

EGFR tyrosine kinase inhibitor receipt and overall survival in metastatic non-small cell lung cancer in Egypt A retrospective descriptive cohort study

Esraa Jamal Moustafa Jamal

Mansoura University

Nourin Ali Sharief

Mansoura University

Manar Hamed Ahmed

Mansoura University

Ahmed M. Elalfy

ahmed.elalfy@ucalgary.ca

Mansoura University

Research Article

Keywords: EGFR mutation, Non-small cell lung cancer, Tyrosine kinase inhibitor, Overall survival, Treatment access, Egypt

Posted Date: October 11th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11097744/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Purpose

To describe epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor (TKI) receipt and overall survival (OS) in an Egyptian institutional cohort with EGFR-mutated metastatic non-small cell lung cancer.

Methods

We reanalysed a patient-level workbook for a retrospective cohort managed at Mansoura University between January 2016 and January 2024. Treatment receipt was described in 99 patients. OS was estimated from recorded durations and vital status using Kaplan–Meier methods. Progression and biomarker analyses were excluded because their source data could not be adequately reconciled.

Results

Seventy-three patients (73.7%) received a TKI, including 29 in the first line, 36 in the second line, and eight in later lines. Available drug records described gefitinib and erlotinib; one treated patient had unresolved drug coding. Retained study summaries attributed nonreceipt to poor performance status or deterioration in 14 patients, refusal in six, and access or reimbursement failure in six. OS was available for 94 patients, with 18 deaths. Estimated OS was 96.2% (95% confidence interval 88.7–98.8%) at 12 months and 74.2% (57.2–85.3%) at 24 months. Median follow-up was 18 months. All five patients with missing OS had received no TKI.

Conclusion

This selected cohort describes heterogeneous TKI receipt across treatment lines. Survival estimates are limited by heavy censoring and missing outcomes concentrated among untreated patients. These data support further study of treatment pathways but cannot establish treatment effects or population-level access rates.

Introduction

Lung cancer remains a major cause of cancer mortality worldwide (Bray et al. 2024). EGFR alterations identify a clinically important subgroup of non-small cell lung cancer (NSCLC), although sensitivity to targeted therapy depends on the specific alteration and drug (Lynch et al. 2004; Mok et al. 2009). Gefitinib and erlotinib established the role of first-generation EGFR tyrosine kinase inhibitors (TKIs), while subsequent trials and recommendations support newer treatment approaches for common sensitizing alterations (Mok et al. 2009; Soria et al. 2018; Hendriks et al. 2025; Rosell et al. 2012).

Routine-care cohorts can describe which patients receive targeted treatment and at what point in their disease course. Treatment nonreceipt may reflect clinical deterioration, patient preferences, financial barriers, or mutation-specific considerations. These explanations should not be conflated: detecting an EGFR alteration does not establish that every available TKI is appropriate for that patient, and a record of nonreceipt does not by itself identify an access failure.

Evidence from settings underrepresented in clinical trials can help define questions for care-delivery research. Real-world work in a Latin American middle-income country has examined genomic testing, targeted therapy, and outcomes (Hernandez-Martinez et al. 2024). Local data from Egypt may similarly inform more detailed evaluation of testing and treatment pathways. We reanalysed an existing institutional cohort to describe recorded TKI exposure and OS, with explicit attention to missingness, coding discrepancies, and the limits of the available data.

Materials and methods

Study setting and original cohort

The source study was a retrospective cohort of adults with documented stage IV EGFR-mutated NSCLC managed at the Oncology Center, Mansoura University, Egypt, between January 2016 and January 2024. The cohort was originally assembled for an exploratory inflammatory-marker study. Its eligibility criteria were retained for this descriptive reanalysis; it was not redefined as a population-based access cohort.

The original study required age at least 18 years, stage IV NSCLC, a documented EGFR alteration, and a differential blood count within two weeks before TKI initiation or, for untreated patients, within two weeks of stage IV diagnosis. Exclusions included concurrent malignancy, active infection or sepsis, autoimmune disease requiring immunosuppression, immunosuppressive corticosteroids equivalent to at least 20 mg prednisolone daily within four weeks, and insufficient haematological data. These requirements describe the original cohort selection; complete values for every derived inflammatory index were not present in the analysis workbook.

