Supplementary Material

1 SUPPLEMENTARY TABLES AND FIGURES

Fixed-effects Model	Random-effects Model
$\beta_k = \beta + e_k$	$\beta_k = \beta + \eta_k + e_k$
$e_k \sim N(0, \sigma_k^2)$	$\eta_k \sim N(0, \tau^2)$
$\beta_k \sim N(\beta, \sigma_k^2)$	$e_k \sim N(0, s_k^2)$
$w_k = \frac{1}{\sigma_k^2}$	$\beta_k \vert \eta_k \sim N(\mu, \tau^2 + s_k^2)$
$\hat{\sigma}_k^2 = s_k^2$ (large L_k)	$w_k = \frac{1}{\tau^2 + s_k^2}$
	η_k and e_k are independent

Table S1. Models to incorporate study results in a meta-analysis.

Factor (number of conditions)	Simulation conditions
number of studies (3)	$K = 3, 7, 10$
Within-study variance (3)	$\sigma^2 = 0.1, 0.45, 1$
Between-studies variances (4)	$\tau^2 = 0.001, 0.2, 0.5, 0.8$
Correlation between study random effects (4)	$\rho_{12} = 0.1, 0.7 \quad \rho_{23} = 0.2 \quad \rho_{13} = 0.6$
Overall effects (3)	$\beta = 0.2, 0.5, 0.8$

Table S2. Summary of the simulated factor conditions. The within-study variance for each study is assumed to be equal.

