Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Kaplan–Meier point picker with genetic optimization enables constraint-rich individual patient data reconstruction for multicenter meta-analysis

Haiping Lu

h.lu@sheffield.ac.uk

University of Sheffield https://orcid.org/0000-0002-0349-2181

Zheqing Zhu

Institute of Big Data Science and Industry, Shanxi University https://orcid.org/0009-0009-9319-2243

Mengdie Zhuang

School of Information, Journalism and Communication, University of Sheffield

Jieru Jia

Institute of Big Data Science and Industry, Shanxi University

Fangfang Shen

Shanxi Hospital Affiliated to Cancer Hospital, Chinese Academy of Medical Sciences

xinyan liang

The Institute of Big Data Science and Industry Shanxi University No.92 Wucheng Road,Taiyuan,Shanxi province 030006 https://orcid.org/0000-0003-2589-5392

Sxe Cheong

School of Medicine and Population Health, University of Sheffield https://orcid.org/0000-0002-8571-2706

Furong Peng

Institute of Big Data Science and Industry, Shanxi University

Feijiang Li

Institute of Big Data Science and Industry, Shanxi University

Yingjie Guo

Institute of Big Data Science and Industry, Shanxi University

Hongren Yan

Institute of Big Data Science and Industry, Shanxi University

Junzhong Miao

School of Cyberspace Science, Harbin Institute of Technology

Jieting Wang

Institute of Big Data Science and Industry, Shanxi University

Mark Sammut
South Yorkshire Cardiothoracic Centre, Sheffield Teaching Hospitals NHS Foundation Trust
Min Zhang
Research Institute of Applied Biology, Shanxi University
Justin Lee
South Yorkshire Cardiothoracic Centre, Sheffield Teaching Hospitals NHS Foundation Trust
Yuhua Qian
Institute of Big Data Science and Industry, Shanxi University

---

Article

Keywords:

Posted Date: June 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8789849/v1

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

Additional Declarations: There is **NO** Competing Interest.

A Kaplan–Meier point picker with genetic optimization enables constraint-rich individual patient data reconstruction for multicenter meta-analysis

Zheqing Zhu[1,2,3], Mengdie Zhuang[4,5], Jieru Jia[1,2,3], Fangfang Shen[6], Xinyan Liang[1,2,3], Sxe Chang Cheong[7,8], Furong Peng[1,2,3], Feijiang Li[1,2,3], Yingjie Guo[1,2,3], Hongren Yan[1,2,3], Junzhong Miao[9], Jieting Wang[1,2,3], Mark Anthony Sammut[8,10], Min Zhang[11], Justin Lee[8,10], Haiping Lu[1,5,12,*], and Yuhua Qian[1,2,3,*]

[1]Institute of Big Data Science and Industry, Shanxi University, Taiyuan, Shanxi, China
[2]Key Laboratory of Evolutionary Science Intelligence of Shanxi Province, Taiyuan, Shanxi, China
[3]School of Artificial Intelligence, Shanxi University, Taiyuan, Shanxi, China
[4]School of Information, Journalism and Communication, University of Sheffield, Sheffield, UK
[5]Centre for Machine Intelligence, University of Sheffield, Sheffield, UK
[6]Shanxi Hospital Affiliated to Cancer Hospital, Chinese Academy of Medical Sciences, Taiyuan, Shanxi, China
[7]School of Medicine and Population Health, University of Sheffield, Sheffield, UK
[8]South Yorkshire Cardiothoracic Centre, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
[9]School of Cyberspace Science, Harbin Institute of Technology, Harbin, China
[10]Division of Clinical Medicine, School of Medicine and Population Health, University of Sheffield, Sheffield, UK
[11]Research Institute of Applied Biology, Shanxi University, Taiyuan, Shanxi, China
[12]School of Computer Science, University of Sheffield, Sheffield, UK
[*]Corresponding authors: Haiping Lu (h.lu@sheffield.ac.uk); Yuhua Qian (jinchengqyh@sxu.edu.cn)

ABSTRACT

Individual patient data (IPD) enable survival analysis and meta-analysis across multicenter studies, supporting evidence-based decision-making and trial design. However, openly available IPD remain scarce, and existing tools that reconstruct IPD from published Kaplan–Meier (KM) curves rely on labor-intensive, error-prone point extraction and exploit only limited subsets of reported information. We present KM-PoPiGo, an open-source, semi-automated IPD reconstruction tool that ensures accurate, quality-controlled point picking and uses genetic optimization to jointly leverage rich reported constraints, including sample size, number of events, numbers at risk, and censoring times, to reconstruct high-fidelity IPD. To standardize evaluation, we construct three benchmark datasets covering ground-truth IPD, KM curves from published studies, and synthetic KM curves, totaling over 3,000 KM curves. Results, including independent blinded clinical validation, show that KM-PoPiGo substantially improves IPD reconstruction fidelity and consistency of derived survival statistics, providing a platform for constraint-rich IPD reconstruction and multicentre meta-analysis.

Introduction

Individual patient data (IPD) refer to participant-level datasets collected in clinical studies. They capture patient characteristics, disease conditions, treatment and testing histories, and observed outcomes. IPD meta-analysis is a powerful approach for multicenter survival analysis, providing critical insights into prognostic factors and treatment effects[1–3]. Such analyses directly inform and accelerate downstream applications such as drug development[4,5] and clinical trial optimization[6,7].

Despite their value, openly available IPD remain scarce, largely due to concerns over patient privacy, data misuse, and institutional policies that limit data sharing[8,9]. In addition, researchers often have limited incentives to make their datasets publicly accessible[10,11]. A cross-sectional study of 487 clinical trials published in JAMA, Lancet, and The New England Journal of Medicine found that fewer than 5% of reviewed trials made their IPD publicly available[12]. To address these limitations, computational methods including iKM[13], IPDfromKM[14], and KMtoIPD[15] have been developed to reverse-engineer patient-level data from published Kaplan–Meier (KM) curves together with reported information that constrain reconstruction (Fig. 1a). Such information is routinely available in clinical publications (Fig. 1b).

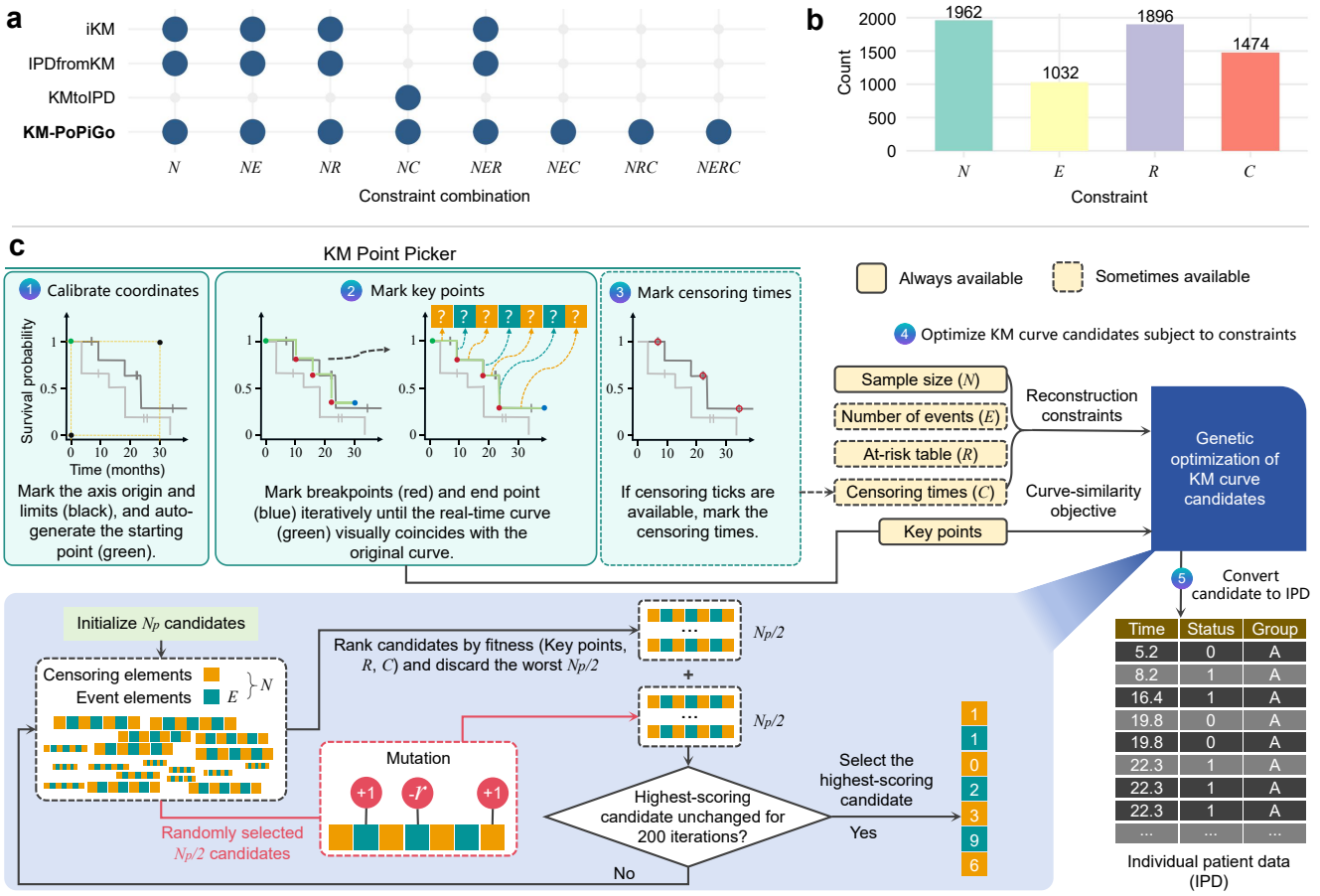


Fig. 1 | KM-PoPiGo: a Kaplan–Meier point picker with genetic optimization for constraint-rich individual patient data reconstruction. a, Comparison of supported constraint combinations across existing IPD reconstruction methods (iKM, IPDfromKM, KMtoIPD) and KM-PoPiGo. Rows correspond to methods and columns to combinations of four commonly reported survival constraints: sample size (N), number of events (E), at-risk table (R), and censoring times (C). Filled circles indicate that a method supports the corresponding constraint combination. KM-PoPiGo is the only method capable of handling all eight combinations, whereas existing approaches are restricted to specific subsets. **b**, Prevalence of reported constraints in the Published KM Dataset ($n = 1,962$ curves) from 280 published studies. Bars show the reporting frequency of N, E, R, and C. While N is reported for all curves, the frequent availability of E, R, and C demonstrates that rich survival constraints are routinely reported in clinical publications, motivating a reconstruction framework that can flexibly exploit available constraints rather than discarding valid information due to methodological limitations. **c**, Overview of the KM-PoPiGo workflow, comprising quality-controlled KM Point Picker (KM-PoPi) followed by genetic optimization enforcing reconstruction constraints. Reconstruction begins by digitally representing the published KM curve using a minimal set of key points. In Step 1, users calibrate the coordinate system by marking the origin and axis limits (black points) of the imported KM curve image to automatically determine the KM curve starting point (green point). In Step 2, users iteratively mark key breakpoints (red points), visually monitoring overlap between the reconstructed and original curves until they coincide. During this process, ineffective points are automatically removed to ensure accurate reconstruction with the fewest possible key points. These key points define intervals used to estimate event and censoring counts for IPD reconstruction. If censoring ticks are shown in the KM curve image, users can mark them (Step 3, conditional) to specify censoring times (C). Together with other reported constraints (N, E, and R, when available), these inputs define the feasible reconstruction space. Genetic optimization is then applied to identify KM curve candidates that simultaneously satisfy the imposed constraints and match the user-marked key points. Each candidate encodes event and censoring counts across M key-point intervals as a vector of $2M - 1$ elements constrained by N and E. An initial population of N_p candidates is generated with random non-negative integer elements. At each iteration, candidates are ranked using a fitness function constructed from the key points and available constraints (R, C); the least-fit $N_p/2$ candidates are then replaced by new candidates generated through mutation of $N_p/2$ randomly selected candidates, while preserving population-level constraints (N, E). This evolutionary process continues until convergence to the highest-scoring candidate, which is subsequently converted into individual patient data using the procedure described in the Methods.

However, existing IPD reconstruction methods are limited in their ability to jointly exploit the full set of reported information that constrains reconstruction. In addition, they typically require manual extraction of large numbers of coordinate points from KM curve images[13–15], a process that is labor-intensive, error-prone, and sensitive to visual challenges such as dense censoring times or overlapping curves[15]. Moreover, the lack of standardized guidance for point selection makes it challenging to identify a set of points that reliably represents the KM curve. Together, these limitations underscore the need for a more intelligent strategy that streamlines the point-selection process, incorporates built-in quality control, and systematically leverages all available constraints. Moreover, most existing methods were validated using only one or two summary statistics[13,16], and the absence of standardized benchmark datasets further hinders direct comparison across these methods.

Here, we present a Kaplan–Meier point picker with genetic optimization (KM-PoPiGo), an open-source tool with a semi-automated, quality-controlled workflow that ensures accurate point picking from KM curve images and leverages all available constraints to enable high-fidelity IPD reconstruction. Our contributions are: (i) a computer-assisted intelligent KM point picker with built-in quality control, (ii) a genetic-optimization algorithm that jointly leverages all available reporting constraints to identify KM curve candidates for IPD reconstruction, and (iii) three comprehensive benchmark datasets totaling over 3,000 KM curves that enable systematic and reproducible comparison across IPD reconstruction methods. These datasets include ground-truth IPD, real-world KM curves from published studies with reported survival statistics and constraints, and synthetic KM curves.

KM-PoPiGo can flexibly accommodate four commonly reported constraints (Fig. 1a): sample size (N), number of events (E), at-risk table (R; numbers at risk beneath the KM curve), and censoring times (C). Results, including independent, blinded clinical validation, show that KM-PoPiGo achieves high reconstruction accuracy across diverse use cases and constraint combinations, with particularly strong gains when all four constraints are jointly enforced. Consequently, KM-PoPiGo serves as a one-stop platform for high-fidelity IPD reconstruction.

Results

KM-PoPiGo workflow and benchmark datasets

Figure 1c presents the KM-PoPiGo workflow for reconstructing IPD from published KM curves. There are two main stages: (i) quality-controlled point picking, in which users digitize KM curves by marking key points and censoring times (when available) under automatically enforced KM curve rules[17–19], and (ii) constrained genetic optimization, which searches over candidate KM curves to identify the best-fitting reconstruction that matches the marked points under available constraints (N, E, R, C).

We developed three benchmark datasets to evaluate IPD reconstruction methods and their downstream applications. The first is a Real IPD Dataset comprising 231 KM curves generated from TCGA[20] and GEO[21], each paired with the original IPD and complete summary-statistic constraints (N, E, R). The second is a Published KM Dataset comprising 981 KM images (1,962 KM curves) from 280 published studies. These curves are accompanied by reported survival statistics (including hazard ratios, median survival times, and survival rates) and available reconstruction constraints (Fig. 1b), representing a cumulative reported sample size of approximately 367,000 patients. The third is a Synthetic KM Dataset consisting of 1,000 curves (500 pairs) with all summary-statistic constraints specified and corresponding survival statistics, enabling fully controlled evaluation.

To ensure consistency, all KM curves across the three datasets were annotated using a unified workflow. Sixteen trained annotators performed the initial markings, which were independently audited by a clinical expert and a computer scientist. The resulting datasets are released as structured JSON files to support reproducible benchmarking.

Reconstructed IPD enables accurate and more complete meta-analysis

Reconstruction of IPD from published KM curves expands the range of studies that can be included in meta-analysis. We evaluated KM-PoPiGo in two complementary settings: first, assessing whether reconstructed IPD retains the fidelity required for accurate pooled analyses; and second, assessing whether reconstructed IPD can recover missing prognostic information, thereby increasing the completeness of multicenter evidence synthesis.

Accuracy of reconstructed IPD for meta-analysis

We used the TCGA subset (21 KM curves) from the Real IPD Dataset to simulate a meta-analysis setting[22,23]. The cohorts were treated as concurrent with homogeneous enrollment and follow-up and divided alphabetically into two synthetic arms: Group A (10 cohorts) and Group B (11 cohorts). Reconstruction fidelity was evaluated under two scenarios: full reconstruction (100% of IPD reconstructed from images) and half reconstruction (mixing 50% original IPD with 50% reconstructed data).

Figure 2a shows that KM curves from both full and half reconstruction scenarios aligned well with the ground-truth curves (log-rank $p = 0.69$ and $p = 0.94$, respectively). Hazard ratios (HRs) from the reconstructed IPD closely matched the ground-truth HR (0.84, 95% CI 0.77-0.90), with the half reconstruction (0.84, 95% CI 0.78-0.91) slightly outperforming the full reconstruction (0.82, 95% CI 0.76-0.88) due to the partial availability of original IPD. Median survival times and 1–5-year survival rates showed similar agreement (Fig. 2b).

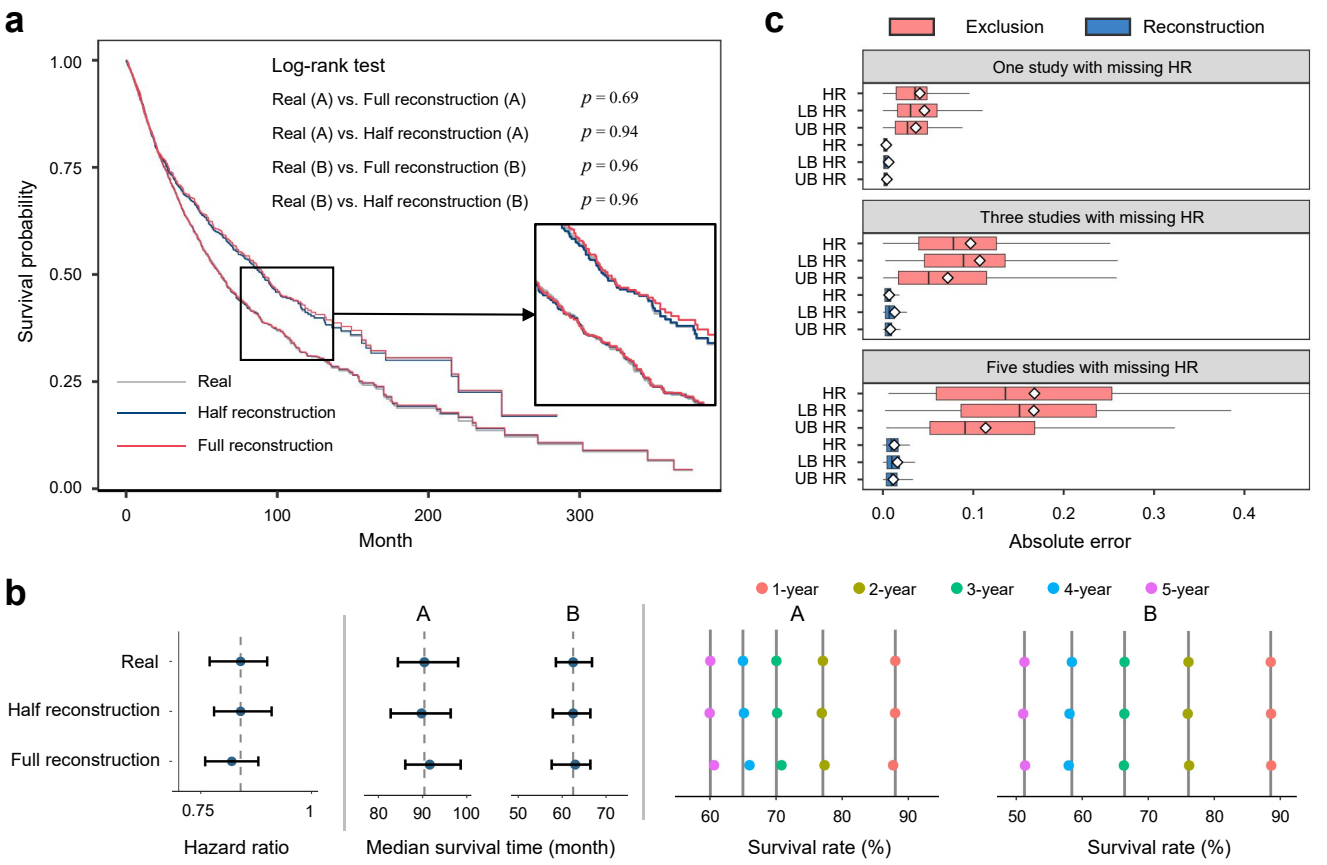


Fig. 2 | KM-PoPiGo accurately reconstructs IPD from KM curves for meta-analysis. a, The alignment of reconstructed survival curves with the ground truth. KM plots compare the "Half" (blue) and "Full" (red) reconstruction scenarios with the ground truth ("Real", grey) on the TCGA subset of the Real IPD Dataset. The zoom-in inset highlights the substantial overlap between the curves. Log-rank tests show no significant differences between the reconstructed and true distributions ($p > 0.05$ for all comparisons), confirming high reconstruction fidelity in both scenarios. The "Half" scenario shows slightly higher concordance ($p = 0.94$ vs. $p = 0.69$ for Group A), reflecting the expected benefits from partial original data availability. **b**, Quantitative comparison of survival statistics across the ground truth and two reconstruction scenarios on the same TCGA subset in **a**, including the hazard ratios (HRs), median survival times, and 1–5-year survival rates. Error bars indicate 95% confidence intervals (CIs). Vertical dashed lines (left and middle left panels) and solid grey lines (right and middle right panel) mark ground-truth values derived from the real IPD, highlighting strong overlap and high fidelity of the reconstructed estimates. **c**, Restoring prognostic accuracy in meta-analyses with missing hazard ratios. Box plots display the distribution of absolute errors for the pooled HR and its 95% CI bounds (LB: lower bound; UB: upper bound) across 100 independent simulations. For each simulation, we sampled a unique subset of 10 studies from the Published KM Dataset and calculated its reference pooled HR using all originally reported data. We then masked the reported HRs for 1, 3, or 5 studies within each subset and calculated the deviation from the reference pooled HR (where absolute error = 0 indicates perfect alignment). The comparison shows the error distributions for meta-analyses performed by excluding studies whose HRs are missing ("Exclusion", pink) versus supplementing the meta-analysis with KM-PoPiGo reconstructions for HR-missing studies ("Reconstruction", blue). Box edges represent the interquartile range (IQR), center lines denote medians, and diamond markers denote means; the shift of the blue distributions towards zero indicates that reconstruction restores estimates close to the reference. See Supplementary Fig. 2 for a representative simulation with 5 studies excluded due to missing HRs.

Reconstructed IPD recovers missing prognostic information

Meta-analyses of prognostic effects in multicenter studies commonly rely on reported aggregate HRs[24,25]. However, HRs are frequently unreported in published studies[26], leading to study exclusion and reduced statistical power. By reconstructing IPD from KM images, KM-PoPiGo enables HR estimation for such studies and their incorporation into pooled analyses.

To evaluate this capability, we performed 600 simulated meta-analyses across 100 randomly sampled sets of ten KM images from the Published KM Dataset, with each image representing one study and containing a pair of KM curves corresponding to

two study arms. In each set, we simulated incomplete reporting by withholding the reported HRs from 1, 3, or 5 studies. We then compared pooled estimates obtained by excluding HR-missing studies against those supplemented with HRs estimated from reconstructed IPD, using the pooled HR computed from all ten studies with the originally reported HRs as the reference.

Across all missing-HR scenarios, supplementing meta-analyses with reconstructed IPD substantially reduced estimation error (Fig. 2c). Reconstruction for a single HR-missing study reduced the mean absolute error (MAE) of the HR by 87.9% ($p < 0.001$), while reconstruction for three or five HR-missing studies yielded MAE reductions of 88.4% and 90.0%, respectively, as shown by the leftward shift of the mean markers (diamonds) with reconstruction in Fig. 2c. Similar improvements were observed for the lower and upper bounds of the pooled HR confidence intervals.

Synthetic controls and cross-trial comparison using reconstructed IPD

Randomized controlled trials (RCTs) remain the gold standard for evaluating therapeutic efficacy. Yet, high costs, lengthy recruitment periods, and ethical concerns about randomization have motivated growing interest in synthetic control groups[27–29]. By reconstructing IPD from published KM curves, KM-PoPiGo provides a practical route to generating synthetic controls and enabling quantitative comparisons across interventions evaluated in separate trials.

Reconstructed IPD accurately emulates real control arms

We first evaluated whether KM-PoPiGo can generate reliable synthetic control groups (Fig. 3a) using all six cohorts from the GEO subset of the Real IPD Dataset that meet the inclusion criteria of fewer than 200 participants and follow-up durations of 50–60 months (Supplementary Table 6), mimicking data-limited scenarios typical of early-phase trials or rare disease studies where external controls are most needed[30]. We assumed that these cohorts represented clinically comparable populations with similar follow-up structures, allowing fair evaluation of synthetic control performance. One cohort (GSE27150[31]) was treated as a single-arm experimental group, and the remaining five served as real control groups. For each of these five controls, we reconstructed IPD from their KM curve images to produce corresponding synthetic control groups. HRs comparing the experimental group against each real versus synthetic control group were nearly identical (Pearson correlation = 0.99, $p < 0.001$; Fig. 3b).

Reconstructed IPD enables cross-trial comparison of multiple interventions

Many diseases have multiple treatment options evaluated across different RCTs, yet direct comparison is often limited because not all interventions are tested head-to-head. KM-PoPiGo addresses this by reconstructing IPD for each intervention arm and pooling IPD from trials involving the same intervention to increase statistical power (Fig. 3c).

To study cross-trial comparison, we identified sixteen independent cohorts from the GEO subset of the Real IPD Dataset with follow-up durations of 50–70 months (centered on the 5-year benchmark[32]). To ensure patient independence, we selected a single survival endpoint per cohort, prioritizing the largest sample size where multiple endpoints qualified[33,34]. These cohorts were aggregated into four simulated intervention arms (Supplementary Table 7), creating a controlled scenario to test whether reconstruction preserves comparative signals. IPD was reconstructed for each arm using the respective KM curve image, N, E, and R. Pooled reconstructed IPD enabled pairwise HR comparisons between interventions, even for intervention pairs never compared directly in a single trial. The resulting HR rankings (Intervention 4 best; Intervention 1 worst) matched those obtained using the real IPD (Fig. 3d).

Impact of constraint combinations on reconstruction fidelity

To quantify how heterogeneous constraint availability affects reconstruction fidelity, we evaluated eight constraint combinations over N, E, R, and C, reflecting common reporting patterns in published studies where N is always available and the remaining constraints are variably reported (Fig. 1b). Reconstruction fidelity was quantified against ground truth using distributional and summary-statistic metrics, including Weibull distribution error[35], HR error[36,37], median survival time error[38], and survival rate error[39].

Evaluation of reconstruction fidelity

On the Real IPD Dataset, we quantified reconstruction fidelity using the geometric mean of the Weibull distribution error, which mitigates the influence of large error magnitudes across curves. We observed a stepwise improvement in reconstruction fidelity as more constraints were incorporated (Fig. 4a). The full constraint combination ($NERC$) achieved the lowest geometric mean error, corresponding to a 45.1% relative improvement over the baseline configuration (N). Constraint combinations including both E and C (NEC and $NERC$) were more frequently ranked among the highest-performing reconstructions than simpler configurations. Visual comparison of fitted distributions on the TCGA subset further confirmed strong distributional agreement, where the reconstructed Weibull distributions across all constraint combinations fall within the 95% confidence interval bands of the ground-truth Weibull distribution (Supplementary Fig. 4).

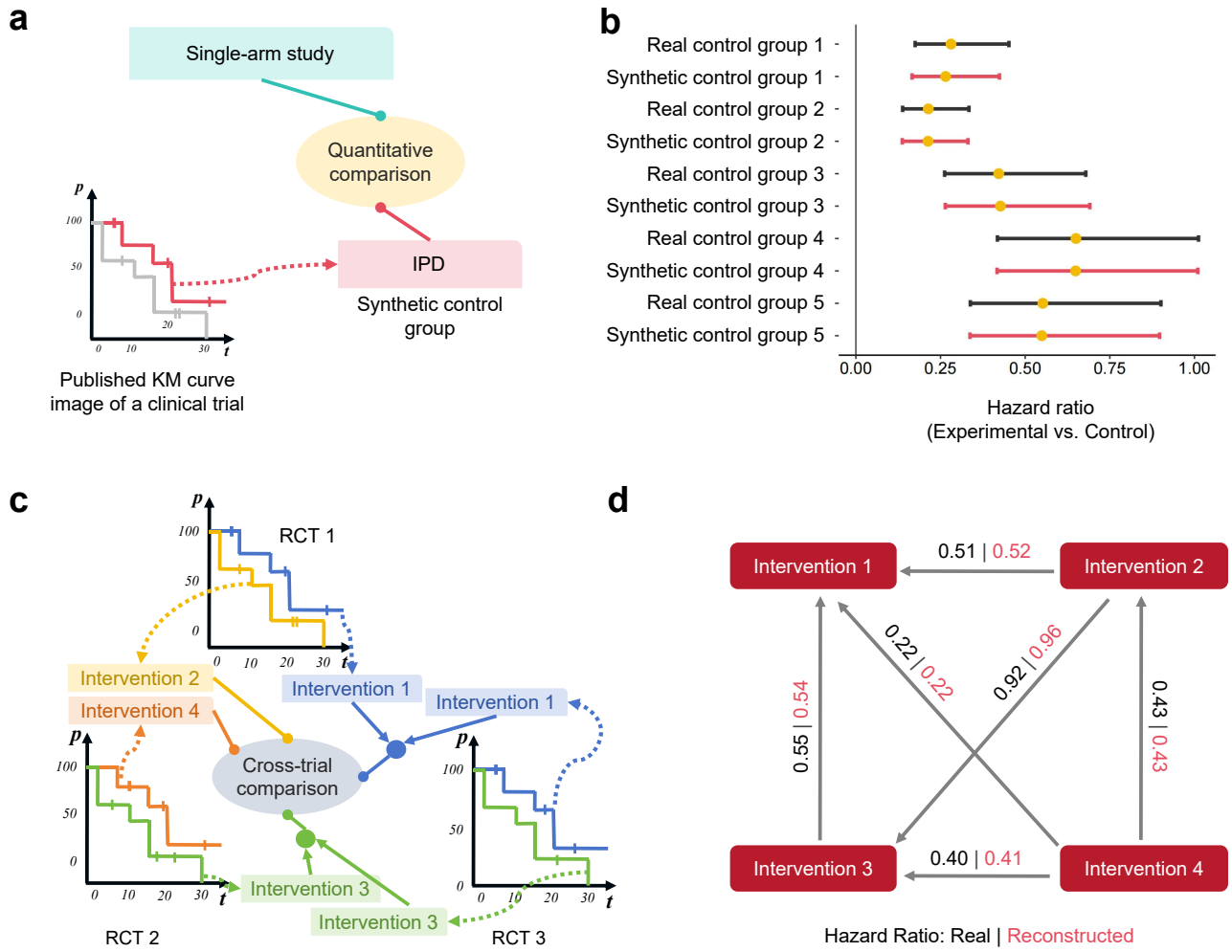


Fig. 3 | Reconstructed IPD enables synthetic controls and cross-trial intervention comparisons. a, Illustration of generating a synthetic control group by reconstructing IPD from a published KM curve image, enabling quantitative comparison with an ongoing single-arm study. **b**, Evaluation of synthetic controls in enabling quantitative comparisons on six cohorts of the GEO subset of the Real IPD Dataset (see Supplementary Table 6 for cohort details). One cohort (GSE27150) serves as the experimental group, and five additional cohorts serve as the real control groups. For each control cohort, IPD was reconstructed from respective KM curve images to form synthetic controls. HRs comparing the experimental group with each real versus synthetic control group show near-identical estimates (yellow dots: HRs; black lines: 95% CI for real controls; red lines: 95% CI for synthetic controls). **c**, Illustration of cross-trial comparison where reconstructed IPD allows pooling data from disparate RCTs (3 trials, 6 arms) to enable quantitative comparison among interventions (4 types), including those never compared head-to-head. **d**, Quantitative cross-trial intervention comparison using sixteen cohorts from the GEO subset of the Real IPD Dataset, aggregated into four simulated intervention arms (Interventions 1–4; see Supplementary Table 7 for arm configurations). Pairwise HR estimates derived from reconstructed IPD exhibited high concordance with those obtained from real IPD, accurately capturing the comparative effectiveness profiles among the four interventions.

Accuracy of reconstructed survival statistics

Using the Synthetic KM Dataset, we analyzed the impact of constraint combinations on the accuracy of reconstructed survival statistics. Figure 4b shows that the log HR error is highest when reconstruction is constrained only by N (MAE: 0.05). The error decreases progressively as additional constraints are incorporated, reaching its minimum after incorporating all constraints (MAE: 0.03). Including R has a relatively modest effect on MAE, whereas including E substantially reduces both the MAE and the variance of HR errors (Fig. 4b).

In contrast, median survival time and survival rate estimates show no systematic improvement across different constraint combinations (Fig. 4c–d). Further analysis indicates that large outliers arise primarily from small local deviations of the

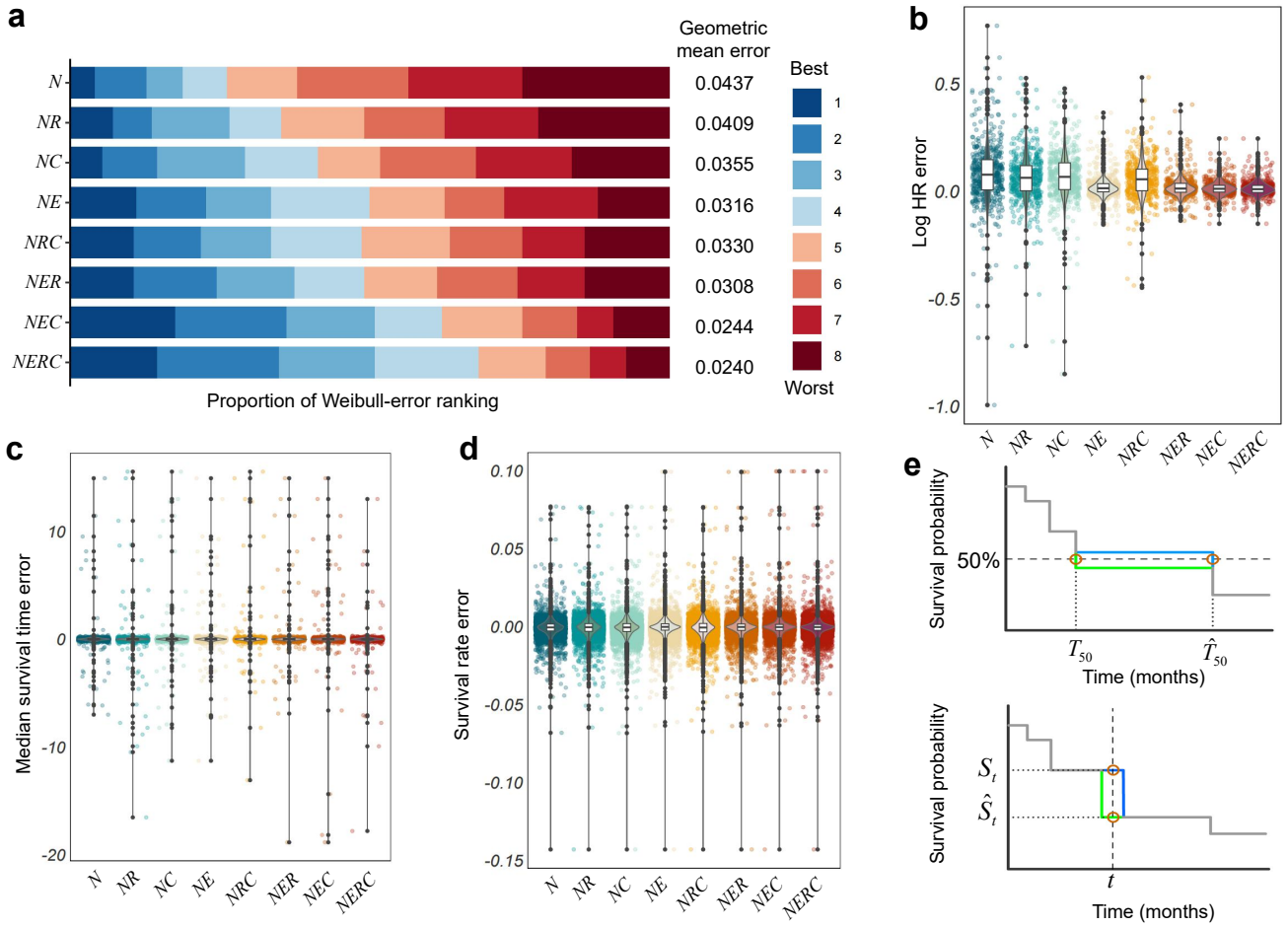


Fig. 4 | Impact of constraint combinations on reconstruction fidelity and survival statistics. a, Weibull distribution error and reconstruction ranking on the Real IPD Dataset. Stacked bars (left) show the distribution of accuracy rankings across constraint combinations, while text labels (right) indicate the geometric mean Weibull distribution error. Increasing the number of constraints reduces the geometric mean error (from 0.0437 for *N* to 0.0240 for *NERC*) and increases the proportion of higher-ranking reconstructions (even with minimal constraints, reconstructions remain within the original 95% CI; see Supplementary Fig. 4 for the TCGA-IPD subset). **b**, HR error analysis on the Synthetic KM Dataset, which enables systematic evaluation across all constraint combinations. The distributions of log HR errors show progressive reduction in error variance as additional constraints are incorporated (from *N* to *NERC*), leading to more precise and robust HR estimates. See Supplementary Fig. 5 for error distributions and Supplementary Table 8 for pairwise statistical comparisons. **c**, Median survival time error analysis on the Synthetic KM Dataset. Error distributions remain relatively stable across constraint combinations, with distributions centered near zero, although notable outliers persist under all configurations. See Supplementary Tables 9 and 10 for pairwise statistical comparisons of survival point estimates. **d**, Survival rate error analysis on the Synthetic KM Dataset. Error distributions are largely comparable across constraint combinations, with isolated large deviations observed in a small subset of cases regardless of the constraints applied. **e**, Mechanisms of large error generation. Representative cases illustrate sources of large errors in median survival time (T_{50} vs. $\hat{T}_{50}$) and survival rate (S_t vs. $\hat{S}_t$). The ground truth is shown in green and the reconstruction in blue. Flat plateaus near the 50% survival threshold amplify errors in median time estimates, while steep step-downs near evaluation times make survival rate estimates highly sensitive to small temporal shifts.

reconstructed KM curves near the 50% survival threshold (for median survival time) or near the evaluation time points (for survival rate estimates) (Fig. 4e). In such regions, even small vertical shifts can induce large errors in median survival time, while small horizontal displacements can markedly affect survival rate estimates.

