fara canzawa.

Haka kuma yana mayar da nauyin tabbatarwa inda ya dace. Nasara a jadawalin gwajin shari'o'in baya dalili ce ta gudanar da gwajin gaba-gaba, ba madadinsa ba. Karatun da ya fi gaskiya shi ne cewa MIRA gayyata ce zuwa bincike na gaba — a auna ta kan marasa lafiya, ba kan gwaji kaɗai ba.

Takaitaccen bayani

MIRA wakiliya ce ta AI mai cin gashin kanta da ke aiki a cikin kundin lafiya na lantarki na gwaji. Tana iya karɓar tarihin mara lafiya, umartar gwaje-gwajen dakin gwaje-gwaje, hoton jiki da ƙwayoyin cuta, fassara sakamakonsu, kai ga ganewar cuta sannan ta rubuta tsarin magani. A shari'o'i 574 na baya daga MIMIC-IV, da suka shafi cututtuka takwas da aka zaɓa tun da farko, ta fi likitoci daidai wajen gano cuta: 88.9% gaba ɗaya, kuma a rukuni na shari'o'i 311 ta samu 87.8% idan aka kwatanta da 78.1% na manyan likitoci. Amma gwajin kwaikwayo ne kan bayanan baya kuma rubutu kaɗai; mafi yawan fifikon ya fito ne daga cututtukan da ke da gwaji mai yanke hukunci; MIRA ta yi amfani da gwajin jini da yawa sosai; wasu ma'aunan magani sun dogara da abin da kundin tarihi ya nuna; kuma kasancewar MIMIC-IV ma'ajin jama'a yana barin yiwuwar wani ɓangare ya shiga bayanan horo. Babban ci gaban shi ne wakili da yake iya tafiyar da aikin gaba ɗaya maimakon amsa tambaya guda. Har yanzu ana buƙatar gwaje-gwajen gaba-gaba a asibitoci na gaske kan aminci, faɗaɗa aiki da tsarin sa ido. Wannan hujja ce mai ƙarfi ta ƙwarewa, ba tabbacin cewa AI ta fi likitoci gano cuta ko cewa ta shirya yin aiki a asibiti ba.

Bincike ba tare da ƙarin gishiri ba

Abin da takardar ta nuna: MIRA, wakiliya mai amfani da samfurin harshe, za ta iya tafiyar da shari'ar mara lafiya daga farko zuwa ƙarshe a cikin EHR na gwaji — tarihin lafiya, gwaje-gwaje, ganewar cuta da umarni — kuma a gwajin shari'o'i 574 na baya a cututtuka takwas ta fi likitoci daidai wajen gano cuta (88.9% gaba ɗaya; 87.8% idan aka kwatanta kai-tsaye da 78.1% na manyan likitoci), ba tare da an ga kuskuren magani mai tsananin haɗari ba.

Abin da yake yiwuwa amma ba a tabbatar ba: Wannan ƙwarewar za ta amfani marasa lafiya na gaske da matsalolin da ba a zaɓa tun da farko ba; fifikon daidaito ya fito ne daga tunani mafi kyau maimakon yawan gwaje-gwaje, daidaituwa da kundin tarihi ko tunawa da bayanan jama'a.

Abin da bai nuna ba: Ba ya nuna cewa MIRA tana aiki a wajen cututtuka takwas, cewa an shirya tura ta asibiti cikin aminci, cewa ta fi likitoci a gwaji mai bayanai iri ɗaya gaba ɗaya, cewa za ta maye gurbin likitoci, ko cewa sakamakon ba shi da wani tasirin haɗuwar bayanan horo.

Babban iyaka: Gwajin na baya ne, na kwaikwayo kuma rubutu kaɗai; cututtuka takwas aka zaɓa; MIRA ta yi

amfani da gwajin jini da yawa (kusan 51% na ma'aunan da ke samuwa, idan aka kwatanta da 28% na manyan likitoci); an auna wasu sakamakon magani da kundin tarihi; kuma MIMIC-IV ma'ajin jama'a ne da zai iya bayyana a bayanan horo. Marubutan da kansu sun kira sakamakon wata yiwuwar iyaka ta sama.

Yaya yawan amincewa ya kamata mai karatu ya yi? Babban amincewa cewa MIRA tana nuna wata sabuwar ƙwarewa ta gaske — wakili da yake *aiwatar da matakai* a cikin kundin, ba chatbot da ke amsa tambaya kawai ba. Karamin amincewa da faɗaɗa wannan zuwa "ta fi likita gano cuta": wannan ya tsaya ne kan wani gwajin baya da wasu abubuwa masu rikitarwa. Mataki na gaba da ya dace shi ne gwaji na gaba-gaba kan marasa lafiya na gaske.

Majiyoyi

Ferber et al., Towards autonomous medical artificial intelligence agents, Nature (2026)
https://doi.org/10.1038/s41586-026-10675-5

PubMed 42310457 — peer-reviewed abstract
https://pubmed.ncbi.nlm.nih.gov/42310457/

Science Media Centre — expert reaction to MIRA and AMIE (2026)
https://www.sciencemediacentre.org/expert-reaction-to-presentation-of-two-new-medical-ai-models-for-

News-Medical (2026) — illustrates the 'outperforms doctors' framing
https://www.news-medical.net/news/20260621/Autonomous-medical-AI-outperforms-doctors-in-simulated-EH
aspx