SUPPLEMENTARY FIGURES

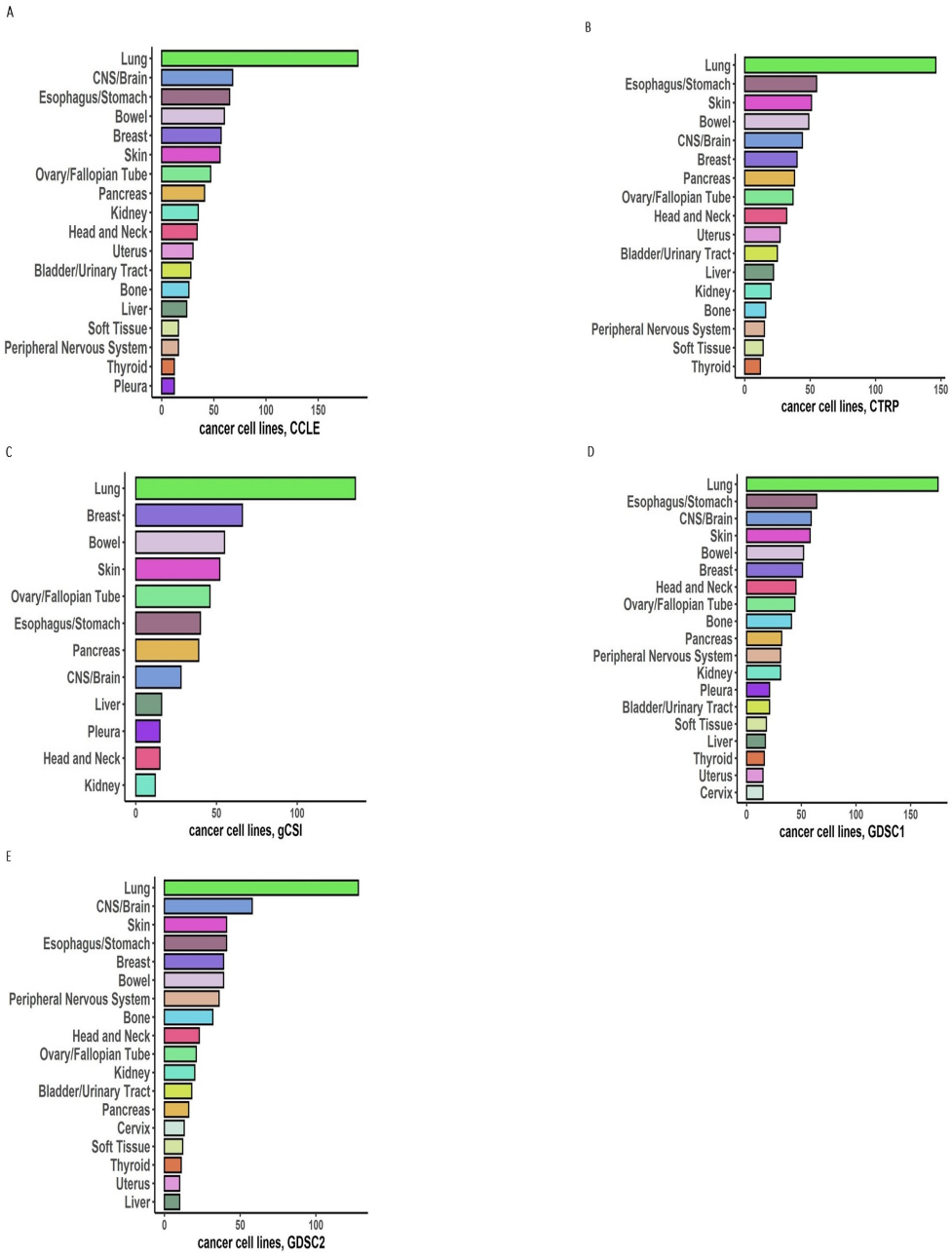


Figure S1. Pan-cancer cell line data. Bar plots illustrate distribution of cell lines across studies and tissue types.