The screening and exclusion counts were retained from the original study summaries because records for screened and excluded patients were not included in the recovered workbook. Stage IV status, the recruitment interval, and original laboratory eligibility could not be independently reconstructed from dates or source reports within that file. We therefore do not infer a population testing rate or claim independent re-verification of the screening process.

Data sources and reconciliation

The recovered workbook contained two patient-data sheets and an earlier statistical-output sheet. The cleaned patient sheet contained 99 unique coded records; the other sheet contained the same records and one additional row with no age, sex, treatment, or survival information. The 99-record sheet defined

the analytic cohort. All OS durations and status codes agreed across the two patient sheets. No previous model coefficients or figure values were reused.

We compared corresponding fields across sheets. One histology entry and one TKI drug entry differed without a supporting source report; these were classified as unknown in the present analysis. A blank EGFR entry in one sheet was coded as unspecified in the other and remained unspecified. One combined exon 19/20 entry was retained as a separate category, resolving a previously omitted category. Exon-level codes were not converted into specific sequence variants. Biomarker values included unexplained additions and changes, without absolute blood counts from which to verify them; consequently, biomarker associations were not estimated. Supplementary Information 1 describes the reconciliation decisions.

Clinical variables and treatment exposure

We summarized age, sex, smoking history, histology, Eastern Cooperative Oncology Group performance status (ECOG PS), metastatic sites, and EGFR code. TKI receipt was defined by a positive recorded treatment-line code; drug coding independently supported receipt except for the single unresolved drug entry. Treatment line referred to the recorded line of systemic therapy in which a TKI was used. The drug categories represented gefitinib, erlotinib, or both; the prior study description identified use of both as sequential. Treatment dates, dose intensity, and duration of access were not available for independent verification.

Reasons for nonreceipt were available only as aggregate categories in the retained study summaries: poor performance status or rapid deterioration, refusal, and access or reimbursement failure. They were not patient-level fields in the recovered workbook and could not be linked to molecular subtype or independently adjudicated. Their denominator was the 26 patients with no recorded TKI. Nonreceipt was not assumed to indicate clinical ineligibility or failed access.

Outcome definition and statistical analysis

OS retained the definition in the original study documentation: time from stage IV diagnosis to death from any cause, with patients alive censored at last known contact. The reanalysis used the recorded durations in months and binary vital-status codes because calendar dates were not supplied. Patients missing both duration and status were excluded from OS estimation and described separately. No missing outcome or event code was imputed.

We estimated OS using the Kaplan–Meier method, with pointwise 95% confidence intervals (CIs) calculated using Greenwood variance and a log–log transformation. Median-survival confidence limits were obtained by inversion of those confidence bands. Follow-up was summarized by reverse Kaplan–Meier estimation, treating censoring observations as events. At each displayed time, the risk set included patients with duration at least that time; cumulative deaths and censorings were counted

strictly before it. Deaths and censorings recorded at the same time contributed to the risk set immediately before that time.

Analyses were descriptive and adopted after methodological review of the earlier analysis, rather than prespecified as an access study. We report counts with explicit denominators, median age with range, and OS at 12 and 24 months. No treatment-group survival comparisons, Cox models, significance tests, or optimized biomarker cutoffs were used. Ever receiving a TKI can occur after diagnosis and would require treatment timing to support appropriate survival comparisons. Calculations used Python 3.12, NumPy 2.3, SciPy 1.17, and Matplotlib 3.10. Kaplan–Meier point estimates were checked against an independent implementation in SciPy.

Progression-free survival (PFS) could not be reconstructed for all 73 treated patients: 25 lacked an event-status code, and the workbook did not provide progression, treatment-start, or death dates needed to resolve the endpoint. Missing status did not establish absence of imaging, because 22 of these patients had a recorded three-month response category. We therefore excluded PFS and time-specific response analyses, rather than restricting the endpoint to the subset with complete progression codes.

OpenAI ChatGPT assisted with data auditing, generation of analysis code, manuscript restructuring, and language editing. Statistical calculations were executed on the supplied workbook; no patient observations were generated or imputed. The authors retain responsibility for the data, analysis decisions, and manuscript.

