

INPLASY

INPLASY2026100109

doi: 10.37766/inplasy2026.10.0109

Received: 26 October 2026

Published: 26 October 2026

Corresponding author:

Mao Sun

381702901@qq.com

Author Affiliation:

Chongqing University cancer Hospital.

Intrathecal chemotherapy added to systemic therapy for leptomeningeal metastasis in EGFR-mutant non-small-cell lung cancer after first-line EGFR-TKI: a systematic review and meta-analysis

Sun, M; Wang, LL; Tao, D; Jiang, Y; Zhou, W; Wu, YZ; Du, L.

ADMINISTRATIVE INFORMATION

Support - National Natural Science Foundation Project (No. 82303673 to LD); Chongqing Young Top Medical Talents Program (No. YXQN202429 to LD).

Review Stage at time of this submission - Completed but not published.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY2026100109

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 26 October 2026 and was last updated on 26 October 2026.

INTRODUCTION

Review question / Objective In adults with advanced non-small-cell lung cancer (NSCLC) harbouring an EGFR sensitising mutation who develop leptomeningeal metastasis (LM) after first-line EGFR tyrosine kinase inhibitor (TKI) therapy, does adding intrathecal chemotherapy (ITC) to systemic therapy, compared with systemic therapy or best supportive care without ITC, improve overall survival, intracranial response and intracranial progression-free survival, and at what cost in grade ≥ 3 adverse events?

Rationale LM is an increasingly frequent site of progression in EGFR-mutant NSCLC, and median survival after diagnosis remains close to one year even with CNS-penetrant TKI. Whether ITC adds survival benefit to systemic therapy is unresolved, and no randomised trial exists in NSCLC. Two propensity-score matched cohorts designed around ITC as the exposure disagree: one reported

13.2 versus 10.1 months without a hazard ratio ($P = 0.488$), the other an adjusted hazard ratio of 0.427 without a confidence interval ($P = 0.022$). Earlier syntheses addressed different questions: a 2016 pooled analysis described response and survival in ITC-treated patients without a comparator, and a 2025 single-arm meta-analysis pooled response proportions only. This review therefore appraises the comparative evidence against a prespecified poolability rule and, separately, quantifies the outcomes achieved by ITC-containing therapy.

Condition being studied Leptomeningeal metastasis from NSCLC, defined by positive cerebrospinal fluid (CSF) cytology, characteristic magnetic resonance imaging (MRI) enhancement, or both, in patients with an EGFR sensitising mutation (exon 19 deletion, L858R, G719X, L861Q or S768I).

METHODS

Participant or population Adults (≥ 18 years) with advanced NSCLC and an EGFR sensitising mutation who received first-line EGFR-TKI therapy and were diagnosed with LM during or after that therapy. Studies enrolling mixed populations were eligible only when an EGFR-mutant or first-line-TKI subgroup was reported separately; cohorts not restricted to EGFR-mutant disease were retained for sensitivity analysis only.

Intervention Intrathecal chemotherapy with pemetrexed, methotrexate or cytarabine, delivered by lumbar puncture, ventricular catheter or Ommaya reservoir, with or without concurrent systemic therapy (EGFR-TKI, chemotherapy, antiangiogenic therapy or radiotherapy).

Comparator Systemic therapy or best supportive care without intrathecal chemotherapy. Studies in which ITC was inseparable from whole-brain radiotherapy were treated as a composite exposure and used for sensitivity analysis only.

Study designs to be included Randomised trials (none identified), prospective or retrospective comparative cohorts including propensity-score matched cohorts, and prospective or retrospective single-arm ITC studies. Comparative cohorts address the question of adding ITC; single-arm studies are used only for proportion synthesis and are never analysed as if they provided a comparator.

Eligibility criteria Inclusion: at least five patients; LM diagnosed by CSF cytology or MRI; ITC as a defined exposure; reporting of overall survival, intracranial progression-free survival, LM response or grade ≥ 3 adverse events. Exclusion: case reports and series with fewer than five patients; reviews, editorials and pooled analyses (used for citation chasing only); populations not enriched for EGFR-mutant disease without a stratified estimate; ITC as one component of a composite exposure inseparable from radiotherapy or surgery (sensitivity analysis only); non-lung primaries; intrathecal agents other than cytotoxic chemotherapy. Two reports of one registered trial were not treated as independent studies; the report with the larger sample entered the pooled models.