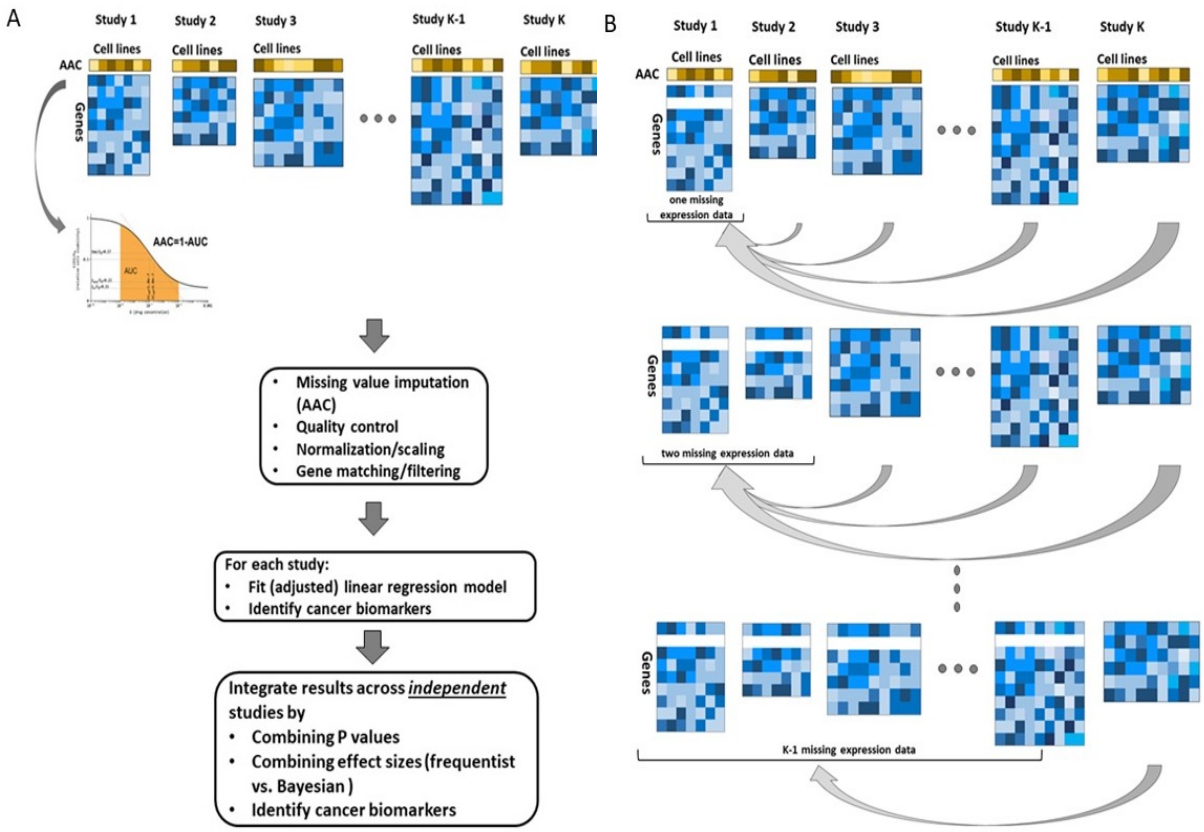


Figure S2. Flow diagram of meta-analysis. Integration analyses using (**A**) independent studies. (**B**) Duplicate partial or full missing expression data.