Ethics

The Medical Research Ethics Committee, Faculty of Medicine, Mansoura University, approved the study (R.25.09.3349). The study was conducted in accordance with the Declaration of Helsinki.

The ethics committee waived individual informed consent for the retrospective study. Direct identifiers were excluded from analysis outputs and publication files.

Results

Cohort and clinical characteristics

The retained screening summary reported 925 screened records, 288 with documented EGFR testing, and 107 with an EGFR alteration. Eight exclusions were reported: missing baseline blood counts in three, active infection or sepsis in two, concurrent cancer in two, and immunosuppressive corticosteroid exposure in one. The recovered analytic cohort contained 99 records. These screening counts were not independently reproducible from the patient workbook.

Median age was 60 years (range 27–81), 61 patients (61.6%) were women, and 83 (83.8%) were never-smokers. ECOG PS was 1 in 44 patients, 2 in 50, and 3 in five; 55 patients (55.6%) therefore had ECOG PS

2–3. After treating the unresolved histology discrepancy as unknown, 95 patients (96.0%) were coded as having adenocarcinoma. EGFR categories comprised exon 19 in 58, exon 21 in 27, exon 20 in four, T790M in two, combined exon 19/20 in one, and unspecified in seven (Table 1).

Table 1
Clinical and molecular characteristics

Characteristic	Category	n (%) or summary
Age, years	Median (range)	60 (27–81)
Sex	Women / men	61 (61.6) / 38 (38.4)
Smoking	Never / ever	83 (83.8) / 16 (16.2)
Histology	Adenocarcinoma	95 (96.0)
	Squamous / adenosquamous	2 (2.0) / 1 (1.0)
	Unresolved coding	1 (1.0)
ECOG PS	1 / 2 / 3	44 (44.4) / 50 (50.5) / 5 (5.1)
Metastatic sites, number	1 / 2 / ≥3	66 (66.7) / 27 (27.3) / 6 (6.1)
Metastatic site	Brain / liver / bone	12 (12.1) / 14 (14.1) / 38 (38.4)
EGFR code	Exon 19	58 (58.6)
	Exon 21	27 (27.3)
	Exon 20	4 (4.0)
	T790M	2 (2.0)
	Combined exon 19/20	1 (1.0)
	Unspecified	7 (7.1)

N = 99. Individual metastatic sites overlap. EGFR codes describe the workbook categories, not independently verified sequence variants or treatment sensitivity. The unresolved histology entry is counted as unknown. ECOG PS, Eastern Cooperative Oncology Group performance status; EGFR, epidermal growth factor receptor.

Treatment receipt

Seventy-three patients (73.7%) had recorded TKI receipt and 26 (26.3%) did not. Among treated patients, 29 received a TKI in the first line, 36 in the second line, and eight in the third or a later line. Drug coding agreed for gefitinib in 45, erlotinib in 22, and both agents in five; one additional treated patient was coded as erlotinib in the cleaned sheet but had no drug specified in the other sheet and was retained as drug unknown (Table 2).

The retained aggregate reasons for nonreceipt were poor performance status or rapid deterioration in 14 patients (53.8% of untreated patients), refusal in six (23.1%), and access or reimbursement failure in six (23.1%). The last category represented 6.1% of the complete cohort. The combined performance-status/deterioration category is not equivalent to baseline ECOG PS 3, which was recorded in five patients. Patient-level linkage of these reasons was unavailable.

Table 2
Recorded TKI exposure and aggregate reasons for nonreceipt

Variable	n / denominator	Percent
Any TKI	73 / 99	73.7
No TKI	26 / 99	26.3
First-line TKI	29 / 73	39.7
Second-line TKI	36 / 73	49.3
Third-line or later TKI	8 / 73	11.0
Gefitinib	45 / 73	61.6
Erlotinib	22 / 73	30.1
Both agents, reported as sequential	5 / 73	6.8
Drug unknown because of discordant coding	1 / 73	1.4
Nonreceipt: poor performance status or deterioration	14 / 26	53.8
Nonreceipt: refusal	6 / 26	23.1
Nonreceipt: access or reimbursement failure	6 / 26	23.1

Exposure and treatment-line counts were recomputed from the workbook. Reasons for nonreceipt were retained from the original aggregate study summary and could not be verified at patient level. Drug categories are mutually exclusive. Percentages may not total 100 because of rounding. TKI, tyrosine kinase inhibitor.