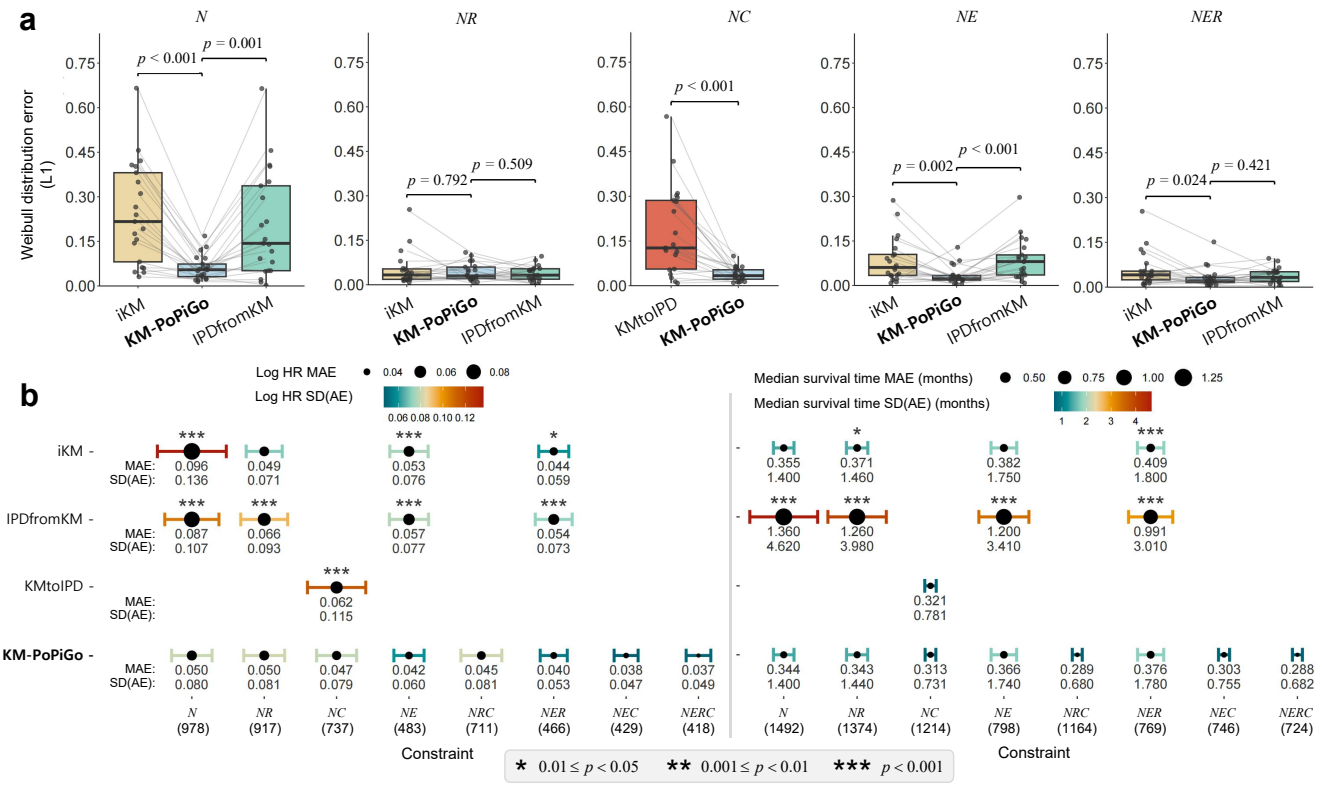


Fig. 5 | Benchmarking KM-PoPiGo's reconstruction fidelity against iKM, IPDfromKM, and KMtoIPD. a, Reconstruction error analysis on the TCGA subset of the Real IPD Dataset. Boxplots show the distribution of Weibull distribution errors (L1 distance) for KM-PoPiGo and existing approaches under identical constraint combinations. Gray lines connect paired errors reconstructed from the same source KM curve. Box limits indicate the interquartile range (IQR); center lines denote medians; *p*-values indicate significance from paired *t*-tests on log-transformed errors. **b**, Survival statistics error analysis on the Published KM Dataset. Evaluation of KM-PoPiGo against existing methods across applicable constraint combinations for hazard ratios (left panel) and median survival times (right panel). In both plots, dot size represents the mean absolute error (MAE), while color intensity and horizontal bar width indicate the standard deviation of absolute errors (SD(AE)). Asterisks denote statistically significant performance differences relative to KM-PoPiGo. Numbers in parentheses below each constraint combination indicate the count of verifiable survival statistics in the common set. See Supplementary Fig. 6 for survival rate results, and Supplementary Table 11 for the sample counts of the common set used for fair benchmarking.

Benchmarking against existing IPD reconstruction methods

We benchmarked KM-PoPiGo against three widely used IPD reconstruction methods: iKM[13], IPDfromKM[14], and KMtoIPD[15], across the constraint combinations supported by each method (Fig. 1a). Benchmarking was restricted accordingly, reflecting method-specific support for different subsets of the four constraints (N, E, R, C).

Comparative analysis of reconstruction fidelity

On the TCGA subset of the Real IPD Dataset, KM-PoPiGo demonstrated comparable or superior reconstruction fidelity across supported constraint combinations based on the geometric mean of the Weibull distribution error (Fig. 5a). Its advantages were most pronounced in scenarios with limited available constraints, particularly when the only reconstruction constraint was N. In this setting, where KMtoIPD is not applicable due to its reliance on C, KM-PoPiGo (error = 0.049) significantly outperformed iKM (error = 0.178, $p < 0.001$, 72.6% improvement) and IPDfromKM (error = 0.116, $p < 0.001$, 57.9% improvement).

In the *NC* scenario, benchmarking was performed against KMtoIPD, the only other method supporting this constraint combination (Fig. 1a). Under this constraint combination, KM-PoPiGo (error = 0.031) achieved a 73.0% improvement over KMtoIPD (error = 0.117, $p < 0.001$), highlighting its effectiveness in leveraging censoring information. In the *NER* scenario, performance differences narrowed: KM-PoPiGo maintained a significant advantage over iKM ($p = 0.024$, 45.2% improvement) while performing comparably to IPDfromKM ($p = 0.421$).

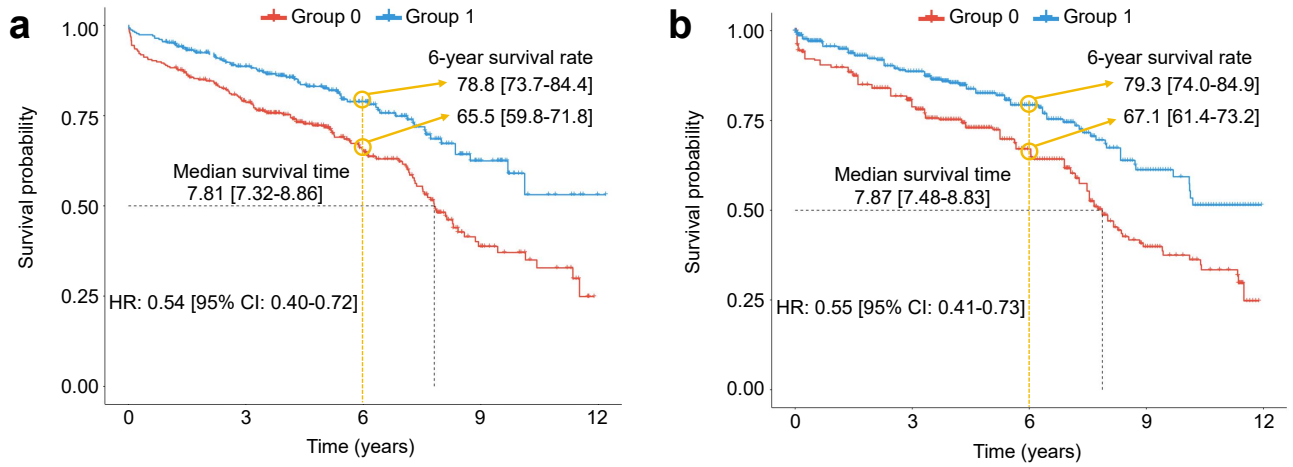


Fig. 6 | Performance of KM-PoPiGo in independent blinded clinical validation. a, KM survival curves derived from the original institutional clinical IPD for 622 patients across two cohorts. Key prognostic statistics, including the hazard ratio (HR) between groups, 6-year survival rates, and median survival times (Group 0), are annotated. **b**, KM curves derived from IPD reconstructed by an independent validator using KM-PoPiGo from KM curve images generated from the original IPD. The close alignment between reconstructed and original curves, together with the consistency of derived survival statistics, demonstrates accurate reconstruction in a blinded, deployment-style validation setting (further details in Supplementary Fig. 7).

Comparative analysis of reconstructed survival statistics

We further assessed the accuracy of derived survival statistics, including HRs, median survival times, and survival rates, using the Published KM Dataset. Figure 5b (left) shows that KM-PoPiGo consistently yields lower HR errors than existing approaches under identical constraint combinations, as measured by both mean absolute error (MAE) and standard deviation of absolute errors (SD(AE)). Under the *NER* constraint, KM-PoPiGo (MAE = 0.040) achieved a 11.5% improvement over iKM (MAE = 0.044, $p = 0.03$) and significantly outperformed IPDfromKM (MAE = 0.054, $p < 0.001$). KM-PoPiGo also attained the lowest MAE overall (0.037) when using the *NERC* combination.

For median survival time, KM-PoPiGo produced errors that were either statistically significantly smaller or statistically indistinguishable from those of competing methods under identical constraint combinations (Fig. 5b, right; specifically *N*, *NE*, *NR*, and *NER*). Similar patterns were observed for survival rate estimates (Supplementary Fig. 6).

Independent blinded clinical validation on unseen data

To assess the usability of KM-PoPiGo under deployment-style clinical settings, we conducted an independent, blinded validation study performed by a clinical research team at the School of Medicine and Population Health, University of Sheffield, and Sheffield Teaching Hospitals NHS Foundation Trust. None of the three clinical research team members was involved in the development of KM-PoPiGo or had prior experience using the software. The evaluation involved IPD reconstruction for 622 patients across two independent clinical cohorts under an approved ethics protocol (University of Sheffield ethics application number 052104).

Two senior clinicians provided institutional clinical IPD for the two cohorts, from which corresponding KM curve images were generated. A senior medical student, independent of the development team, then used KM-PoPiGo to reconstruct IPD from these KM images without developer assistance; the KM-PoPiGo development team had no access to the original IPD.

To ensure standardized and reproducible evaluation under deployment-style conditions, the development team provided a Validation Kit comprising a user manual and a standardized R benchmarking script. Users were instructed to format the original IPD using a predefined schema, reconstruct IPD from the corresponding KM images using KM-PoPiGo without developer assistance, and run a standardized benchmarking script to compute evaluation metrics. This workflow enabled automated generation of evaluation outputs, including KM curves, Weibull distributions, hazard ratios, median survival times, and survival rates.

Figure 6 shows that KM curves reconstructed by the independent validator are closely aligned with those derived from the original IPD. Derived Weibull distributions and key prognostic indicators, including hazard ratios, median survival times, and survival rates, were consistent with the corresponding ground-truth values (Supplementary Fig. 7). These results demonstrate that KM-PoPiGo can be used by new users and supports deployment in routine clinical analysis workflows.

Discussion

We introduced KM-PoPiGo, an open-source framework that combines quality-controlled point picking with global optimization that jointly incorporates multiple reported constraints to address the inherent ambiguity in reconstructing IPD from Kaplan–Meier curves[40]. This design enables recovery of high-fidelity IPD consistent with observed survival curves and reported constraints, even under heterogeneous constraint availability across studies (Fig. 1b).

We advocate survival reporting practices to include censoring times (C) alongside the number of events (E) as our systematic evaluation identifies them as key constraints governing reconstruction fidelity (Fig. 4b). Hazard ratio (HR) analysis (Supplementary Table 8) shows that incorporating E alongside sample size (NE) substantially reduces HR estimation error, with further improvement when adding censoring times (NEC). In contrast, adding the at-risk table (NER) does not yield additional gains over NE alone.

We found that the utility of different constraints depends on the intended analysis. For descriptive statistics such as median survival times and survival rates, the sample size constraint alone often yields good accuracy in our studies, whereas adding other constraints does not improve accuracy much. There is also geometric sensitivity inherent to the Kaplan–Meier estimator, whereby small vertical shifts in the reconstructed curve can induce disproportionate errors in median survival times, and subtle horizontal displacements can result in substantial deviations in survival rates.

KM-PoPiGo helps address the "data silo" problem in medical research by enabling the reuse of clinical evidence where direct IPD sharing is complicated, restricted, or infeasible due to data protection regulations[41–43]. By leveraging KM curve images together with rich, routinely reported survival constraints, KM-PoPiGo facilitates the assembly of large-scale, standardized IPD resources capturing substantially richer historical evidence. Such resources can summarize outcomes from hundreds of thousands of patients reported in the literature to construct synthetic control arms in single-arm trials. This is particularly useful in rare diseases and pediatric oncology, where recruitment is particularly challenging[28,44,45]. These reconstructed datasets offer a robust evidence base for clinical decision-making and trial design[46] without the ethical, logistical, and financial burden of recruiting new control groups.

In contrast with sequential estimation in existing methods[13–15], KM-PoPiGo enforces KM-curve rules and jointly incorporates reported constraints within genetic optimization to achieve high-fidelity reconstruction. This approach can be adapted to digitize other longitudinal biomedical data, such as pharmacokinetic profiles or epidemiological time series. Moreover, while the genetic optimization in KM-PoPiGo focuses on point estimation, its ability to explore multiple plausible reconstructions provides a foundation for future extensions toward uncertainty-aware evidence synthesis.

KM-PoPiGo executes entirely on the client side and retains full traceability of user annotations and reconstruction outputs, improving reproducibility while enabling privacy-preserving reuse of published KM curves. More broadly, KM-PoPiGo illustrates how static graphical outputs can be transformed into analyzable, reproducible data, unlocking decades of historical clinical evidence for modern, transparent statistical analysis.

Methods

Semi-automated point selection

Mathematical representation of KM curves

A Kaplan-Meier (KM) curve is geometrically defined as a monotonic step function. We refer to the collection of step heads and the end point as the key points (Supplementary Fig. 8a), the minimal set of coordinates required to describe the curve's geometry. This set comprises three point types: the starting point, the breakpoints (representing event occurrences), and the end point (representing the final follow-up).

The curve is partitioned into M intervals by the key points. We denote the coordinates of the key point set as a sequence $\{(t_i, p_i)\}_{i=0}^{M}$, where t and p represent time and survival probability, respectively. Specifically, $(t_0, p_0) = (0, 1)$ corresponds to the starting point, (t_i, p_i) for $i \in \{1, \ldots, M-1\}$ represents the i-th breakpoint, and (t_M, p_M) denotes the end point. Two boundary scenarios exist for the end point: either $t_M > t_{M-1}$ (survival plateau, Supplementary Fig. 8a) or $t_M = t_{M-1}$ (event at termination, Supplementary Fig. 8b, c). Collectively, $\boldsymbol{t} = [t_0, \cdots, t_M]^T$ and $\boldsymbol{p} = [p_0, \cdots, p_M]^T$ represent the vectors of key times and survival probabilities, respectively. Additionally, the minimum number of censoring times marked within each interval constitutes the censoring constraint vector $\boldsymbol{c}^{min} \in \mathbb{R}^M$ (Supplementary Fig. 8d).

Automated quality control based on KM properties

We extract the key points using the KM Point Picker, a semi-automated module within the KM-PoPiGo framework. This module includes an automated quality control system that maintains data integrity by enforcing fundamental geometric constraints, consisting of the following geometric rules[17,18]:

- **Initialization:** The starting point is automatically locked to $(0, 1)$ in the coordinate system.

- **Uniqueness:** Key points must have distinct t-coordinates (time), and both the starting point and breakpoints must have distinct p-coordinates (survival probability).
- **Coordinate ranges:** All coordinates must satisfy $t \geq 0$ and $0 \leq p \leq 1$.
- **Monotonicity:** The p-coordinate of any breakpoint must be strictly lower than that of its preceding neighbor.
- **Continuity:** The p-coordinate of the end point must be equal to that of the last breakpoint (representing the final horizontal step).

Upon completion of marking, the rightmost point is designated as the end point. If a user places an end point with a p-coordinate lower than the preceding breakpoint (implying a vertical drop at the end), the algorithm automatically inserts a necessary breakpoint to maintain structural integrity (Supplementary Fig. 8b, c).

To address inaccuracies in manual cursor positioning and maintain valid curve topology, we build an auto-correction mechanism, a real-time rectification algorithm that maps digitized points onto a non-increasing step function. We introduce tolerance thresholds α and β (empirically set to 1% of the t- and p-axis ranges, respectively) as resolution constraints. These thresholds define the minimum distinguishable scale, preventing over-segmentation where point density exceeds the actual event count. The mechanism operates as follows:

- **Boundary correction:** For each point marked by the user, coordinates with $t_i \leq 0$ are corrected to α to ensure a positive time increment relative to the starting point, while $p_i < 0$ is clamped to 0.
- **Monotonicity enforcement:** When a new point (t_i, p_i) is marked (provided it is not the rightmost point), the mechanism designates it as the step-head for that survival level. It removes any preceding points with $p < p_i + \beta$ and any succeeding points with $p > p_i - \beta$. This greedy filtering allows users to correct the curve geometry so that each selected point represents a distinct, non-increasing step.
- **End point alignment:** When the added point is the rightmost one, any pre-existing end point is removed. If its p-coordinate is lower than the preceding key point (implying a vertical drop), this point is designated as the end point, while a breakpoint is simultaneously inserted at the same location. This creates a zero-duration interval between the breakpoint and the end point to capture potential censoring counts at this final instant. Conversely, if its p-coordinate is greater than or equal to that of the preceding key point (indicating the user's intent to mark a plateau), the added point becomes the end point, with its p-coordinate adjusted to match the preceding key point.

Throughout the marking process, the system generates a semi-transparent green curve from the marked key points, overlaying the original curve in real time for visual verification. Once the key points are finalized, users can also mark censoring times (C) and input summary-statistic constraints (N, E, R). Users can then export the final annotation data, including coordinates, constraints, and the Base64-encoded source image, as a structured JSON file. This design promotes data portability, allowing these files to serve as a digital archive for collaboration while enabling others to reproduce the annotation session to enhance transparency and reproducibility.

Genetic optimization for constraint-rich IPD reconstruction

The reconstruction aims to estimate the censoring counts $c_i^{'}$, $i \in \{1, \cdots, M\}$, in each time interval of $\boldsymbol{t}$ and the event count $e_j^{'}$, $j \in \{1, \cdots, M-1\}$, at the breakpoints. Combining the sample size N reported in the published studies and the $\boldsymbol{p}$ obtained from the key point set, we can construct an underdetermined system of equations (Supplementary Note 9). This means multiple sets of IPDs can satisfy the constraints imposed by N and $\boldsymbol{p}$, resulting in multiple possible solutions, each solution is referred to as a KM curve candidate.

To search for the best-fitting reconstruction, we incorporate additional available constraints, including E, R, and C, with a dual strategy: (1) For structural hard constraints (N, and E if available), the candidate's vector representation and a constrained mutation operator enforce N and E to ensure structural validity during evolution; (2) For fitness-based soft constraints (R, C) and curve similarity objective ($\boldsymbol{p}$), a composite fitness function incorporates R and/or C alongside $\boldsymbol{p}$. This function guides the search by penalizing candidates that deviate from the available constraints.

Representation of the candidate

The KM curve solution vector (representing one candidate arm) to be optimized by the genetic algorithm is denoted as:

$$\boldsymbol{\theta} = (c_1^{'}, e_1^{'}, \cdots, c_{M-1}^{'}, e_{M-1}^{'}, c_M^{'}).$$

Here, $c_i^{'} \geq 0$, $i \in \{1, \cdots, M\}$, represents the censoring count during the period from t_{i-1} to t_i. The element $e_j^{'}$ represents the event count occurring at time t_j, $j \in \{1, \cdots, M-1\}$. Since t_j corresponds to a vertical drop (breakpoint) in the KM curve, $e_j^{'}$ is

constrained to be a positive integer ($e'_j \geq 1$). In this representation, the positions of $c^{'}$ elements constitute the censoring field, while those of $e^{'}$ elements constitute the event field. The sum of elements in the KM curve candidate, given by $\sum_{i=1}^{M} c'_i + \sum_{j=1}^{M-1} e'_j$, must be equal to N. If E is available, the sum of event elements $\sum_{j=1}^{M-1} e'_j$ must be equal to E.

Constrained mutation

The crossover operator is not used to preserve hard constraints (N, E), since standard crossover disrupts the sum invariants. Candidates evolve through a constrained mutation process. To redistribute values in a closed loop, the value subtracted from one element is added to other elements to keep the sum invariant. This procedure involves two steps (Supplementary Algorithm 3):

Reductive step: Select a random position k and subtract an integer r from the element θ_k, determined as follows:

- If k corresponds to a censoring element, to balance local optimization with global exploration, we set $r = \min(1, \theta_k)$ with 50% probability; otherwise, r is drawn uniformly from $\{0, \ldots, \theta_k\}$, ensuring that the updated element value remains non-negative ($\theta_k - r \geq 0$).
- If k corresponds to an event element, to balance local optimization with global exploration, we set $r = \min(1, \theta_k - 1)$ with 50% probability; otherwise, r is drawn uniformly from $\{1, \ldots, \theta_k - 1\}$, ensuring that the updated element value remains at least one ($\theta_k - r \geq 1$).

Restoration step: To restore the sum to N, the subtracted value r is redistributed by incrementing r randomly selected positions by 1.

- **Scenario A** (E is available): To keep the sum of event elements equal to E, the redistribution is field-restricted. If the reduced element at position k was a censoring element, the r increments are applied to the censoring field; if it was an event element, they are applied to the event field.
- **Scenario B** (E is unavailable): The r increments are applied to positions chosen uniformly at random from the entire vector.

Candidate fitness function

While the constrained mutation ensures adherence to structural hard constraints (N,E), the fidelity of the reconstruction is evaluated via a composite fitness function $f(\boldsymbol{\theta})$ aggregating up to three terms: a primary objective term for curve similarity (L_{curve}) and two soft-constraint penalty terms for C (L_C) and R (L_R), normalized by their scaling factors:

$$f(\boldsymbol{\theta}) = L_{\text{curve}} + \frac{1}{M} L_C + \frac{1}{MJ} L_R, \tag{1}$$

where M is the number of intervals partitioned by the key points and J is the number of reported time points in the at-risk table. The normalization by M and MJ converts the violations into normalized averages, so that the magnitudes of the soft constraint penalties are commensurate with the primary fitting objective. If C (or R) is unavailable, L_C (or L_R) is set to 0.

The curve similarity objective L_{curve} quantifies the discrepancy between the survival probabilities derived from a candidate and those extracted from the published KM curve to evaluate the fidelity of the survival characteristics. Each candidate contains the information to reconstruct a KM curve. We calculate the survival probability at each breakpoint using the candidate vector. According to the KM definition, we set the initial survival probability $p^{'}_0 = 1$, the initial number at risk $l_{R_0} = N$, and the initial event count $e^{'}_0 = 0$. The number at risk at the i-th time point is updated recursively:

$$l_{R_i} = l_{R_{i-1}} - e^{'}_{i-1} - c^{'}_i, \quad \text{for } i \in \{1, \ldots, M-1\}. \tag{2}$$

The survival probability at the i-th breakpoint is calculated as:

$$p^{'}_i = p^{'}_{i-1} \frac{l_{R_i} - e^{'}_i}{l_{R_i}}. \tag{3}$$

We compare the survival probabilities derived from the KM candidate $\{p^{'}_i\}$ with those extracted from the published KM curve $\{p_i\}$ using the sum of absolute differences as:

$$L_{curve} = \sum_{i=1}^{M-1} |p_i - p^{'}_i|. \tag{4}$$

The at-risk table constraint L_R evaluates the consistency between the candidate's risk profile and the reported number at risk at specific time points, computing the cumulative reduction and the residual risk magnitude, penalizing deviations from the constraints. Given the at-risk table data $\{(\hat{t}_j, \hat{n}_j)\}_{j=1}^{J}$ (where $\hat{t}_j$ is the time and $\hat{n}_j$ is the number at risk), we align the KM candidate with these reported time points by segregating the time indices into two disjoint sets of "past" and "future" relative to each $\hat{t}_j$:

$$\begin{cases} \mathbf{I}_<(j) &= \{i \mid t_i < \hat{t}_j\} \\ \mathbf{I}_\geq(j) &= \{i \mid t_i \geq \hat{t}_j\}. \end{cases} \tag{5}$$

Following the principles of the KM estimator, the censoring and event elements associated with these partitions satisfy:

$$\begin{cases} \sum_{i=1}^{k} (c_i' + e_i') \leq N - \hat{n}_j, & \forall k \in \mathbf{I}_<(j) \\ \sum_{i=k}^{M-1} (c_{i+1}' + e_i') \leq \hat{n}_j, & \forall k \in \mathbf{I}_\geq(j). \end{cases} \tag{6}$$

These inequalities enforce structural consistency across the timeline. The first constraint dictates that for any time point $t_k < \hat{t}_j$, the sum of events and censorings accumulated up to t_k does not exceed the allowable loss $(N - \hat{n}_j)$. The second constraint bounds the aggregate of events and censorings from t_k onwards by the reported risk count $\hat{n}_j$. To enforce these constraints, L_R aggregates the violations across the past and future periods for each reported time point $\hat{t}_j$:

$$L_R = \sum_{j=1}^{J} \left[\sum_{k \in \mathbf{I}_<(j)} \max\left(\sum_{i=1}^{k} (c_i' + e_i') - (N - \hat{n}_j), 0 \right) + \sum_{k \in \mathbf{I}_\geq(j)} \max\left(\sum_{i=k}^{M-1} (c_{i+1}' + e_i') - \hat{n}_j, 0 \right) \right]. \tag{7}$$

This formulation aims for global consistency rather than sequential estimation to mitigate the error propagation in iterative reconstruction methods[13] and reduce tail instability arising from small risk sets[47]. The max function ensures that penalties are incurred when constraints are violated.

The censoring constraint L_C uses the count of marked censoring times to constrain the candidate solution. Given visual occlusion where censoring ticks may be superimposed, this count constitutes a lower bound rather than a total. The constraint enforces that the censoring count in each interval is no less than the count identified in that interval. The censoring times C are aggregated into interval-specific minimum counts c_i^{min} (Supplementary Note 8), where c_i^{min} denotes the count of censoring times falling within the i-th interval. To enforce this lower bound, L_C is formulated as the shortfall in terms of the corresponding censoring count c_i':

$$L_C \triangleq \sum_{i=1}^{M} \max\left[c_i^{min} - c_i^{'}, 0 \right]. \tag{8}$$

The max operator penalizes deficits where c_i' falls below c_i^{min}, while incurring no penalty for counts exceeding this lower bound.

Evolutionary optimization

The genetic algorithm is initialized by generating a population of 100 random KM candidates (Supplementary Algorithm 2) that satisfy the structural hard constraints. The fitness of each candidate is evaluated using the composite fitness function $f(\boldsymbol{\theta})$.

In each evolutionary iteration, the population is ranked based on $f(\boldsymbol{\theta})$. We select 50 candidates uniformly at random to undergo constrained mutation. The resulting mutated offspring then replace the 50 least-fit candidates (those with the highest loss values) in the population. This replacement strategy[48,49] ensures that while the algorithm explores the solution space via random selection, it drives the population toward optimal solutions by eliminating inferior solutions. The algorithm maintains a record of the highest-scoring candidate, updating it when an improved offspring is generated, until convergence, where this record remains unchanged for 200 consecutive cycles (Supplementary Algorithm 1).

Conversion of the highest-scoring candidate to IPD

Upon termination of the genetic optimization, the highest-scoring candidate is converted into the reconstructed IPD. As the candidate vector contains the censoring count $c_i^{'}$ for each interval $(t_{i-1}, t_i]$ and event count $e_i^{'}$ for each breakpoint t_i, we assign time values to these aggregated counts using the following strategies to reconstruct patient-level data:

Event time assignment: The element $e_i^{'}$ in the candidate's event field represents the event count occurring at the breakpoint t_i. Therefore, for the $e_i^{'}$ patients allocated to this position, the observed time is set to t_i, and the status is recorded as 1 (event).

Censoring time assignment: For the $c_i^{'}$ censored patients estimated within the interval $(t_{i-1}, t_i]$, the status is recorded as 0 (censored), while the assignment of observed time depends on the availability of censoring constraint:

- **Uniform censoring** (without C): If C are not available, we assume a non-informative censoring distribution. The observed times for the $c_i^{'}$ patients are distributed uniformly across the interval $[t_{i-1}, t_i]$.
- **Constrained censoring** (with C): If C are available, they define the temporal window of censoring events. Within each interval $(t_{i-1}, t_i]$ delimited by the key points, given that a single censoring time may represent multiple overlapping patients and the censoring count $c_i^{'}$ often exceeds the number of marked censoring times, we identify the earliest ($t_i^{\min}$) and latest ($t_i^{\max}$) censoring times contained in that interval. The observed times for $c_i^{'}$ patients are then distributed uniformly over this bounded range $[t_i^{\min}, t_i^{\max}]$, aligning the reconstructed data with the visible censoring span.

Following these assignment steps, the observed times and event status for all patients are consolidated to form the final reconstructed IPD, enabling subsequent analysis. While previous studies[13,14] have relied on the uniform censoring assumption, our study on the Synthetic KM Dataset showed that constrained censoring resulted in lower HR errors (Supplementary Table 12). Therefore, we use constrained censoring when C is available and uniform censoring otherwise.

Benchmark datasets construction

We constructed the Real IPD, Published KM, and Synthetic KM Datasets to enable standardized evaluation of IPD reconstruction methods and applications. These datasets draw from established repositories, including The Cancer Genome Atlas (TCGA)[20], the Gene Expression Omnibus (GEO)[21], and published studies. Supplementary Note 1 details the data sources.

The Real IPD Dataset

The Real IPD Dataset integrates participant-level time-to-event data from TCGA and GEO. We refer to these sub-datasets as TCGA-IPD and GEO-IPD. Each entry comprises three variables: patient identifier (`id`), survival time in months (`time`), and event status (`status`).

For TCGA-IPD, we retrieved pre-processed clinical data for 21 cancer types from OncoLnc[50]. To assemble the patient cohorts, we employed a gene-based query strategy (e.g., using *TP53* as a retrieval mechanism) to retrieve and merge survival data across complementary high- and low-expression subgroups from TCGA, obtaining IPD for each cancer type. For GEO-IPD, we referenced the survival-related GEO series (GSE) accessions identified by PESSA[51] and retrieved the corresponding MINiML XML files from the GEO database to extract clinical metadata. We selected studies that reported survival outcomes and event status, resulting in a final subset of 147 studies encompassing 210 distinct survival analyses across various outcomes (e.g., OS, PFS). We parsed the structured XML files to extract relevant clinical variables. Given the structural heterogeneity of GEO XML files, variable naming conventions and schema definitions for survival outcomes and event indicators are not fully standardized across studies. As a result, minor discrepancies may arise between extracted summary quantities (e.g., N or E) and those reported in the original associated publications. Importantly, the extracted IPD are used here as internally consistent numerical reference data for benchmarking reconstruction fidelity rather than for independent clinical re-analysis. This heterogeneity therefore provides a realistic and challenging testbed for assessing algorithm robustness under non-uniform reporting conditions. Data from both sources were harmonized: original public identifiers were retained for traceability, survival times were standardized to months, and status labels were recoded into binary values (0 for censored/alive, 1 for event/dead).

Combining these sources, the final Real IPD Dataset comprises 231 distinct survival analyses (21 from TCGA-IPD and 210 from GEO-IPD). For each analysis, we generated the ground-truth KM curve using the R `survival` package[52]. To establish the benchmarking ground truth for reconstruction, three trained graduate annotators utilized the KM Point Picker tool to manually annotate key coordinates and censoring times on all 231 images. Subsequently, all annotations underwent an audit by two independent reviewers to correct potential errors and verify data consistency. During this process, summary statistic constraints, including N, E, and R, were recorded and exported as structured data files. The final dataset thus integrates original IPD, paired KM images, and structured annotations. Detailed data specifications are provided in Supplementary Note 1.2.

The Published KM Dataset

We constructed the Published KM Dataset by curating KM curve images from studies identified via Google Scholar and PubMed. Ten trained graduate annotators searched using keywords: "prognosis", "Kaplan–Meier", "RCT", "meta-analysis", and "randomized controlled trial". After removing duplicates, studies were screened based on the availability of extractable KM curves and reported hazard ratios (HRs) with confidence intervals (CIs). To expand coverage, the annotators screened reference lists of identified prognostic meta-analyses to retrieve additional eligible studies.

The search identified 280 studies featuring comparative experimental and control arms, encompassing 981 KM images and 1,962 individual KM curves representing approximately 367,000 patients. Through manual cross-referencing of full texts and tables, we recorded study arm names (to define HR direction) and extracted 1,962 sample sizes (N). We further retrieved other available constraints, yielding 1,032 numbers of events (E), 1,896 at-risk tables (R), and 1,474 curves with censoring times (C), alongside a validation set of survival statistics comprising 981 HRs, 1,575 median survival times, and 661 survival rates.

Each source KM image was assigned a unique identifier (`km_image_id`) and linked to the reported HR. Each KM curve within the image was assigned a unique identifier (`km_curve_id`). Finally, ten annotators used KM-PoPiGo to annotate each KM curve in the image, and all annotations audited by two independent reviewers. To adhere to copyright policies, we do not redistribute the original KM images. Instead, we provide the complete set of annotated coordinates and constraints as structured data files. These files contain the data required to reconstruct the IPD, ensuring reproducibility without infringing on third-party copyrights. Furthermore, the inclusion of the full source study title for each record enables users to uniquely locate the original publications for visual verification. Detailed data specifications are provided in Supplementary Note 1.2.

The Synthetic KM Dataset

To provide a controlled benchmarking environment with complete constraints, we constructed a Synthetic KM Dataset by pooling the IPD from TCGA-IPD into a unified pool. From this pool, we performed 500 iterations of random sampling to generate synthetic cohorts with sizes ranging from 30 to 1,000. In each iteration, block randomization was applied to assign subjects to experimental or control arms, maintaining an approximate 1:1 allocation ratio. KM curve images were generated using the R `survival` package[52]. Each simulation iteration generated one KM image containing a pair of curves (representing two treatment arms, identified by a unique `km_image_id`), alongside two corresponding individual curve images (each identified by a unique `km_curve_id`).

For each iteration, the HR was calculated using a Cox proportional hazards model and linked to `km_image_id`. At the arm level, we recorded summary-statistic constraints (N, E, and R) alongside survival statistics (median survival times and 1–5-year survival rates), all mapped to the corresponding `km_curve_id`.

To finalize the dataset, three trained graduate annotators utilized KM-PoPiGo to extract coordinate data and censoring times C from the single-arm images, with all annotations subsequently audited by two independent reviewers. The exported structured data were linked to their `km_curve_ids`. The final dataset comprises 1,000 KM curves with complete constraints, featuring 1,000 instances of each constraint (N, E, R, C). Survival statistics include 500 HRs, 1,000 median survival times, and 4,885 survival rate records. Detailed data specifications are provided in Supplementary Note 1.3.

Weibull reconstruction metric

To quantify the disparity between the reconstructed survival distribution and the ground truth, we compute the Weibull reconstruction metric (L_1 distance), defined as:

$$L_1(p,q) = \int_0^\infty |(p(t) - q(t)|dt,$$

where $p(t)$ and $q(t)$ denote the probability density functions of the reconstructed and ground-truth Weibull distributions. This integral represents the area between the two density curves. The L_1 values range from 0 (identical distributions) to 2 (disjoint distributions), providing a scale for assessing distributional fidelity.