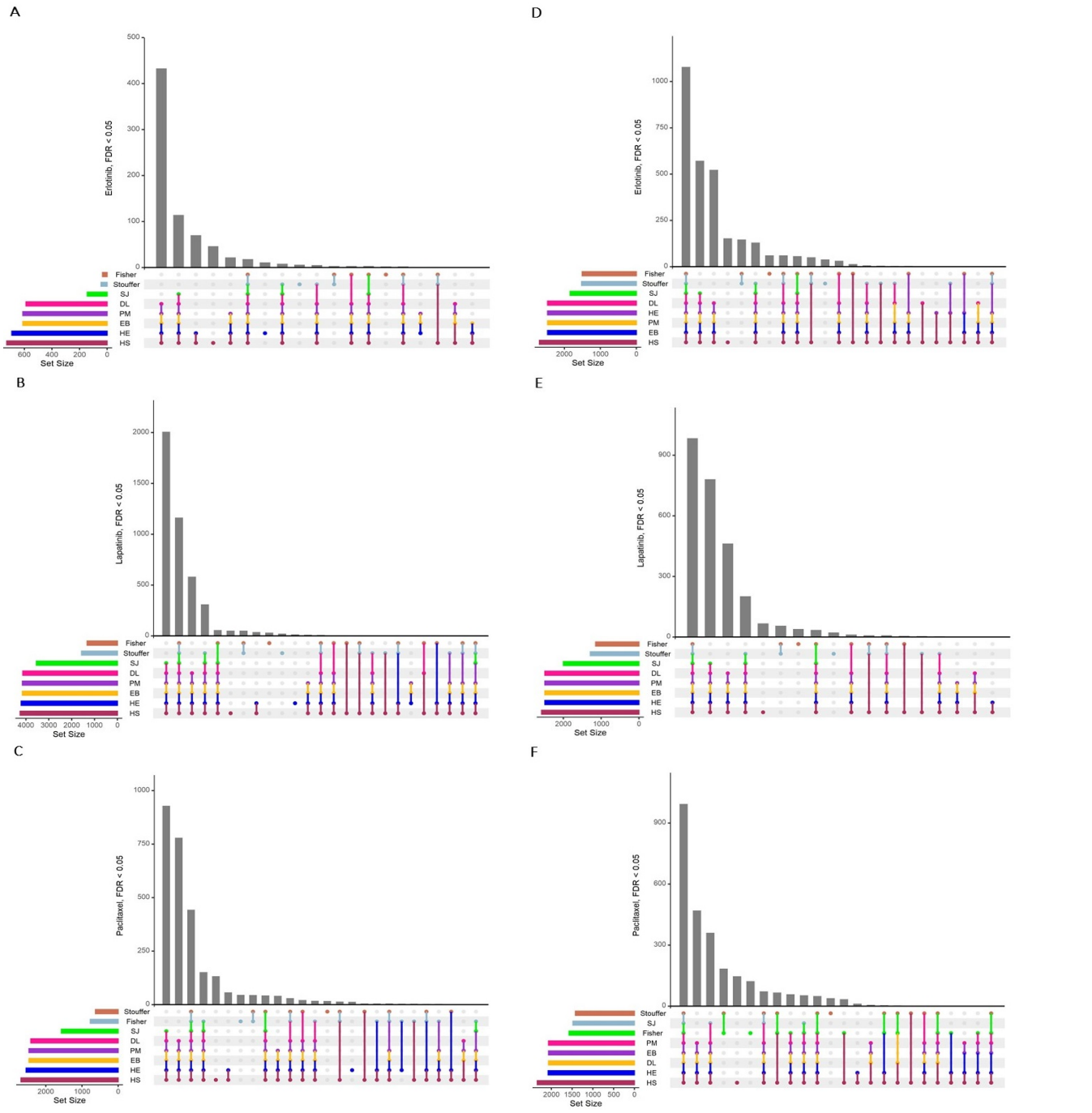


Figure S3. Comparison of different meta-analysis approaches. Upset diagrams represent the number of statistically significant genes associated with drug response (FDR < 0.05) using various meta-analysis methods for (**A-C**) breast and (**D-F**) pan-cancer data.

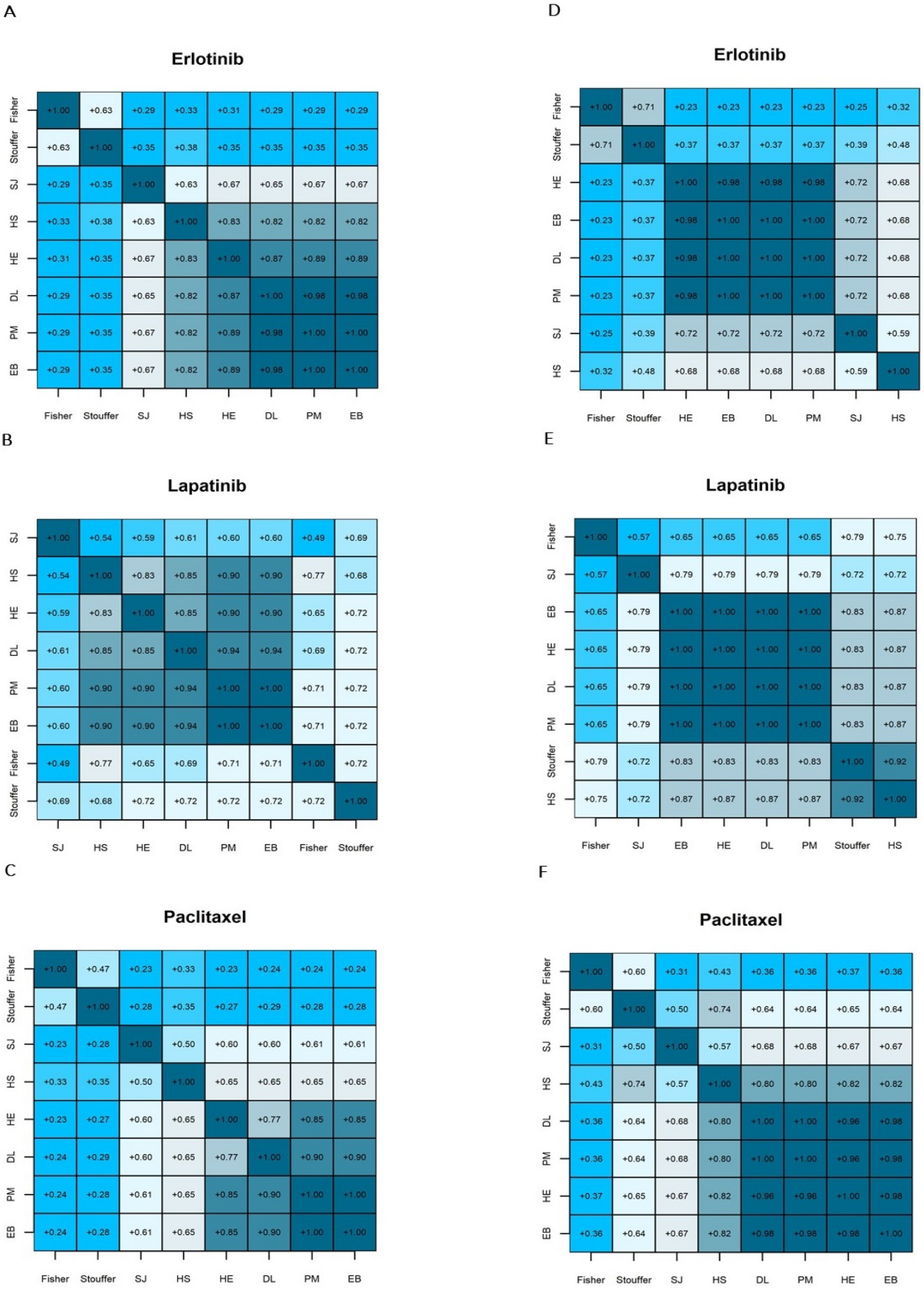
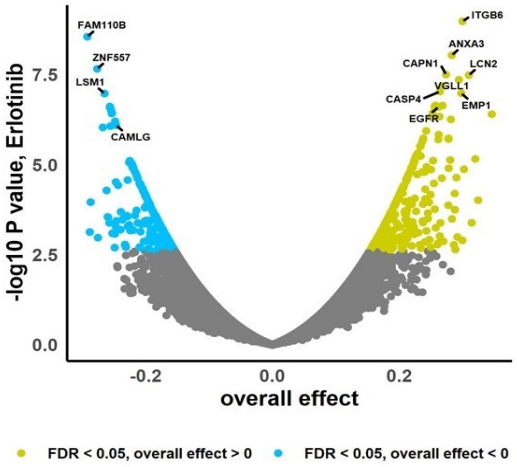