Overall survival and missing outcomes

OS duration and vital status were recorded for 94 patients (94.9%); 18 deaths and 76 censored observations were present. All five patients missing OS were among the 26 without a TKI, corresponding to 19.2% of that group. All 73 treated patients had OS data. The estimated survival curve therefore reflects an available-case population with differential outcome completeness.

Estimated OS was 96.2% (95% CI 88.7–98.8%) at 12 months and 74.2% (57.2–85.3%) at 24 months. Reverse Kaplan–Meier median follow-up was 18 months (95% CI 13–19). Median OS was 42 months (95% CI 26 months to not estimable), but only four patients were at risk immediately before the event

that crossed the median. This late estimate is highly unstable. Table 3 and Fig. 1 display the estimates and risk-set information.

Table 3
Overall survival in 94 patients with recorded outcomes

Measure	Estimate	95% CI
Deaths / censored	18 / 76	—
12-month OS	96.2%	88.7–98.8%
24-month OS	74.2%	57.2–85.3%
Median OS	42 months	26 months to not estimable
Median follow-up, reverse Kaplan–Meier	18 months	13–19 months

Numbers at risk immediately before 0, 12, 24, 36, 48, and 60 months were 94, 64, 21, 4, 2, and 1. Only four patients remained at risk when the median was reached at 42 months. Five patients without recorded OS were excluded; none had received a TKI. CI, confidence interval; OS, overall survival.

Discussion

This reanalysis describes a selected institutional cohort in which nearly three quarters of patients received an EGFR TKI, usually after the first line of systemic therapy. The exposure profile was dominated by gefitinib and erlotinib. Retained study summaries described several reasons for nonreceipt, including access or reimbursement failure, but the absence of patient-level reason fields limits the strength of that observation. The study supports more detailed characterization of local treatment pathways rather than a population estimate of unmet need.

The treatment-line distribution is central to interpretation: 44 of the 73 treated patients received a TKI in the second or a later line. Any comparison of survival from diagnosis by eventual TKI receipt would mix different clinical pathways and require accounting for the timing of treatment. The present data contain no treatment dates with which to address that problem. We therefore make no claim that observed OS estimates quantify the effectiveness of either drug or the consequences of nonreceipt.

Survival estimation also requires caution independently of treatment selection. Five missing outcomes were concentrated entirely among untreated patients, and censoring was extensive among the available cases. If missingness or loss of contact is related to deterioration or access, the survival curve may not describe the complete cohort. The 42-month median is mathematically estimable but depends on a risk set of four patients; it should not be interpreted as evidence of unusually favourable treatment outcomes. Reporting confidence intervals and risk counts makes this uncertainty visible but does not remove potential bias from incomplete follow-up.

The recorded exposure pattern must be interpreted in its treatment era and molecular context. Contemporary recommendations differ from the first-generation TKI approaches represented in this dataset (Soria et al. 2018; Hendriks et al. 2025). However, the workbook does not establish when newer agents became locally available or why a particular patient did not receive one. Its exon-level mutation codes are also insufficient to classify all patients as equally eligible for gefitinib or erlotinib. Consequently, neither TKI receipt nor nonreceipt can serve as a stand-alone measure of guideline-concordant care.

The data audit changed several conclusions that would otherwise have been carried forward. A combined mutation code explained the missing category in the previous molecular summary. More importantly, a missing progression code was often accompanied by a recorded response category, so the missing codes could not be equated with absence of post-baseline imaging. Reconstructing a valid PFS endpoint would require the underlying dates and inclusion of death as an event, including deaths without documented progression. Re-labelling a selected subset would not repair this gap. Unexplained differences in inflammatory indices similarly prevented a defensible prognostic analysis.

Limitations and implications

The study is retrospective, single-centre, and based on a cohort originally selected for a different scientific question. Blood-count eligibility and exclusions related to infection, malignancy, and immunosuppression may select patients with different illness trajectories or treatment opportunities. Screening counts and reasons for nonreceipt were available only as retained summaries. Source dates, exact mutation reports, longitudinal treatment details, and explanations for several sheet-to-sheet edits were unavailable. Conservative unknown categories prevent unsupported resolution of two clinical discrepancies but cannot substitute for source-record verification.