Because the L_1 errors between the reconstructed and ground-truth Weibull distributions vary across the D KM curves (due to varying time scales and distribution shapes), arithmetic means can be dominated by curves with larger errors. To aggregate performance and mitigate biases associated with averaging results spanning different scales[53], we computed the geometric mean error G_k for each constraint combination k over the D KM curves, defined as:

$$G_k = \left(\prod_{d=1}^{D} L_{d,k} \right)^{\frac{1}{D}} = \exp\left(\frac{1}{D} \sum_{d=1}^{D} \ln L_{d,k} \right),$$

where $L_{d,k}$ denotes the L_1 error between the reconstructed and ground-truth distributions for the d-th KM curve under the k-th constraint combination.

To quantify the performance gain yielded by the genetic optimization, we defined the relative improvement rate (Δ_k) for each constraint combination k relative to the baseline (G_{base}) to reflect the percentage reduction in aggregate error:

$$\Delta_k = \left(\frac{G_{\text{base}} - G_k}{G_{\text{base}}} \right) \times 100\%.$$

A positive Δ_k indicates a reduction in reconstruction error relative to the baseline.

Statistical analysis

Statistical computations and analyses were performed using the R statistical computing environment (version 4.3.3; R Core Team). We utilized the R `survival` package[52] to derive descriptive statistics: Weibull distribution parameters were estimated using parametric survival models (`survreg`), while median survival times and landmark survival rates were calculated via Kaplan-Meier estimators. Hazard ratios (HRs) and associated 95% CIs were estimated by fitting Cox proportional-hazards models to the reconstructed IPD.

For the meta-analysis experiments, HRs were pooled using random-effects models (estimating heterogeneity via the Restricted Maximum Likelihood method) implemented in the R `meta` package[54]. To assess differences in reconstruction errors between methods, paired t-tests were performed on log-transformed error values, consistent with the geometric mean metric. For comparisons involving limited sample sizes ($n < 10$), linear association was assessed using Pearson correlation coefficients. Differences between survival distributions (e.g., reconstructed vs. original) were evaluated using the log-rank test. Statistical tests were two-sided, with a significance threshold set at $p < 0.05$.

To quantify performance gains, we calculated the relative error reduction (η) using the following formula:

$$\eta = \left(\frac{\varepsilon_{\text{base}} - \varepsilon_{\text{new}}}{\varepsilon_{\text{base}}} \right) \times 100\%, \tag{9}$$

where $\varepsilon_{\text{base}}$ and ε_{new} represent the aggregate error metrics of the comparator method and the proposed method.

To evaluate algorithmic robustness, we conducted sensitivity analyses across boundary conditions, including subsets with small sample sizes ($N < 100$) and low hazard ratios (HR < 0.3) (Supplementary Fig. 9).

Data inclusion criteria for comparative benchmarking

To ensure a fair and consistent evaluation, we aligned the benchmarking scope with the capabilities of each algorithm. While KM-PoPiGo supports a set of eight constraint combinations, the baseline methods are limited in scope: iKM and IPDfromKM support only four constraint combinations (N, NE, NR, NER), whereas KMtoIPD supports only the NC constraint (Fig. 1b). Consequently, comparisons were limited to those specific constraint combinations.

As comparative benchmarking was conducted on the TCGA-IPD subset (from the Real IPD Dataset) and the Published KM Dataset, the data inclusion strategy was tailored to the characteristics of these two evaluation datasets. For the TCGA-IPD subset, where the objective was to evaluate the fidelity of the reconstructed Weibull distributions, baseline methods were able to generate IPD reconstructions under their supported constraint combinations. No exclusion based on reconstruction failure was performed for this subset.

In contrast, for the Published KM Dataset for survival statistics evaluation (HR, median survival time, and survival rate), the availability of valid outcomes varied across methods due to real-world complexity (Supplementary Table 11). To ensure fair comparison in this heterogeneous context, we applied an intersection-based inclusion criterion. We restricted the comparative analysis of reconstructed survival statistics to the common subset of KM curves where both the proposed method and the baseline were able to generate valid estimates for the target metric under identical constraints. This criterion ensures that the reported performance differences reflect methodological efficacy by reducing biases from sample inclusion.

Code availability

The KM-PoPiGo web-based platform is freely available at `https://kmpopigo.github.io/`. To ensure data privacy, the platform employs a client-side architecture where all data processing occurs locally within the user's browser, eliminating the need to upload KM images to a server. For verification and validation, or for offline execution or custom deployment, the source code is available at `https://github.com/kmpopigo/kmpopigo.github.io`.

Data availability

All benchmark datasets generated and analyzed during this study are available in the Zenodo repository at `https://doi.org/10.5281/zenodo.18320575`. The Real IPD Dataset was compiled from public data in The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO); detailed series information for the GEO-IPD subset is provided in Supplementary Table 13. Its repository contains the processed, de-identified IPD CSV files, along with the corresponding author-generated KM curve images and structured CSV annotation files required for benchmarking (see Supplementary Note 1.1 for detailed descriptions of these data components). The Published KM Dataset was curated from published studies (sources listed in Supplementary Table 14). Due to third-party copyright restrictions, the original KM curve images extracted from journals are not included. Instead, we provide the paper titles of the included studies, along with collected survival statistics (HR, median survival time, and survival rates), summary constraints (N, E, R), and structured CSV annotation files (see Supplementary Note 1.2 for detailed descriptions of these data components). These files contain all the necessary information to reproduce the reconstruction results. The Synthetic KM Dataset is a freely available, fully simulated dataset that includes generated KM curve images, ground-truth constraints, and validation metrics (see Supplementary Note 1.3).

Acknowledgements

The results shown here are, in whole or in part, based on data generated by the TCGA Research Network: https://www.cancer.gov/tcga. We thank J. Yang, H. Xie, F. Bian, Y. Zeng, L. Wang, J. Zhou, J. Gao, Y. Nan, X. Zan, N. Sun, L. Duan, Z. Wang, L. Qi, X. Cao, J. Yan and R. Wang for their assistance with data collection and annotation. This work was supported by the National Natural Science Foundation of China Major Project (T2495251 to Y.Q.), the National Natural Science Foundation of China (grant nos. 62306170 to J.W., and 62441239 to F.L.), and the Special Fund for Science and Technology Innovation Teams of Shanxi Province (grant no. 202304051001001). M.S. and J.L. would like to acknowledge the financial support of AtriCure Europe B.V. to generate the independent clinical validation data. H.L. is supported by the EPSRC (grant UKRI396). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Z.Z., H.L. and Y.Q. conceived the study and designed the experiments. X.L., J.W. and F.L. established the mathematical framework for IPD reconstruction. Z.Z. built the core algorithm, developed the software, generated the synthetic datasets, performed the statistical analysis, and created the visualizations. Z.Z. and F.L. implemented the algorithms, with guidance on genetic optimization from X.L. H.Y. provided the Weibull distribution error metric. Z.Z., J.J., F.S. and F.P. designed and constructed the benchmark datasets. J.J. and F.S. audited the data annotations. Me.Z., H.L. and Y.G. formulated the experimental validation strategy. S.C.C., M.A.S. and J.L. conducted the independent blinded clinical validation on unseen data. Z.Z. wrote the original draft with assistance from Me.Z. and J.M. Subsequently, H.L., J.J., F.P., Mi.Z., S.C.C., M.A.S. and J.L. contributed to the revision and editing of the manuscript. H.L. and Y.Q. supervised the project. All authors reviewed and approved the final manuscript.

Additional information

Supplementary Information is available for this paper.

References

1. Riley, R. D. *et al.* A guide to systematic review and meta-analysis of prognostic factor studies. *BMJ* **364** (2019).
2. Zeymer, U. *et al.* Venoarterial extracorporeal membrane oxygenation in patients with infarct-related cardiogenic shock: an individual patient data meta-analysis of randomised trials. *The Lancet* **402**, 1338–1346 (2023).
3. Inker, L. A. *et al.* A meta-analysis of GFR slope as a surrogate endpoint for kidney failure. *Nat. Medicine* **29**, 1867–1876 (2023).
4. Subbiah, V. The next generation of evidence-based medicine. *Nat. Medicine* **29**, 49–58 (2023).
5. Popat, S. *et al.* Addressing challenges with real-world synthetic control arms to demonstrate the comparative effectiveness of Pralsetinib in non-small cell lung cancer. *Nat. Commun.* **13**, 3500 (2022).
6. Tierney, J. F. *et al.* How individual participant data meta-analyses have influenced trial design, conduct, and analysis. *J. Clin. Epidemiol.* **68**, 1325–1335 (2015).
7. Debray, T. P. A. *et al.* Individual participant data (IPD) meta-analyses of diagnostic and prognostic modeling studies: guidance on their use. *PLOS Medicine* **12**, e1001886 (2015).
8. Ursin, G. *et al.* Sharing data safely while preserving privacy. *The Lancet* **394**, 1902 (2019).
9. Rathi, V. *et al.* Sharing of clinical trial data among trialists: a cross sectional survey. *BMJ* **345** (2012).
10. Van Noorden, R. Data-sharing: Everything on display. *Nature* **500**, 243–245 (2013).
11. International Consortium of Investigators for Fairness in Trial Data Sharing. Toward fairness in data sharing. *New Engl. J. Medicine* **375**, 405–407 (2016).
12. Danchev, V., Min, Y., Borghi, J., Baiocchi, M. & Ioannidis, J. P. Evaluation of data sharing after implementation of the International Committee of Medical Journal Editors data sharing statement requirement. *JAMA Netw. Open* **4**, e2033972 (2021).
13. Guyot, P., Ades, A. E., Ouwens, M. J. N. M. & Welton, N. J. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan–Meier survival curves. *BMC Med. Res. Methodol.* **12**, 9 (2012).

14. Liu, N., Zhou, Y. & Lee, J. J. IPDfromKM: reconstruct individual patient data from published Kaplan–Meier survival curves. *BMC Med. Res. Methodol.* **21**, 111 (2021).

15. Rogula, B., Lozano-Ortega, G. & Johnston, K. M. A method for reconstructing individual patient data from Kaplan–Meier survival curves that incorporate marked censoring times. *MDM Policy & Pract.* **7**, 23814683221077643 (2022).

16. Saluja, R., Cheng, S., delos Santos, K. A. & Chan, K. K. Estimating hazard ratios from published Kaplan–Meier survival curves: a methods validation study. *Res. Synth. Methods* **10**, 465–475 (2019).

17. Kaplan, E. L. & Meier, P. Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.* **53**, 457–481 (1958).

18. Bentzen, S. M. & Vogelius, I. R. Using and understanding survival statistics—or how we learned to stop worrying and love the Kaplan–Meier estimate. *Int. J. Radiat. Oncol. Biol. Phys.* **115**, 839–846 (2023).

19. Dey, T. *et al.* Survival analysis—time-to-event data and censoring. *Nat. Methods* **19**, 906–908, DOI: 10.1038/s41592-022-01563-7 (2022).

20. Tomczak, K., Czerwińska, P. & Wiznerowicz, M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp. Oncol. (Poznan, Poland)* **19**, A68–A77 (2015).

21. Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* **41**, D991–D995 (2013).

22. Jhund, P. S. *et al.* Dapagliflozin across the range of ejection fraction in patients with heart failure: a patient-level, pooled meta-analysis of DAPA-HF and DELIVER. *Nat. Medicine* **28**, 1956–1964 (2022).

23. Riley, R. D., Ensor, J., Hattle, M., Papadimitropoulou, K. & Morris, T. P. Two-stage or not two-stage? That is the question for IPD meta-analysis projects. *Res. Synth. Methods* **14**, 903–910 (2023).

24. Banna, G. L. *et al.* Neoadjuvant chemo-immunotherapy for early-stage non–small cell lung cancer: a systematic review and meta-analysis. *JAMA Netw. Open* **7**, e246837 (2024).

25. Benjafield, A. V. *et al.* Positive airway pressure therapy and all-cause and cardiovascular mortality in people with obstructive sleep apnoea: a systematic review and meta-analysis of randomised controlled trials and confounder-adjusted, non-randomised controlled studies. *The Lancet Respir. Medicine* **13**, 403–413 (2025).

26. Wang, F. *et al.* A systematic review and meta-analysis of 90 cohort studies of social isolation, loneliness and mortality. *Nat. Hum. Behav.* **7**, 1307–1319 (2023).

27. Dempsey, T. M. & Limper, A. H. The added power of synthetic control groups: Challenging conventional wisdom and trial design in idiopathic pulmonary fibrosis research. *Am. J. Respir. Critical Care Medicine* **208**, 515–516 (2023).

28. Serrano, C. *et al.* Rethinking placebos: embracing synthetic control arms in clinical trials for rare tumors. *Nat. Medicine* **29**, 2689–2692 (2023).

29. Yin, X. *et al.* The validity of a synthetic control arm derived from historical multiple myeloma clinical trials. *Blood* **142**, 6628 (2023).

30. Thorlund, K., Dron, L., Park, J. J. H. & Mills, E. J. Synthetic and external controls in clinical trials—a primer for researchers. *Clin. Epidemiol.* **12**, 457–467 (2020).

31. Zhou, Q. *et al.* Transcriptional profile of human liver tissues: hepatocellular carcinoma vs. matched noncancerous liver tissue. NCBI Gene Expression Omnibus (2011). GSE27150.

32. Howlader, N. *et al.* (eds.) *SEER Cancer Statistics Review, 1975-2017* (National Cancer Institute, Bethesda, MD, 2020). Based on November 2019 SEER data submission, posted April 2020.

33. Higgins, J. P. T. *et al.* (eds.) *Cochrane Handbook for Systematic Reviews of Interventions* (Wiley-Blackwell, Hoboken, NJ, 2019), 2nd edn.

34. Schoenfeld, D. A. Sample-size formula for the proportional-hazards regression model. *Biometrics* **39**, 499–503 (1983).

35. Plana, D., Fell, G., Alexander, B. M., Palmer, A. C. & Sorger, P. K. Cancer patient survival can be parametrized to improve trial precision and reveal time-dependent therapeutic effects. *Nat. Commun.* **13**, 873 (2022).

36. Zhou, C. *et al.* Interim survival analysis of the randomized phase III GEMSTONE-302 trial: sugemalimab or placebo plus chemotherapy as first-line treatment for metastatic NSCLC. *Nat. Cancer* **4**, 860–871 (2023).

37. Dey, T. *et al.* Regression modeling of time-to-event data with censoring. *Nat. Methods* **19**, 1513–1515, DOI: 10.1038/s41592-022-01689-8 (2022).

38. Zheng, M.-M. *et al.* Cerebrospinal fluid circulating tumor DNA profiling for risk stratification and matched treatment of central nervous system metastases. *Nat. Medicine* **31**, 1547–1556 (2025).

39. Foersch, S. *et al.* Multistain deep learning for prediction of prognosis and therapy response in colorectal cancer. *Nat. Medicine* **29**, 430–439 (2023).

40. Tarantola, A. *Inverse Problem Theory and Methods for Model Parameter Estimation* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2005).

41. European Parliament and Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council (General Data Protection Regulation) (2016).

42. US Congress. Health Insurance Portability and Accountability Act of 1996 (1996).

43. Price, W. N. & Cohen, I. G. Privacy in the age of medical big data. *Nat. Medicine* **25**, 37–43 (2019).

44. Aymé, S. *et al.* Common elements in rare kidney diseases: conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) controversies conference. *Kidney Int.* **92**, 796–808 (2017).

45. Spreafico, A., Hansen, A. R., Abdul Razak, A. R., Bedard, P. L. & Siu, L. L. The future of clinical trial design in oncology. *Cancer Discov.* **11**, 822–837 (2021).

46. Sherman, R. E. *et al.* Real-world evidence—what is it and what can it tell us. *New Engl. J. Medicine* **375**, 2293–2297 (2016).

47. Wei, Y. & Royston, P. Reconstructing time-to-event data from published Kaplan–Meier curves. *The Stata J.* **17**, 786–802 (2017).

48. Whitley, D. The GENITOR algorithm and selection pressure: why rank-based allocation of reproductive trials is best. In *Proceedings of the Third International Conference on Genetic Algorithms*, 116–121 (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1989).

49. Syswerda, G. Uniform crossover in genetic algorithms. In *Proceedings of the 3rd International Conference on Genetic Algorithms*, 2–9 (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1989).

50. Anaya, J. OncoLnc: linking TCGA survival data to mRNAs, miRNAs, and lncRNAs. *PeerJ Comput. Sci.* **2**, e67 (2016).

51. Yang, H. *et al.* PESSA: A web tool for pathway enrichment score-based survival analysis in cancer. *PLOS Comput. Biol.* **20**, e1012024 (2024).

52. Therneau, T. M., Lumley, T., Atkinson, E. & Crowson, C. *survival: Survival Analysis* (2024). R package version 3.8-3.

53. Fleming, P. J. & Wallace, J. J. How not to lie with statistics: the correct way to summarize benchmark results. *Commun. ACM* **29**, 218–221 (1986).

54. Schwarzer, G. *et al.* meta: an R package for meta-analysis. *R News* **7**, 40–45 (2007).

Supplementary information

Table of Contents

1. Description of benchmark datasets 2
2. Recovering unreported HRs via IPD reconstruction 6
3. Experimental design and data sources for synthetic control simulation 7
4. Experimental design and data sources for the pooled intervention simulation 7
5. Detailed performance benchmarking and statistical validation across constraint combinations 9
6. Benchmarking of IPD reconstruction methods on the Published KM Dataset 12
7. Supplementary results of independent clinical validation 12
8. KM curve topologies and derivation of censoring constraints 13
9. Mathematical formulation: IPD reconstruction as an underdetermined system 14
10. Pseudocode for the genetic optimization algorithm 15
11. Efficacy of the constrained censoring strategy 17
12. Sensitivity analysis: Impact of sample size and effect size on reconstruction fidelity 17
13. Source data and provenance 19

Supplementary note

1. Description of benchmark datasets

Supplementary Table 1 summarizes the characteristics of the Real IPD, Published KM, and Synthetic KM datasets.

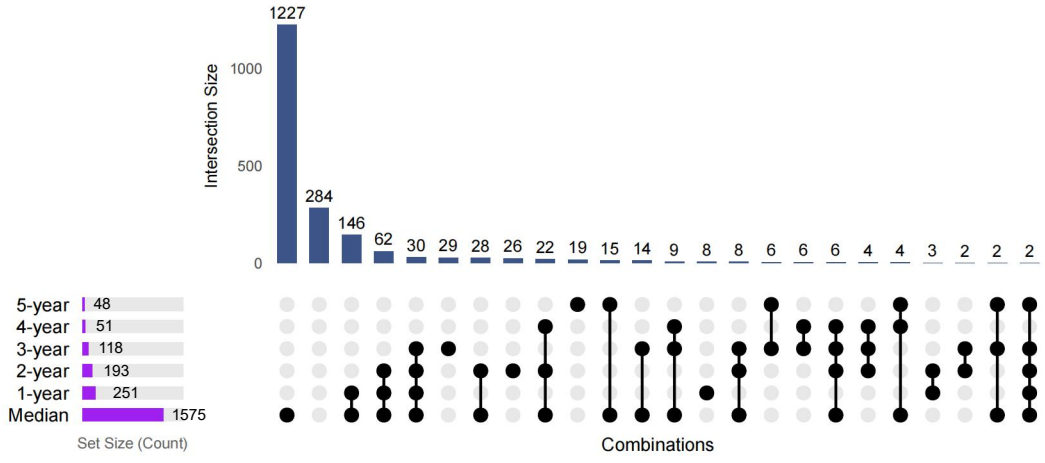
The Real IPD Dataset comprises 231 Kaplan-Meier (KM) curves derived from patient-level time-to-event data from the TCGA[1] and GEO[2] databases, stratified into the TCGA-IPD ($n = 21$) and GEO-IPD ($n = 210$) subsets (Supplementary Table 13). Serving as the ground truth, this dataset establishes the reference standard for quantifying reconstruction accuracy. The TCGA-IPD subset was used for IPD meta-analysis validation and benchmarking, while the GEO-IPD subset validated synthetic control arms and cross-trial comparisons; both subsets supported Weibull distribution error analysis.

The Published KM Dataset encompasses 1,962 KM curves derived from 280 published studies (Supplementary Table 14). From the texts and tables of these sources, we retrieved sample size (N) constraints for all 1,962 curves, the number of events

Supplementary Table 1 Detailed specifications of benchmark datasets

Dataset	Source	Total no. of curves	Available constraints	Evaluation scenario	Analyzed subset
Real IPD Dataset	TCGA	21	N (21), E (21), R (21), C (21)	IPD meta-analysis validation	21
				Weibull distribution error analysis	21
				Comparative benchmarking	21
	GEO	210	N (210), E (210), R (210), C (210)	Validation of synthetic control arms	6
				Cross-trial comparison	16
				Weibull distribution error analysis	210
Published KM Dataset	Published studies	1,962	N (1,962), E (1,032), R (1,896), C (1,474)	Validation of recovering excluded prognostic evidence	1,962
				Benchmarking of reconstruction accuracy (four methods)	1,962
Synthetic KM Dataset	Simulated data	1,000	N (1,000), E (1,000), R (1,000), C (1,000)	Impact assessment of constraint combinations	1,000

Note: "Available constraints" lists the constraints utilized for reconstruction (N: Sample size, E: Number of events, R: At-risk table, C: Censoring times); parenthetical values indicate the number of curves associated with each constraint. "Evaluation scenario" details the specific experiments performed.



Supplementary Fig. 1 Reporting profile of survival statistics in the Published KM Dataset. UpSet plot illustrating the co-occurrence of median survival times and 1- to 5-year survival rates in the Published KM Dataset. Statistics were defined as "reported" if documented, including numerical values and "Not Reached" or "Not Estimable" annotations. Vertical bars denote intersection sizes (count of KM curves per combination), while horizontal bars indicate the total set size for each statistic.

(E) for 1,032, and at-risk tables (R) for 1,896. In addition, censoring times (C) were obtained by annotating the KM curve images using KM-PoPiGo for 1,474 cases. For validation purposes, we collated available survival statistics, specifically 981 HRs, 1,575 median survival times (including those reported as "Not Reached" or "Not Estimable"), and 661 survival rates (reporting availability detailed in Supplementary Fig. 1).

This dataset enables the benchmarking of reconstruction accuracy of four methods in real-world scenarios by evaluating the concordance of reconstructed HRs, median survival times, and survival rates with their reported values. Furthermore, reported HRs for all 981 comparisons enabled the validation of the recovery of excluded prognostic evidence.

The Synthetic KM Dataset is a simulated dataset containing 1,000 curves with complete constraints (N, E, R, C), designed to quantify the impact of constraint combinations on the accuracy of reconstructed survival statistics (HRs, median survival times, and survival rates).

1.1 Description of the Real IPD Dataset

The dataset package is organized into the following files and directories:

- `RIPD_subGroup.csv`: Metadata file containing summary-statistic constraints and identifiers for 231 KM curves (encompassing both TCGA-IPD and GEO-IPD subsets) (Supplementary Table 2).
- `RIPD_markerInfo.csv`: Comprehensive annotation file recording the detailed annotation data extracted from the KM images (Supplementary Table 3).
- `km_curve_images`: Directory containing KM curve images generated from the real IPD for the 231 KM curves; filenames correspond to the unique `km_curve_id`.
- `IPD`: Directory containing the ground-truth IPD files (CSV format) for each KM curve. Each file includes columns for patient identifier (`id`), survival time (`time`), and event status (`status`).

1.2 Description of the Published KM Dataset

In contrast to the Real IPD Dataset, which focuses on single KM curves, the Published KM Dataset is derived from KM images containing paired curves (typically representing Treatment vs. Control arms) extracted from published studies. Due to copyright restrictions associated with published images, the original images are not included. Instead, we provide the annotation data and metadata, enabling users to reconstruct the analyses or locate the original studies. The dataset comprises:

- **`RA_KM.csv`**: Study-level metadata documenting the source study titles and statistical characteristics of the paired KM curves, including reported HRs and 95% CIs. It also records binary indicators regarding the availability of constraints (Supplementary Table 4).
- **`RA_subGroup.csv`**: Summary-statistic constraints (N, E, R) and reported survival statistics (e.g., median survival times, 1- to 5-year survival rates) for each group in the paired analysis (Supplementary Table 5).
- **`RA_markerInfo.csv`**: Annotation data extracted from published KM images. The file structure is consistent with `RIPD_markerInfo.csv`, utilizing `km_curve_id` to uniquely identify each KM curve within the paired KM images.

Supplementary Table 2 Variable descriptions for `RIPD_subGroup.csv`

Variable	Label	Description
`km_curve_id`	KM curve ID	Unique identifier for the KM curve image, where each image contains a single KM curve. This ID corresponds to the source image file in the `km_curve_images/` directory.
`group`	Group label	Identifier for the specific study cohort or dataset series (e.g., TCGA-LUAD or GSE10300_RFS) from which the KM curve and its corresponding IPD were obtained.
`sample_size`	Sample size (N)	Total number of patients in the KM curve.
`number_of_events`	Number of events (E)	Total number of patients with observed events.
`at_risk_table`	At-risk table (R)	JSON-formatted string of time points and corresponding number at risk (format: `[[time, number], ...]`).

Supplementary Table 3 Variable descriptions for `RIPD_markerInfo.csv`

Variable	Label	Description
`annotation_id`	Annotation ID	Unique identifier for the annotation record of a KM curve.
`key_point_coordinates`	Marked curve coordinates	Array of (x, y) pixel coordinates representing the marked breakpoints and the end point of the KM curve, mapped onto a standard 1600×800 canvas for consistency.
`censoring_time_coordinates`	Censoring time coordinates	Array of (x, y) pixel coordinates representing the ticks of marked censoring times on the KM curve, normalized to a 1600×800 canvas.
`survival_probabilities`	Survival probability	Array of survival probabilities ($\boldsymbol{p}$) corresponding to the starting point and breakpoints, mapped to the data coordinate system (value range $[0, 1]$).
`time_values`	Time array	Array of time values corresponding to the starting point and breakpoints, mapped to the data coordinate system (actual time units).
`end_time`	End time	The time value corresponding to the end point (t_M) of the KM curve.
`min_censoring_counts`	Censoring counts ($\boldsymbol{c}^{min}$)	Array of counts of user-marked censoring times within each interval defined by the successive start, break, and end points of the KM curve.
`coordinate_origin`	Coordinate origin	Pixel coordinates in the normalized image coordinate system representing the axes origin $(0, 0)$ of the data coordinate system for the KM curve, as marked by the annotator.
`axis_limit_coordinates`	Axis limit coordinate	Pixel coordinates in the normalized image coordinate system representing the point $(time_axis_max_scale, 1.0)$ in the data coordinate system, used to establish the mapping between image pixels and survival analysis scales as marked by the annotator.
`origin_data_values`	Origin data values	The numerical values in the data coordinate system corresponding to the origin of the data plot axes $(0, 0)$.
`axis_limit_data_values`	Axis limit data values	The numerical values in the data coordinate system corresponding to the maximum identifiable Time axis tick ($(time_axis_max_scale, 1.0)$) and the maximum Survival probability axis tick (1.0) of the data plot axes.
`km_image_id`	KM image ID	Unique identifier linking this annotation record to the metadata in `RIPD_subGroup.csv` and the source image in the `km_img` directory.

1.3 Description of the Synthetic KM Dataset

The Synthetic KM Dataset shares an identical file structure with the Published KM Dataset and was generated to quantify the impact of constraint combinations (N, E, R, C) on reconstruction accuracy. This ensures the availability of ground truth for all constraints, facilitating the evaluation of algorithmic performance under varying conditions.

- `SA_KM.csv` and `SA_subGroup.csv`: Metadata and statistics adapted from the Published KM Dataset schemas (refer to Supplementary Tables 4 and 5). Fields regarding source identification (e.g., `source_study_title`) and constraint availability (e.g., `event_available`, `at_risk_available`, `censoring_available`) are excluded, as these methodological constraints are consistently provided for the synthetic curves.
- `SA_markerInfo.csv`: Records the annotation data for the synthetic KM curves. The schema is consistent with that of `RA_markerInfo.csv` (indexed by `km_curve_id`).
- `km_paired_images/` and `km_curve_images/`: Directories containing images of paired KM curves (indexed by `km_image_id`) and images of individual KM curves (indexed by `km_curve_id`), respectively.

Supplementary Table 4 Variable descriptions for `RA_KM.csv`

Variable	Label	Description
`km_image_id`	KM image ID	Unique identifier for the source KM image, where each image contains a pair of KM curves.
`source_study_title`	Source study title	Full title of the source study from which the KM image was extracted.
`hr`	Hazard ratio (HR)	The hazard ratio (HR) reported in the source study specifically corresponding to this KM curve image.
`hr_ci_lower`	HR lower CI	Reported lower bound of the 95% confidence interval for the hazard ratio, specifically corresponding to this KM curve image.
`hr_ci_upper`	HR upper CI	Reported upper bound of the 95% confidence interval for the hazard ratio, specifically corresponding to this KM curve image.
`reference_group`	Reference group	Label of the reference group (exactly matches the `group` column in `RA_subGroup.csv`), serving as the denominator in the calculation of the hazard ratio (*HR*).
`experimental_group`	Experimental group	Label of the experimental group (exactly matches the `group` column in `RA_subGroup.csv`), serving as the numerator in the calculation of the hazard ratio (*HR*).
`event_available`	Availability of number of events (E)	Binary indicator (1 = available; 0 = unavailable) representing whether the total number of events (E) is reported in the source study.
`at_risk_available`	Availability of at-risk table (R)	Binary indicator (1 = available; 0 = unavailable) representing whether the at-risk table (R) is reported in the source study.
`censoring_available`	Availability of censoring times (C)	Binary indicator (1 = available; 0 = unavailable) representing whether censoring times (C) are marked on the KM curve image.

Supplementary Table 5 Variable descriptions for `RA_subGroup.csv`

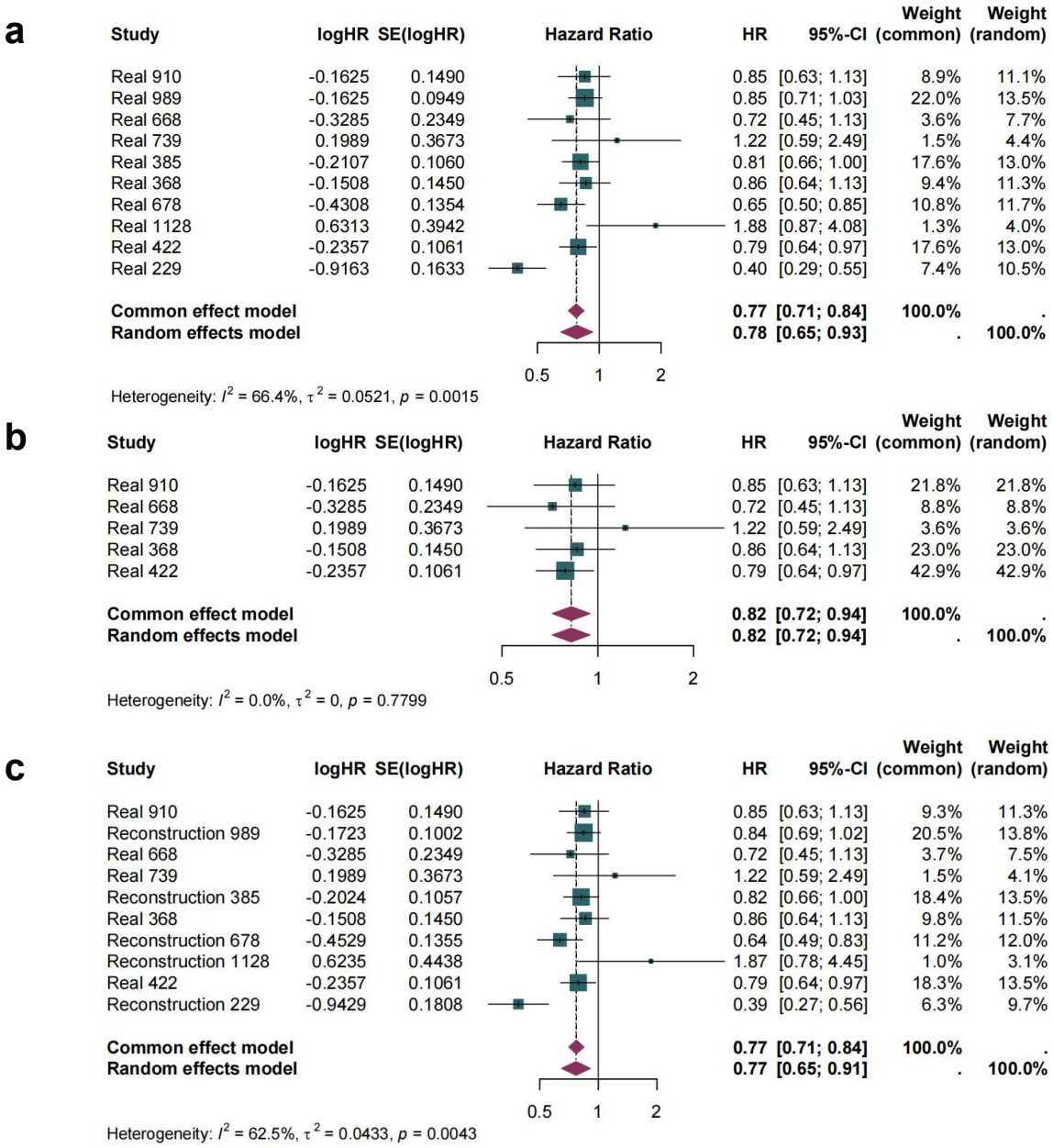
Variable	Label	Description
`km_curve_id`	KM curve ID	Unique identifier for each KM curve within a KM image, used to distinguish between the pair of curves present in the image.
`km_image_id`	KM image ID	Identifier linking to the source KM image record in `RA_KM.csv`.
`group`	Group label	Identifier for the specific treatment arm, matching either the `reference_group` or `experimental_group` field in the corresponding `RA_KM.csv` file.
`sample_size`	Sample size (N)	Total number of patients in the group.
`number_of_events`	Number of events (E)	Total number of patients with observed events in the group. Recorded as NULL if unavailable.
`at_risk_table`	At-risk table (R)	JSON-formatted string of time points and corresponding number at risk (format: `[[time, number], ...]`). Recorded as NULL if unavailable.
`median_survival_time`	Median survival time	String containing reported median survival time and 95% CI. Recorded as NULL if unavailable.
`one_year_survival_rate`, ..., `five_year_survival_rate`	1-year to 5-year survival rates	Five separate variables recording the survival probability ($S(t)$) at years 1, 2, 3, 4, and 5. Recorded as NULL if unavailable.

2. Recovering unreported HRs via IPD reconstruction

Supplementary to Main Text Results subsection: "Reconstructed IPD recovers missing prognostic information".

Conventional meta-analyses of time-to-event outcomes rely on reported HRs. Excluding studies lacking these data[3,4] introduces reporting bias that limits the generalizability of pooled estimates[5]. Kaplan–Meier curves encode the individual patient data (IPD) required to estimate HRs. Recovering this information enables the integration of studies that lack reported statistics, enhancing the utilization of available evidence.

To illustrate evidence recovery using KM-PoPiGo, we aggregated 10 randomly selected studies under the assumption of a



Supplementary Fig. 2 Case study: Validation of IPD reconstruction in recovering excluded prognostic evidence. a, Forest plot representing the ground truth meta-analysis of the 10 real studies using reported HRs. **b**, Meta-analysis results under a simulated reporting bias scenario, where 5 studies were randomly excluded to mimic unreported data. **c**, Meta-analysis results incorporating HRs derived from IPD reconstructed by KM-PoPiGo for the five excluded studies. The pooled estimate aligns with the ground truth in **a**, indicating the recovery of prognostic information.

common intervention to establish the ground truth. We compared this ground truth against a biased scenario (excluding five studies) and a restored scenario where excluded studies were recovered via KM-PoPiGo.

In the ground truth scenario, pooled HRs were 0.77 (95% CI: 0.71–0.84) and 0.78 (95% CI: 0.65–0.93) under fixed-effect[6] and random-effects models[7] (Supplementary Fig. 2a). In the biased scenario, pooled estimates deviated from the ground truth (Supplementary Fig. 2b). In the restored scenario, results under the fixed-effect model were identical to the ground truth; the random-effects pooled HR differed by 0.01, with a 0.02 deviation in the 95% CI upper limit (Supplementary Fig. 2c). These results indicate that KM-PoPiGo can mitigate reporting bias caused by unreported HRs.

3. Experimental design and data sources for synthetic control simulation

Supplementary to Main Text Results subsection: "Synthetic controls and cross-trial comparison using reconstructed IPD".

Single-arm clinical trials, conducted for rare diseases or during early-phase development, lack a concurrent internal comparator. In such contexts, reconstructed IPD for time-to-event outcomes from published clinical trials involving clinically homogeneous populations can be used to construct an external synthetic control arm (SCA) for comparative effectiveness assessment.