Figure S4. Stability metric to compare different meta-analysis approaches. Stability of identified 100 top-ranked genes associated with drugs by applying Jaccard similarity index using (**A-C**) breast and (**D-F**) pan-cancer data.



Erlotinib/EGFR	Effect	SE	95% CI Lower	95% CI Upper	P value
GDSC2, n = 39	0.09	0.1638	-0.23	0.41	5.92e-01
UHNBreast, n = 80	0.15	0.1120	-0.07	0.36	1.94e-01
CTRP, n = 40	0.17	0.1597	-0.14	0.49	2.76e-01
GDSC1, n = 51	0.22	0.1394	-0.05	0.49	1.14e-01
gCSI, n = 66	0.25	0.1210	0.01	0.49	3.72e-02
CCLE, n = 57	0.42	0.1223	0.18	0.66	5.81e-04
GRAY, n = 56	0.43	0.1225	0.19	0.68	3.86e-04

Total (fixed effect, 95% CI)
Random effect
Heterogeneity: Tau^2 = 0.0009; Chi^2 = 6.30, df = 6 (P = 3.91e-01); I^2 = 5%
Test for overall effect (fixed effect): Z = 5.30 (P = 1.14e-07)
Test for overall effect (random effects): Z = 5.16 (P = 2.47e-07)

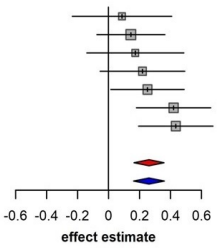


Figure S5. Breast cancer independent meta-analyses not limited to common cell lines across studies. Volcano plots show genes associated with drug response using RE meta-analysis model and forest plots illustrate overall effect estimate using FE (red diamond) and RE (blue diamond) meta-analysis models. DL approach was applied to estimate heterogeneity across studies. pan-cancer data.