Information sources PubMed (MEDLINE), Embase (Elsevier) and the Cochrane Library (Cochrane Central Register of Controlled Trials, CENTRAL) were searched from database inception to 25 October 2026 without language or

publication-type restrictions. The three databases were searched through a single literature platform that indexes them simultaneously, so individual records could not be attributed to a specific database and per-database yields cannot be separated. Chinese-language studies were identified from Chinese journals indexed in these databases. Reference lists of all included studies and of previously published systematic reviews and pooled analyses of intrathecal therapy for leptomeningeal metastasis were hand-searched for additional records. Trial registries, grey literature sources, conference abstract repositories and direct contact with investigators were not searched, and no additional update search was performed after 25 October 2026.

Main outcome(s) The primary outcome was overall survival (OS). OS was analysed as the median OS reported by each study, measured from the time origin defined by that study, that is from the diagnosis of leptomeningeal metastasis (LM) in some cohorts and from the first intrathecal chemotherapy (ITC) dose in others. The time origin of each study was recorded where reported; when a study reported only a survival rate at a fixed time point rather than a median, this was recorded as such and the study was excluded from the median OS synthesis. For the comparative question, OS was the outcome used to judge whether adding ITC provided benefit; for the single-arm question, OS described survival achieved within ITC-treated populations. No hierarchy across alternative survival instruments was required because all studies reported OS as defined here.

Additional outcome(s) Secondary outcomes were intracranial or LM progression-free survival (PFS), LM response, and grade ≥ 3 adverse events. Intracranial or LM PFS was taken as reported, with the assessment instrument and the starting point recorded per study. LM response was analysed in two prespecified, non-overlapping strata. The first comprised radiological objective response assessed with Response Assessment in Neuro-Oncology (RANO) criteria for LM or an intracranial objective response rate derived from imaging. The second comprised disease control or clinical response, a stratum that combined clinical response, symptom improvement, intracranial disease control rate and disease control rate as labelled by the contributing studies; the label used by each study was recorded alongside its estimate, and the two strata were never combined. Grade ≥ 3 adverse events were extracted with each study's own definition and grading system, and studies reporting only any-grade myelosuppression without a Common Terminology

Criteria for Adverse Events grade were not entered into the pooled safety estimate. Safety was assessed within ITC-treated populations only, because no comparative study reported adverse events by arm.

Data management Records retrieved from all searches were pooled and de-duplicated before screening. Two reviewers independently screened all titles and abstracts against the eligibility criteria; records judged potentially eligible by either reviewer were obtained in full and assessed independently. Disagreements were resolved by discussion and, when unresolved, by a third reviewer. Two reviewers independently extracted data into a piloted standardised form and reconciled discrepancies against the source record. Data were extracted from the full record where available and from the structured abstract where a full text could not be accessed; abstract-level data were retained only when the required outcome was reported explicitly. Unreported values were recorded as not reported and were not imputed, recalculated or converted, so the number of studies contributing to each analysis varies by outcome. Multiple reports of one study were linked and were not treated as independent studies; where two reports described the same registered trial, only the report with the larger sample entered the pooled models and the other contributed descriptive data only. Full-text exclusions were recorded with reasons. Screening, extraction and reconciliation were performed in Microsoft Excel [version], and all statistical analyses were run in Python 3.11 with numpy and scipy. The analysis script and the study-level extraction table are available from the corresponding author on request, and the extracted data are reported in the published review.

Quality assessment / Risk of bias analysis

Comparative cohorts were assessed with the Newcastle-Ottawa Scale (NOS), scoring selection, comparability and outcome ascertainment to a maximum of nine stars, applied domain by domain. Single-arm studies were assessed with NOS items adapted to a design without a comparator, covering representativeness of the series, ascertainment of ITC exposure, outcome definition, completeness of follow-up and adequacy of follow-up duration; domain judgements for these studies are reported descriptively rather than as a total score, because numerical quality scores obscure which domain failed and the inter-rater reliability of the NOS is limited. A domain-based instrument for non-randomised studies of interventions (ROBINS-I) was not applied to the single-arm set, because such instruments

presuppose a defined comparator and estimate the bias of an effect. Risk of bias was judged at study level, summarised narratively, and was not used to exclude studies from the review. For comparative cohorts, comparability was scored 2 for the two propensity-score matched cohorts, in which performance status entered the multivariable model, and 1 for the study that presented no adjusted comparison and for the study whose adjustment variables were not reported. Certainty of evidence was assessed with GRADE separately for each pooled estimate and for the comparative question, starting at low certainty for single-arm proportions and for the observational comparison and downgrading for risk of bias, inconsistency, indirectness and imprecision. Publication bias was not assessed because fewer than ten studies contributed to every model, and it was therefore not used as a reason to downgrade certainty; small-study effects remain an unassessed risk.