To evaluate the performance of KM-PoPiGo in constructing SCAs, we simulated a single-arm trial scenario using the Real IPD Dataset. We selected six independent cohorts with comparable follow-up durations (maximum follow-up: 50–60 months) and analyzed them as a clinically homogeneous population under the assumption of a common survival endpoint (Supplementary Table 6). The cohort with the smallest sample size (GSE27150, n = 81) was designated as the surrogate treatment arm, while the remaining five cohorts served as external controls. The accuracy of the synthetic controls reconstructed via KM-PoPiGo was assessed by comparing the HRs from "treatment vs. synthetic control" analyses against the ground-truth HRs derived from "treatment vs. real control" comparisons.

The KM curve of the reconstructed SCA (GSE76427) was superimposed onto its ground truth, with the surrogate treatment arm (GSE27150) included for reference (Supplementary Fig. 3). The reconstructed curve aligned with the ground truth, and the Log-rank test showed no significant difference ($p = 0.88$). These results indicate that the reconstructed SCA reflects the survival characteristics of the original cohort.

4. Experimental design and data sources for the pooled intervention simulation

Supplementary to Main Text Results subsection: "Synthetic controls and cross-trial comparison using reconstructed IPD".

This section details the experimental design and cohort selection to assess the performance of KM-PoPiGo-reconstructed IPD in pooled intervention analyses. To evaluate if reconstructed data produce estimates consistent with real IPD within a meta-analysis framework, we selected 16 independent cohorts from the Real IPD Dataset. Cohorts with maximum follow-up durations of 50–70 months were selected to align with the 5-year (60-month) clinical benchmark[32], ensuring temporal comparability across studies. To maintain patient independence, inclusion was restricted to a single survival endpoint per cohort. When multiple endpoints within a cohort met the criteria, the outcome with the largest sample size (N) was selected to improve the statistical power and representation of the reference data[33,34]. The 16 cohorts were assigned to four simulated intervention arms (Interventions 1–4; Supplementary Table 7). This design assesses the statistical concordance between pooled estimates derived from KM-PoPiGo-reconstructed IPD and those from real IPD.

Supplementary Table 6 Characteristics of the six cohorts included in the synthetic control arm simulation

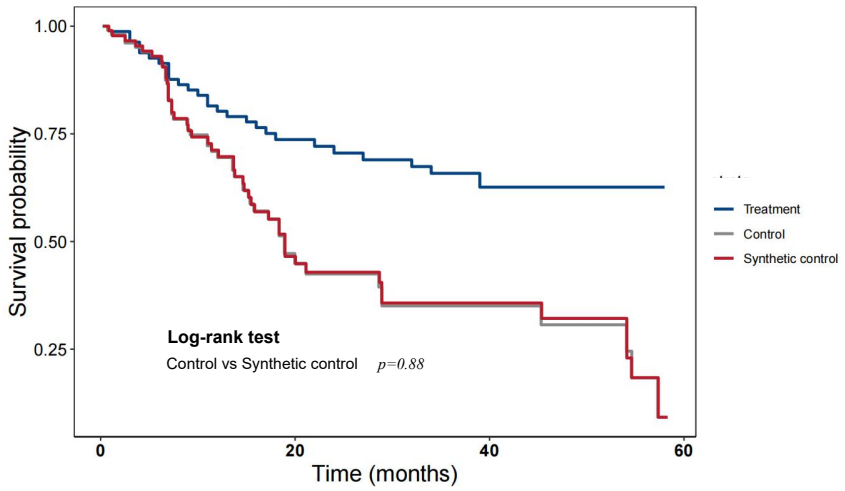
GEO accession*	Sample size (N)	Max follow-up (months)	Role in simulation
GSE27150	81	58.00	Surrogate treatment arm
GSE28735	84	51.00	Control 1
GSE41994	97	56.00	Control 2
GSE76427	108	58.32	Control 3
GSE53622	120	58.20	Control 4
GSE24550	154	59.64	Control 5

Note: All cohorts were analyzed under the assumption of a common survival endpoint, irrespective of original reporting (e.g., OS or PFS).

Supplementary Table 7 Characteristics of cohorts included in the pooled intervention analysis

GEO accession	Sample size	Max follow-up (months)	Assigned intervention
GSE76427	108	58.32	1
GSE69795	38	66.60	1
GSE72969	56	66.59	1
GSE53622	120	58.20	1
GSE48390	81	69.00	2
GSE45725	334	52.27	2
GSE41994	97	56.00	2
GSE28735	84	51.00	2
GSE27150	81	58.00	3
GSE24550	154	59.64	3
GSE205209	60	62.90	3
GSE159216	170	60.00	3
GSE146558	106	69.60	4
GSE12945	62	64.00	4
GSE117973	77	60.57	4
GSE109710	162	61.93	4

All cohorts were analyzed under the assumption of a unified survival endpoint, irrespective of original clinical reporting (e.g., OS, PFS, or DFS).



Supplementary Fig. 3 Validation of synthetic control reconstruction. Comparison between the real control (grey solid line) and the reconstructed synthetic control (red solid line) for the GSE76427 cohort. The surrogate treatment arm (GSE27150) is included for reference. The log-rank test ($p = 0.88$) shows no significant difference between the real and reconstructed control data, indicating the accuracy of the reconstruction.

5. Detailed performance benchmarking and statistical validation across constraint combinations

Supplementary to Main Text Results subsection: "Impact of different constraint combinations on reconstruction fidelity".

5.1 Evaluation of distributional fidelity in the TCGA-IPD subset across constraint combinations

Weibull distributions were fitted to the TCGA-IPD subset (21 KM curves), with reference curves and 95% CIs derived from the real IPD. Weibull distributions were fitted to the IPD reconstructed under each of the eight constraint combinations. Reconstruction fidelity was assessed by overlaying these curves onto the ground-truth plots for visual comparison.

Across all 21 cohorts, the distribution curves reconstructed under all constraint combinations remained within the 95% CIs of the ground truth fits (Supplementary Fig. 4). These observations indicate that KM-PoPiGo-reconstructed IPD exhibits distributional alignment with ground truth, with stable performance observed regardless of the constraint combinations applied.

5.2 Evaluation of HR estimation accuracy across constraint combinations

The eight constraint combinations were evaluated using 500 pairs of KM curves from the Synthetic KM Dataset. Estimation accuracy was quantified by the Mean Absolute Error (MAE) of the log-transformed hazard ratio ($\ln \mathrm{HR}$) relative to ground-truth values.

Impact across constraint combinations followed a hierarchy (Main Text Fig. 4b), with the highest MAE observed for sample size (N) alone (MAE = 0.123). While adding censoring times (C) or at-risk tables (R) improved accuracy, the inclusion of the number of events (E) provided the largest error reduction. For instance, supplementing the NR combination with E reduced the MAE from 0.097 to 0.035 ($p < 0.001$; Supplementary Fig. 5). However, incorporating R into combinations already containing E resulted in minimal marginal improvement. The fully constrained $NERC$ combination (MAE = 0.0264) performed comparably to the NEC combination (MAE = 0.0269, $p = 0.51$; Supplementary Table 8). These results indicate that while E and C are the primary determinants for HR estimation, the incremental benefit of adding at-risk tables (R) is non-significant when the number of events is available.

5.3 Evaluation of survival point estimate stability across constraint combinations

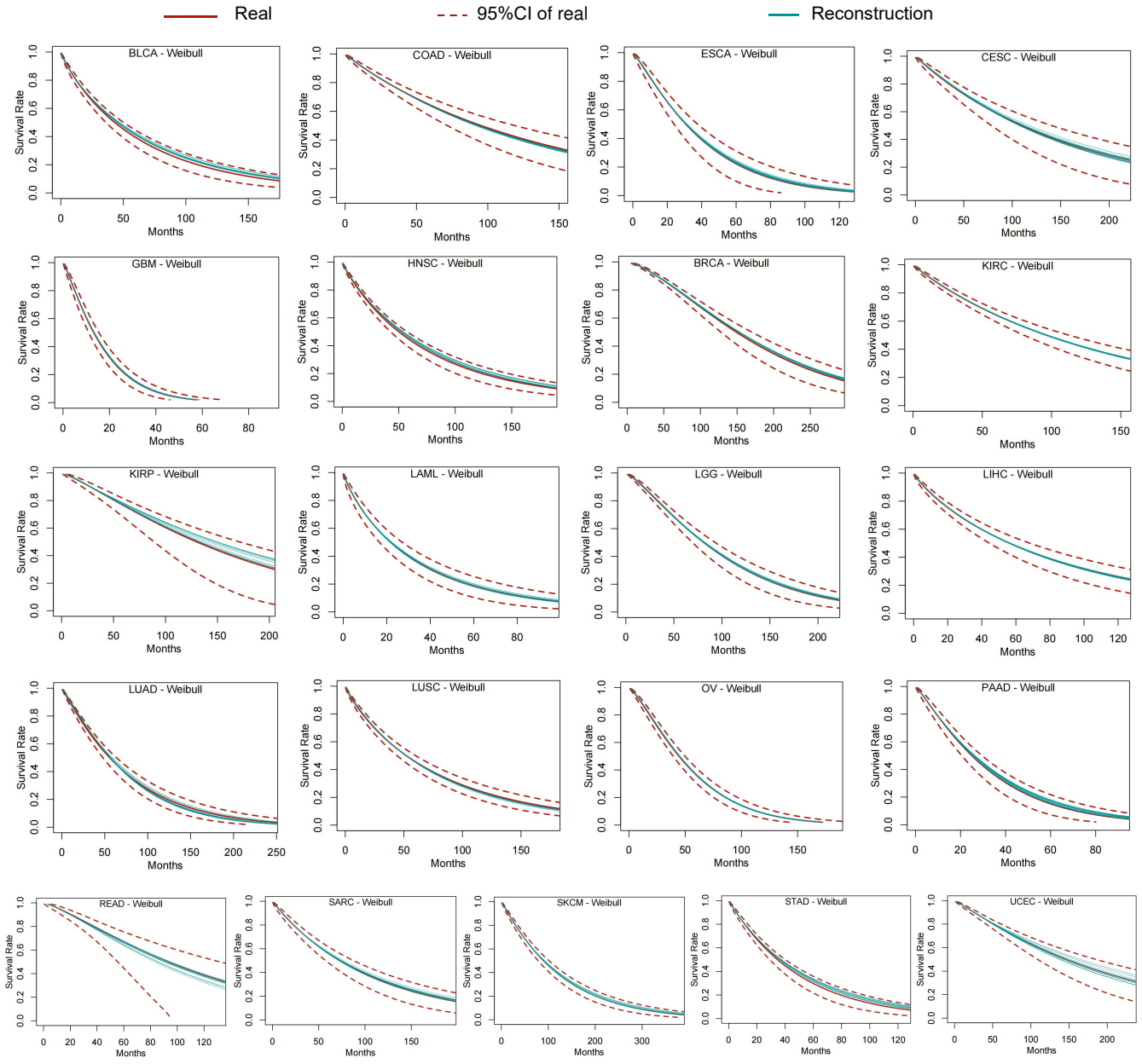
In contrast to HRs, survival point estimates—including median survival and survival rates—are coordinate-dependent metrics determined by the survivor function $S(t)$. Their accuracy depends on the fidelity of curve extraction rather than the reconstruction of the censorship distribution. To validate this geometric dependency, we assessed the impact of varying constraint combinations on reconstruction accuracy using 1,000 KM curves from the Synthetic KM Dataset.

Consistent with this hypothesis, increasing constraint complexity yielded no significant improvement in point estimate accuracy. Pairwise comparisons indicated no statistically significant difference between complex constraint combinations and the combination constrained solely by sample size (N) (Supplementary Tables 9 and 10). For median survival time, the combination constrained solely by sample size (N) achieved a Mean Absolute Error (MAE) of 0.2804, showing no significant difference from the fully constrained $NERC$ combination ($p = 0.72$; Supplementary Table 9). A similar trend was observed for survival rates, where the N-only combination (MAE = 0.00356) was statistically equivalent to the $NERC$ combination ($p = 0.95$; Supplementary Table 10).

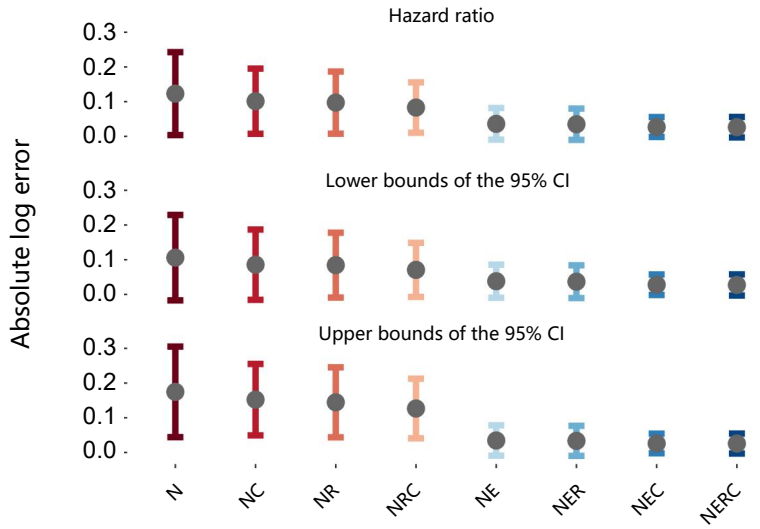
These findings indicate that for research objectives restricted to point estimates, curve extraction constrained by sample size (N) is sufficient for accuracy reconstruction. While improving HR estimation accuracy, additional constraints (E, R, C) provide limited incremental benefit for these coordinate-dependent metrics.

Supplementary Table 8 Pairwise validation of HR estimation accuracy across constraint combinations. MAE of log-transformed HRs and p-values from pairwise t-tests ($n = 500$ pairs of KM curves). Results validate the findings in Main Text Fig. 4b and Supplementary Fig. 5

	MAE	*NERC*	*NEC*	*NER*	*NE*	*NRC*	*NR*	*NC*	*N*
NERC	0.0264	1	0.51	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
NEC	0.0269	-	1	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
NER	0.0349	-	-	1	0.29	<0.001	<0.001	<0.001	<0.001
NE	0.0360	-	-	-	1	<0.001	<0.001	<0.001	<0.001
NRC	0.0830	-	-	-	-	1	<0.001	<0.001	<0.001
NR	0.0972	-	-	-	-	-	1	0.19	<0.001
NC	0.1013	-	-	-	-	-	-	1	<0.001
N	0.1231	-	-	-	-	-	-	-	1



Supplementary Fig. 4 Distributional fidelity of IPD reconstructed under eight constraint combinations in the TCGA-IPD subset. Weibull distributions were fitted to real and reconstructed IPD across 21 cancer types. Red solid lines: ground-truth distributions fitted to real IPD; red dashed lines: 95% confidence intervals (CIs). Blue solid lines ($n = 8$ per plot): distributions fitted to IPD reconstructed under each of the eight constraint combinations. The tight clustering of the eight reconstruction lines, all confined within the 95% CIs, demonstrates the stability of distributional alignment across varying constraints.



Supplementary Fig. 5 Evaluation of estimation accuracy for the hazard ratio and its 95% confidence interval across eight constraint combinations. Mean absolute error (MAE) and standard deviation of absolute errors for the log-transformed hazard ratio ($\ln$HR) and the lower and upper bounds of the 95% CI, evaluated for $n = 500$ pairs of KM curves. Results correspond to the error distributions in Main Text Fig. 4b.

Supplementary Table 9 Evaluation of median survival time estimation stability. Mean absolute error (MAE) and p-values from pairwise t-tests across constraint combinations ($n = 1,000$ KM curves).

	MAE	*NERC*	*NEC*	*NER*	*NE*	*NRC*	*NR*	*NC*	*N*
NERC	0.3096	1	0.10	0.92	0.75	0.49	0.36	0.50	0.72
NEC	0.5273	-	1	0.10	0.05	0.03	0.31	0.29	0.06
NER	0.3180	-	-	1	0.66	0.42	0.43	0.59	0.64
NE	0.2887	-	-	-	1	0.55	0.18	0.34	0.90
NRC	0.2611	-	-	-	-	1	0.07	0.19	0.77
NR	0.3918	-	-	-	-	-	1	0.87	0.05
NC	0.3745	-	-	-	-	-	-	1	0.35
N	0.2804	-	-	-	-	-	-	-	1

Supplementary Table 10 Evaluation of survival rate estimation stability. Mean absolute error (MAE) and p-values from pairwise t-tests across constraint combinations ($n = 1,000$ KM curves).

	MAE	*NERC*	*NEC*	*NER*	*NE*	*NRC*	*NR*	*NC*	*N*
NERC	0.00356	1	0.90	0.12	0.92	<0.001	<0.001	<0.001	0.95
NEC	0.00356	-	1	0.10	0.98	<0.001	<0.001	<0.001	0.98
NER	0.00363	-	-	1	0.09	<0.001	0.01	<0.001	0.24
NE	0.00405	-	-	-	1	<0.001	<0.001	<0.001	0.99
NRC	0.00356	-	-	-	-	1	<0.001	<0.001	<0.001
NR	0.00387	-	-	-	-	-	1	0.08	<0.001
NC	0.00379	-	-	-	-	-	-	1	<0.001
N	0.00356	-	-	-	-	-	-	-	1

6. Benchmarking of IPD reconstruction methods on the Published KM Dataset

Supplementary to Main Text Results subsection: "Benchmarking of reconstruction methods".

To benchmark KM-PoPiGo against existing reconstruction methods, we analyzed the Published KM Dataset, evaluating accuracy across hazard ratios, median survival times, and survival rates. Evaluation protocols were stratified by method compatibility: iKM[10] and IPDfromKM[11] were assessed under N, NE, NR, and NER combinations, while KMtoIPD[12] was evaluated under the NC combination.

KM-PoPiGo was compared with existing methods across applicable constraint combinations (N through $NERC$). Supplementary Table 11 lists the counts of survival statistics recovered by methods within their supported constraint frameworks. To ensure comparability, benchmarking was restricted to the "Common set"—defined as the strict intersection of these valid extractions, comprising only KM curves where all compared methods generated output for the same source survival statistics. The resulting counts are provided in the bottom row of each section in Supplementary Table 11.

Supplementary Fig. 6 illustrates the reconstruction accuracy of survival rates across evaluated methods and applicable constraint combinations. Error distributions under combinations lacking event information (N, NR, and NC) exhibited notable dispersion compared to fully constrained configurations, a trend driven primarily by the presence of outliers (Main Text Fig. 5e). While our large-scale synthetic validation ($n > 4,000$) established the equivalence of these combinations, the reduced sample size of the real-world cohort ($n \approx 400$–600) amplifies the visual impact of these stochastic outliers on error dispersion.

Comparative analysis indicates that across N, NE, NR, and NER combinations, IPDfromKM yielded higher Mean Absolute Errors (MAE) for survival rates relative to iKM and KM-PoPiGo. Under the NC combination, KM-PoPiGo exhibited a higher MAE than KMtoIPD (0.017 vs. 0.016; $0.001 < p < 0.01$). This minor discrepancy reflects distinct optimization objectives: KMtoIPD prioritizes local geometric accuracy, whereas KM-PoPiGo optimizes for global multi-constraint consistency. The standard deviation of absolute errors (SD(AE)) was identical for both methods (0.156), confirming that despite the shift in mean error, the error dispersion and algorithmic stability remained indistinguishable.

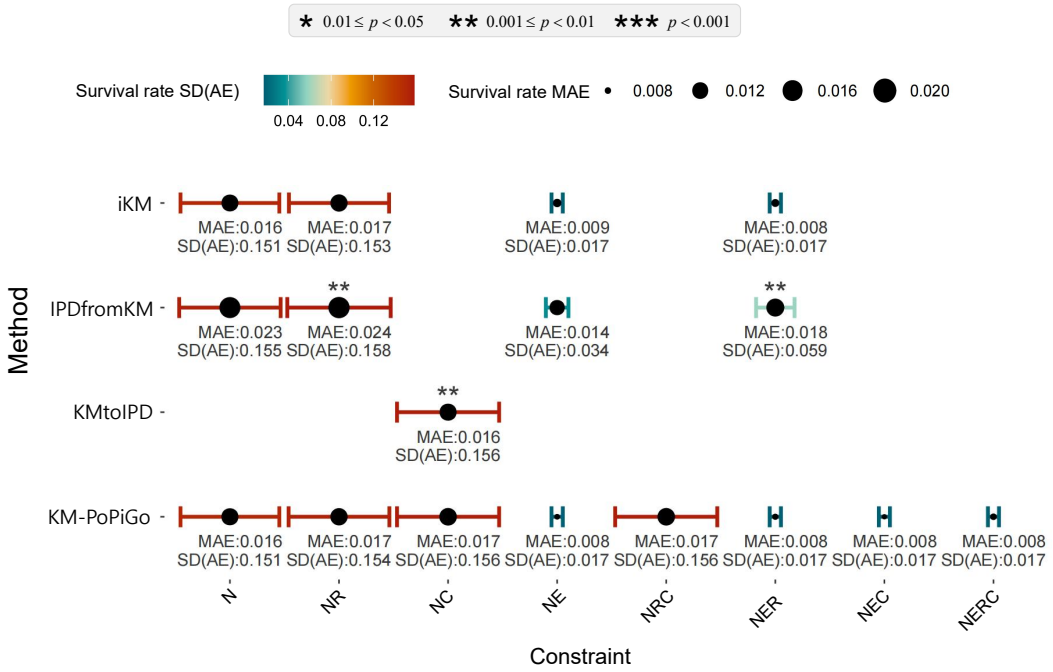
7. Supplementary results of independent clinical validation

Supplementary to Main Text Results subsection: "Independent blinded clinical validation on unseen data".

To ensure objective assessment, independent clinical practitioners conducted a blinded validation using proprietary clinical datasets. They reconstructed IPD from KM curve image derived from two distinct patient cohorts according to the KM-PoPiGo workflow.

Supplementary Table 11 Counts of verifiable survival statistics for comparative benchmarking. Values represent the subset of survival statistics available in the source KM curves that were recovered by each method within its supported constraint framework. The "Common set" defines the intersection of these recovered instances across all applicable methods, ensuring that performance comparisons are conducted on a common set of verifiable KM curves.

Metric	Method	*N*	*NE*	*NR*	*NC*	*NER*	*NEC*	*NRC*	*NERC*
HR	iKM	981	483	920	-	466	-	-	-
	IPDfromKM	978	516	943	-	496	-	-	-
	KMtoIPD	-	-	-	737	-	-	-	-
	KM-PoPiGo	981	516	948	737	498	429	711	418
	Common set	**978**	**483**	**917**	**737**	**466**	**429**	**711**	**418**
Median survival time	iKM	1511	801	1393	-	775	-	-	-
	IPDfromKM	1493	861	1418	-	822	-	-	-
	KMtoIPD	-	-	-	1214	-	-	-	-
	KM-PoPiGo	1511	869	1444	1216	833	746	1164	724
	Common set	**1492**	**798**	**1374**	**1214**	**769**	**746**	**1164**	**724**
Survival rate	iKM	648	417	628	-	433	-	-	-
	IPDfromKM	598	403	597	-	409	-	-	-
	KMtoIPD	-	-	-	560	-	-	-	-
	KM-PoPiGo	646	448	647	566	448	406	566	406
	Common set	**595**	**386**	**560**	**401**	**397**	**406**	**566**	**406**



Supplementary Fig. 6 Benchmarking survival rate reconstruction on the Published KM Dataset. Error analysis of four reconstruction methods across their applicable constraint configurations. Points represent the Mean Absolute Error (MAE), while the width of the error bars represent the Standard Deviation of Absolute Errors (SD(AE)).

As illustrated in Supplementary Fig. 7, the KM curves generated from the reconstructed IPD approximated the ground truth, a finding corroborated by Log-rank tests showing no statistically significant differences in either cohort (Group 0: $p = 0.89$; Group 1: $p = 0.98$). Furthermore, the fitted Weibull distributions remained within the 95% confidence intervals of the ground-truth distributions.

The fitted Weibull distributions aligned with the ground truth. The reconstructed trajectories remained within the 95% confidence intervals of the ground-truth distributions throughout the observed timeframe (0–12 years), suggesting its utility as a robust framework for clinical researchers to facilitate effective and high-fidelity data reconstruction.

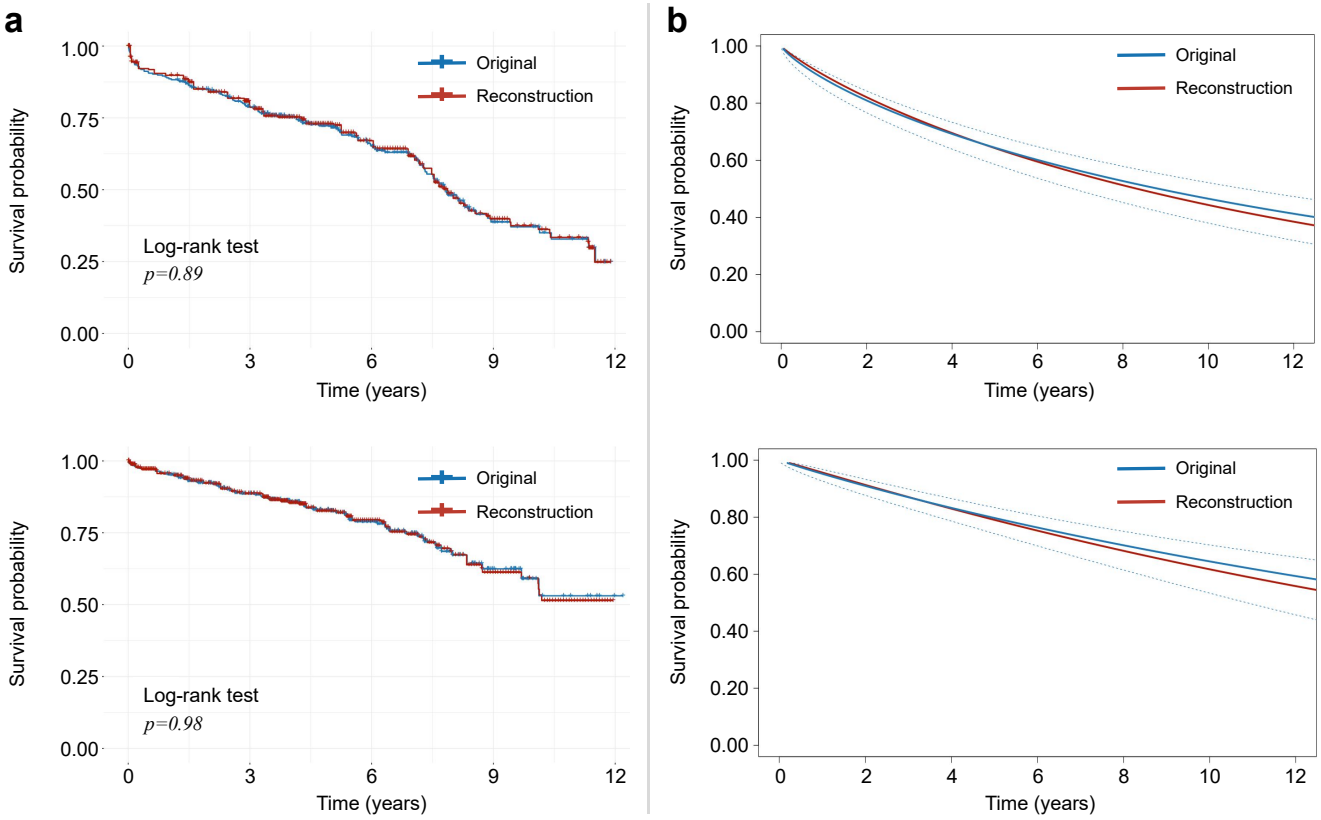
8. KM curve topologies and derivation of censoring constraints

As defined in the Methods subsection "Mathematical representation of KM curves", the curve's geometry corresponds to the survival data. Vertical drops represent event counts, while the horizontal segments between them represent time intervals where censoring counts are situated. This geometric logic also applies to the curve's termination, leading to two distinct topologies:

(1) Plateau (Right-censored tail): As shown in Supplementary Fig. 8a, if the curve ends with a horizontal line, this final segment spans the time from the last event to the last follow-up. This defines a valid interval where censoring counts are accounted for.

(2) Vertical drop (Event-terminated tail): As shown in Supplementary Fig. 8b and c, the curve ends with a vertical drop, indicating the observation stopped at an event. To handle this, we define a zero-duration interval at this final point. The censoring count is determined by the optimization process. The objective function naturally guides the evolution: for curves ending above zero (Supplementary Fig. 8b), the process tends to retain a positive censoring count to match the terminal survival probability; conversely, when the curve drops to zero (Supplementary Fig. 8c), the algorithm favors a zero censoring count, as solutions implying remaining survivors are typically eliminated due to poor fit.

Matching this geometric structure, the dimension of the constraint vector $\boldsymbol{c}^{min}$ equals the number of intervals partitioned by the key points. This vector is constructed based on the user-marked censoring times C: each element of $\boldsymbol{c}^{min}$ represents the number of censoring times falling within each interval. In the example (Supplementary Fig. 8d), the counts 1, 0, 2, and 1 constitute the vector $\boldsymbol{c}^{min} = [1, 0, 2, 1]$.



Supplementary Fig. 7 Independent blind validation results from clinical practitioners. a, Analysis of Kaplan-Meier survival curves derived from proprietary clinical datasets (original, blue) versus IPD reconstructed by practitioners using KM-PoPiGo (reconstructed, red) across two distinct cohorts. **b**, Comparison of fitted Weibull distribution models. The original models are shown with 95% confidence intervals (blue dashed lines), while reconstructed models are shown as red solid lines. These reconstructed models remain within the 95% CI of the original models throughout follow-up (0–12 years), reflecting the reconstruction's alignment with the clinical source data.

9. Mathematical formulation: IPD reconstruction as an underdetermined system

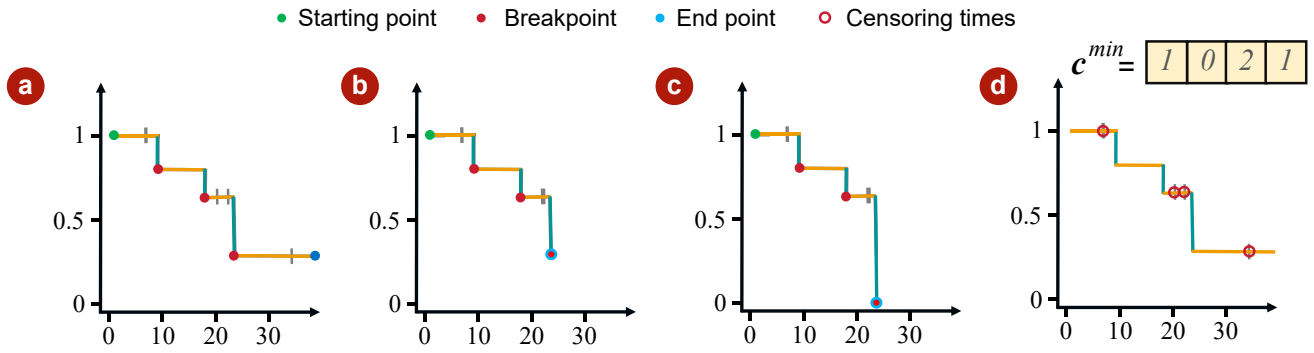
The geometry of the KM curve is characterized by the sequence of key points $\{(t_i, p_i)\}_{i=0}^{M}$. From this sequence, we define two vectors, both of dimension $M+1$: the time vector $\boldsymbol{t} = [t_0, t_1, \ldots, t_M]^T$, corresponding to the t-coordinates, and the survival probability vector $\boldsymbol{p} = [p_0, p_1, \ldots, p_M]^T$, corresponding to the p-coordinates. These vectors map the curve's stepwise trajectory.

Based on this geometric structure, the reconstruction task involves determining two sequences of non-negative integers that define the risk set dynamics:

- e_i: the number of events at the key time point t_i, for $i \in \{1, \ldots, M\}$. (Note: $e_i > 0$ implies a vertical drop at t_i).
- c_i: the censoring count within the i-th interval $[t_{i-1}, t_i)$, for $i \in \{1, \ldots, M\}$.

Let n_i denote the number of patients at risk prior to the event time t_i. These variables follow the recursive relationship defined by the Kaplan-Meier estimator:

$$\begin{cases} n_1 = N - c_1 \\ n_i = n_{i-1} - e_{i-1} - c_i \quad \text{for } i = 2, \ldots, M \\ p_i = p_{i-1} \cdot \dfrac{n_i - e_i}{n_i} \quad \text{for } i = 1, \ldots, M \\ N = \displaystyle\sum_{i=1}^{M} e_i + \sum_{j=1}^{M} c_j \end{cases}$$



Supplementary Fig. 8 KM curve topologies and constraint construction. a, Key points and plateau topology. Visualization of the key points (starting point, breakpoints and end point) and censoring times. The horizontal tail (plateau) spans the time from the last event to the last follow-up, defining an interval where censoring can occur. **b, c**, Event-terminated topologies (vertical drops). Examples where the curve terminates with a vertical drop, either ending at a non-zero probability (**b**) or dropping to zero (**c**). In both scenarios, the final event is instantaneous. Thus, a zero-duration interval is defined to maintain structural consistency. The lowest marker with a red center and blue outline denotes a terminal event, indicating that the observation period ended precisely at the time of an event. **d**, Construction of the censoring constraint vector $\boldsymbol{c}^{min}$. Each element represents the count of user-marked censoring times (C) falling within the corresponding interval. In this example, the counts of 1, 0, 2, and 1 form the constraint vector $\boldsymbol{c}^{min} = [1,0,2,1]$.

where $p_0 = 1$, N is the sample size, and the sequences e and c are defined over the full range $i \in \{1,\ldots,M\}$ to accommodate both plateau and event-terminated topologies.

This formulation seeks to resolve the unknowns subject to the observed constraints:

- **Unknown variables:** The set of unknowns comprises both the event counts and censoring counts: $\{e_1,\ldots,e_M,c_1,\ldots,c_M\}$. The total number of unknown integer variables is $M+M=2M$.
- **Constraints (equations):** The formulation incorporates M independent equations derived from the observed survival probabilities $(p_1,\ldots,p_M)$ at the key time points. While the sample size N is fixed, the system remains mathematically underdetermined because the number of unknowns ($2M$) exceeds the number of constraints (M), creating M degrees of freedom that motivate the optimization approach described below.

To address this underdeterminacy, we employ genetic optimization within KM-PoPiGo. By incorporating the available constraints—specifically the censoring lower bounds $\boldsymbol{c}^{min}$ derived from user-marked censoring times C, total events E, and at-risk table R—the algorithm narrows the feasible region to identify a solution consistent with both the curve geometry and the observed data constraints.

10. Pseudocode for the genetic optimization algorithm

Three pseudocode modules describe the genetic optimization-based IPD reconstruction. Algorithm 1 details the iterative framework, which adopts steady-state genetic algorithm framework[13,14] to maintain selection pressure. Algorithm 2 handles the initialization, ensuring the population complies with constraints on sample size N and event count E. Algorithm 3 specifies the constraint-preserving mutation operations used to navigate the search space.