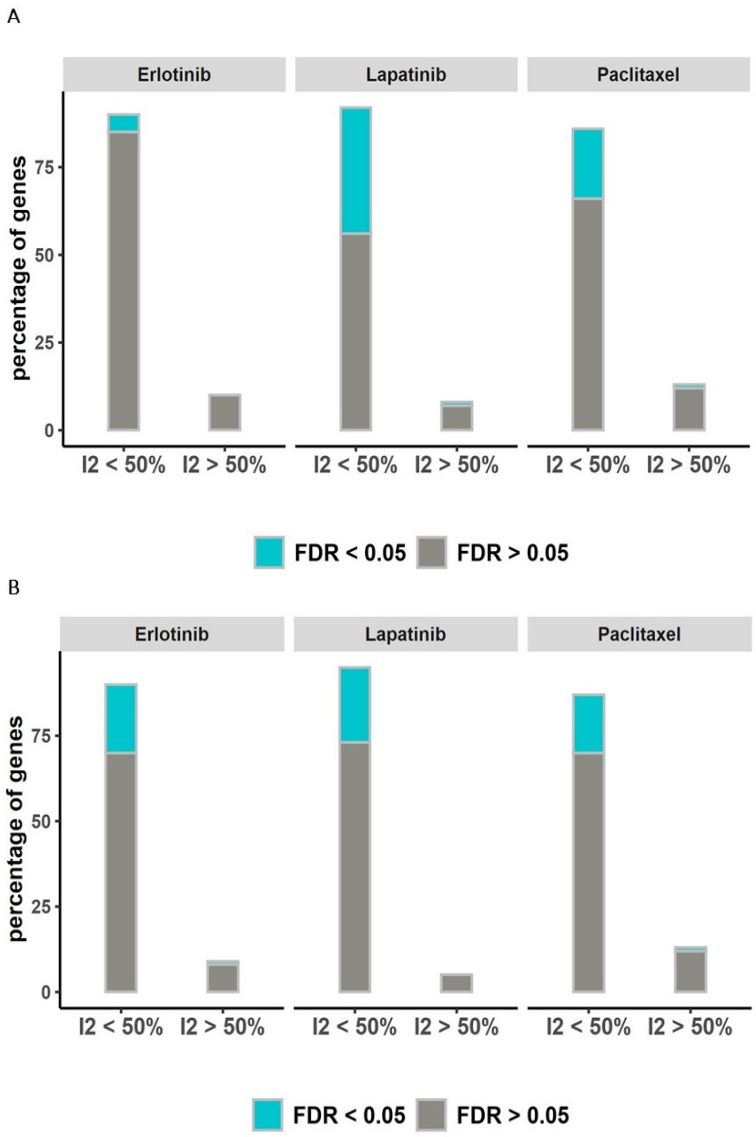


Figure S6. Gene-drug association meta-analyses and estimated heterogeneity. Bar plots present the percentage of gene-drug association meta-analyses using RE meta-analysis across estimated heterogeneity using (**A**) breast and (**B**) pan-cancer data.