Outcome limitations include missing OS in five untreated patients, only 18 observed deaths, possible informative censoring, and inability to verify the recorded durations against calendar dates. No adjusted or causal estimates are presented. PFS, response rates, and biomarker associations have been removed from the present report because the available data do not support their reliable reconstruction. These restrictions reduce the scope and novelty of the study but define what can currently be reproduced.

Future data collection should link exact molecular findings with the dates of treatment recommendation, approval, dispensing, initiation, progression, death, and last contact. Reasons for nonreceipt should be collected at patient level, with clinical, financial, and preference-related explanations allowed to overlap. Such work would permit clearer separation of treatment suitability, delays, and access barriers and would support more reliable outcome estimation.

Conclusion

Among 99 selected patients with EGFR-mutated metastatic NSCLC at an Egyptian oncology centre, 73 had recorded TKI receipt, commonly in later treatment lines. OS could be reproduced for 94 patients but

was limited by extensive censoring and missing outcomes concentrated among untreated patients. The findings describe institutional treatment patterns and identify priorities for improved data capture; they do not establish treatment effects, validated biomarkers, or population-level access rates.

Declarations

Competing interests

The authors declare that they have no competing interests.

Ethics approval

The Medical Research Ethics Committee, Faculty of Medicine, Mansoura University, approved the study (R.25.09.3349). The study was conducted in accordance with the Declaration of Helsinki.

Consent to participate

The ethics committee waived individual informed consent for the retrospective study.

Consent to publish

Not applicable; the manuscript contains no identifiable individual participant data.

Funding

No funding was received for this study.

Author Contribution

EJMN contributed to data collection, formal analysis, and writing of the original draft. NAS contributed to data collection and clinical review. MHA contributed to data collection and clinical review. AME contributed to conceptualization, methodology, formal analysis, interpretation, supervision, writing of the original draft, and review and editing. All authors read and approved the final manuscript and agree to be accountable for the work.

Acknowledgements

The authors thank the staff of the Oncology Center, Mansoura University, for their support with data collection and record review.

Data availability

The clinical workbook contains direct identifiers and is not suitable for public release. Access to a de-identified extract may be requested from the corresponding author, subject to institutional and ethics approval. The aggregate survival table and the code used for this reanalysis have been retained by the corresponding author. Screening records, source dates, and patient-level reasons for nonreceipt were not available in the recovered workbook.

References

1. Bray F, Laversanne M, Sung H et al (2024) Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 74:229–263. <https://doi.org/10.3322/caac.21834>
2. Hendriks LEL, Cortiula F, Martins-Branco D et al (2025) Updated treatment recommendations for systemic treatment: from the ESMO oncogene-addicted metastatic NSCLC living guideline. *Ann Oncol* 36:1227–1231. <https://doi.org/10.1016/j.annonc.2025.06.007>
3. Hernandez-Martinez JM, Guijosa A, Flores-Estrada D et al (2024) Real-world survival outcomes in non-small cell lung cancer: the impact of genomic testing and targeted therapies in a Latin American middle-income country. *JCO Glob Oncol* 10:e2400338. <https://doi.org/10.1200/GO-24-00338>
4. Lynch TJ, Bell DW, Sordella R et al (2004) Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 350:2129–2139. <https://doi.org/10.1056/NEJMoa040938>
5. Mok TS, Wu YL, Thongprasert S et al (2009) Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 361:947–957. <https://doi.org/10.1056/NEJMoa0810699>
6. Rosell R, Carcereny E, Gervais R et al (2012) Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. *Lancet Oncol* 13:239–246. [https://doi.org/10.1016/S1470-2045\(11\)70393-X](https://doi.org/10.1016/S1470-2045(11)70393-X)
7. Soria JC, Ohe Y, Vansteenkiste J et al (2018) Osimertinib in untreated EGFR-mutated advanced non-small-cell lung cancer. *N Engl J Med* 378:113–125. <https://doi.org/10.1056/NEJMoa1713137>