Strategy of data synthesis Synthesis was planned as a systematic review with, where the data allowed, meta-analysis of proportions and of medians. Proportions were pooled after logit transformation, with the per-study variance taken as $1/(n \cdot p \cdot (1-p))$ on the logit scale, using random-effects models with the DerSimonian-Laird estimator of τ^2 ; estimates were back-transformed to the proportion scale. Median OS was pooled on the log scale, with the standard error of each median calculated as $(\text{upper} - \text{lower})/3.92$ from the reported 95% confidence interval; only studies reporting a median with a confidence interval could enter this model, and studies reporting medians without a confidence interval were summarised descriptively. Heterogeneity was quantified with Cochran's Q , I^2 and τ^2 on the transformed scale, and a 95% prediction interval was derived from τ^2 and the standard error of the pooled estimate using a t distribution on $k - 1$ degrees of freedom. All models were refitted with the Hartung-Knapp variance correction as a sensitivity analysis. Pooling of comparative effect estimates was governed by a prespecified rule: a hazard ratio would be pooled only if at least three comparative studies provided both a point estimate and a measure of variance. Only one study met this requirement, because the remaining comparative studies reported either no hazard ratio or a hazard ratio without a confidence interval and without a reconstructed standard error acceptable under the rule; comparative evidence was therefore synthesised narratively. Where studies could not be combined because they measured different quantities, they were reported separately rather than pooled. Network meta-analysis was not planned. No meta-regression was performed,

because too few studies contributed to each model. Clinical and methodological diversity across studies was described alongside the quantitative results, and the direction of comparative estimates was reported with their precision rather than as a pooled effect.

Subgroup analysis Three subgroup analyses were prespecified: by intrathecal agent (intrathecal pemetrexed versus intrathecal methotrexate), by whether ITC was combined with a third-generation EGFR-TKI, and by the definition of LM response. Subgroup analyses were designated as exploratory and descriptive. Formal tests of interaction and subgroup-specific pooled estimates were not planned, and none were performed, because the number of studies and events within each subgroup was small, subgroup membership was confounded with response definition, cohort size and line of therapy, and most single-arm studies used a single intrathecal agent. Subgroup observations are reported as differences in the spread of study-level values and are not interpreted as effect modification.

Sensitivity analysis Four sensitivity analyses were prespecified or performed. First, all pooled models were refitted with the Hartung-Knapp variance correction, which is more conservative when the number of studies is small. Second, leave-one-out refitting was used to examine whether any pooled estimate was carried by a single study. Third, because two single-arm reports described one registered trial, the pooled models were also refitted with both reports included, to assess whether the overlap influenced the direction or size of the estimates. Fourth, comparative studies that could not answer the question as posed, that is those in which the population was not restricted to EGFR-mutant disease or in which the exposure was a composite of intrathecal therapy with or without whole-brain radiotherapy, were restricted to sensitivity consideration and were not entered into any pooled comparative estimate. Publication bias was not assessed, because fewer than ten studies contributed to every model; no funnel plot was generated and no Egger test was performed. A formal forest plot of the comparative effect could not be produced because no comparative effect estimate with a variance was available for pooling.

Language restriction English. The protocol, the search strategy documentation and the review will be prepared and published in English. No language restriction was applied to eligible studies; non-English records were assessed through.

Country(ies) involved China.

Other relevant information Registration is retrospective. Screening, data extraction, risk-of-bias assessment and analysis were completed before the protocol was registered, and the dates of those activities are recorded in field 04; no methodological decision was altered after the study findings were known. Four protocol deviations are reported in the published review: the review was not registered before screening; no subgroup meta-regression was performed; publication bias was not assessed because fewer than ten studies contributed to each model; and no forest plot of a comparative effect estimate was generated because no poolable comparative estimate existed. Several limitations shape the interpretation of the results and are stated in the protocol for transparency: no randomised or prospective comparative study exists in this population; only two of the four comparative cohorts were designed around ITC as the exposure; ITC was not allocated at random and in the matched cohorts the survival difference was concentrated in the worst performance-status stratum, so indication bias cannot be separated from an effect of treatment; no included cohort required LM after first-line TKI therapy as its sole eligibility condition; response definitions differed across studies and radiological and clinical response were kept apart; time origins for median OS differed between cohorts; and several studies were extracted at abstract level, which limited adjustment variables and subgroup results. Two single-arm reports describe one registered trial (ChiCTR1800016615) and were not treated as independent studies. The review is reported according to PRISMA 2020.

Keywords Leptomeningeal metastasis; Non-small-cell lung cancer; EGFR mutation; Intrathecal chemotherapy; Pemetrexed; Meta-analysis.

Contributions of each author

Author 1 - Mao sun
Email: sunmao4379@gmail.com
Author 2 - Lu-Lu Wang
Author 3 - Dan Tao
Author 4 - Yong Jiang
Author 5 - Wei Zhou
Author 6 - Yong-Zhong Wu
Author 7 - Liang Du
Email: i.du@cqu.edu.cn