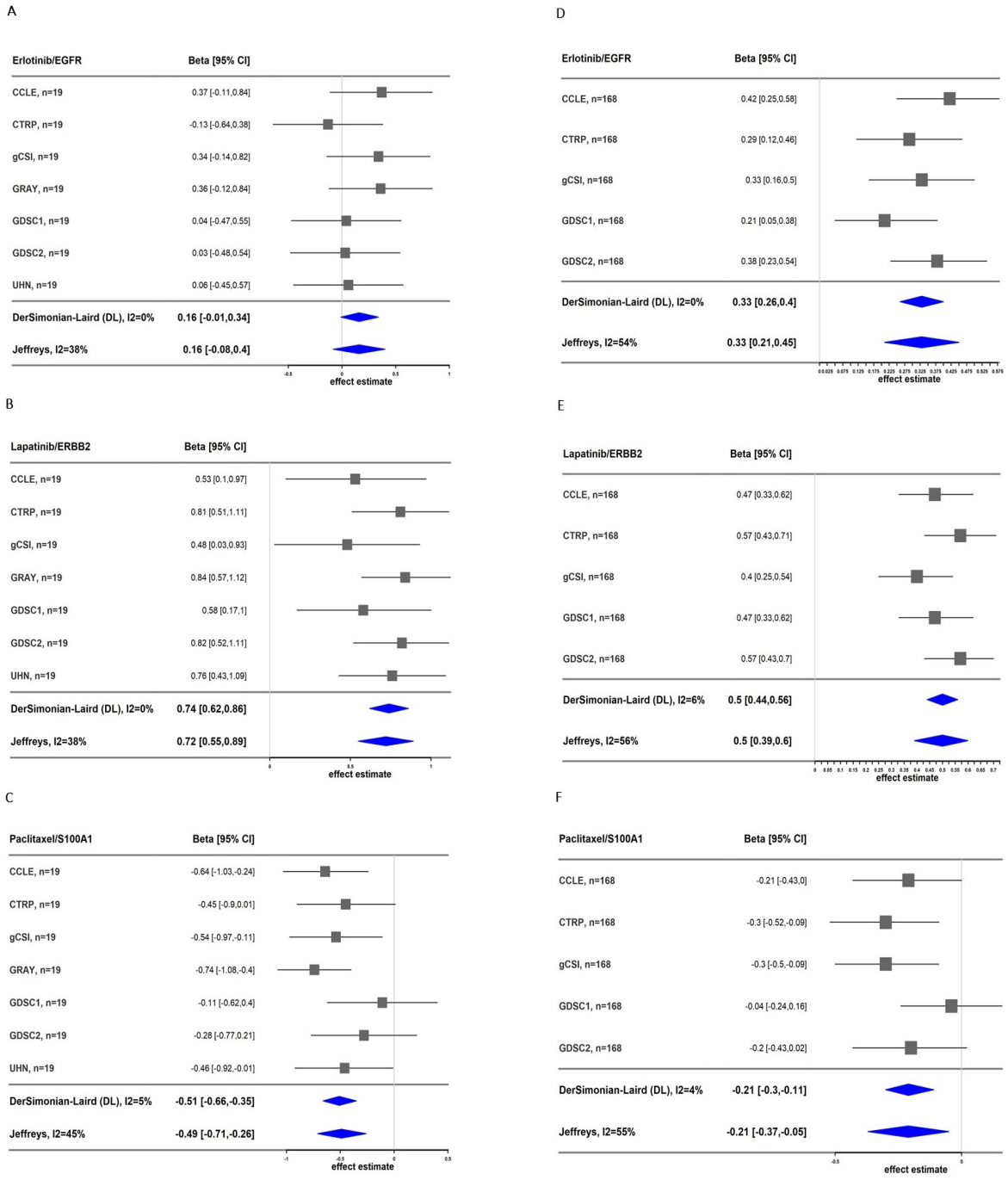


Figure S7. Bayesian versus classical independent meta-analyses. Forest plots illustrate overall effect estimate and 95% confidence and credible intervals along with heterogeneity estimate I^2 using DL and Bayesian Jeffreys procedures using (**A-C**) breast and (**D-F**) pan-cancer data.