Figures

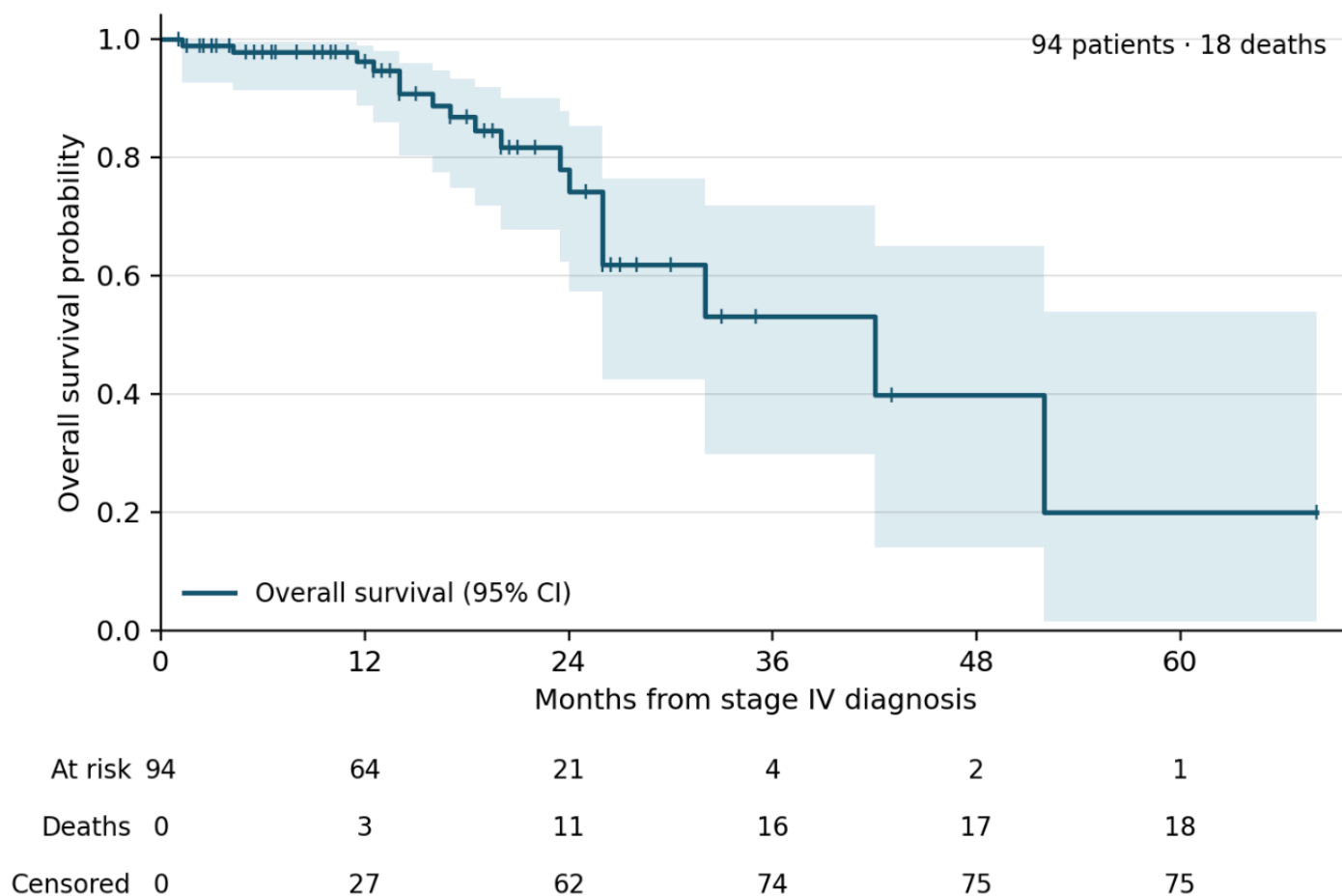


Figure 1

Overall survival from recorded time since stage IV diagnosis. The shaded region denotes pointwise 95% log–log confidence intervals and tick marks indicate censoring. The risk table gives the number at risk immediately before each displayed time and cumulative deaths and censorings strictly before that time. The analysis includes 94 patients, 18 deaths, and 76 censored observations. The last censoring occurs at 68 months; cumulative counts in the table extend only to 60 months. The sparse tail should not be interpreted as a precise long-term prognosis.