Algorithm 1. Genetic optimization framework

procedure EVOLUTIONARYOPTIMIZATION
  Input:
    N (Sample size), E (Number of events), M (Number of intervals)
    $\boldsymbol{p}$ (Key survival probabilities), C (Censoring times), R (At-risk table)
    $N_P = 100$ (Population size)
  Output: Highest-scoring candidate $\boldsymbol{\theta}_{\text{best}}$
  Step 1: Population initialization
  $\mathscr{P} = \{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_{N_P}\} \leftarrow$ INITIALIZEPOPULATION(N,E,M,N_P) ▷ See Algorithm 2
    where each $\boldsymbol{\theta} = (c_1^{'},\ldots,c_M^{'};e_1^{'},\ldots,e_{M-1}^{'}) \in \mathbb{Z}_{\geq 0}^{2M-1}$

Step 2: Initialize best candidate
Calculate fitness $f(\boldsymbol{\theta})$ for all $\boldsymbol{\theta} \in \mathscr{P}$ using $\boldsymbol{p}, C, R$
$\boldsymbol{\theta}_{\text{best}} \leftarrow \arg\min_{\boldsymbol{\theta} \in \mathscr{P}} f(\boldsymbol{\theta})$
$k \leftarrow 0$ ▷ Convergence counter
while $k < 200$ **do**
  Step 3: Ranking
  Sort candidates in $\mathscr{P}$ by $f(\boldsymbol{\theta})$ in ascending order
  Step 4: Selection
  Select 50 candidates **uniformly at random** from $\mathscr{P}$ to obtain $\mathscr{P}_{\text{select}}$
  Step 5: Mutation
  $\mathscr{P}_{\text{offspring}} \leftarrow \emptyset$
  for each $\boldsymbol{\theta}_i$ in $\mathscr{P}_{\text{select}}$ **do**
    $\boldsymbol{\theta}_i' \leftarrow \textsc{Mutate}(\boldsymbol{\theta}_i, E, M)$ ▷ See Algorithm 3
    Add $\boldsymbol{\theta}_i'$ to $\mathscr{P}_{\text{offspring}}$
  end for
  Step 6: Replacement (steady-state)
  Replace the 50 worst candidates (indices $51 \ldots 100$) in $\mathscr{P}$ with $\mathscr{P}_{\text{offspring}}$
  Step 7: Check convergence
  $\boldsymbol{\theta}_{\text{best}}^{\text{new}} \leftarrow \arg\min_{\boldsymbol{\theta} \in \mathscr{P}} f(\boldsymbol{\theta})$
  if $f(\boldsymbol{\theta}_{\text{best}}^{\text{new}}) < f(\boldsymbol{\theta}_{\text{best}})$ **then**
    $\boldsymbol{\theta}_{\text{best}} \leftarrow \boldsymbol{\theta}_{\text{best}}^{\text{new}}$
    $k \leftarrow 0$ ▷ Reset counter
  else
    $k \leftarrow k + 1$
  end if
end while
return $\boldsymbol{\theta}_{\text{best}}$
end procedure

Algorithm 2. Population initialization

procedure INITIALIZEPOPULATION(N, E, M, N_P)
  $\mathscr{P} \leftarrow \emptyset$
  Let $\mathscr{I}_{\text{event}}$ be the indices of event elements in $\boldsymbol{\theta}$ ▷ Positions $2, 4, \ldots, 2M-2$
  Let $\mathscr{I}_{\text{censor}}$ be the indices of censoring elements in $\boldsymbol{\theta}$ ▷ Positions $1, 3, \ldots, 2M-1$
  for $k = 1$ to N_P **do**
    Initialize vector $\boldsymbol{\theta}$ of length $2M-1$ with zeros
    if E is provided **then** ▷ Scenario A: E known
      Set $\boldsymbol{\theta}[j] \leftarrow 1$ for all $j \in \mathscr{I}_{\text{event}}$ ▷ Satisfy structural constraint $e_i' \geq 1$
      $E_{\text{resid}} \leftarrow E - (M-1)$
      Distribute integer E_{resid} randomly across indices $\mathscr{I}_{\text{event}}$
      Distribute integer $N - E$ randomly across indices $\mathscr{I}_{\text{censor}}$
    else ▷ Scenario B: E unknown
      Set $\boldsymbol{\theta}[j] \leftarrow 1$ for all $j \in \mathscr{I}_{\text{event}}$ ▷ Satisfy structural constraint $e_i' \geq 1$
      $N_{\text{resid}} \leftarrow N - (M-1)$
      Distribute integer N_{resid} randomly across all indices $\{1, \ldots, 2M-1\}$
    end if
    Add $\boldsymbol{\theta}$ to $\mathscr{P}$
  end for
  return $\mathscr{P}$
end procedure

Algorithm 3. Constraint-preserving mutation

procedure MUTATE($\boldsymbol{\theta}, E, M$)
  Let $\mathscr{I}_{\text{censor}}$ be odd indices $\{1, 3, \ldots, 2M-1\}$
  Let $\mathscr{I}_{\text{event}}$ be even indices $\{2, 4, \ldots, 2M-2\}$

$k \leftarrow$ Randomly select an index from $\{1, \ldots, 2M-1\}$
$val \leftarrow \boldsymbol{\theta}[k]$
$r \leftarrow 0$
// Phase 1: Reductive step
if $k \in \mathscr{I}_{\text{censor}}$ **then**
    $limit \leftarrow val$ ▷ Can reduce down to 0
    if $limit > 0$ **then**
        With 50% prob: $r \leftarrow \min(1, limit)$
        Else: $r \leftarrow \text{RandomInteger}(0, limit)$
    end if
else ▷ $k \in \mathscr{I}_{\text{event}}$
    $limit \leftarrow val - 1$ ▷ Must keep at least 1
    if $limit > 0$ **then**
        With 50% prob: $r \leftarrow \min(1, limit)$
        Else: $r \leftarrow \text{RandomInteger}(1, limit)$
    end if
end if
$\boldsymbol{\theta}[k] \leftarrow \boldsymbol{\theta}[k] - r$
// Phase 2: Compensatory step
if $r > 0$ **then**
    if E is provided **then** ▷ Scenario A: Maintain field-specific sums
        if $k \in \mathscr{I}_{\text{censor}}$ **then**
            Target region $\Delta \leftarrow \mathscr{I}_{\text{censor}}$
        else
            Target region $\Delta \leftarrow \mathscr{I}_{\text{event}}$
        end if
    else ▷ Scenario B: Maintain only global sum N
        Target region $\Delta \leftarrow \{1, \ldots, 2M-1\}$
    end if
    for $j = 1$ to r **do**
        $p \leftarrow$ Randomly select index from Δ
        $\boldsymbol{\theta}[p] \leftarrow \boldsymbol{\theta}[p] + 1$
    end for
end if
return $\boldsymbol{\theta}$
end procedure

11. Efficacy of the constrained censoring strategy

Clinical trial protocols mandate fixed-interval assessments, resulting in the temporal clustering of censoring times. Consequently, a single visual censoring tick on a KM curve may represent an aggregate of multiple patients, or several ticks may coincide within the same interval. This reality contradicts the uniform censoring assumption—a common simplification in conventional reconstruction strategies where censored patients are assumed to be spread evenly between inter-event intervals.

In contrast, KM-PoPiGo employs a "Constrained censoring" strategy (see Methods: Conversion of highest-scoring candidate to IPD). By utilizing the coordinates of marked censoring times, this approach localizes the censoring occurrences within a refined temporal window $[t^{\min}, t^{\max}]$ for each interval. Under this model, censored patients are assumed to be uniformly distributed within these observed bounds, rather than across the inter-event interval.

Validation using 500 KM curve pairs from the Synthetic KM Dataset compared uniform censoring with the constrained approach across four constraint combinations ($NC, NRC, NEC, NERC$). The constrained approach yielded lower mean absolute log HR errors and reduced variability, indicating a reduction in estimation bias (Supplementary Table 12).

12. Sensitivity analysis: Impact of sample size and effect size on reconstruction fidelity

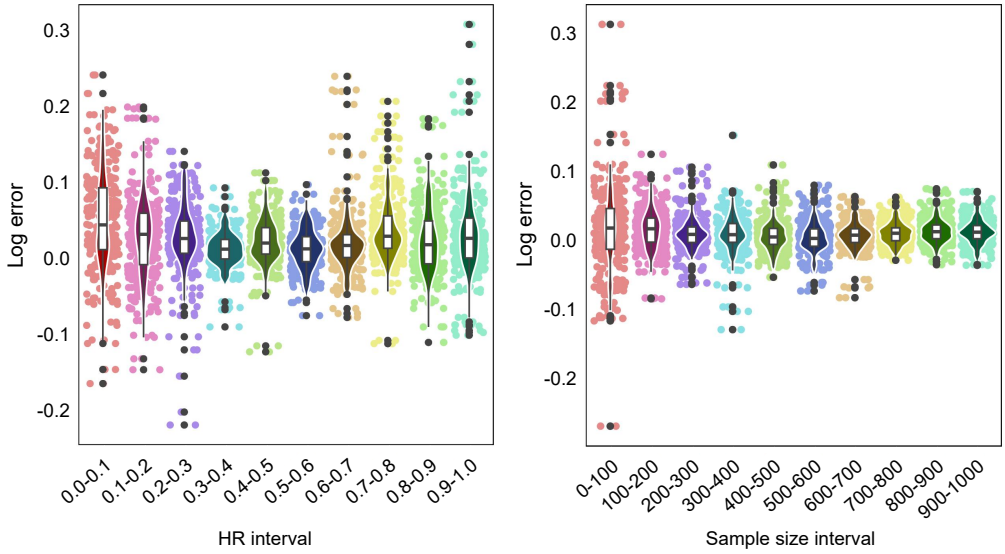
Sensitivity analysis evaluated the robustness of KM-PoPiGo against data-intrinsic factors using synthetic cohorts from the TCGA-IPD subset. Following simulation guidelines[15], reconstruction accuracy was quantified across various sample sizes (N) and effect sizes (HR) using a resampling-based strategy.

We assessed the sensitivity of KM-PoPiGo to Hazard Ratio (HR) magnitude by fixing the sample size at $N = 500$ and generating synthetic cohorts across 10 HR intervals (ranging from 0.0–0.1 to 0.9–1.0). To maintain an analysis range of $HR \in (0, 1]$, the

Supplementary Table 12 Comparison of hazard ratio estimation between constrained and uniform censoring strategies. Results are based on 500 KM curve pairs from the Synthetic KM Dataset evaluated across four constraint combinations.

Constraint Scenario	Strategy	MAE (log HR)	SD(AE)
NC	Constrained	**0.101**	**0.094**
	Uniform	0.114	0.105
NRC	Constrained	**0.083**	**0.073**
	Uniform	0.095	0.086
NEC	Constrained	**0.025**	**0.029**
	Uniform	0.034	0.042
NERC	Constrained	**0.026**	**0.029**
	Uniform	0.034	0.043

Note: MAE, mean absolute error; SD(AE), standard deviation of absolute errors. "Constrained" denotes the strategy distributing censored patients within observed bounds; "Uniform" assumes distribution across the inter-event interval. Bold values indicate the lowest error.



Supplementary Fig. 9 Sensitivity analysis of reconstruction fidelity. a, Distribution of log HR errors across stratified HR intervals (fixed $N = 500$). The reconstruction error exhibits a U-shaped pattern, with lower error magnitudes within the HR range of 0.3–0.7. **b**, Distribution of log HR errors across stratified sample size intervals (fixed HR ≈ 0.5). Reconstruction variability and error magnitude decline after the interval of $30 \leq N < 100$, stabilizing for sample sizes $N \geq 100$.

cohort with longer survival served as the reference. To test the influence of sample size, we fixed the effect size at a moderate level (HR ≈ 0.5) and varied N from 30 to 1000 across 10 intervals. For each condition, 20 KM images were synthesized, each containing a pair of KM curves, providing 400 images in total. By comparing the reconstructed IPD with the simulation ground truth, the algorithm's accuracy was quantified across the specified effect and sample size configurations.

As illustrated in Supplementary Fig. 9a, the reconstruction error follows a characteristic U-shaped trajectory relative to the effect size. Accuracy remains stable within the HR interval of $[0.3, 0.7]$. In contrast, error rates increase substantially at extreme HR values (< 0.3 or > 0.7; $P < 0.001$, Student's t-test). This degradation likely stems from visual overlap in similar curves or information sparsity in cases of extreme separation. Reconstruction accuracy also demonstrated sensitivity to sample size (Supplementary Fig. 9b). Patient cohorts with smaller sizes ($30 \leq N < 100$) yielded higher error magnitudes than those with $N \geq 100$ ($P < 0.001$, Student's t-test). These findings indicate the sensitivity of small patient samples to reconstruction bias, suggesting that HR estimates from such curves should be interpreted with caution.

13. Source data and provenance

Supplementary Table 13 Characteristics of the GEO-IPD subset within the Real IPD Dataset

Index	Dataset ID	Endpoint	Reference	Sample size (N)	Number of events (E)
1	GSE1456	OS	Pawitan et al.[16]	318	80
2	GSE1456	RFS	Pawitan et al.[16]	318	80
3	GSE1992	RFS	Hu et al.[17]	119	34
4	GSE1992	OS	Hu et al.[17]	119	26
5	GSE2607	RFS	Perreard et al.[18]	90	28
6	GSE2607	OS	Perreard et al.[18]	90	24
7	GSE2990	RFS	Sotiriou et al.[19]	122	46
8	GSE3141	OS	Bild et al.[20]	111	58
9	GSE3143	OS	Bild et al.[20]	158	50
10	GSE4475	OS	Hummel et al.[21]	159	75
11	GSE4922	DFS	Ivshina et al.[22]	484	166
12	GSE5327	MFS	Minn et al.[23]	58	29
13	GSE6128	RFS	Hoadley et al.[24]	38	14
14	GSE6128	OS	Hoadley et al.[24]	38	8
15	GSE6130	RFS	Mullins et al.[25]	108	31
16	GSE6130	OS	Mullins et al.[25]	107	23
17	GSE7378	DFS	Zhou et al.[26]	54	9
18	GSE8842	OS	Marchini et al.[27]	82	15
19	GSE9195	RFS	Loi et al.[28]	77	10
20	GSE10300	RFS	Cohen et al.[29]	44	30
21	GSE10884	RFS	Herschkowitz et al.[30]	43	17
22	GSE10884	OS	Herschkowitz et al.[30]	43	14
23	GSE10886	RFS	Parker et al.[31]	101	16
24	GSE10886	OS	Parker et al.[31]	101	10
25	GSE11117	OS	Baty et al.[32]	41	23
26	GSE11121	DMFS	Schmidt et al.[33]	200	46
27	GSE11969	OS	Takeuchi et al.[34]	149	74
28	GSE12093	DFS	Zhang et al.[35]	136	20
29	GSE12945	TFS	Staub et al.[36]	51	5
30	GSE12945	OS	Staub et al.[36]	62	12
31	GSE13041	OS	Lee et al.[37]	267	246
32	GSE13507	OS	Kim et al.[38]	165	69
33	GSE14764	OS	Denkert et al.[39]	80	21
34	GSE14814	OS	Zhu et al.[40]	133	60
35	GSE14814	DSS	Zhu et al.[40]	90	29
36	GSE16125	OS	Reid et al.[41]	74	48
37	GSE16560	OS	Sboner et al.[42]	281	206
38	GSE17260	PFS	Yoshihara et al.[43]	110	76
39	GSE17260	OS	Yoshihara et al.[43]	110	46
40	GSE17536	OS	Smith et al.[44]	177	73
41	GSE17536	DFS	Smith et al.[44]	145	36
42	GSE17537	OS	Smith et al.[44]	55	20
43	GSE17537	DFS	Smith et al.[44]	42	6
44	GSE17710	RFS	Wilkerson et al.[45]	56	36
45	GSE17710	OS	Wilkerson et al.[45]	56	33
46	GSE18229	RFS	Prat et al.[46]	254	78
47	GSE18229	OS	Prat et al.[46]	253	59
48	GSE18655	RFS	Barwick et al.[47]	139	33
49	GSE19161	OS	Konstantinopoulos et al.[48]	61	32
50	GSE19423	OS	Kim et al.[49]	48	16
51	GSE19783	DFS	Enerly et al.[50]	211	87
52	GSE19829	OS	Konstantinopoulos et al.[51]	70	40
53	GSE20685	OS	Kao et al.[52]	327	83

Continued on next page

Table 13 – continued from previous page

Index	Dataset ID	Endpoint	Reference	Sample size (N)	Number of events (E)
54	GSE20711	RFS	Dedeurwaerder et al.[53]	86	37
55	GSE20711	OS	Dedeurwaerder et al.[53]	88	25
56	GSE21653	DFS	Sabatier et al.[54]	248	79
57	GSE22133	OS	Jönsson et al.[55]	700	320
58	GSE22226	OS	Esserman et al.[56]	149	28
59	GSE23501	PFS	Shaknovich et al.[57]	69	17
60	GSE23501	OS	Shaknovich et al.[57]	69	13
61	GSE24549	DFS	Sveen et al.[58]	166	86
62	GSE24550	DFS	Sveen et al.[58]	154	38
63	GSE24556	OS	Végran et al.[59]	41	18
64	GSE24556	DFS	Végran et al.[59]	41	21
65	GSE25055	DRFS	Hatzis et al.[60]	309	65
66	GSE25065	DRFS	Hatzis et al.[60]	198	45
67	GSE25307	OS	Jönsson et al.[61]	551	228
68	GSE25727	DFS	Fountzilas et al.[62]	55	17
69	GSE26193	PFS	Gentric et al.[63]	107	80
70	GSE26193	OS	Gentric et al.[63]	107	76
71	GSE26253	RFS	Lee et al.[64]	432	177
72	GSE26304	OS	Muggerud et al.[65]	109	21
73	GSE26338	RFS	Harrell et al.[66]	244	71
74	GSE26338	OS	Harrell et al.[66]	243	54
75	GSE26712	OS	Bonome et al.[67]	185	129
76	GSE26939	OS	Wilkerson et al.[68]	115	66
77	GSE27020	DFS	Fountzilas et al.[69]	109	34
78	GSE27150	OS	Zhou et al.[70]	81	28
79	GSE28722	RFS	Loboda et al.[71]	125	33
80	GSE28722	OS	Loboda et al.[71]	125	55
81	GSE28735	OS	Zhang et al.[72]	84	58
82	GSE29013	PFS	Xie et al.[73]	55	28
83	GSE29013	OS	Xie et al.[73]	55	18
84	GSE29609	OS	Edeline et al.[74]	39	17
85	GSE29621	OS	Chen et al.[75]	65	25
86	GSE29621	DFS	Chen et al.[75]	53	9
87	GSE30161	OS	Ferriss et al.[76]	58	36
88	GSE30219	OS	Rousseaux et al.[77]	289	196
89	GSE30219	DFS	Rousseaux et al.[77]	274	114
90	GSE30378	DFS	Agesen et al.[78]	88	43
91	GSE30929	DRFS	Gobble et al.[79]	140	49
92	GSE31245	OS	Spentzos et al.[80]	58	24
93	GSE31428	PFS	Blumenschein et al.[81]	37	34
94	GSE31448	DFS	Sabatier et al.[82]	246	79
95	GSE37745	RFS	Botling et al.[83]	96	49
96	GSE37745	OS	Botling et al.[83]	196	145
97	GSE31684	RFS	Riester et al.[84]	93	39
98	GSE32063	PFS	Yoshihara et al.[85]	40	27
99	GSE32063	OS	Yoshihara et al.[85]	40	22
100	GSE32603	RFS	Yoshihara et al.[85]	248	79
101	GSE32894	OS	Sjödah et al.[86]	224	25
102	GSE32918	OS	Barrans et al.[87]	244	132
103	GSE37181	DFS	Callari et al.[88]	123	59
104	GSE38832	DSS	Tripathi et al.[89]	122	28
105	GSE38832	DFS	Tripathi et al.[89]	92	9
106	GSE39582	RFS	Marisa et al.[90]	536	147
107	GSE39582	OS	Marisa et al.[90]	573	190
108	GSE41994	DFS	Jansen et al.[91]	97	80
109	GSE42127	OS	Tang et al.[92]	173	61

Continued on next page

Table 13 – continued from previous page

Index	Dataset ID	Endpoint	Reference	Sample size (N)	Number of events (E)
110	GSE42568	RFS	Clarke et al.[93]	104	48
111	GSE42568	OS	Clarke et al.[93]	104	35
112	GSE43615	DFS	Medico et al.[94]	86	40
113	GSE44001	DFS	Lee et al.[95]	300	38
114	GSE45725	RFS	Wang et al.[96]	334	20
115	GSE46563	DMFS	Jonsdottir et al.[97]	94	23
116	GSE46602	BCR	Mortensen et al.[98]	36	22
117	GSE47115	OS	Yu et al.[99]	46	23
118	GSE48075	OS	Choi et al.[100]	73	45
119	GSE48276	OS	Choi et al.[100]	73	34
120	GSE48390	OS	Huang et al.[101]	81	11
121	GSE49997	OS	Pils et al.[102]	194	57
122	GSE49997	PFS	Pils et al.[102]	194	124
123	GSE50081	DFS	Der et al.[103]	181	51
124	GSE50081	OS	Der et al.[103]	181	75
125	GSE51088	OS	Karlan et al.[104]	152	112
126	GSE53031	RFS	Azim et al.[105]	167	42
127	GSE53622	OS	Li et al.[106]	120	66
128	GSE53624	OS	Li et al.[106]	238	146
129	GSE53963	OS	Konecny et al.[107]	174	153
130	GSE57495	OS	Chen et al.[108]	63	42
131	GSE58812	OS	Jézéquel et al.[109]	107	29
132	GSE58812	MFS	Jézéquel et al.[109]	107	31
133	GSE58984	DDFS	Fumagalli et al.[110]	94	11
134	GSE61304	DFS	Grinchuk et al.[111]	57	19
135	GSE61676	OS	Baty et al.[112]	86	68
136	GSE62452	OS	Yang et al.[113]	66	50
137	GSE63885	OS	Lisowska et al.[114]	75	66
138	GSE65904	DMFS	Cirenajwis et al.[115]	150	83
139	GSE65904	DSS	Cirenajwis et al.[115]	210	102
140	GSE66166	FFS	Pastore et al.[116]	106	47
141	GSE66166	OS	Pastore et al.[116]	106	32
142	GSE69031	RFS	Chin et al.[117]	124	34
143	GSE69049	OS	Sha et al.[118]	75	30
144	GSE69051	OS	Sha et al.[119]	218	51
145	GSE69795	OS	McConkey et al.[120]	38	14
146	GSE72968	OS	McConkey et al.[120]	68	49
147	GSE72968	PFS	McConkey et al.[120]	68	62
148	GSE72969	OS	Del et al.[121]	56	43
149	GSE72969	PFS	Del et al.[121]	56	52
150	GSE73403	OS	Feng et al.[122]	69	28
151	GSE75535	DSS	Levine et al.[123]	62	27
152	GSE75535	PFS	Levine et al.[123]	61	41
153	GSE76427	OS	Grinchuk et al.[124]	115	23
154	GSE76427	RFS	Grinchuk et al.[124]	108	48
155	GSE84426	OS	Yoon et al.[125]	76	35
156	GSE85047	OS	Rajbhandari et al.[126]	269	75
157	GSE85047	PFS	Rajbhandari et al.[126]	268	97
158	GSE85916	OS	Puleo et al.[127]	79	57
159	GSE86166	OS	Prabhakaran et al.[128]	366	259
160	GSE86166	RFS	Prabhakaran et al.[128]	355	71
161	GSE87211	OS	Hu et al.[129]	342	49
162	GSE87211	DFS	Hu et al.[129]	316	59
163	GSE87371	OS	Dubois et al.[130]	221	168
164	GSE87371	PFS	Dubois et al.[130]	221	150
165	GSE88770	DRFS	Metzger-Filho et al.[131]	117	26

Continued on next page

Table 13 – continued from previous page

Index	Dataset ID	Endpoint	Reference	Sample size (N)	Number of events (E)
166	GSE88770	OS	Metzger-Filho et al.[131]	117	28
167	GSE92921	DFS	Gotoh et al.[132]	59	6
168	GSE93984	PFS	Gotoh et al.[132]	88	12
169	GSE101929	OS	Mitchell et al.[133]	66	40
170	GSE102073	OS	Ducie et al.[134]	84	19
171	GSE102073	PFS	Ducie et al.[134]	85	57
172	GSE103479	OS	Allen et al.[135]	155	60
173	GSE103479	PFS	Allen et al.[135]	154	41
174	GSE103745	RFS	Sjöström et al.[136]	336	144
175	GSE106584	DFS	Zhu et al.[137]	156	52
176	GSE106584	OS	Zhu et al.[137]	156	78
177	GSE109710	DFS	Ignatiadis et al.[138]	162	26
178	GSE113863	DMFS	Yang et al.[139]	301	35
179	GSE115577	RFS	Kensler et al.[140]	1569	234
180	GSE116918	BCR	Jain et al.[141]	248	56
181	GSE116918	MFS	Jain et al.[141]	248	22
182	GSE117973	DSS	Mock et al.[142]	77	16
183	GSE117973	PFS	Mock et al.[142]	77	22
184	GSE124647	OS	Sinn et al.[143]	140	130
185	GSE124647	PFS	Sinn et al.[143]	140	130
186	GSE132342	OS	Millstein et al.[144]	3769	2475
187	GSE133057	OS	Ferrandon et al.[145]	33	11
188	GSE135565	OS	Kim et al.[146]	83	7
189	GSE135820	OS	Talhouk et al.[147]	3972	2661
190	GSE140082	OS	Kommoss et al.[148]	380	96
191	GSE140082	PFS	Kommoss et al.[148]	380	235
192	GSE143985	DFS	Shinto et al.[149]	91	15
193	GSE146558	RFS	Chen et al.[150]	106	12
194	GSE151335	OS	Hu et al.[151]	143	73
195	GSE151335	PFS	Hu et al.[151]	134	94
196	GSE152430	OS	Buechler et al.[152]	48	17
197	GSE152430	DMFS	Buechler et al.[152]	47	7
198	GSE157009	OS	Bueno et al.[153]	248	148
199	GSE157010	OS	Bueno et al.[153]	235	116
200	GSE159216	OS	Eide et al.[154]	170	107
201	GSE159956	OS	Regua et al.[155]	295	79
202	GSE161158	DFS	Szeglin et al.[156]	191	46
203	GSE162228	OS	Chen et al.[150]	133	31
204	GSE162228	RFS	Chen et al.[150]	127	16
205	GSE165808	OS	Kang et al.[157]	51	11
206	GSE167573	OS	Sun et al.[158]	57	10
207	GSE175692	OS	Brasó-Maristany et al.[159]	180	104
208	GSE192667	OS	Sugawara et al.[160]	89	72
209	GSE205209	OS	Jones et al.[161]	60	14
210	GSE205209	RFS	Jones et al.[161]	53	14

Abbreviations: OS, Overall Survival; DFS, Disease-Free Survival; DRFS, Distant Relapse-Free Survival; DDFS, Distant Disease-Free Survival; DSS, Disease-Specific Survival; PFS, Progression-Free Survival; RFS, Recurrence-Free Survival; DMFS, Distant Metastasis-Free Survival; MFS, Metastasis-Free Survival; BCR, Biochemical Recurrence.

Supplementary Table 14 Mapping of dataset identifiers to original studies

Index	Reference	Year	Journal
001	Perren et al.[162]	2011	*N. Engl. J. Med.*
002	Park et al.[163]	2017	*Eur. J. Cancer*

Continued on next page

Table 14 – continued from previous page

Index	Reference	Year	Journal
003	Fizazi et al.[164]	2017	*N. Engl. J. Med.*
004	James et al.[165]	2016	*Lancet*
005	Oza et al.[166]	2020	*Clin. Cancer Res.*
006	Zhou et al.[167]	2023	*Clin. Cancer Res.*
007	Park et al.[168]	2020	*Ann. Oncol.*
008	Goueffic et al.[169]	2020	*JACC Cardiovasc. Interv.*
009	Ozaka et al.[170]	2023	*Eur. J. Cancer*
010	Martin et al.[171]	2017	*Ann. Oncol.*
011	Matsumura et al.[172]	2022	*Langenbecks Arch. Surg.*
012	Dingemans et al.[173]	2015	*Ann. Oncol.*
013	Kato et al.[174]	2004	*N. Engl. J. Med.*
014	Reck et al.[175]	2013	*Lung Cancer*
015	Bossi et al.[176]	2017	*Ann. Oncol.*
016	Johnston et al.[177]	2021	*NPJ Breast Cancer*
017	Johnston et al.[178]	2023	*Lancet Oncol.*
018	Neven et al.[179]	2021	*Breast Cancer Res.*
019	Tolaney et al.[180]	2020	*Lancet Oncol.*
020	James et al.[181]	2017	*N. Engl. J. Med.*
021	Robertson et al.[182]	2009	*J. Clin. Oncol.*
022	De et al.[183]	2018	*Lancet Oncol.*
023	Mileshkin et al.[184]	2023	*Lancet Oncol.*
024	Tempero et al.[185]	2023	*J. Clin. Oncol.*
025	Herbst et al.[186]	2023	*J. Clin. Oncol.*
026	Chi et al.[187]	2023	*Clin. Cancer Res.*
027	Peng et al.[188]	2018	*BMC Cancer*
028	Pugh et al.[189]	2017	*Oncotarget*
029	West et al.[190]	2019	*Lancet Oncol.*
030	Oaknin et al.[191]	2024	*Lancet*
031	Reck et al.[192]	2019	*Lancet Respir. Med.*
032	Ko et al.[193]	2023	*Oncologist*
033	Galsky et al.[194]	2020	*Lancet*
034	Gutzmer et al.[195]	2020	*Lancet*
035	Pujade et al.[196]	2012	*J. Clin. Oncol.*
036	Cardone et al.[197]	2020	*Eur. J. Cancer*
037	Yoshino et al.[198]	2017	*Eur. J. Cancer*
038	Turner et al.[199]	2019	*Ann. Oncol.*
039	Coleman et al.[200]	2017	*Lancet Oncol.*
040	Tewari et al.[201]	2017	*Lancet*
041	Zielinski et al.[202]	2016	*Lancet Oncol.*
042	Zhou et al.[203]	2021	*ESMO Open*
043	Fukuoka et al.[204]	2011	*J. Clin. Oncol.*
044	Yoshino et al.[205]	2019	*Ann. Oncol.*
045	Yang et al.[206]	2023	*J. Thorac. Oncol.*
046	Camidge et al.[207]	2020	*J. Clin. Oncol.*
047	Camidge et al.[208]	2021	*J. Thorac. Oncol.*
048	Camidge et al.[209]	2018	*N. Engl. J. Med.*
049	Campone et al.[210]	2018	*Eur. J. Cancer*
050	Baselga et al.[211]	2017	*Lancet Oncol.*
051	Ren et al.[212]	2022	*J. Thorac. Oncol.*
052	Zhou et al.[213]	2021	*Lancet Respir. Med.*
053	Qin et al.[214]	2023	*Lancet*
054	Huang et al.[215]	2020	*Lancet Oncol.*
055	Yang et al.[216]	2021	*Lancet Oncol.*
056	Schmid et al.[217]	2020	*J. Clin. Oncol.*
057	Williams et al.[218]	2023	*Clin. Cancer Res.*
058	Herbst et al.[219]	2018	*Lancet Oncol.*

Continued on next page

Table 14 – continued from previous page

Index	Reference	Year	Journal
059	Sweeney et al.[220]	2015	*N. Engl. J. Med.*
060	Vansteenkiste et al.[221]	2015	*Ann. Oncol.*
061	Pellini et al.[222]	2023	*Clin. Cancer Res.*
062	Hu et al.[223]	2015	*Lancet Oncol.*
063	Xu et al.[224]	2022	*Nat. Commun.*
064	Carvajal et al.[225]	2022	*Nat. Med.*
065	Wang et al.[226]	2023	*Eur. J. Cancer*
066	Shi et al.[227]	2019	*Cancer Commun.*
067	DeMichele et al.[228]	2021	*Breast Cancer Res.*
068	Smith et al.[229]	2017	*J. Clin. Oncol.*
069	Chen et al.[230]	2019	*J. Clin. Oncol.*
070	Huang et al.[231]	2014	*Clin. Colorectal Cancer*
071	Lorch et al.[232]	2011	*J. Clin. Oncol.*
072	Zhang et al.[233]	2021	*Front. Oncol.*
073	Reiman et al.[234]	2012	*Ann. Oncol.*
074	Zhou et al.[235]	2021	*Front. Oncol.*
075	Xu et al.[236]	2021	*Nat. Med.*
076	Zhang et al.[237]	2023	*Lancet Oncol.*
077	Kim et al.[238]	2023	*Exp. Hematol. Oncol.*
078	Mirza et al.[239]	2023	*N. Engl. J. Med.*
079	Jassem et al.[240]	2009	*Oncol. Res. Treat.*
080	Modest et al.[241]	2013	*Acta Oncol.*
081	Trukhin et al.[242]	2021	*BioDrugs*
082	Shitara et al.[243]	2023	*J. Cancer Res. Clin. Oncol.*
083	Kogure et al.[244]	2021	*Lancet Healthy Longev.*
084	Balar et al.[245]	2023	*Ann. Oncol.*
085	Ramalingam et al.[246]	2018	*Clin. Lung Cancer*
086	Qin et al.[247]	2018	*J. Clin. Oncol.*
087	Wang et al.[248]	2021	*J. Clin. Oncol.*
088	Reck et al.[249]	2006	*Lung Cancer*
089	Tabernero et al.[250]	2021	*J. Clin. Oncol.*
090	Goto et al.[251]	2012	*J. Thorac. Oncol.*
091	Loriot et al.[252]	2023	*N. Engl. J. Med.*
092	Baselga et al.[253]	2012	*N. Engl. J. Med.*
093	Modest et al.[254]	2017	*Oncotarget*
094	Song et al.[255]	2021	*BMC Pulm. Med.*
095	Price et al.[256]	2016	*Eur. J. Cancer*
096	Monk et al.[257]	2016	*Gynecol. Oncol.*
097	Ciardiello et al.[258]	2022	*Clin. Colorectal Cancer*
098	Takahashi et al.[259]	2022	*Int. J. Clin. Oncol.*
099	Zhou et al.[260]	2023	*N. Engl. J. Med.*
100	García-Alfonso et al.[261]	2021	*Drugs Aging*
101	Stintzing et al.[262]	2012	*Ann. Oncol.*
102	Stintzing et al.[263]	2016	*Lancet Oncol.*
103	Lorenzen et al.[264]	2022	*Eur. J. Cancer*
104	Brodowicz et al.[265]	2013	*Ann. Oncol.*
105	Bokemeyer et al.[266]	2015	*Eur. J. Cancer*
106	Robertson et al.[267]	2016	*Lancet*
107	Robertson et al.[268]	2012	*Breast Cancer Res. Treat.*
108	Ellis et al.[269]	2015	*J. Clin. Oncol.*
109	Jones et al.[270]	2020	*Lancet Oncol.*
110	Howell et al.[271]	2022	*Lancet Oncol.*
111	Cristofanilli et al.[272]	2016	*Lancet Oncol.*
112	Howell et al.[273]	2011	*Clin. Breast Cancer*
113	Fiteni et al.[274]	2016	*Eur. Respir. J.*
114	Lorusso et al.[275]	2023	*Eur. J. Cancer*

Continued on next page

Table 14 – continued from previous page

Index	Reference	Year	Journal
115	Monk et al.[276]	2023	*Lancet Oncol.*
116	Perera et al.[277]	2022	*Clin. Cancer Res.*
117	You et al.[278]	2023	*Int. J. Radiat. Oncol. Biol. Phys.*
118	Stintzing et al.[279]	2017	*Eur. J. Cancer*
119	Swisher et al.[280]	2022	*Gynecol. Oncol.*
120	Modest et al.[281]	2015	*J. Clin. Oncol.*
121	Aghajanian et al.[282]	2022	*Gynecol. Oncol.*
122	Von et al.[283]	2013	*N. Engl. J. Med.*
123	Scher et al.[284]	2012	*N. Engl. J. Med.*
124	Geoffrois et al.[285]	2018	*J. Clin. Oncol.*
125	Luo et al.[286]	2022	*BMC Cancer*
126	Sastre et al.[287]	2021	*ESMO Open*
127	Adams et al.[288]	2011	*Lancet Oncol.*
128	Kim et al.[289]	2017	*Lancet Oncol.*
129	Matsubara et al.[290]	2023	*Cancer Med.*
130	Satoh et al.[291]	2014	*J. Clin. Oncol.*
131	Goetz et al.[292]	2023	*Ann. Oncol.*
132	Von et al.[293]	2014	*J. Cancer Res. Clin. Oncol.*
133	Makker et al.[294]	2022	*N. Engl. J. Med.*
134	Yonemori et al.[295]	2022	*Cancer Sci.*
135	Bradley et al.[296]	2020	*J. Clin. Oncol.*
136	Monk et al.[297]	2020	*J. Clin. Oncol.*
137	Sledge et al.[298]	2017	*J. Clin. Oncol.*
138	Zhang et al.[299]	2020	*Ther. Adv. Med. Oncol.*
139	Ye et al.[300]	2014	*Mol. Ther.*
140	Weiss et al.[301]	2019	*Ann. Oncol.*
141	Young et al.[302]	2018	*Asia Pac. J. Clin. Oncol.*
142	Dai et al.[303]	2022	*Ther. Adv. Med. Oncol.*
143	Tang et al.[304]	2023	*Ann. Oncol.*
144	Ferron et al.[305]	2023	*Gynecol. Oncol.*
145	Forde et al.[306]	2022	*N. Engl. J. Med.*
146	Hirsh et al.[307]	2015	*BioDrugs*
147	Van et al.[308]	2023	*N. Engl. J. Med.*
148	Haddad et al.[309]	2023	*J. Clin. Oncol.*
149	Saba et al.[310]	2019	*Oral Oncol.*
150	Kasper et al.[311]	2021	*Eur. J. Cancer*
151	Aghajanian et al.[312]	2012	*J. Clin. Oncol.*
152	Arend et al.[313]	2024	*J. Clin. Oncol.*
153	Bang et al.[314]	2017	*Lancet Oncol.*
154	Pfisterer et al.[315]	2023	*J. Clin. Oncol.*
155	Fan et al.[316]	2024	*Lancet Oncol.*
156	Wu et al.[317]	2020	*N. Engl. J. Med.*
157	Soria et al.[318]	2018	*N. Engl. J. Med.*
158	Mok et al.[319]	2017	*N. Engl. J. Med.*
159	Papadimitrakopoulou et al.[320]	2020	*Ann. Oncol.*
160	Tolaney et al.[321]	2024	*Clin. Cancer Res.*
161	Nathan et al.[322]	2021	*N. Engl. J. Med.*
162	Piulats et al.[323]	2024	*Ann. Oncol.*
163	Finn et al.[324]	2020	*Breast Cancer Res. Treat.*
164	Ascierto et al.[325]	2023	*Lancet Oncol.*
165	Ramalingam et al.[326]	2020	*N. Engl. J. Med.*
166	Lu et al.[327]	2022	*The Innovation*
167	Zhang et al.[328]	2022	*Cancer Commun.*
168	Kitagawa et al.[329]	2015	*J. Clin. Oncol.*
169	Icon et al.[330]	2003	*Lancet*
170	Martin et al.[331]	2021	*Ann. Oncol.*

