MADRID 2026 ESMO congress

LBA19

Large language model-assisted chemotherapy side effect management in solid tumor patients: The ALLIANCE randomized controlled trial

Y-J. Kang[1], J.M. Song[1], J.Y. Park[2], Y-J. Yang[3], M.J. Kim[1], Y.C. Yoon[4], H. Kim[2], S. Kim[2], H. Choi[1], J.H. Byun[5], S.J. Oh[6]

[1] Breast Surgical Oncology, The Catholic University of Korea, Incheon St. Mary's Hospital, Incheon, Republic of Korea, [2] Breast Surgery, The Catholic University of Korea, Incheon St. Mary's Hospital, Incheon, Republic of Korea, [3] Department of Convergence Science, Catholic Kwandong University College of Medicine, Incheon, Republic of Korea, [4] Department of Surgery, The Catholic University of Korea - Incheon St. Mary's Hospital, Incheon, Republic of Korea, [5] Medical Oncology, The Catholic University of Korea - Incheon St. Mary's Hospital, Incheon, Republic of Korea[6] Breast Surgery, The Catholic University of Incheon - St. Mary's Hospital, Incheon, Republic of Korea

Background

ESMO ELCAP requires systematic validation of clinician-facing (Type 2) large language model (LLM) use. Of 4,609 clinical LLM studies, 1,048 used real patient data and 19 were prospective randomised trials; Type 2 use in oncology lacks randomised evidence.

Methods

Single-centre, 1:1 randomised two-stage trial; adults with breast or colorectal cancer starting ≥3 months of chemotherapy, ECOG 0-2. Intervention-arm clinicians could consult ChatGPT Plus with de-identified symptom data at visits; all decisions remained theirs. Patients, not clinicians, were blinded. Primary endpoint: EORTC QLQ-C30 global health status (GHS) change, baseline to end of treatment. This pre-specified 60-patient interim analysis addressed safety and feasibility; enrolment then stopped (cut-off 22 Jul 2026).

Results

Of 60 enrolled, 56 were analysed, 28 per arm. The LLM was used at 47 of 99 monthly assessments (47.5%); 7 of 28 patients never used it, "not needed" at 51 of 52 unused assessments. Recorded adverse events nearly doubled (150 vs 78; 5.4 vs 2.8 per patient, p=0.006). In a post hoc classification the excess lay in toxicities with no routine management: 48 vs 27 distinct terms, against 29 vs 22 for routine toxicities. Advice was adopted in full at 8 of 47 uses and partly at 29; 5 of the 8 full adoptions concerned toxicities without a routine. Dose reduction was unchanged (12/28 vs 14/28). Output was appropriate at 34 of 47 and never inappropriate; hallucination was absent at 44 and mild at 3. Adjusted GHS difference was -5.3 (95% CI -18.2 to 7.7) at mid-treatment (n=48) and -18.4 (-31.2 to -5.5) at end of treatment (n=31, 55%).

Conclusions

For toxicities clinicians already have routines for, LLM advice was adopted mostly in part. What it added was direction on managing toxicities that have no routine and are easily passed over: dry eye, dysgeusia, herpes zoster, sinus bradycardia. Dose modification, the one action traded against tumour control, was unchanged. No patient-level benefit was shown. Until it is, clinician-facing LLM support belongs where clinicians have no routine, as a reference rather than a quality-of-life intervention, and should be trialled by targeting it there.

Clinical trial identification

NCT07198581 (ALLIANCE trial).

Legal entity responsible for the study

The Catholic University of Korea, Incheon St. Mary's Hospital.

Funding

Institute for Bio-Medical Convergence, Incheon St. Mary's Hospital, The Catholic University of Korea; National Research Foundation of Korea (RS-2025-19643006).

Disclosure

All authors have declared no conflicts of interest.

© European Society for Medical Oncology