Continued on next page

Table 14 – continued from previous page

Index	Reference	Year	Journal
171	Xu et al.[332]	2022	*Eur. J. Cancer*
172	Lorusso et al.[333]	2019	*Int. J. Gynecol. Cancer*
173	Yoh et al.[334]	2020	*Curr. Med. Res. Opin.*
174	Powles et al.[335]	2021	*Lancet Oncol.*
175	Eskander et al.[336]	2023	*N. Engl. J. Med.*
176	Borghaei et al.[337]	2020	*Cancer*
177	Diaz et al.[338]	2022	*Lancet Oncol.*
178	Cao et al.[339]	2022	*ESMO Open*
179	Mok et al.[340]	2019	*Lancet*
180	Cohen et al.[341]	2019	*Lancet*
181	Harrington et al.[342]	2023	*Oral Oncol.*
182	Fuchs et al.[343]	2022	*Gastric Cancer*
183	O'Brien et al.[344]	2022	*Lancet Oncol.*
184	Luke et al.[345]	2022	*Lancet*
185	Powles et al.[346]	2022	*Lancet Oncol.*
186	Wakelee et al.[347]	2023	*N. Engl. J. Med.*
187	Hadoux et al.[348]	2023	*N. Engl. J. Med.*
188	Carvajal et al.[349]	2022	*J. Clin. Oncol.*
189	Reckamp et al.[350]	2022	*J. Clin. Oncol.*
190	Slamon et al.[351]	2018	*J. Clin. Oncol.*
191	Ishigami et al.[352]	2018	*J. Clin. Oncol.*
192	Tveit et al.[353]	2012	*J. Clin. Oncol.*
193	Baselga et al.[354]	2014	*Ann. Oncol.*
194	Fribbens et al.[355]	2016	*J. Clin. Oncol.*
195	De et al.[356]	2010	*Lancet*
196	Lin et al.[357]	2023	*Eur. J. Cancer*
197	Miles et al.[358]	2021	*Ann. Oncol.*
198	Fang et al.[359]	2022	*Front. Cell Dev. Biol.*
199	Trillsch et al.[360]	2016	*Ann. Oncol.*
200	Tejpar et al.[361]	2017	*JAMA Oncol.*
201	Sinicrope et al.[362]	2013	*J. Clin. Oncol.*
202	Moretto et al.[363]	2020	*Eur. J. Cancer*
203	Aokage et al.[364]	2021	*Lung Cancer*
204	Zinner et al.[365]	2015	*J. Thorac. Oncol.*
205	Láng et al.[366]	2013	*Eur. J. Cancer*
206	Al-Batran et al.[367]	2016	*Ann. Oncol.*
207	Kudo et al.[368]	2020	*J. Gastroenterol.*
208	De et al.[369]	2019	*Future Oncol.*
209	Yoon et al.[370]	2016	*Ann. Oncol.*
210	Nishio et al.[371]	2020	*Cancer Sci.*
211	Petrylak et al.[372]	2017	*Lancet*
212	Nakagawa et al.[373]	2019	*Lancet Oncol.*
213	Garon et al.[374]	2023	*ESMO Open*
214	Wilke et al.[375]	2014	*Lancet Oncol.*
215	Zhu et al.[376]	2015	*Lancet Oncol.*
216	Tabernero et al.[377]	2015	*Lancet Oncol.*
217	Sakai et al.[378]	2020	*Eur. J. Cancer*
218	Patel et al.[379]	2021	*Radiology*
219	Smith et al.[380]	2014	*J. Clin. Oncol.*
220	Albanell et al.[381]	2022	*Eur. J. Cancer*
221	Curigliano et al.[382]	2013	*Breast*
222	Baselga et al.[383]	2013	*J. Clin. Oncol.*
223	Kalinsky et al.[384]	2023	*J. Clin. Oncol.*
224	Herzog et al.[385]	2023	*Gynecol. Oncol.*
225	García-Carbonero et al.[386]	2017	*Oncologist*
226	Nakajima et al.[387]	2020	*Gastric Cancer*

Continued on next page

Table 14 – continued from previous page

Index	Reference	Year	Journal
227	Kojima et al.[388]	2020	*J. Clin. Oncol.*
228	Ang et al.[389]	2014	*J. Clin. Oncol.*
229	Langer et al.[390]	2014	*J. Clin. Oncol.*
230	Sclafani et al.[391]	2014	*Eur. J. Cancer*
231	Aix et al.[392]	2021	*Cancer Treat. Res. Commun.*
232	Holch et al.[393]	2019	*Eur. J. Cancer*
233	Segelov et al.[394]	2016	*J. Clin. Oncol.*
234	Di et al.[395]	2010	*J. Clin. Oncol.*
235	Okunaka et al.[396]	2020	*BMC Cancer*
236	Slamon et al.[397]	2021	*Ann. Oncol.*
237	O'shaughnessy et al.[398]	2018	*Breast Cancer Res. Treat.*
238	Wang et al.[399]	2023	*Front. Immunol.*
239	Wang et al.[400]	2019	*Ann. Oncol.*
240	Shi et al.[401]	2021	*Lancet Oncol.*
241	Huemer et al.[402]	2018	*BMC Cancer*
242	Huber et al.[403]	2006	*J. Clin. Oncol.*
243	Stahler et al.[404]	2020	*Eur. J. Cancer*
244	Lu et al.[405]	2022	*BMJ*
245	Shitara et al.[406]	2016	*Gastric Cancer*
246	Obermannová et al.[407]	2016	*Ann. Oncol.*
247	Chung et al.[408]	2018	*Asia Pac. J. Clin. Oncol.*
248	Fernandez-Martinez et al.[409]	2020	*J. Clin. Oncol.*
249	Bidard et al.[410]	2022	*Lancet Oncol.*
250	Pfeiffer et al.[411]	2020	*Lancet Oncol.*
251	Takashima et al.[412]	2016	*Lancet Oncol.*
252	Tesileanu et al.[413]	2022	*Clin. Cancer Res.*
253	Möbus et al.[414]	2018	*Ann. Oncol.*
254	Quoix et al.[415]	2016	*Lancet Oncol.*
255	Finn et al.[416]	2015	*Lancet Oncol.*
256	Trullas et al.[417]	2021	*ESMO Open*
257	Alarfi et al.[418]	2017	*BMC Res. Notes*
258	Weidhaas et al.[419]	2023	*Cancer Res. Commun.*
259	Okada et al.[420]	2022	*Clin. Cancer Res.*
260	Hassel et al.[421]	2023	*N. Engl. J. Med.*
261	Shen et al.[422]	2022	*J. Clin. Oncol.*
262	Lian et al.[423]	2022	*Ann. Oncol.*
263	Xu et al.[424]	2021	*Technol. Cancer Res. Treat.*
264	Mai et al.[425]	2021	*Nat. Med.*
265	Yan et al.[426]	2024	*Ann. Oncol.*
266	Jiang et al.[427]	2024	*Nat. Med.*
267	Gershenson et al.[428]	2022	*Lancet*
268	Paz-Ares et al.[429]	2017	*Lung Cancer*
269	Goodwin et al.[430]	2010	*Ann. Oncol.*
270	Shitara et al.[431]	2018	*Lancet Oncol.*
271	Cascinu et al.[432]	2021	*Oncologist*
272	Leal et al.[433]	2023	*Lancet Oncol.*
273	Neven et al.[434]	2023	*Breast Cancer Res.*
274	Hortobagyi et al.[435]	2018	*Ann. Oncol.*
275	Diéras et al.[436]	2020	*Lancet Oncol.*
276	Mizuno et al.[437]	2023	*Int. J. Clin. Oncol.*
277	Lindeman et al.[438]	2022	*Clin. Cancer Res.*
278	Cortot et al.[439]	2020	*Eur. J. Cancer*
279	Ngoi et al.[440]	2022	*Cochrane Database Syst. Rev.*
280	Spigel et al.[441]	2017	*Clin. Lung Cancer*

Supplementary References

[S1] Tomczak, K., Czerwińska, P. & Wiznerowicz, M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp. Oncol. (Poznan, Poland)* **19**, A68–A77 (2015).

[S2] Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* **41**, D991–D995 (2013).

[S3] Tierney, J. F., Stewart, L. A., Ghersi, D., Burdett, S. & Sydes, M. R. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* **8**, 16 (2007).

[S4] Wang, F. *et al.* A systematic review and meta-analysis of 90 cohort studies of social isolation, loneliness and mortality. *Nat. Hum. Behav.* **7**, 1307–1319 (2023).

[S5] Dwan, K., Gamble, C., Williamson, P. R., Kirkham, J. J. & Group, R. B. Systematic review of the empirical evidence of study publication bias and outcome reporting bias—an updated review. *PLOS ONE* **8**, e66844 (2013).

[S6] Higgins, J. P. *et al.* *Cochrane Handbook for Systematic Reviews of Interventions* (Wiley-Blackwell, Hoboken, NJ, 2019), 2nd edn.

[S7] DerSimonian, R. & Laird, N. Meta-analysis in clinical trials. *Control. Clin. Trials* **7**, 177–188 (1986).

[S8] Howlader, N. *et al.* (eds.) *SEER Cancer Statistics Review, 1975-2017* (National Cancer Institute, Bethesda, MD, 2020). Based on November 2019 SEER data submission, posted April 2020.

[S9] Schoenfeld, D. A. Sample-size formula for the proportional-hazards regression model. *Biometrics* 499–503 (1983).

[S10] Guyot, P., Ades, A. E., Ouwens, M. J. N. M. & Welton, N. J. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan–Meier survival curves. *BMC Med. Res. Methodol.* **12**, 9 (2012).

[S11] Liu, N., Zhou, Y. & Lee, J. J. IPDfromKM: reconstruct individual patient data from published Kaplan-Meier survival curves. *BMC Med. Res. Methodol.* **21**, 111 (2021).

[S12] Rogula, B., Lozano-Ortega, G. & Johnston, K. M. A method for reconstructing individual patient data from Kaplan-Meier survival curves that incorporate marked censoring times. *MDM Policy & Pract.* **7**, 23814683221077643 (2022).

[S13] Whitley, D. The GENITOR algorithm and selection pressure: why rank-based allocation of reproductive trials is best. In *Proceedings of the Third International Conference on Genetic Algorithms*, 116–121 (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1989).

[S14] Syswerda, G. Uniform crossover in genetic algorithms. In *Proceedings of the 3rd International Conference on Genetic Algorithms*, 2–9 (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1989).

[S15] Morris, T. P., White, I. R. & Crowther, M. J. Using simulation studies to evaluate statistical methods. *Stat. Medicine* **38**, 2074–2102 (2019).

[S16] Hall, P. *et al.* Hormone-replacement therapy influences gene expression profiles and is associated with breast-cancer prognosis: a cohort study. *BMC Medicine* **4**, 16 (2006).

[S17] Hu, Z. *et al.* The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* **7**, 96 (2006).

[S18] Perreard, L. *et al.* Classification and risk stratification of invasive breast carcinomas using a real-time quantitative RT-PCR assay. *Breast Cancer Res.* **8**, R23 (2006).

[S19] Sotiriou, C. *et al.* Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J. Natl. Cancer Inst.* **98**, 262–272 (2006).

[S20] Bild, A. H. *et al.* Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* **439**, 353–357 (2006).

[S21] Hummel, M. *et al.* A biologic definition of burkitt's lymphoma from transcriptional and genomic profiling. *New Engl. J. Medicine* **354**, 2419–2430 (2006).

[S22] Ivshina, A. V. *et al.* Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. *Cancer Res.* **66**, 10292–10301 (2006).

[S23] Minn, A. J. *et al.* Lung metastasis genes couple breast tumor size and metastatic spread. *Proc. Natl. Acad. Sci.* **104**, 6740–6745 (2007).

[S24] Hoadley, K. A. *et al.* EGFR associated expression profiles vary with breast tumor subtype. *BMC Genomics* **8**, 258 (2007).

[S25] Mullins, M. *et al.* Agreement in breast cancer classification between microarray and quantitative reverse transcription PCR from fresh-frozen and formalin-fixed, paraffin-embedded tissues. *Clin. Chem.* **53**, 1273–1279 (2007).

[S26] Zhou, Y. *et al.* Enhanced NFκB and AP-1 transcriptional activity associated with antiestrogen resistant breast cancer. *BMC Cancer* **7**, 59 (2007).

[S27] Marchini, S. *et al.* Analysis of gene expression in early-stage ovarian cancer. *Clin. Cancer Res.* **14**, 7850–7860 (2008).

[S28] Loi, S. *et al.* Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. *BMC Genomics* **9**, 239 (2008).

[S29] Cohen, E. E. *et al.* A feed-forward loop involving protein kinase Cα and microRNAs regulates tumor cell cycle. *Cancer Res.* **69**, 65–74 (2009).

[S30] Herschkowitz, J. I., He, X., Fan, C. & Perou, C. M. The functional loss of the retinoblastoma tumour suppressor is a common event in basal-like and luminal B breast carcinomas. *Breast Cancer Res.* **10**, R75 (2008).

[S31] Parker, J. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).

[S32] Baty, F. *et al.* Gene profiling of clinical routine biopsies and prediction of survival in non–small cell lung cancer. *Am. J. Respir. Critical Care Medicine* **181**, 181–188 (2010).

[S33] Schmidt, M. *et al.* The humoral immune system has a key prognostic impact in node-negative breast cancer. *Cancer Res.* **68**, 5405–5413 (2008).

[S34] Takeuchi, T. *et al.* Expression profile–defined classification of lung adenocarcinoma shows close relationship with underlying major genetic changes and clinicopathologic behaviors. *J. Clin. Oncol.* **24**, 1679–1688 (2006).

[S35] Zhang, Y. *et al.* The 76-gene signature defines high-risk patients that benefit from adjuvant tamoxifen therapy. *Breast Cancer Res. Treat.* **116**, 303–309 (2009).

[S36] Staub, E. *et al.* An expression module of WIPF1-coexpressed genes identifies patients with favorable prognosis in three tumor types. *J. Mol. Medicine* **87**, 633–644 (2009).

[S37] Lee, Y. *et al.* Gene expression analysis of glioblastomas identifies the major molecular basis for the prognostic benefit of younger age. *BMC Med. Genomics* **1**, 52 (2008).

[S38] Kim, W.-J. *et al.* Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. *Mol. Cancer* **9**, 3 (2010).

[S39] Denkert, C. *et al.* A prognostic gene expression index in ovarian cancer—validation across different independent data sets. *The J. Pathol. A J. Pathol. Soc. Gt. Br. Irel.* **218**, 273–280 (2009).

[S40] Zhu, C.-Q. *et al.* Prognostic and predictive gene signature for adjuvant chemotherapy in resected non–small-cell lung cancer. *J. Clin. Oncol.* **28**, 4417–4424 (2010).

[S41] Reid, J. F. *et al.* Integrative approach for prioritizing cancer genes in sporadic colon cancer. *Genes, Chromosom. Cancer* **48**, 953–962 (2009).

[S42] Sboner, A. *et al.* Molecular sampling of prostate cancer: a dilemma for predicting disease progression. *BMC Med. Genomics* **3**, 8 (2010).

[S43] Yoshihara, K. *et al.* Gene expression profile for predicting survival in advanced-stage serous ovarian cancer across two independent datasets. *PLOS ONE* **5**, e9615 (2010).

[S44] Smith, J. J. *et al.* Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology* **138**, 958–968 (2010).

[S45] Wilkerson, M. D. *et al.* Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically important, and correspond to normal cell types. *Clin. Cancer Res.* **16**, 4864–4875 (2010).

[S46] Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* **12**, R68 (2010).

[S47] Barwick, B. *et al.* Prostate cancer genes associated with TMPRSS2–ERG gene fusion and prognostic of biochemical recurrence in multiple cohorts. *Br. J. Cancer* **102**, 570–576 (2010).

[S48] Konstantinopoulos, P. A. *et al.* Integrated analysis of multiple microarray datasets identifies a reproducible survival predictor in ovarian cancer. *PLOS ONE* **6**, e18202 (2011).

[S49] Kim, Y.-J. *et al.* Gene signatures for the prediction of response to Bacillus Calmette-Guerin immunotherapy in primary pT1 bladder cancers. *Clin. Cancer Res.* **16**, 2131–2137 (2010).

[S50] Enerly, E. *et al.* miRNA-mRNA integrated analysis reveals roles for miRNAs in primary breast tumors. *PLOS ONE* **6**, e16915 (2011).

[S51] Konstantinopoulos, P. A. *et al.* Gene expression profile of BRCA ness that correlates with responsiveness to chemotherapy and with outcome in patients with epithelial ovarian cancer. *J. Clin. Oncol.* **28**, 3555–3561 (2010).

[S52] Kao, K.-J., Chang, K.-M., Hsu, H.-C. & Huang, A. T. Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: implications for treatment optimization. *BMC Cancer* **11**, 143 (2011).

[S53] Dedeurwaerder, S. *et al.* DNA methylation profiling reveals a predominant immune component in breast cancers. *EMBO Mol. Medicine* **3**, 726–741 (2011).

[S54] Sabatier, R. *et al.* A gene expression signature identifies two prognostic subgroups of basal breast cancer. *Breast Cancer Res. Treat.* **126**, 407–420 (2011).

[S55] Jönsson, G. *et al.* Genomic subtypes of breast cancer identified by array-comparative genomic hybridization display distinct molecular and clinical characteristics. *Breast Cancer Res.* **12**, R42 (2010).

[S56] Esserman, L. J. *et al.* Chemotherapy response and recurrence-free survival in neoadjuvant breast cancer depends on biomarker profiles: results from the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657). *Breast Cancer Res. Treat.* **132**, 1049–1062 (2012).

[S57] Shaknovich, R. *et al.* DNA methylation signatures define molecular subtypes of diffuse large B-cell lymphoma. *Blood, The J. Am. Soc. Hematol.* **116**, e81–e89 (2010).

[S58] Sveen, A. *et al.* Transcriptome instability in colorectal cancer identified by exon microarray analyses: Associations with splicing factor expression levels and patient survival. *Genome Medicine* **3**, 32 (2011).

[S59] Végran, F. *et al.* Apoptosis gene signature of survivin and its splice variant expression in breast carcinoma. *Endocrine-Related Cancer* **18**, 783–792 (2011).

[S60] Hatzis, C. *et al.* A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. *JAMA* **305**, 1873–1881 (2011).

[S61] Jönsson, G. *et al.* The retinoblastoma gene undergoes rearrangements in BRCA1-deficient basal-like breast cancer. *Cancer Res.* **72**, 4028–4036 (2012).

[S62] Fountzilas, E. *et al.* Identification and validation of gene expression models that predict clinical outcome in patients with early-stage laryngeal cancer. *Annals Oncol.* **23**, 2146–2153 (2012).

[S63] Gentric, G. *et al.* PML-regulated mitochondrial metabolism enhances chemosensitivity in human ovarian cancers. *Cell Metab.* **29**, 156–173 (2019).

[S64] Lee, J. *et al.* Nanostring-based multigene assay to predict recurrence for gastric cancer patients after surgery. *PLOS ONE* **9**, e90133 (2014).

[S65] Muggerud, A. A. *et al.* Molecular diversity in ductal carcinoma in situ (DCIS) and early invasive breast cancer. *Mol. Oncol.* **4**, 357–368 (2010).

[S66] Harrell, J. C. *et al.* Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer Res. Treat.* **132**, 523–535 (2012).

[S67] Bonome, T. *et al.* A gene signature predicting for survival in suboptimally debulked patients with ovarian cancer. *Cancer Res.* **68**, 5478–5486 (2008).

[S68] Wilkerson, M. D. *et al.* Differential pathogenesis of lung adenocarcinoma subtypes involving sequence mutations, copy number, chromosomal instability, and methylation. *PLOS ONE* **7**, e36530 (2012).

[S69] Fountzilas, E. *et al.* Identification and validation of a multigene predictor of recurrence in primary laryngeal cancer. *PLOS ONE* **8**, e70429 (2013).

[S70] Zhou, Q. *et al.* Transcriptional profile of human liver tissues: hepatocellular carcinoma vs. matched noncancerous liver tissue. NCBI Gene Expression Omnibus (2011). GSE27150.

[S71] Loboda, A. *et al.* EMT is the dominant program in human colon cancer. *BMC Med. Genomics* **4**, 9 (2011).

[S72] Zhang, G. *et al.* DPEP1 inhibits tumor cell invasiveness, enhances chemosensitivity and predicts clinical outcome in pancreatic ductal adenocarcinoma. *PLOS ONE* **7**, e31507 (2012).

[S73] Xie, Y. *et al.* Robust gene expression signature from formalin-fixed paraffin-embedded samples predicts prognosis of non–small-cell lung cancer patients. *Clin. Cancer Res.* **17**, 5705–5714 (2011).

[S74] Edeline, J. *et al.* Description of 2 angiogenic phenotypes in clear cell renal cell carcinoma. *Hum. Pathol.* **43**, 1982–1990 (2012).

[S75] Chen, D.-T. *et al.* Complementary strand microRNAs mediate acquisition of metastatic potential in colonic adenocarcinoma. *J. Gastrointest. Surg.* **16**, 905–913 (2012).

[S76] Ferriss, J. S. *et al.* Multi-gene expression predictors of single drug responses to adjuvant chemotherapy in ovarian carcinoma: predicting platinum resistance. *PLOS ONE* **7**, e30550 (2012).

[S77] Rousseaux, S. *et al.* Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Sci. Transl. Medicine* **5**, 186ra66–186ra66 (2013).

[S78] Ågesen, T. H. *et al.* ColoGuideEx: a robust gene classifier specific for stage II colorectal cancer prognosis. *Gut* **61**, 1560–1567 (2012).

[S79] Gobble, R. M. *et al.* Expression profiling of liposarcoma yields a multigene predictor of patient outcome and identifies genes that contribute to liposarcomagenesis. *Cancer Res.* **71**, 2697–2705 (2011).

[S80] Spentzos, D. *et al.* Unique gene expression profile based on pathologic response in epithelial ovarian cancer. *J. Clin. Oncol.* **23**, 7911–7918 (2005).

[S81] Blumenschein Jr, G. R. *et al.* Comprehensive biomarker analysis and final efficacy results of sorafenib in the BATTLE trial. *Clin. Cancer Res.* **19**, 6967–6975 (2013).

[S82] Sabatier, R. *et al.* Down-regulation of ECRG4, a candidate tumor suppressor gene, in human breast cancer. *PLOS ONE* **6**, e27656 (2011).

[S83] Botling, J. *et al.* Biomarker discovery in non–small cell lung cancer: integrating gene expression profiling, meta-analysis, and tissue microarray validation. *Clin. Cancer Res.* **19**, 194–204 (2013).

[S84] Riester, M. *et al.* Combination of a novel gene expression signature with a clinical nomogram improves the prediction of survival in high-risk bladder cancer. *Clin. Cancer Res.* **18**, 1323–1333 (2012).

[S85] Yoshihara, K. *et al.* High-risk ovarian cancer based on 126-gene expression signature is uniquely characterized by downregulation of antigen presentation pathway. *Clin. Cancer Res.* **18**, 1374–1385 (2012).

[S86] Sjödahl, G. *et al.* A molecular taxonomy for urothelial carcinoma. *Clin. Cancer Res.* **18**, 3377–3386 (2012).

[S87] Barrans, S. L. *et al.* Whole genome expression profiling based on paraffin embedded tissue can be used to classify diffuse large b-cell lymphoma and predict clinical outcome. *Br. J. Haematol.* **159**, 441–453 (2012).

[S88] Callari, M. *et al.* Subtype-dependent prognostic relevance of an interferon-induced pathway metagene in node-negative breast cancer. *Mol. Oncol.* **8**, 1278–1289 (2014).

[S89] Tripathi, M. K. *et al.* Nuclear factor of activated T-cell activity is associated with metastatic capacity in colon cancer. *Cancer Res.* **74**, 6947–6957 (2014).

[S90] Marisa, L. *et al.* Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Medicine* **10**, e1001453 (2013).

[S91] Jansen, M. P. *et al.* Hallmarks of aromatase inhibitor drug resistance revealed by epigenetic profiling in breast cancer. *Cancer Res.* **73**, 6632–6641 (2013).

[S92] Tang, H. *et al.* A 12-gene set predicts survival benefits from adjuvant chemotherapy in non–small cell lung cancer patients. *Clin. Cancer Res.* **19**, 1577–1586 (2013).

[S93] Clarke, C. *et al.* Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. *Carcinogenesis* **34**, 2300–2308 (2013).

[S94] Medico, E., Isella, C., Renzulli, T. & De Bortoli, M. IRCC breast cancer dataset. NCBI Gene Expression Omnibus (2019). GSE43615.

[S95] Lee, Y.-Y. *et al.* Genetic profiling to predict recurrence of early cervical cancer. *Gynecol. Oncol.* **131**, 650–654 (2013).

[S96] Wang, D.-Y., Done, S. J., Mc Cready, D. R. & Leong, W. L. Validation of the prognostic gene portfolio, ClinicoMolecular Triad Classification, using an independent prospective breast cancer cohort and external patient populations. *Breast Cancer Res.* **16**, R71 (2014).

[S97] Jonsdottir, K. *et al.* Prognostic value of gene signatures and proliferation in lymph-node-negative breast cancer. *PLOS ONE* **9**, e90642 (2014).

[S98] Mortensen, M. M. *et al.* Expression profiling of prostate cancer tissue delineates genes associated with recurrence after prostatectomy. *Sci. Reports* **5**, 16018 (2015).

[S99] Yu, G. *et al.* Matrix metalloproteinase-19 promotes metastatic behavior in vitro and is associated with increased mortality in non–small cell lung cancer. *Am. J. Respir. Critical Care Medicine* **190**, 780–790 (2014).

[S100] Choi, W. *et al.* Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. *Cancer Cell* **25**, 152–165 (2014).

[S101] Huang, C.-C. *et al.* Concurrent gene signatures for han chinese breast cancers. *PLOS ONE* **8**, e76421 (2013).

[S102] Pils, D. *et al.* Validating the impact of a molecular subtype in ovarian cancer on outcomes: a study of the OVCAD consortium. *Cancer Sci.* **103**, 1334–1341 (2012).

[S103] Der, S. D. *et al.* Validation of a histology-independent prognostic gene signature for early-stage, non–small-cell lung cancer including stage IA patients. *J. Thorac. Oncol.* **9**, 59–64 (2014).

[S104] Karlan, B. Y. *et al.* POSTN/TGFBI-associated stromal signature predicts poor prognosis in serous epithelial ovarian cancer. *Gynecol. Oncol.* **132**, 334–342 (2014).

[S105] Azim, J., Hatem A. *et al.* Biology of breast cancer during pregnancy using genomic profiling. *Endocrine-Related Cancer* **21**, 545–554 (2014).

[S106] Li, J. *et al.* LncRNA profile study reveals a three-lncRNA signature associated with the survival of patients with oesophageal squamous cell carcinoma. *Gut* **63**, 1700–1710 (2014).

[S107] Konecny, G. E. *et al.* Prognostic and therapeutic relevance of molecular subtypes in high-grade serous ovarian cancer. *J. Natl. Cancer Inst.* **106**, dju249 (2014).

[S108] Chen, D.-T. *et al.* Prognostic fifteen-gene signature for early stage pancreatic ductal adenocarcinoma. *PLOS ONE* **10**, e0133562 (2015).

[S109] Jézéquel, P. *et al.* Gene-expression molecular subtyping of triple-negative breast cancer tumours: importance of immune response. *Breast Cancer Res.* **17**, 43 (2015).

[S110] Fumagalli, D. *et al.* Gene profiling of HER2 breast cancer patients treated with adjuvant trastuzumab. NCBI Gene Expression Omnibus (2014). GSE58984.

[S111] Grinchuk, O. V. *et al.* Sense-antisense gene-pairs in breast cancer and associated pathological pathways. *Oncotarget* **6**, 42197 (2015).

[S112] Baty, F. *et al.* 24h-gene variation effect of combined bevacizumab/erlotinib in advanced non-squamous non-small cell lung cancer using exon array blood profiling. *J. Transl. Medicine* **15**, 66 (2017).

[S113] Yang, S. *et al.* A novel MIF signaling pathway drives the malignant character of pancreatic cancer by targeting NR3C2. *Cancer Res.* **76**, 3838–3850 (2016).

[S114] Lisowska, K. M. *et al.* Gene expression analysis in ovarian cancer–faults and hints from DNA microarray study. *Front. Oncol.* **4**, 6 (2014).

[S115] Cirenajwis, H. *et al.* Molecular stratification of metastatic melanoma using gene expression profiling: Prediction of survival outcome and benefit from molecular targeted therapy. *Oncotarget* **6**, 12297 (2015).

[S116] Pastore, A. *et al.* Integration of gene mutations in risk prognostication for patients receiving first-line immunochemotherapy for follicular lymphoma: a retrospective analysis of a prospective clinical trial and validation in a population-based registry. *The Lancet Oncol.* **16**, 1111–1122 (2015).

[S117] Chin, K. *et al.* Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell* **10**, 529–541 (2006).

[S118] Sha, C. *et al.* Molecular high-grade B-cell lymphoma: defining a poor-risk group that requires different approaches to therapy. *J. Clin. Oncol.* **37**, 202–212 (2019).

[S119] Sha, C. *et al.* Transferring genomics to the clinic: distinguishing Burkitt and diffuse large B cell lymphomas. *Genome Medicine* **7**, 64 (2015).

[S120] McConkey, D. J. *et al.* A prognostic gene expression signature in the molecular classification of chemotherapy-naive urothelial cancer is predictive of clinical outcomes from neoadjuvant chemotherapy: a phase 2 trial of dose-dense methotrexate, vinblastine, doxorubicin, and cisplatin with bevacizumab in urothelial cancer. *Eur. Urol.* **69**, 855–862 (2016).

[S121] Del Rio, M. *et al.* Molecular subtypes of metastatic colorectal cancer are associated with patient response to irinotecan-based therapies. *Eur. J. Cancer* **76**, 68–75 (2017).

[S122] Feng, L. *et al.* A network-based method for identifying prognostic gene modules in lung squamous carcinoma. *Oncotarget* **7**, 18006 (2016).

[S123] Levine, E. A. *et al.* Prognostic molecular subtypes of low-grade cancer of the appendix. *J. Am. Coll. Surg.* **222**, 493–503 (2016).

[S124] Grinchuk, O. V. *et al.* Tumor-adjacent tissue co-expression profile analysis reveals pro-oncogenic ribosomal gene signature for prognosis of resectable hepatocellular carcinoma. *Mol. Oncol.* **12**, 89–113 (2018).

[S125] Yoon, S.-J. *et al.* Deconvolution of diffuse gastric cancer and the suppression of CD34 on the BALB/c nude mice model. *BMC Cancer* **20**, 314 (2020).

[S126] Rajbhandari, P. *et al.* Cross-cohort analysis identifies a TEAD4–MYCN positive feedback loop as the core regulatory element of high-risk neuroblastoma. *Cancer Discov.* **8**, 582–599 (2018).

[S127] Puleo, F. *et al.* Stratification of pancreatic ductal adenocarcinomas based on tumor and microenvironment features. *Gastroenterology* **155**, 1999–2013 (2018).

[S128] Prabhakaran, S. *et al.* Evaluation of invasive breast cancer samples using a 12-chemokine gene expression score: correlation with clinical outcomes. *Breast Cancer Res.* **19**, 71 (2017).

[S129] Hu, Y. *et al.* Colorectal cancer susceptibility loci as predictive markers of rectal cancer prognosis after surgery. *Genes, Chromosom. Cancer* **57**, 140–149 (2018).

[S130] Dubois, S. *et al.* Biological and clinical relevance of associated genomic alterations in MYD88 L265P and non-L265P–mutated diffuse large B-cell lymphoma: analysis of 361 cases. *Clin. Cancer Res.* **23**, 2232–2244 (2017).

[S131] Metzger-Filho, O. *et al.* Genomic grade adds prognostic value in invasive lobular carcinoma. *Annals Oncol.* **24**, 377–384 (2013).

[S132] Gotoh, K. *et al.* Prognostic model of stage II/III colon cancer constructed using gene expression subtypes and KRAS mutation status. *J. Clin. & Exp. Oncol.* **7**, 2 (2018).

[S133] Mitchell, K. A., Zingone, A., Toulabi, L., Boeckelman, J. & Ryan, B. M. Comparative transcriptome profiling reveals coding and noncoding RNA differences in NSCLC from African Americans and European Americans. *Clin. Cancer Res.* **23**, 7412–7425 (2017).

[S134] Ducie, J. *et al.* Molecular analysis of high-grade serous ovarian carcinoma with and without associated serous tubal intra-epithelial carcinoma. *Nat. Commun.* **8**, 990 (2017).

[S135] Allen, W. L. *et al.* Transcriptional subtyping and CD8 immunohistochemistry identifies poor prognosis stage II/III colorectal cancer patients who benefit from adjuvant chemotherapy. *JCO Precis. Oncol.* **2**, PO.17.00241 (2018).

[S136] Sjöström, M. *et al.* Identification and validation of single-sample breast cancer radiosensitivity gene expression predictors. *Breast Cancer Res.* **20**, 64 (2018).

[S137] Zhu, J. *et al.* Evaluation of frozen tissue-derived prognostic gene expression signatures in FFPE colorectal cancer samples. *Sci. Reports* **6**, 33273 (2016).

[S138] Ignatiadis, M. *et al.* Tumor-infiltrating lymphocytes in patients receiving trastuzumab/pertuzumab-based chemotherapy: a TRYPHAENA substudy. *JNCI: J. Natl. Cancer Inst.* **111**, 69–77 (2019).

[S139] Yang, B. *et al.* An assessment of prognostic immunity markers in breast cancer. *NPJ Breast Cancer* **4**, 35 (2018).

[S140] Kensler, K. H. *et al.* PAM50 molecular intrinsic subtypes in the nurses' health study cohorts. *Cancer Epidemiol. Biomarkers & Prev.* **28**, 798–806 (2019).

[S141] Jain, S. *et al.* Validation of a Metastatic Assay using biopsies to improve risk stratification in patients with prostate cancer treated with radical radiation therapy. *Annals Oncol.* **29**, 215–222 (2018).

[S142] Mock, A. *et al.* EGFR and PI3K pathway activities might guide drug repurposing in HPV-negative head and neck cancers. *Front. Oncol.* **11**, 678966 (2021).

[S143] Sinn, B. V. *et al.* SETER/PR: a robust 18-gene predictor for sensitivity to endocrine therapy for metastatic breast cancer. *NPJ Breast Cancer* **5**, 16 (2019).

[S144] Millstein, J. *et al.* Prognostic gene expression signature for high-grade serous ovarian cancer. *Annals Oncol.* **31**, 1240–1250 (2020).

[S145] Ferrandon, S. *et al.* CoA synthase (COASY) mediates radiation resistance via PI3K signaling in rectal cancer. *Cancer Res.* **80**, 334–346 (2020).

[S146] Kim, S.-K. *et al.* Genomic signature of the standardized uptake value in 18F-fluorodeoxyglucose positron emission tomography in breast cancer. *Cancers* **12**, 497 (2020).

[S147] Talhouk, A. *et al.* Development and validation of the gene expression predictor of high-grade serous ovarian carcinoma molecular SubTYPE (PrOTYPE). *Clin. Cancer Res.* **26**, 5411–5423 (2020).

[S148] Kommoss, S. *et al.* Bevacizumab may differentially improve ovarian cancer outcome in patients with proliferative and mesenchymal molecular subtypes. *Clin. Cancer Res.* **23**, 3794–3801 (2017).

[S149] Shinto, E. *et al.* Clinical significance of a gene signature generated from tumor budding grade in colon cancer. *Annals Surg. Oncol.* **27**, 4044–4054 (2020).

[S150] Chen, Y.-J. *et al.* Molecular subtyping of breast cancer intrinsic taxonomy with oligonucleotide microarray and NanoString nCounter. *Biosci. Reports* **41**, BSR20211428 (2021).

[S151] Hu, Z. *et al.* The oxford classic links epithelial-to-mesenchymal transition to immunosuppression in poor prognosis ovarian cancers. *Clin. Cancer Res.* **27**, 1570–1579 (2021).

[S152] Buechler, S. A. *et al.* ColoType: a forty gene signature for consensus molecular subtyping of colorectal cancer tumors using whole-genome assay or targeted RNA-sequencing. *Sci. Reports* **10**, 12123 (2020).

[S153] Bueno, R. *et al.* Multi-institutional prospective validation of prognostic mRNA signatures in early stage squamous lung cancer (alliance). *J. Thorac. Oncol.* **15**, 1748–1757 (2020).

[S154] Eide, P. W. *et al.* Metastatic heterogeneity of the consensus molecular subtypes of colorectal cancer. *NPJ Genomic Medicine* **6**, 59 (2021).

[S155] Regua, A. *et al.* ABI1-based expression signature predicts breast cancer metastasis and survival. *Mol. Oncol.* **16**, 2632–2657 (2022).

[S156] Szeglin, B. C. *et al.* A SMAD4-modulated gene profile predicts disease-free survival in stage II and III colorectal cancer. *Cancer Reports* **5**, e1423 (2022).

[S157] Kang, H. *et al.* USP19 and RPL23 as candidate prognostic markers for advanced-stage high-grade serous ovarian carcinoma. *Cancers* **13**, 3976 (2021).

[S158] Sun, G. *et al.* Integrated exome and RNA sequencing of TFE3-translocation renal cell carcinoma. *Nat. Commun.* **12**, 5262 (2021).

[S159] Brasó-Maristany, F. *et al.* Gene expression profiles of breast cancer metastasis according to organ site. *Mol. Oncol.* **16**, 69–87 (2022).

[S160] Sugawara, T. *et al.* Immune subtypes and neoantigen-related immune evasion in advanced colorectal cancer. *iScience* **25** (2022).

[S161] Jones, W., Tait, D., Livasy, C., Ganapathi, M. & Ganapathi, R. PLK3 amplification and tumor immune microenvironment of metastatic tumors are linked to adjuvant treatment outcomes in uterine serous cancer. *NAR Cancer* **4**, zcac026 (2022).

[S162] Perren, T. J. *et al.* A phase 3 trial of bevacizumab in ovarian cancer. *New Engl. J. Medicine* **365**, 2484–2496 (2011).

[S163] Park, Y. H. *et al.* Phase II, multicentre, randomised trial of eribulin plus gemcitabine versus paclitaxel plus gemcitabine as first-line chemotherapy in patients with HER2-negative metastatic breast cancer. *Eur. J. Cancer* **86**, 385–393 (2017).

[S164] Fizazi, K. *et al.* Abiraterone plus prednisone in metastatic, castration-sensitive prostate cancer. *New Engl. J. Medicine* **377**, 352–360 (2017).

[S165] James, N. D. *et al.* Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. *The Lancet* **387**, 1163–1177 (2016).

[S166] Oza, A. M. *et al.* A biomarker-enriched, randomized phase II trial of adavosertib (AZD1775) plus paclitaxel and carboplatin for women with platinum-sensitive TP53-mutant ovarian cancer. *Clin. Cancer Res.* **26**, 4767–4776 (2020).

[S167] Zhou, C. *et al.* A phase 1/2 multicenter randomized trial of local ablation plus toripalimab versus toripalimab alone for previously treated unresectable hepatocellular carcinoma. *Clin. Cancer Res.* **29**, 2816–2825 (2023).

[S168] Park, S. *et al.* A phase II, multicenter, two cohort study of 160 mg osimertinib in EGFR T790M-positive non-small-cell lung cancer patients with brain metastases or leptomeningeal disease who progressed on prior EGFR TKI therapy. *Annals Oncol.* **31**, 1397–1404 (2020).

[S169] Gouëffic, Y. *et al.* A polymer-free paclitaxel-eluting stent versus a bare-metal stent for de novo femoropopliteal lesions: the BATTLE trial. *JACC: Cardiovasc. Interv.* **13**, 447–457 (2020).

[S170] Ozaka, M. *et al.* A randomised phase II study of modified FOLFIRINOX versus gemcitabine plus nab-paclitaxel for locally advanced pancreatic cancer (JCOG1407). *Eur. J. Cancer* **181**, 135–144 (2023).

[S171] Martín, M. *et al.* A randomized adaptive phase II/III study of buparlisib, a pan-class I PI3K inhibitor, combined with paclitaxel for the treatment of HER2–advanced breast cancer (BELLE-4). *Annals Oncol.* **28**, 313–320 (2017).

[S172] Matsumura, M. *et al.* A randomized controlled trial of surgery and postoperative modified FOLFOX6 versus surgery and perioperative modified FOLFOX6 plus cetuximab in patients with KRAS wild-type resectable colorectal liver metastases: EXPERT study. *Langenbeck's Arch. Surg.* **407**, 1345–1356 (2022).

[S173] Dingemans, A.-M. *et al.* A randomized phase II study comparing paclitaxel–carboplatin–bevacizumab with or without nitroglycerin patches in patients with stage IV nonsquamous nonsmall-cell lung cancer: NVALT12 (NCT01171170). *Annals Oncol.* **26**, 2286–2293 (2015).

[S174] Kato, H. *et al.* A randomized trial of adjuvant chemotherapy with uracil–tegafur for adenocarcinoma of the lung. *New Engl. J. Medicine* **350**, 1713–1721 (2004).

[S175] Reck, M. *et al.* A randomized, double-blind, placebo-controlled phase 2 study of tigatuzumab (CS-1008) in combination with carboplatin/paclitaxel in patients with chemotherapy-naive metastatic/unresectable non-small cell lung cancer. *Lung Cancer* **82**, 441–448 (2013).

[S176] Bossi, P. *et al.* A randomized, phase 2 study of cetuximab plus cisplatin with or without paclitaxel for the first-line treatment of patients with recurrent and/or metastatic squamous cell carcinoma of the head and neck. *Annals Oncol.* **28**, 2820–2826 (2017).

[S177] Johnston, S. *et al.* Abemaciclib as initial therapy for advanced breast cancer: MONARCH 3 updated results in prognostic subgroups. *NPJ Breast Cancer* **7**, 80 (2021).

[S178] Johnston, S. R. *et al.* Abemaciclib plus endocrine therapy for hormone receptor-positive, HER2-negative, node-positive, high-risk early breast cancer (monarchE): results from a preplanned interim analysis of a randomised, open-label, phase 3 trial. *The Lancet Oncol.* **24**, 77–90 (2023).

[S179] Neven, P. *et al.* Abemaciclib plus fulvestrant in hormone receptor-positive, human epidermal growth factor receptor 2-negative advanced breast cancer in premenopausal women: subgroup analysis from the MONARCH 2 trial. *Breast Cancer Res.* **23**, 87 (2021).

[S180] Tolaney, S. M. *et al.* Abemaciclib plus trastuzumab with or without fulvestrant versus trastuzumab plus standard-of-care chemotherapy in women with hormone receptor-positive, HER2-positive advanced breast cancer (monarcHER): a randomised, open-label, phase 2 trial. *The Lancet Oncol.* **21**, 763–775 (2020).

[S181] James, N. D. *et al.* Abiraterone for prostate cancer not previously treated with hormone therapy. *New Engl. J. Medicine* **377**, 338–351 (2017).

[S182] Robertson, J. F. *et al.* Activity of fulvestrant 500 mg versus anastrozole 1 mg as first-line treatment for advanced breast cancer: results from the FIRST study. *J. Clin. Oncol.* **27**, 4530–4535 (2009).

[S183] De Boer, S. M. *et al.* Adjuvant chemoradiotherapy versus radiotherapy alone for women with high-risk endometrial cancer (PORTEC-3): final results of an international, open-label, multicentre, randomised, phase 3 trial. *The Lancet Oncol.* **19**, 295–309 (2018).

[S184] Mileshkin, L. R. *et al.* Adjuvant chemotherapy following chemoradiotherapy as primary treatment for locally advanced cervical cancer versus chemoradiotherapy alone (outback): an international, open-label, randomised, phase 3 trial. *The Lancet Oncol.* **24**, 468–482 (2023).

[S185] Tempero, M. A. *et al.* Adjuvant nab-paclitaxel+ gemcitabine in resected pancreatic ductal adenocarcinoma: results from a randomized, open-label, phase III trial. *J. Clin. Oncol.* **41**, 2007–2019 (2023).

[S186] Herbst, R. S. *et al.* Adjuvant osimertinib for resected EGFR-mutated stage IB-IIIA non–small-cell lung cancer: updated results from the phase III randomized ADAURA trial. *J. Clin. Oncol.* **41**, 1830–1840 (2023).

[S187] Chi, Y. *et al.* Anlotinib in locally advanced or metastatic radioiodine-refractory differentiated thyroid carcinoma: a randomized, double-blind, multicenter phase II trial. *Clin. Cancer Res.* **29**, 4047–4056 (2023).

[S188] Peng, H. *et al.* Anti-EGFR targeted therapy delivered before versus during radiotherapy in locoregionally advanced nasopharyngeal carcinoma: a big-data, intelligence platform-based analysis. *BMC Cancer* **18**, 323 (2018).

[S189] Pugh, S. *et al.* Association between miR-31-3p expression and cetuximab efficacy in patients with KRAS wild-type metastatic colorectal cancer: a post-hoc analysis of the New EPOC trial. *Oncotarget* **8**, 93856 (2017).

[S190] West, H. *et al.* Atezolizumab in combination with carboplatin plus nab-paclitaxel chemotherapy compared with chemotherapy alone as first-line treatment for metastatic non-squamous non-small-cell lung cancer (IMpower130): a multicentre, randomised, open-label, phase 3 trial. *The Lancet Oncol.* **20**, 924–937 (2019).

[S191] Oaknin, A. *et al.* Atezolizumab plus bevacizumab and chemotherapy for metastatic, persistent, or recurrent cervical cancer (BEATcc): a randomised, open-label, phase 3 trial. *The Lancet* **403**, 31–43 (2024).

[S192] Reck, M. *et al.* Atezolizumab plus bevacizumab and chemotherapy in non-small-cell lung cancer (IMpower150): key subgroup analyses of patients with EGFR mutations or baseline liver metastases in a randomised, open-label phase 3 trial. *The Lancet Respir. Medicine* **7**, 387–401 (2019).

[S193] Ko, A. H. *et al.* Atezolizumab plus PEGPH20 versus chemotherapy in advanced pancreatic ductal adenocarcinoma and gastric cancer: MORPHEUS phase Ib/II umbrella randomized study platform. *The Oncol.* **28**, 553–e472 (2023).

[S194] Galsky, M. D. *et al.* Atezolizumab with or without chemotherapy in metastatic urothelial cancer (IMvigor130): a multicentre, randomised, placebo-controlled phase 3 trial. *The Lancet* **395**, 1547–1557 (2020).

[S195] Gutzmer, R. *et al.* Atezolizumab, vemurafenib, and cobimetinib as first-line treatment for unresectable advanced BRAFV600 mutation-positive melanoma (IMspire150): primary analysis of the randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet* **395**, 1835–1844 (2020).

[S196] Pujade-Lauraine, E. *et al.* AURELIA: A randomized phase III trial evaluating bevacizumab combined with chemotherapy for platinum-resistant recurrent ovarian cancer. *J Clin Oncol* **30**, 327s (2012).

[S197] Cardone, C. *et al.* AXL is a predictor of poor survival and of resistance to anti-EGFR therapy in RAS wild-type metastatic colorectal cancer. *Eur. J. Cancer* **138**, 1–10 (2020).

[S198] Yoshino, T. *et al.* Baseline carcinoembryonic antigen as a predictive factor of ramucirumab efficacy in RAISE, a second-line metastatic colorectal carcinoma phase III trial. *Eur. J. Cancer* **78**, 61–69 (2017).

[S199] Turner, N. *et al.* BEECH: a dose-finding run-in followed by a randomised phase II study assessing the efficacy of AKT inhibitor capivasertib (AZD5363) combined with paclitaxel in patients with estrogen receptor-positive advanced or metastatic breast cancer, and in a PIK3CA mutant sub-population. *Annals Oncol.* **30**, 774–780 (2019).

[S200] Coleman, R. L. *et al.* Bevacizumab and paclitaxel–carboplatin chemotherapy and secondary cytoreduction in recurrent, platinum-sensitive ovarian cancer (NRG Oncology/Gynecologic Oncology Group study GOG-0213): a multicentre, open-label, randomised, phase 3 trial. *The Lancet Oncol.* **18**, 779–791 (2017).

[S201] Tewari, K. S. *et al.* Bevacizumab for advanced cervical cancer: final overall survival and adverse event analysis of a randomised, controlled, open-label, phase 3 trial (Gynecologic Oncology Group 240). *The Lancet* **390**, 1654–1663 (2017).

[S202] Zielinski, C. *et al.* Bevacizumab plus paclitaxel versus bevacizumab plus capecitabine as first-line treatment for HER2-negative metastatic breast cancer (TURANDOT): primary endpoint results of a randomised, open-label, non-inferiority, phase 3 trial. *The Lancet Oncol.* **17**, 1230–1239 (2016).

[S203] Zhou, T. *et al.* Bevacizumab versus placebo in combination with paclitaxel and carboplatin as first-line treatment for recurrent or metastatic nasopharyngeal carcinoma: a multicentre, randomised, open-label, phase II trial. *ESMO Open* **6**, 100313 (2021).

[S204] Fukuoka, M. *et al.* Biomarker analyses and final overall survival results from a phase III, randomized, open-label, first-line study of gefitinib versus carboplatin/paclitaxel in clinically selected patients with advanced non–small-cell lung cancer in Asia (IPASS). *J. Clin. Oncol.* **29**, 2866–2874 (2011).

[S205] Yoshino, T. *et al.* Biomarker analysis beyond angiogenesis: RAS/RAF mutation status, tumour sidedness, and second-line ramucirumab efficacy in patients with metastatic colorectal carcinoma from RAISE—a global phase III study. *Annals Oncol.* **30**, 124–131 (2019).

[S206] Yang, J. C.-H. *et al.* Brigatinib versus alectinib in ALK-positive NSCLC after disease progression on crizotinib: results of phase 3 ALTA-3 trial. *J. Thorac. Oncol.* **18**, 1743–1755 (2023).

[S207] Camidge, D. R. *et al.* Brigatinib versus crizotinib in advanced ALK inhibitor–naive ALK-positive non–small cell lung cancer: second interim analysis of the phase III ALTA-1L trial. *J. Clin. Oncol.* **38**, 3592–3603 (2020).

[S208] Camidge, D. R. *et al.* Brigatinib versus crizotinib in ALK inhibitor–naive advanced ALK-positive NSCLC: final results of phase 3 ALTA-1L trial. *J. Thorac. Oncol.* **16**, 2091–2108 (2021).

[S209] Camidge, D. R. *et al.* Brigatinib versus crizotinib in ALK-positive non–small-cell lung cancer. *New Engl. J. Medicine* **379**, 2027–2039 (2018).

[S210] Campone, M. *et al.* Buparlisib plus fulvestrant versus placebo plus fulvestrant for postmenopausal, hormone receptor-positive, human epidermal growth factor receptor 2-negative, advanced breast cancer: Overall survival results from BELLE-2. *Eur. J. Cancer* **103**, 147–154 (2018).

[S211] Baselga, J. *et al.* Buparlisib plus fulvestrant versus placebo plus fulvestrant in postmenopausal, hormone receptor-positive, HER2-negative, advanced breast cancer (BELLE-2): a randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Oncol.* **18**, 904–916 (2017).

[S212] Ren, S. *et al.* Camrelizumab plus carboplatin and paclitaxel as first-line treatment for advanced squamous NSCLC (CameL-Sq): a phase 3 trial. *J. Thorac. Oncol.* **17**, 544–557 (2022).

[S213] Zhou, C. *et al.* Camrelizumab plus carboplatin and pemetrexed versus chemotherapy alone in chemotherapy-naive patients with advanced non-squamous non-small-cell lung cancer (CameL): a randomised, open-label, multicentre, phase 3 trial. *The Lancet Respir. Medicine* **9**, 305–314 (2021).

[S214] Qin, S. *et al.* Camrelizumab plus rivoceranib versus sorafenib as first-line therapy for unresectable hepatocellular carcinoma (CARES-310): a randomised, open-label, international phase 3 study. *The Lancet* **402**, 1133–1146 (2023).

[S215] Huang, J. *et al.* Camrelizumab versus investigator's choice of chemotherapy as second-line therapy for advanced or metastatic oesophageal squamous cell carcinoma (ESCORT): a multicentre, randomised, open-label, phase 3 study. *The Lancet Oncol.* **21**, 832–842 (2020).

[S216] Yang, Y. *et al.* Camrelizumab versus placebo in combination with gemcitabine and cisplatin as first-line treatment for recurrent or metastatic nasopharyngeal carcinoma (CAPTAIN-1st): a multicentre, randomised, double-blind, phase 3 trial. *The Lancet Oncol.* **22**, 1162–1174 (2021).

[S217] Schmid, P. *et al.* Capivasertib plus paclitaxel versus placebo plus paclitaxel as first-line therapy for metastatic triple-negative breast cancer: the PAKT trial. *J. Clin. Oncol.* **38**, 423–433 (2020).

[S218] Williams, T. M. *et al.* Caveolin gene expression predicts clinical outcomes for early-stage HER2-negative breast cancer treated with paclitaxel-based chemotherapy in the GeparSepto trial. *Clin. Cancer Res.* **29**, 3384–3394 (2023).

[S219] Herbst, R. S. *et al.* Cetuximab plus carboplatin and paclitaxel with or without bevacizumab versus carboplatin and paclitaxel with or without bevacizumab in advanced NSCLC (SWOG S0819): a randomised, phase 3 study. *The Lancet Oncol.* **19**, 101–114 (2018).

[S220] Sweeney, C. J. *et al.* Chemohormonal therapy in metastatic hormone-sensitive prostate cancer. *New Engl. J. Medicine* **373**, 737–746 (2015).

[S221] Vansteenkiste, J. *et al.* Cilengitide combined with cetuximab and platinum-based chemotherapy as first-line treatment in advanced non-small-cell lung cancer (NSCLC) patients: results of an open-label, randomized, controlled phase II study (CERTO). *Annals Oncol.* **26**, 1734–1740 (2015).

[S222] Pellini, B. *et al.* Circulating tumor DNA monitoring on chemo-immunotherapy for risk stratification in advanced non–small cell lung cancer. *Clin. Cancer Res.* **29**, 4596–4605 (2023).

[S223] Hu, X.-C. *et al.* Cisplatin plus gemcitabine versus paclitaxel plus gemcitabine as first-line therapy for metastatic triple-negative breast cancer (CBCSG006): a randomised, open-label, multicentre, phase 3 trial. *The Lancet Oncol.* **16**, 436–446 (2015).

[S224] Xu, J. *et al.* Clinical and biomarker analyses of sintilimab versus chemotherapy as second-line therapy for advanced or metastatic esophageal squamous cell carcinoma: a randomized, open-label phase 2 study (ORIENT-2). *Nat. Commun.* **13**, 857 (2022).

[S225] Carvajal, R. D. *et al.* Clinical and molecular response to tebentafusp in previously treated patients with metastatic uveal melanoma: a phase 2 trial. *Nat. Medicine* **28**, 2364–2373 (2022).

[S226] Wang, J. *et al.* Clinical efficacy of fulvestrant versus exemestane as first-line therapies for Chinese postmenopausal oestrogen-receptor positive/human epidermal growth factor receptor 2-advanced breast cancer (FRIEND study). *Eur. J. Cancer* **184**, 73–82 (2023).

[S227] Shi, Y. *et al.* CMAB009 plus irinotecan versus irinotecan-only as second-line treatment after fluoropyrimidine and oxaliplatin failure in KRAS wild-type metastatic colorectal cancer patients: promising findings from a prospective, open-label, randomized, phase III trial. *Cancer Commun.* **39**, 1–13 (2019).

[S228] DeMichele, A. *et al.* Comparative effectiveness of first-line palbociclib plus letrozole versus letrozole alone for HR+/HER2- metastatic breast cancer in US real-world clinical practice. *Breast Cancer Res.* **23**, 37 (2021).

[S229] Smith, I. *et al.* Comparative efficacy and safety of adjuvant letrozole versus anastrozole in postmenopausal patients with hormone receptor–positive, node-positive early breast cancer: Final results of the randomized phase III femara versus anastrozole clinical evaluation (face) trial. *J. Clin. Oncol.* **35**, 1041–1048 (2017).

[S230] Chen, Y. *et al.* Comparing paclitaxel plus fluorouracil versus cisplatin plus fluorouracil in chemoradiotherapy for locally advanced esophageal squamous cell cancer: a randomized, multicenter, phase III clinical trial. *J. Clin. Oncol.* **37**, 1695–1703 (2019).

[S231] Huang, J. *et al.* Comparison of FOLFIRI with or without cetuximab in patients with resected stage III colon cancer; NCCTG (Alliance) intergroup trial N0147. *Clin. Color. Cancer* **13**, 100–109 (2014).

[S232] Lorch, A. *et al.* Conventional-dose versus high-dose chemotherapy as first salvage treatment in male patients with metastatic germ cell tumors: evidence from a large international database. *J. Clin. Oncol.* **29**, 2178–2184 (2011).

[S233] Zhang, Q., Wu, P., He, X., Ding, Y. & Shu, Y. Cost-effectiveness analysis of camrelizumab vs. placebo added to chemotherapy as first-line therapy for advanced or metastatic esophageal squamous cell carcinoma in china. *Front. Oncol.* **11**, 790373 (2021).

[S234] Reiman, T. *et al.* Cross-validation study of class III beta-tubulin as a predictive marker for benefit from adjuvant chemotherapy in resected non-small-cell lung cancer: analysis of four randomized trials. *Annals Oncol.* **23**, 86–93 (2012).

[S235] Zhou, F. *et al.* CTHRC1 is a prognostic biomarker and correlated with immune infiltrates in kidney renal papillary cell carcinoma and kidney renal clear cell carcinoma. *Front. Oncol.* **10**, 570819 (2021).

[S236] Xu, B. *et al.* Dalpiciclib or placebo plus fulvestrant in hormone receptor-positive and HER2-negative advanced breast cancer: a randomized, phase 3 trial. *Nat. Medicine* **27**, 1904–1909 (2021).

[S237] Zhang, P. *et al.* Dalpiciclib plus letrozole or anastrozole versus placebo plus letrozole or anastrozole as first-line treatment in patients with hormone receptor-positive, HER2-negative advanced breast cancer (DAWNA-2): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Oncol.* **24**, 646–657 (2023).

[S238] Kim, S. *et al.* DCF versus doublet chemotherapy as first-line treatment of advanced squamous anal cell carcinoma: a multicenter propensity score-matching study. *Exp. Hematol. & Oncol.* **12**, 63 (2023).

[S239] Mirza, M. R. *et al.* Dostarlimab for primary advanced or recurrent endometrial cancer. *New Engl. J. Medicine* **388**, 2145–2158 (2023).

[S240] Jassem, J. *et al.* Doxorubicin and paclitaxel versus fluorouracil, doxorubicin and cyclophosphamide as first-line therapy for women with advanced breast cancer: long-term analysis of the previously published trial. *Oncol. Res. Treat.* **32**, 468–472 (2009).

[S241] Modest, D. P. *et al.* Early tumor shrinkage in patients with metastatic colorectal cancer receiving first-line treatment with cetuximab combined with either CAPIRI or CAPOX: an analysis of the German AIO KRK 0104 trial. *Acta Oncol.* **52**, 956–962 (2013).

[S242] Trukhin, D. *et al.* Efficacy, safety and immunogenicity of MB02 (bevacizumab biosimilar) versus reference bevacizumab in advanced non-small cell lung cancer: a randomized, double-blind, phase III study (STELLA). *BioDrugs* **35**, 429–444 (2021).

[S243] Shitara, K. *et al.* Effects of prior therapies on outcomes with trifluridine/tipiracil in patients with metastatic gastric/gastroesophageal junction cancer in a randomized phase III trial (tags). *J. Cancer Res. Clin. Oncol.* **149**, 9361–9374 (2023).

[S244] Kogure, Y. *et al.* Efficacy and safety of carboplatin with nab-paclitaxel versus docetaxel in older patients with squamous non-small-cell lung cancer (CAPITAL): a randomised, multicentre, open-label, phase 3 trial. *The Lancet Heal. Longev.* **2**, e791–e800 (2021).

[S245] Balar, A. *et al.* Efficacy and safety of pembrolizumab in metastatic urothelial carcinoma: results from KEYNOTE-045 and KEYNOTE-052 after up to 5 years of follow-up. *Annals Oncol.* **34**, 289–299 (2023).

[S246] Ramalingam, S. S. *et al.* Efficacy and safety of ramucirumab with docetaxel versus placebo with docetaxel as second-line treatment of advanced non–small-cell lung cancer: a subgroup analysis according to patient age in the REVEL trial. *Clin. Lung Cancer* **19**, 270–279 (2018).

[S247] Qin, S. *et al.* Efficacy and tolerability of first-line cetuximab plus leucovorin, fluorouracil, and oxaliplatin (FOLFOX-4) versus FOLFOX-4 in patients with RAS wild-type metastatic colorectal cancer: the open-label, randomized, phase III TAILOR trial. *J. Clin. Oncol.* **36**, 3031–3039 (2018).

[S248] Wang, F.-H. *et al.* Efficacy, safety, and correlative biomarkers of toripalimab in previously treated recurrent or metastatic nasopharyngeal carcinoma: a phase II clinical trial (POLARIS-02). *J. Clin. Oncol.* **39**, 704–712 (2021).

[S249] Reck, M. *et al.* Efficient palliation in patients with small-cell lung cancer by a combination of paclitaxel, etoposide and carboplatin: quality of life and 6-years'-follow-up results from a randomised phase III trial. *Lung Cancer* **53**, 67–75 (2006).

[S250] Tabernero, J. *et al.* Encorafenib plus cetuximab as a new standard of care for previously treated BRAF V600E–mutant metastatic colorectal cancer: updated survival results and subgroup analyses from the BEACON study. *J. Clin. Oncol.* **39**, 273–284 (2021).

[S251] Goto, K. *et al.* Epidermal growth factor receptor mutation status in circulating free DNA in serum: from IPASS, a phase III study of gefitinib or carboplatin/paclitaxel in non-small cell lung cancer. *J. Thorac. Oncol.* **7**, 115–121 (2012).

[S252] Loriot, Y. *et al.* Erdafitinib or chemotherapy in advanced or metastatic urothelial carcinoma. *New Engl. J. Medicine* **389**, 1961–1971 (2023).

[S253] Baselga, J. *et al.* Everolimus in postmenopausal hormone-receptor–positive advanced breast cancer. *New Engl. J. Medicine* **366**, 520–529 (2012).

[S254] Modest, D. P. *et al.* Exploring the effect of primary tumor sidedness on therapeutic efficacy across treatment lines in patients with metastatic colorectal cancer: analysis of FIRE-3 (AIOKRK0306). *Oncotarget* **8**, 105749 (2017).

[S255] Song, L. *et al.* Expression signature, prognosis value and immune characteristics of cathepsin F in non-small cell lung cancer identified by bioinformatics assessment. *BMC Pulm. Medicine* **21**, 420 (2021).

[S256] Price, T. *et al.* Final results and outcomes by prior bevacizumab exposure, skin toxicity, and hypomagnesaemia from ASPECCT: randomized phase 3 non-inferiority study of panitumumab versus cetuximab in chemorefractory wild-type KRAS exon 2 metastatic colorectal cancer. *Eur. J. Cancer* **68**, 51–59 (2016).

[S257] Monk, B. J. *et al.* Final results of a phase 3 study of trebananib plus weekly paclitaxel in recurrent ovarian cancer (TRINOVA-1): long-term survival, impact of ascites, and progression-free survival-2. *Gynecol. Oncol.* **143**, 27–34 (2016).

[S258] Ciardiello, D. *et al.* Final results of the CAVE trial in RAS wild type metastatic colorectal cancer patients treated with cetuximab plus avelumab as rechallenge therapy: Neutrophil to lymphocyte ratio predicts survival. *Clin. Color. Cancer* **21**, 141–148 (2022).

[S259] Takahashi, S. *et al.* First-line pembrolizumab±chemotherapy for recurrent/metastatic head and neck cancer: Japanese subgroup of KEYNOTE-048. *Int. J. Clin. Oncol.* **27**, 1805–1817 (2022).

[S260] Zhou, C. *et al.* First-line selpercatinib or chemotherapy and pembrolizumab in RET fusion–positive NSCLC. *New Engl. J. Medicine* **389**, 1839–1850 (2023).

[S261] García-Alfonso, P. *et al.* First-line biological agents plus chemotherapy in older patients with metastatic colorectal cancer: a retrospective pooled analysis. *Drugs & Aging* **38**, 219–231 (2021).

[S262] Stintzing, S. *et al.* FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer–subgroup analysis of patients with KRAS: mutated tumours in the randomised German AIO study KRK-0306. *Annals Oncol.* **23**, 1693–1699 (2012).

[S263] Stintzing, S. *et al.* FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab for metastatic colorectal cancer (FIRE-3): a post-hoc analysis of tumour dynamics in the final RAS wild-type subgroup of this randomised open-label phase 3 trial. *The Lancet Oncol.* **17**, 1426–1434 (2016).

[S264] Lorenzen, S. *et al.* FOLFIRI plus ramucirumab versus paclitaxel plus ramucirumab as second-line therapy for patients with advanced or metastatic gastroesophageal adenocarcinoma with or without prior docetaxel–results from the phase II RAMIRIS study of the german gastric cancer study group at AIO. *Eur. J. Cancer* **165**, 48–57 (2022).

[S265] Brodowicz, T. *et al.* FOLFOX4 plus cetuximab administered weekly or every second week in the first-line treatment of patients with KRAS wild-type metastatic colorectal cancer: a randomized phase II CECOG study. *Annals Oncol.* **24**, 1769–1777 (2013).

[S266] Bokemeyer, C. *et al.* FOLFOX4 plus cetuximab treatment and RAS mutations in colorectal cancer. *Eur. J. Cancer* **51**, 1243–1252 (2015).

[S267] Robertson, J. F. *et al.* Fulvestrant 500 mg versus anastrozole 1 mg for hormone receptor-positive advanced breast cancer (FALCON): an international, randomised, double-blind, phase 3 trial. *The Lancet* **388**, 2997–3005 (2016).

[S268] Robertson, J. F. *et al.* Fulvestrant 500 mg versus anastrozole 1 mg for the first-line treatment of advanced breast cancer: follow-up analysis from the randomized 'FIRST' study. *Breast Cancer Res. Treat.* **136**, 503–511 (2012).

[S269] Ellis, M. J. *et al.* Fulvestrant 500 mg versus anastrozole 1 mg for the first-line treatment of advanced breast cancer: overall survival analysis from the phase II FIRST study. *J. Clin. Oncol.* **33**, 3781–3787 (2015).

[S270] Jones, R. H. *et al.* Fulvestrant plus capivasertib versus placebo after relapse or progression on an aromatase inhibitor in metastatic, oestrogen receptor-positive breast cancer (FAKTION): a multicentre, randomised, controlled, phase 2 trial. *The Lancet Oncol.* **21**, 345–357 (2020).

[S271] Howell, S. J. *et al.* Fulvestrant plus capivasertib versus placebo after relapse or progression on an aromatase inhibitor in metastatic, oestrogen receptor-positive, HER2-negative breast cancer (FAKTION): overall survival, updated progression-free survival, and expanded biomarker analysis from a randomised, phase 2 trial. *The Lancet Oncol.* **23**, 851–864 (2022).

[S272] Cristofanilli, M. *et al.* Fulvestrant plus palbociclib versus fulvestrant plus placebo for treatment of hormone-receptor-positive, HER2-negative metastatic breast cancer that progressed on previous endocrine therapy (PALOMA-3): final analysis of the multicentre, double-blind, phase 3 randomised controlled trial. *The Lancet Oncol.* **17**, 425–439 (2016).

[S273] Howell, A. & Sapunar, F. Fulvestrant revisited: efficacy and safety of the 500-mg dose. *Clin. Breast Cancer* **11**, 204–210 (2011).

[S274] Fiteni, F. *et al.* Health-related quality of life in elderly patients with advanced non-small cell lung cancer comparing carboplatin and weekly paclitaxel doublet chemotherapy with monotherapy. *Eur. Respir. J.* **48**, 861–872 (2016).

[S275] Lorusso, D. *et al.* Health-related quality of life in patients with advanced endometrial cancer treated with lenvatinib plus pembrolizumab or treatment of physician's choice. *Eur. J. Cancer* **186**, 172–184 (2023).

[S276] Monk, B. J. *et al.* Health-related quality of life with pembrolizumab or placebo plus chemotherapy with or without bevacizumab for persistent, recurrent, or metastatic cervical cancer (KEYNOTE-826): a randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Oncol.* **24**, 392–402 (2023).

[S277] Perera, S. *et al.* hENT1 expression predicts response to gemcitabine and nab-paclitaxel in advanced pancreatic ductal adenocarcinoma. *Clin. Cancer Res.* **28**, 5115–5120 (2022).

[S278] You, J. *et al.* High-dose versus standard-dose intensity-modulated radiotherapy with concurrent paclitaxel plus carboplatin for patients with thoracic esophageal squamous cell carcinoma: a randomized, multicenter, open-label, phase 3 superiority trial. *Int. J. Radiat. Oncol. Biol. Phys.* **115**, 1129–1137 (2023).

[S279] Stintzing, S. *et al.* Impact of BRAF and RAS mutations on first-line efficacy of FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab: analysis of the FIRE-3 (AIO KRK-0306) study. *Eur. J. Cancer* **79**, 50–60 (2017).

[S280] Swisher, E. M. *et al.* Impact of homologous recombination status and responses with veliparib combined with first-line chemotherapy in ovarian cancer in the phase 3 VELIA/GOG-3005 study. *Gynecol. Oncol.* **164**, 245–253 (2022).

[S281] Modest, D. P. *et al.* Impact of subsequent therapies on outcome of the FIRE-3/aio KRK0306 trial: first-line therapy with FOLFIRI plus cetuximab or bevacizumab in patients with KRAS wild-type tumors in metastatic colorectal cancer. *J. Clin. Oncol.* **33**, 3718–3726 (2015).

[S282] Aghajanian, C. *et al.* Impact of veliparib, paclitaxel dosing regimen, and germline BRCA status on the primary treatment of serous ovarian cancer–an ancillary data analysis of the VELIA trial. *Gynecol. Oncol.* **164**, 278–287 (2022).

[S283] Von Hoff, D. D. *et al.* Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. *New Engl. J. Medicine* **369**, 1691–1703 (2013).

[S284] Scher, H. I. *et al.* Increased survival with enzalutamide in prostate cancer after chemotherapy. *New Engl. J. Medicine* **367**, 1187–1197 (2012).

[S285] Geoffrois, L. *et al.* Induction chemotherapy followed by cetuximab radiotherapy is not superior to concurrent chemoradiotherapy for head and neck carcinomas: results of the GORTEC 2007-02 phase III randomized trial. *J. Clin. Oncol.* **36**, 3077–3083 (2018).

[S286] Luo, X. *et al.* Induction TPF followed by concurrent chemoradiotherapy versus concurrent chemoradiotherapy alone in locally advanced hypopharyngeal cancer: a preliminary analysis of a randomized phase 2 trial. *BMC Cancer* **22**, 1235 (2022).

[S287] Sastre, J. *et al.* Influence of BRAF and PIK3CA mutations on the efficacy of FOLFIRI plus bevacizumab or cetuximab as first-line therapy in patients with RAS wild-type metastatic colorectal carcinoma and< 3 baseline circulating tumour cells: the randomised phase II VISNÚ-2 study. *ESMO Open* **6**, 100062 (2021).

[S288] Adams, R. A. *et al.* Intermittent versus continuous oxaliplatin and fluoropyrimidine combination chemotherapy for first-line treatment of advanced colorectal cancer: results of the randomised phase 3 MRC COIN trial. *The Lancet Oncol.* **12**, 642–653 (2011).

[S289] Kim, S.-B. *et al.* Ipatasertib plus paclitaxel versus placebo plus paclitaxel as first-line therapy for metastatic triple-negative breast cancer (LOTUS): a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. *The Lancet Oncol.* **18**, 1360–1372 (2017).

[S290] Matsubara, N. *et al.* Japanese subgroup analysis of EV-301: an open-label, randomized phase 3 study to evaluate enfortumab vedotin versus chemotherapy in subjects with previously treated locally advanced or metastatic urothelial carcinoma. *Cancer Medicine* **12**, 2761–2771 (2023).

[S291] Satoh, T. *et al.* Lapatinib plus paclitaxel versus paclitaxel alone in the second-line treatment of HER2-amplified advanced gastric cancer in asian populations: TyTAN—a randomized, phase III study. *J. Clin. Oncol.* **32**, 2039–2049 (2014).

[S292] Goetz, M. *et al.* Lasofoxifene versus fulvestrant for ER+/HER2- metastatic breast cancer with an ESR1 mutation: results from the randomized, phase II ELAINE 1 trial. *Annals Oncol.* **34**, 1141–1151 (2023).

[S293] Von Einem, J. *et al.* Left-sided primary tumors are associated with favorable prognosis in patients with KRAS codon 12/13 wild-type metastatic colorectal cancer treated with cetuximab plus chemotherapy: an analysis of the AIO KRK-0104 trial. *J. Cancer Res. Clin. Oncol.* **140**, 1607–1614 (2014).

[S294] Makker, V. *et al.* Lenvatinib plus pembrolizumab for advanced endometrial cancer. *New Engl. J. Medicine* **386**, 437–448 (2022).

[S295] Yonemori, K. *et al.* Lenvatinib plus pembrolizumab in Japanese patients with endometrial cancer: Results from Study 309/KEYNOTE-775. *Cancer Sci.* **113**, 3489–3497 (2022).

[S296] Bradley, J. D. *et al.* Long-term results of NRG oncology RTOG 0617: standard-versus high-dose chemoradiotherapy with or without cetuximab for unresectable stage III non–small-cell lung cancer. *J. Clin. Oncol.* **38**, 706–714 (2020).

[S297] Monk, B. J. *et al.* MILO/ENGOT-ov11: binimetinib versus physician's choice chemotherapy in recurrent or persistent low-grade serous carcinomas of the ovary, fallopian tube, or primary peritoneum. *J. Clin. Oncol.* **38**, 3753–3762 (2020).

[S298] Sledge Jr, G. W. *et al.* MONARCH 2: abemaciclib in combination with fulvestrant in women with HR+/HER2- advanced breast cancer who had progressed while receiving endocrine therapy. *J. Clin. Oncol.* **35**, 2875–2884 (2017).

[S299] Zhang, Q. Y. *et al.* MONARCH plus: abemaciclib plus endocrine therapy in women with HR+/HER2–advanced breast cancer: the multinational randomized phase III study. *Ther. Adv. Med. Oncol.* **12**, 1758835920963925 (2020).

[S300] Ye, W. *et al.* Multicenter randomized phase 2 clinical trial of a recombinant human endostatin adenovirus in patients with advanced head and neck carcinoma. *Mol. Ther.* **22**, 1221–1229 (2014).

[S301] Weiss, J. M. *et al.* Myelopreservation with the CDK4/6 inhibitor trilaciclib in patients with small-cell lung cancer receiving first-line chemotherapy: a phase Ib/randomized phase II trial. *Annals Oncol.* **30**, 1613–1621 (2019).

[S302] Young, R. *et al.* nab-paclitaxel plus gemcitabine in metastatic pancreatic adenocarcinoma: Australian subset analyses of the phase III mpact trial. *Asia-Pacific J. Clin. Oncol.* **14**, e325–e331 (2018).

[S303] Dai, Y.-H. *et al.* Nab-paclitaxel plus S-1 versus oxaliplatin plus S-1 as first-line treatment in advanced gastric cancer: results of a multicenter, randomized, phase III trial (GAPSO study). *Ther. Adv. Med. Oncol.* **14**, 17588359221118020 (2022).

[S304] Tang, H. *et al.* Neoadjuvant chemoradiotherapy versus neoadjuvant chemotherapy followed by minimally invasive esophagectomy for locally advanced esophageal squamous cell carcinoma: a prospective multicenter randomized clinical trial. *Annals Oncol.* **34**, 163–172 (2023).

[S305] Ferron, G. *et al.* Neoadjuvant chemotherapy with or without nintedanib for advanced epithelial ovarian cancer: lessons from the GINECO double-blind randomized phase II CHIVA trial. *Gynecol. Oncol.* **170**, 186–194 (2023).

[S306] Forde, P. M. *et al.* Neoadjuvant nivolumab plus chemotherapy in resectable lung cancer. *New Engl. J. Medicine* **386**, 1973–1985 (2022).

[S307] Hirsh, V. Next-generation covalent irreversible kinase inhibitors in NSCLC: focus on afatinib. *BioDrugs* **29**, 167–183 (2015).

[S308] Van Der Heijden, M. S. *et al.* Nivolumab plus gemcitabine–cisplatin in advanced urothelial carcinoma. *New Engl. J. Medicine* **389**, 1778–1789 (2023).

[S309] Haddad, R. I. *et al.* Nivolumab plus ipilimumab versus EXTREME regimen as first-line treatment for recurrent/metastatic squamous cell carcinoma of the head and neck: the final results of CheckMate 651. *J. Clin. Oncol.* **41**, 2166–2180 (2023).

[S310] Saba, N. F. *et al.* Nivolumab versus investigator's choice in patients with recurrent or metastatic squamous cell carcinoma of the head and neck: Efficacy and safety in CheckMate 141 by age. *Oral Oncol.* **96**, 7–14 (2019).

[S311] Kasper, S. *et al.* Noninferiority of cetuximab every-2-weeks versus standard once-weekly administration schedule for the first-line treatment of RAS wild-type metastatic colorectal cancer. *Eur. J. Cancer* **144**, 291–301 (2021).

[S312] Aghajanian, C. *et al.* OCEANS: a randomized, double-blind, placebo-controlled phase III trial of chemotherapy with or without bevacizumab in patients with platinum-sensitive recurrent epithelial ovarian, primary peritoneal, or fallopian tube cancer. *J. Clin. Oncol.* **30**, 2039–2045 (2012).

[S313] Arend, R. C. *et al.* Ofranergene obadenovec (Ofra-Vec, VB-111) with weekly paclitaxel for platinum-resistant ovarian cancer: randomized controlled phase III trial (OVAL study/GOG 3018). *J. Clin. Oncol.* **42**, 170–179 (2024).

[S314] Bang, Y.-J. *et al.* Olaparib in combination with paclitaxel in patients with advanced gastric cancer who have progressed following first-line therapy (GOLD): a double-blind, randomised, placebo-controlled, phase 3 trial. *The Lancet Oncol.* **18**, 1637–1651 (2017).

[S315] Pfisterer, J. *et al.* Optimal treatment duration of bevacizumab as front-line therapy for advanced ovarian cancer: AGO-OVAR 17 BOOST/GINECO OV118/ENGOT Ov-15 open-label randomized phase III trial. *J. Clin. Oncol.* **41**, 893–902 (2023).

[S316] Fan, L. *et al.* Optimising first-line subtyping-based therapy in triple-negative breast cancer (FUTURE-SUPER): a multi-cohort, randomised, phase 2 trial. *The Lancet Oncol.* **25**, 184–197 (2024).

[S317] Wu, Y.-L. *et al.* Osimertinib in resected EGFR-mutated non–small-cell lung cancer. *New Engl. J. Medicine* **383**, 1711–1723 (2020).

[S318] Soria, J.-C. *et al.* Osimertinib in untreated EGFR-mutated advanced non–small-cell lung cancer. *New Engl. J. Medicine* **378**, 113–125 (2018).

[S319] Mok, T. S. *et al.* Osimertinib or platinum–pemetrexed in EGFR T790M–positive lung cancer. *New Engl. J. Medicine* **376**, 629–640 (2017).

[S320] Papadimitrakopoulou, V. *et al.* Osimertinib versus platinum–pemetrexed for patients with EGFR T790M advanced NSCLC and progression on a prior EGFR-tyrosine kinase inhibitor: AURA3 overall survival analysis. *Annals Oncol.* **31**, 1536–1544 (2020).

[S321] Tolaney, S. M. *et al.* Overall survival and exploratory biomarker analyses of abemaciclib plus trastuzumab with or without fulvestrant versus trastuzumab plus chemotherapy in HR+, HER2+ metastatic breast cancer patients. *Clin. Cancer Res.* **30**, 39–49 (2024).

[S322] Nathan, P. *et al.* Overall survival benefit with tebentafusp in metastatic uveal melanoma. *New Engl. J. Medicine* **385**, 1196–1206 (2021).

[S323] Piulats, J. M. *et al.* Overall survival from tebentafusp versus nivolumab plus ipilimumab in first-line metastatic uveal melanoma: a propensity score-weighted analysis. *Annals Oncol.* **35**, 317–326 (2024).

[S324] Finn, R. S. *et al.* Overall survival results from the randomized phase 2 study of palbociclib in combination with letrozole versus letrozole alone for first-line treatment of ER+/HER2- advanced breast cancer (PALOMA-1, TRIO-18). *Breast Cancer Res. Treat.* **183**, 419–428 (2020).

[S325] Ascierto, P. A. *et al.* Overall survival with first-line atezolizumab in combination with vemurafenib and cobimetinib in BRAFV600 mutation-positive advanced melanoma (IMspire150): second interim analysis of a multicentre, randomised, phase 3 study. *The Lancet Oncol.* **24**, 33–44 (2023).

[S326] Ramalingam, S. S. *et al.* Overall survival with osimertinib in untreated, EGFR-mutated advanced NSCLC. *New Engl. J. Medicine* **382**, 41–50 (2020).

[S327] Lu, Z. *et al.* Paclitaxel and cisplatin with or without cetuximab in metastatic esophageal squamous cell carcinoma: a randomized, multicenter phase II trial. *The Innov.* **3** (2022).

[S328] Zhang, J. *et al.* Paclitaxel liposome for injection (Lipusu) plus cisplatin versus gemcitabine plus cisplatin in the first-line treatment of locally advanced or metastatic lung squamous cell carcinoma: A multicenter, randomized, open-label, parallel controlled clinical study. *Cancer Commun.* **42**, 3–16 (2022).

[S329] Kitagawa, R. *et al.* Paclitaxel plus carboplatin versus paclitaxel plus cisplatin in metastatic or recurrent cervical cancer: the open-label randomized phase III trial JCOG0505. *J. Clin. Oncol.* **33**, 2129–2135 (2015).

[S330] Icon, Collaborators, A. *et al.* Paclitaxel plus platinum-based chemotherapy versus conventional platinum-based chemotherapy in women with relapsed ovarian cancer: the ICON4/AGO-OVAR-2.2 trial. *The Lancet* **361**, 2099–2106 (2003).

[S331] Martin, M. *et al.* Palbociclib in combination with endocrine therapy versus capecitabine in hormonal receptor-positive, human epidermal growth factor 2-negative, aromatase inhibitor-resistant metastatic breast cancer: a phase III randomised controlled trial—PEARL. *Annals Oncol.* **32**, 488–499 (2021).

[S332] Xu, B. *et al.* Palbociclib plus letrozole versus placebo plus letrozole in Asian postmenopausal women with oestrogen receptor–positive/human epidermal growth factor receptor 2–negative advanced breast cancer: Primary results from PALOMA-4. *Eur. J. Cancer* **175**, 236–245 (2022).

[S333] Lorusso, D. *et al.* Patient-reported outcomes and final overall survival results from the randomized phase 3 PENELOPE trial evaluating pertuzumab in low tumor human epidermal growth factor receptor 3 (HER3) mRNA-expressing platinum-resistant ovarian cancer. *Int. J. Gynecol. Cancer* **29**, 1141–1147 (2019).

[S334] Yoh, K. *et al.* Patient-reported outcomes in RELAY, a phase 3 trial of ramucirumab plus erlotinib versus placebo plus erlotinib in untreated EGFR-mutated metastatic non-small-cell lung cancer. *Curr. Med. Res. Opin.* **36**, 1667–1675 (2020).

[S335] Powles, T. *et al.* Pembrolizumab alone or combined with chemotherapy versus chemotherapy as first-line therapy for advanced urothelial carcinoma (KEYNOTE-361): a randomised, open-label, phase 3 trial. *The Lancet Oncol.* **22**, 931–945 (2021).

[S336] Eskander, R. N. *et al.* Pembrolizumab plus chemotherapy in advanced endometrial cancer. *New Engl. J. Medicine* **388**, 2159–2170 (2023).

[S337] Borghaei, H. *et al.* Pembrolizumab plus chemotherapy versus chemotherapy alone in patients with advanced non–small cell lung cancer without tumor PD-L1 expression: A pooled analysis of 3 randomized controlled trials. *Cancer* **126**, 4867–4877 (2020).

[S338] Diaz, L. A. *et al.* Pembrolizumab versus chemotherapy for microsatellite instability-high or mismatch repair-deficient metastatic colorectal cancer (KEYNOTE-177): final analysis of a randomised, open-label, phase 3 study. *The Lancet Oncol.* **23**, 659–670 (2022).

[S339] Cao, Y. *et al.* Pembrolizumab versus chemotherapy for patients with esophageal squamous cell carcinoma enrolled in the randomized KEYNOTE-181 trial in Asia. *ESMO Open* **7**, 100341 (2022).

[S340] Mok, T. S. *et al.* Pembrolizumab versus chemotherapy for previously untreated, PD-L1-expressing, locally advanced or metastatic non-small-cell lung cancer (KEYNOTE-042): a randomised, open-label, controlled, phase 3 trial. *The Lancet* **393**, 1819–1830 (2019).

[S341] Cohen, E. E. *et al.* Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. *The Lancet* **393**, 156–167 (2019).

[S342] Harrington, K. J. *et al.* Pembrolizumab with or without chemotherapy in recurrent or metastatic head and neck squamous cell carcinoma: updated results of the phase III KEYNOTE-048 study. *J. Clin. Oncol.* **41**, 790–802 (2023).

[S343] Fuchs, C. S. *et al.* Pembrolizumab versus paclitaxel for previously treated PD-L1-positive advanced gastric or gastroesophageal junction cancer: 2-year update of the randomized phase 3 KEYNOTE-061 trial. *Gastric Cancer* **25**, 197–206 (2022).

[S344] O'Brien, M. *et al.* Pembrolizumab versus placebo as adjuvant therapy for completely resected stage IB–IIIA non-small-cell lung cancer (PEARLS/KEYNOTE-091): an interim analysis of a randomised, triple-blind, phase 3 trial. *The Lancet Oncol.* **23**, 1274–1286 (2022).

[S345] Luke, J. J. *et al.* Pembrolizumab versus placebo as adjuvant therapy in completely resected stage IIB or IIC melanoma (KEYNOTE-716): a randomised, double-blind, phase 3 trial. *The Lancet* **399**, 1718–1729 (2022).

[S346] Powles, T. *et al.* Pembrolizumab versus placebo as post-nephrectomy adjuvant therapy for clear cell renal cell carcinoma (KEYNOTE-564): 30-month follow-up analysis of a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Oncol.* **23**, 1133–1144 (2022).

[S347] Wakelee, H. *et al.* Perioperative pembrolizumab for early-stage non–small-cell lung cancer. *New Engl. J. Medicine* **389**, 491–503 (2023).

[S348] Hadoux, J. *et al.* Phase 3 trial of selpercatinib in advanced RET-mutant medullary thyroid cancer. *New Engl. J. Medicine* **389**, 1851–1861 (2023).

[S349] Carvajal, R. D. *et al.* Phase I study of safety, tolerability, and efficacy of tebentafusp using a step-up dosing regimen and expansion in patients with metastatic uveal melanoma. *J. Clin. Oncol.* **40**, 1939–1948 (2022).

[S350] Reckamp, K. L. *et al.* Phase II randomized study of ramucirumab and pembrolizumab versus standard of care in advanced non–small-cell lung cancer previously treated with immunotherapy—lung-MAP S1800A. *J. Clin. Oncol.* **40**, 2295–2307 (2022).

[S351] Slamon, D. J. *et al.* Phase III randomized study of ribociclib and fulvestrant in hormone receptor–positive, human epidermal growth factor receptor 2–negative advanced breast cancer: MONALEESA-3. *J. Clin. Oncol.* **36**, 2465–2472 (2018).

[S352] Ishigami, H. *et al.* Phase III trial comparing intraperitoneal and intravenous paclitaxel plus S-1 versus cisplatin plus S-1 in patients with gastric cancer with peritoneal metastasis: PHOENIX-GC trial. *J. Clin. Oncol.* **36**, 1922–1929 (2018).

[S353] Tveit, K. M. *et al.* Phase III trial of cetuximab with continuous or intermittent fluorouracil, leucovorin, and oxaliplatin (Nordic FLOX) versus FLOX alone in first-line treatment of metastatic colorectal cancer: the NORDIC-VII study. *J. Clin. Oncol.* **30**, 1755–1762 (2012).

[S354] Baselga, J. *et al.* Phase III trial of nonpegylated liposomal doxorubicin in combination with trastuzumab and paclitaxel in HER2-positive metastatic breast cancer. *Annals Oncol.* **25**, 592–598 (2014).

[S355] Fribbens, C. *et al.* Plasma ESR1 mutations and the treatment of estrogen receptor–positive advanced breast cancer. *J. Clin. Oncol.* **34**, 2961–2968 (2016).

[S356] De Bono, J. S. *et al.* Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment: a randomised open-label trial. *The Lancet* **376**, 1147–1154 (2010).

[S357] Lin, Q. *et al.* Preoperative chemotherapy prior to primary tumour resection for asymptomatic synchronous unresectable colorectal liver-limited metastases: The RECUT multicenter randomised controlled trial. *Eur. J. Cancer* **191**, 112961 (2023).

[S358] Miles, D. *et al.* Primary results from IMpassion131, a double-blind, placebo-controlled, randomised phase III trial of first-line paclitaxel with or without atezolizumab for unresectable locally advanced/metastatic triple-negative breast cancer. *Annals Oncol.* **32**, 994–1004 (2021).

[S359] Fang, G. & Wang, X. Prognosis-related genes participate in immunotherapy of renal clear cell carcinoma possibly by targeting dendritic cells. *Front. Cell Dev. Biol.* **10**, 892616 (2022).

[S360] Trillsch, F. *et al.* Prognostic and predictive effects of primary versus secondary platinum resistance for bevacizumab treatment for platinum-resistant ovarian cancer in the AURELIA trial. *Annals Oncol.* **27**, 1733–1739 (2016).

[S361] Tejpar, S. *et al.* Prognostic and predictive relevance of primary tumor location in patients with RAS wild-type metastatic colorectal cancer: retrospective analyses of the CRYSTAL and FIRE-3 trials. *JAMA Oncol.* **3**, 194–201 (2017).

[S362] Sinicrope, F. A. *et al.* Prognostic impact of deficient DNA mismatch repair in patients with stage III colon cancer from a randomized trial of FOLFOX-based adjuvant chemotherapy. *J. Clin. Oncol.* **31**, 3664–3672 (2013).

[S363] Moretto, R. *et al.* Prognostic impact of immune-microenvironment in colorectal liver metastases resected after triplets plus a biologic agent: A pooled analysis of five prospective trials. *Eur. J. Cancer* **135**, 78–88 (2020).

[S364] Aokage, K. *et al.* Prognostic influence of epidermal growth factor receptor mutation and radiological ground glass appearance in patients with early-stage lung adenocarcinoma. *Lung Cancer* **160**, 8–16 (2021).

[S365] Zinner, R. G. *et al.* PRONOUNCE: randomized, open-label, phase III study of first-line pemetrexed+ carboplatin followed by maintenance pemetrexed versus paclitaxel+ carboplatin+ bevacizumab followed by maintenance bevacizumab in patients ith advanced nonsquamous non–small-cell lung cancer. *J. Thorac. Oncol.* **10**, 134–142 (2015).

[S366] Láng, I. *et al.* Quality of life analysis in patients with KRAS wild-type metastatic colorectal cancer treated first-line with cetuximab plus irinotecan, fluorouracil and leucovorin. *Eur. J. Cancer* **49**, 439–448 (2013).

[S367] Al-Batran, S.-E. *et al.* Quality-of-life and performance status results from the phase III RAINBOW study of ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated gastric or gastroesophageal junction adenocarcinoma. *Annals Oncol.* **27**, 673–679 (2016).

[S368] Kudo, M. *et al.* Ramucirumab after prior sorafenib in patients with advanced hepatocellular carcinoma and elevated alpha-fetoprotein: Japanese subgroup analysis of the REACH-2 trial. *J. Gastroenterol.* **55**, 627–639 (2020).

[S369] De Vita, F. *et al.* Ramucirumab and paclitaxel in patients with gastric cancer and prior trastuzumab: subgroup analysis from RAINBOW study. *Futur. Oncol.* **15**, 2723–2731 (2019).

[S370] Yoon, H. H. *et al.* Ramucirumab combined with FOLFOX as front-line therapy for advanced esophageal, gastroesophageal junction, or gastric adenocarcinoma: a randomized, double-blind, multicenter phase II trial. *Annals Oncol.* **27**, 2196–2203 (2016).

[S371] Nishio, M. *et al.* Ramucirumab or placebo plus erlotinib in EGFR-mutated, metastatic non-small-cell lung cancer: East Asian subset of RELAY. *Cancer Sci.* **111**, 4510–4525 (2020).

[S372] Petrylak, D. P. *et al.* Ramucirumab plus docetaxel versus placebo plus docetaxel in patients with locally advanced or metastatic urothelial carcinoma after platinum-based therapy (RANGE): a randomised, double-blind, phase 3 trial. *The Lancet* **390**, 2266–2277 (2017).

[S373] Nakagawa, K. *et al.* Ramucirumab plus erlotinib in patients with untreated, EGFR-mutated, advanced non-small-cell lung cancer (RELAY): a randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Oncol.* **20**, 1655–1669 (2019).

[S374] Garon, E. B. *et al.* Ramucirumab plus erlotinib versus placebo plus erlotinib in previously untreated EGFR-mutated metastatic non-small-cell lung cancer (RELAY): exploratory analysis of next-generation sequencing results. *ESMO Open* **8**, 101580 (2023).

[S375] Wilke, H. *et al.* Ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (RAINBOW): a double-blind, randomised phase 3 trial. *The Lancet Oncol.* **15**, 1224–1235 (2014).

[S376] Zhu, A. X. *et al.* Ramucirumab versus placebo as second-line treatment in patients with advanced hepatocellular carcinoma following first-line therapy with sorafenib (REACH): a randomised, double-blind, multicentre, phase 3 trial. *The Lancet Oncol.* **16**, 859–870 (2015).

[S377] Tabernero, J. *et al.* Ramucirumab versus placebo in combination with second-line FOLFIRI in patients with metastatic colorectal carcinoma that progressed during or after first-line therapy with bevacizumab, oxaliplatin, and a fluoropyrimidine (RAISE): a randomised, double-blind, multicentre, phase 3 study. *The Lancet Oncol.* **16**, 499–508 (2015).

[S378] Sakai, D. *et al.* Randomised phase II study of panitumumab plus irinotecan versus cetuximab plus irinotecan in patients with KRAS wild-type metastatic colorectal cancer refractory to fluoropyrimidine, irinotecan and oxaliplatin (WJOG 6510G). *Eur. J. Cancer* **135**, 11–21 (2020).

[S379] Patel, A. *et al.* Randomized controlled trial comparing drug-coated balloon angioplasty versus conventional balloon angioplasty for treating below-the-knee arteries in critical limb ischemia: the SINGA-PACLI trial. *Radiology* **300**, 715–724 (2021).

[S380] Smith, M. R. *et al.* Randomized controlled trial of early zoledronic acid in men with castration-sensitive prostate cancer and bone metastases: results of CALGB 90202 (alliance). *J. Clin. Oncol.* **32**, 1143–1150 (2014).

[S381] Albanell, J. *et al.* Randomized phase II study of fulvestrant plus palbociclib or placebo in endocrine-sensitive, hormone receptor-positive/HER2–advanced breast cancer: GEICAM/2014–12 (FLIPPER). *Eur. J. Cancer* **161**, 26–37 (2022).

[S382] Curigliano, G. *et al.* Randomized phase II study of sunitinib versus standard of care for patients with previously treated advanced triple-negative breast cancer. *The Breast* **22**, 650–656 (2013).

[S383] Baselga, J. *et al.* Randomized phase II study of the anti–epidermal growth factor receptor monoclonal antibody cetuximab with cisplatin versus cisplatin alone in patients with metastatic triple-negative breast cancer. *J. Clin. Oncol.* **31**, 2586–2592 (2013).

[S384] Kalinsky, K. *et al.* Randomized phase II trial of endocrine therapy with or without ribociclib after progression on cyclin-dependent kinase 4/6 inhibition in hormone receptor–positive, human epidermal growth factor receptor 2–negative metastatic breast cancer: MAINTAIN trial. *J. Clin. Oncol.* **41**, 4004–4013 (2023).

[S385] Herzog, T. J. *et al.* Randomized phase II trial of farletuzumab plus chemotherapy versus placebo plus chemotherapy in low CA-125 platinum-sensitive ovarian cancer. *Gynecol. Oncol.* **170**, 300–308 (2023).

[S386] García-Carbonero, R. *et al.* Randomized phase II trial of parsatuzumab (anti-EGFL7) or placebo in combination with FOLFOX and bevacizumab for first-line metastatic colorectal cancer. *The Oncol.* **22**, 375–e30 (2017).

[S387] Nakajima, T. E. *et al.* Randomized phase II/III study of 5-fluorouracil/l-leucovorin versus 5-fluorouracil/l-leucovorin plus paclitaxel administered to patients with severe peritoneal metastases of gastric cancer (JCOG1108/WJOG7312G). *Gastric Cancer* **23**, 677–688 (2020).

[S388] Kojima, T. *et al.* Randomized phase III KEYNOTE-181 study of pembrolizumab versus chemotherapy in advanced esophageal cancer. *J. Clin. Oncol.* **38**, 4138–4148 (2020).

[S389] Ang, K. K. *et al.* Randomized phase III trial of concurrent accelerated radiation plus cisplatin with or without cetuximab for stage III to IV head and neck carcinoma: RTOG 0522. *J. Clin. Oncol.* **32**, 2940–2950 (2014).

[S390] Langer, C. J. *et al.* Randomized, phase III trial of first-line figitumumab in combination with paclitaxel and carboplatin versus paclitaxel and carboplatin alone in patients with advanced non–small-cell lung cancer. *J. Clin. Oncol.* **32**, 2059–2066 (2014).

[S391] Sclafani, F. *et al.* RAS mutations and cetuximab in locally advanced rectal cancer: results of the EXPERT-C trial. *Eur. J. Cancer* **50**, 1430–1436 (2014).

[S392] Aix, S. P. *et al.* RELAY, ramucirumab plus erlotinib versus placebo plus erlotinib in patients with untreated, EGFR-mutated, metastatic non-small cell lung cancer: Europe/United States subset analysis. *Cancer Treat. Res. Commun.* **27**, 100378 (2021).

[S393] Holch, J. *et al.* Relevance of baseline carcinoembryonic antigen for first-line treatment against metastatic colorectal cancer with FOLFIRI plus cetuximab or bevacizumab (FIRE-3 trial). *Eur. J. Cancer* **106**, 115–125 (2019).

[S394] Segelov, E. *et al.* Response to cetuximab with or without irinotecan in patients with refractory metastatic colorectal cancer harboring the KRAS G13D mutation: Australasian Gastro-Intestinal Trials Group ICECREAM study. *J. Clin. Oncol.* **34**, 2258–2264 (2016).

[S395] Di Leo, A. *et al.* Results of the CONFIRM phase III trial comparing fulvestrant 250 mg with fulvestrant 500 mg in postmenopausal women with estrogen receptor–positive advanced breast cancer. *J. Clin. Oncol.* **28**, 4594–4600 (2010).

[S396] Okunaka, M. *et al.* Retrospective cohort study of nanoparticle albumin-bound paclitaxel plus ramucirumab versus paclitaxel plus ramucirumab as second-line treatment in patients with advanced gastric cancer. *BMC Cancer* **20**, 1111 (2020).

[S397] Slamon, D. *et al.* Ribociclib plus fulvestrant for postmenopausal women with hormone receptor-positive, human epidermal growth factor receptor 2-negative advanced breast cancer in the phase III randomized MONALEESA-3 trial: updated overall survival. *Annals Oncol.* **32**, 1015–1024 (2021).

[S398] O'shaughnessy, J. *et al.* Ribociclib plus letrozole versus letrozole alone in patients with de novo HR+, HER2- advanced breast cancer in the randomized MONALEESA-2 trial. *Breast Cancer Res. Treat.* **168**, 127–134 (2018).

[S399] Wang, Y. *et al.* Safety and feasibility of toripalimab plus lenvatinib with or without radiotherapy in advanced BTC. *Front. Immunol.* **14**, 1084843 (2023).

[S400] Wang, F. *et al.* Safety, efficacy and tumor mutational burden as a biomarker of overall survival benefit in chemo-refractory gastric cancer treated with toripalimab, a PD-1 antibody in phase Ib/II clinical trial NCT02915432. *Annals Oncol.* **30**, 1479–1486 (2019).

[S401] Shi, T. *et al.* Secondary cytoreduction followed by chemotherapy versus chemotherapy alone in platinum-sensitive relapsed ovarian cancer (SOC-1): a multicentre, open-label, randomised, phase 3 trial. *The Lancet Oncol.* **22**, 439–449 (2021).

[S402] Huemer, F. *et al.* Sidedness and TP53 mutations impact OS in anti-EGFR but not anti-VEGF treated mCRC-an analysis of the KRAS registry of the AGMT (Arbeitsgemeinschaft Medikamentöse Tumortherapie). *BMC Cancer* **18**, 11 (2018).

[S403] Huber, R. M. *et al.* Simultaneous chemoradiotherapy compared with radiotherapy alone after induction chemotherapy in inoperable stage IIIA or IIIB non–small-cell lung cancer: study CTRT99/97 by the bronchial carcinoma therapy group. *J. Clin. Oncol.* **24**, 4397–4404 (2006).

[S404] Stahler, A. *et al.* Single-nucleotide variants, tumour mutational burden and microsatellite instability in patients with metastatic colorectal cancer: Next-generation sequencing results of the FIRE-3 trial. *Eur. J. Cancer* **137**, 250–259 (2020).

[S405] Lu, Z. *et al.* Sintilimab versus placebo in combination with chemotherapy as first line treatment for locally advanced or metastatic oesophageal squamous cell carcinoma (ORIENT-15): multicentre, randomised, double blind, phase 3 trial. *BMJ* **377** (2022).

[S406] Shitara, K. *et al.* Subgroup analyses of the safety and efficacy of ramucirumab in japanese and western patients in RAINBOW: a randomized clinical trial in second-line treatment of gastric cancer. *Gastric Cancer* **19**, 927–938 (2016).

[S407] Obermannová, R. *et al.* Subgroup analysis in RAISE: a randomized, double-blind phase III study of irinotecan, folinic acid, and 5-fluorouracil (FOLFIRI) plus ramucirumab or placebo in patients with metastatic colorectal carcinoma progression. *Annals Oncol.* **27**, 2082–2089 (2016).

[S408] Chung, H. C. *et al.* Subgroup analysis of East Asian patients in REGARD: A phase III trial of ramucirumab and best supportive care for advanced gastric cancer. *Asia-Pacific J. Clin. Oncol.* **14**, 204–209 (2018).

[S409] Fernandez-Martinez, A. *et al.* Survival, pathologic response, and genomics in CALGB 40601 (Alliance), a neoadjuvant phase III trial of paclitaxel-trastuzumab with or without lapatinib in HER2-positive breast cancer. *J. Clin. Oncol.* **38**, 4184–4193 (2020).

[S410] Bidard, F.-C. *et al.* Switch to fulvestrant and palbociclib versus no switch in advanced breast cancer with rising ESR1 mutation during aromatase inhibitor and palbociclib therapy (PADA-1): a randomised, open-label, multicentre, phase 3 trial. *The Lancet Oncol.* **23**, 1367–1377 (2022).

[S411] Pfeiffer, P. *et al.* TAS-102 with or without bevacizumab in patients with chemorefractory metastatic colorectal cancer: an investigator-initiated, open-label, randomised, phase 2 trial. *The Lancet Oncol.* **21**, 412–420 (2020).

[S412] Takashima, T. *et al.* Taxanes versus S-1 as the first-line chemotherapy for metastatic breast cancer (SELECT BC): an open-label, non-inferiority, randomised phase 3 trial. *The Lancet Oncol.* **17**, 90–98 (2016).

[S413] Tesileanu, C. M. S. *et al.* Temozolomide and radiotherapy versus radiotherapy alone in patients with glioblastoma, IDH-wildtype: post hoc analysis of the EORTC randomized phase III CATNON trial. *Clin. Cancer Res.* **28**, 2527–2535 (2022).

[S414] Möbus, V. *et al.* Ten-year results of intense dose-dense chemotherapy show superior survival compared with a conventional schedule in high-risk primary breast cancer: final results of AGO phase III iddEPC trial. *Annals Oncol.* **29**, 178–185 (2018).

[S415] Quoix, E. *et al.* TG4010 immunotherapy and first-line chemotherapy for advanced non-small-cell lung cancer (TIME): results from the phase 2b part of a randomised, double-blind, placebo-controlled, phase 2b/3 trial. *The Lancet Oncol.* **17**, 212–223 (2016).

[S416] Finn, R. S. *et al.* The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. *The Lancet Oncol.* **16**, 25–35 (2015).

[S417] Trullas, A. *et al.* The EMA assessment of encorafenib in combination with cetuximab for the treatment of adult patients with metastatic colorectal carcinoma harbouring the BRAFV600E mutation who have received prior therapy. *ESMO Open* **6**, 100031 (2021).

[S418] Alarfi, H. *et al.* The impact of baseline body mass index on clinical outcomes in metastatic breast cancer: a prospective study. *BMC Res. Notes* **10**, 550 (2017).

[S419] Weidhaas, J. B. *et al.* The inherited KRAS-variant as a biomarker of cetuximab response in NSCLC. *Cancer Res. Commun.* **3**, 2074–2081 (2023).

[S420] Okada, M. *et al.* Three-year follow-up and response–survival relationship of nivolumab in previously treated patients with advanced esophageal squamous cell carcinoma (ATTRACTION-3). *Clin. Cancer Res.* **28**, 3277–3286 (2022).

[S421] Hassel, J. C. *et al.* Three-year overall survival with tebentafusp in metastatic uveal melanoma. *New Engl. J. Medicine* **389**, 2256–2266 (2023).

[S422] Shen, L. *et al.* Tislelizumab versus chemotherapy as second-line treatment for advanced or metastatic esophageal squamous cell carcinoma (RATIONALE-302): a randomized phase III study. *J. Clin. Oncol.* **40**, 3065–3076 (2022).

[S423] Lian, B. *et al.* Toripalimab (anti-PD-1) versus high-dose interferon-α2b as adjuvant therapy in resected mucosal melanoma: a phase II randomized trial. *Annals Oncol.* **33**, 1061–1070 (2022).

[S424] Xu, Y.-J. *et al.* Toripalimab combined with hepatic arterial infusion chemotherapy versus lenvatinib for advanced hepatocellular carcinoma. *Technol. Cancer Res. & Treat.* **20**, 15330338211063848 (2021).

[S425] Mai, H.-Q. *et al.* Toripalimab or placebo plus chemotherapy as first-line treatment in advanced nasopharyngeal carcinoma: a multicenter randomized phase 3 trial. *Nat. Medicine* **27**, 1536–1543 (2021).

[S426] Yan, X. *et al.* Toripalimab plus axitinib versus sunitinib as first-line treatment for advanced renal cell carcinoma: RENOTORCH, a randomized, open-label, phase III study. *Annals Oncol.* **35**, 190–199 (2024).

[S427] Jiang, Z. *et al.* Toripalimab plus nab-paclitaxel in metastatic or recurrent triple-negative breast cancer: a randomized phase 3 trial. *Nat. Medicine* **30**, 249–256 (2024).

[S428] Gershenson, D. M. *et al.* Trametinib versus standard of care in patients with recurrent low-grade serous ovarian cancer (GOG 281/LOGS): an international, randomised, open-label, multicentre, phase 2/3 trial. *The Lancet* **399**, 541–553 (2022).

[S429] Paz-Ares, L. G. *et al.* Treatment outcomes by histology in REVEL: a randomized phase III trial of ramucirumab plus docetaxel for advanced non-small cell lung cancer. *Lung Cancer* **112**, 126–133 (2017).

[S430] Goodwin, R. *et al.* Treatment-emergent hypertension and outcomes in patients with advanced non-small-cell lung cancer receiving chemotherapy with or without the vascular endothelial growth factor receptor inhibitor cediranib: NCIC Clinical Trials Group Study BR24. *Annals Oncol.* **21**, 2220–2226 (2010).

[S431] Shitara, K. *et al.* Trifluridine/tipiracil versus placebo in patients with heavily pretreated metastatic gastric cancer (TAGS): a randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Oncol.* **19**, 1437–1448 (2018).

[S432] Cascinu, S. *et al.* Tumor response and symptom palliation from RAINBOW, a phase III trial of ramucirumab plus paclitaxel in previously treated advanced gastric cancer. *The Oncol.* **26**, e414–e424 (2021).

[S433] Leal, T. *et al.* Tumor treating fields therapy with standard systemic therapy versus standard systemic therapy alone in metastatic non-small-cell lung cancer following progression on or after platinum-based therapy (LUNAR): a randomised, open-label, pivotal phase 3 study. *The Lancet Oncol.* **24**, 1002–1017 (2023).

[S434] Neven, P. *et al.* Updated overall survival from the MONALEESA-3 trial in postmenopausal women with HR+/HER2- advanced breast cancer receiving first-line ribociclib plus fulvestrant. *Breast Cancer Res.* **25**, 103 (2023).

[S435] Hortobagyi, G. N. *et al.* Updated results from MONALEESA-2, a phase III trial of first-line ribociclib plus letrozole versus placebo plus letrozole in hormone receptor-positive, HER2-negative advanced breast cancer. *Annals Oncol.* **29**, 1541–1547 (2018).

[S436] Diéras, V. *et al.* Veliparib with carboplatin and paclitaxel in BRCA-mutated advanced breast cancer (BROCADE3): a randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Oncol.* **21**, 1269–1282 (2020).

[S437] Mizuno, M. *et al.* Veliparib with frontline chemotherapy and as maintenance in Japanese women with ovarian cancer: a subanalysis of efficacy, safety, and antiemetic use in the phase 3 VELIA trial. *Int. J. Clin. Oncol.* **28**, 163–174 (2023).

[S438] Lindeman, G. J. *et al.* VERONICA: randomized phase II study of fulvestrant and venetoclax in ER-positive metastatic breast cancer post-CDK4/6 inhibitors–efficacy, safety, and biomarker results. *Clin. Cancer Res.* **28**, 3256–3267 (2022).

[S439] Cortot, A. B. *et al.* Weekly paclitaxel plus bevacizumab versus docetaxel as second-or third-line treatment in advanced non-squamous non–small-cell lung cancer: Results of the IFCT-1103 ULTIMATE study. *Eur. J. Cancer* **131**, 27–36 (2020).

[S440] Ngoi, N. Y. *et al.* Weekly versus tri-weekly paclitaxel with carboplatin for first-line treatment in women with epithelial ovarian cancer. *Cochrane Database Syst. Rev.* (2022).

[S441] Spigel, D. R. *et al.* An open-label, randomized, controlled phase II study of paclitaxel-carboplatin chemotherapy with necitumumab versus paclitaxel-carboplatin alone in first-line treatment of patients with stage IV squamous non–small-cell lung cancer. *Clin. Lung Cancer* **18**, 480–488 (2017).