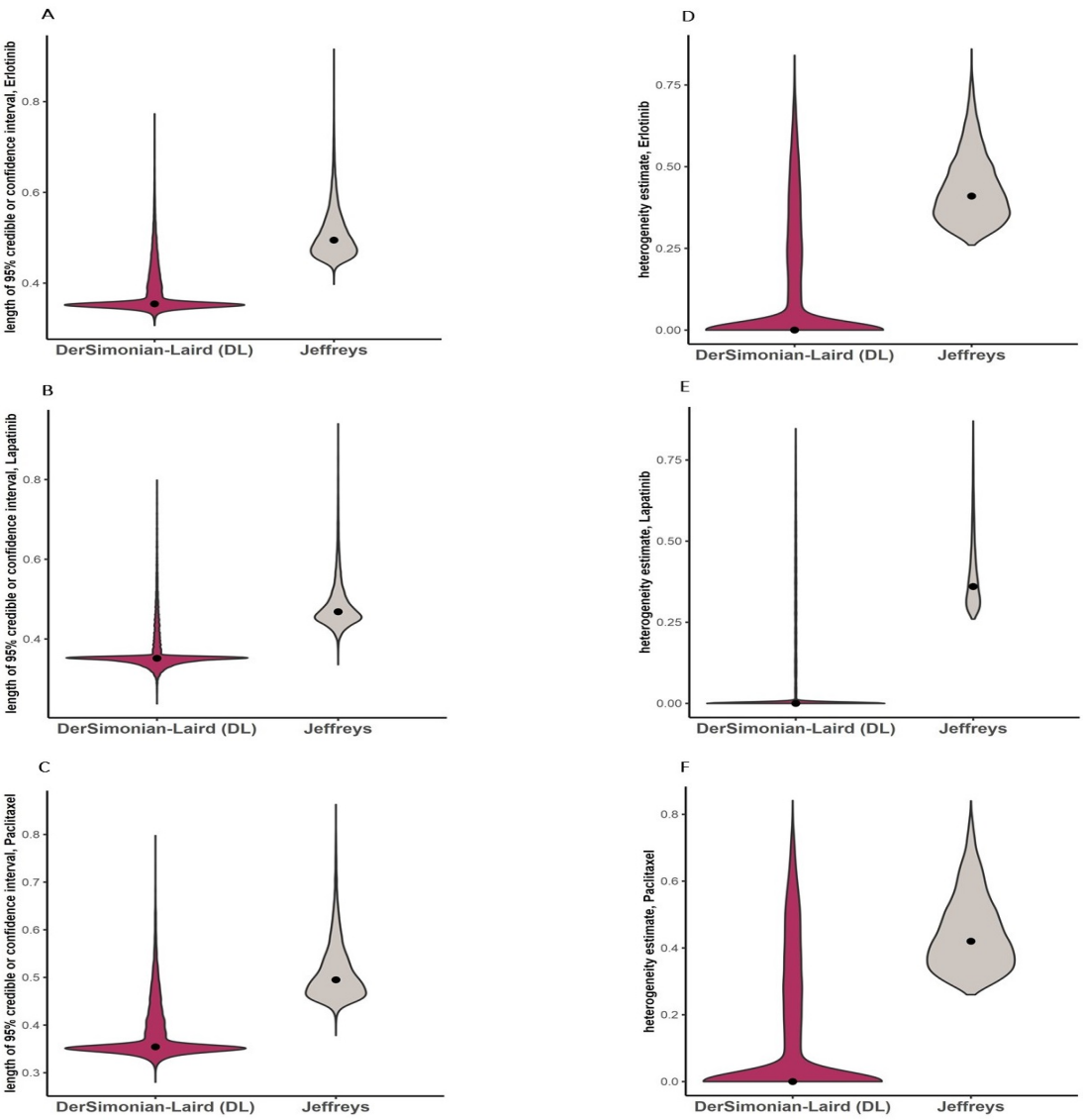


Figure S8. Bayesian versus classical independent to assess gene-drug association using breast cancer data. Violin plots illustrate (**A-C**) length of 95% confidence and credible intervals and (**D-F**) heterogeneity estimate I^2 using DL and Bayesian Jeffreys procedures. Black dot represents the median.

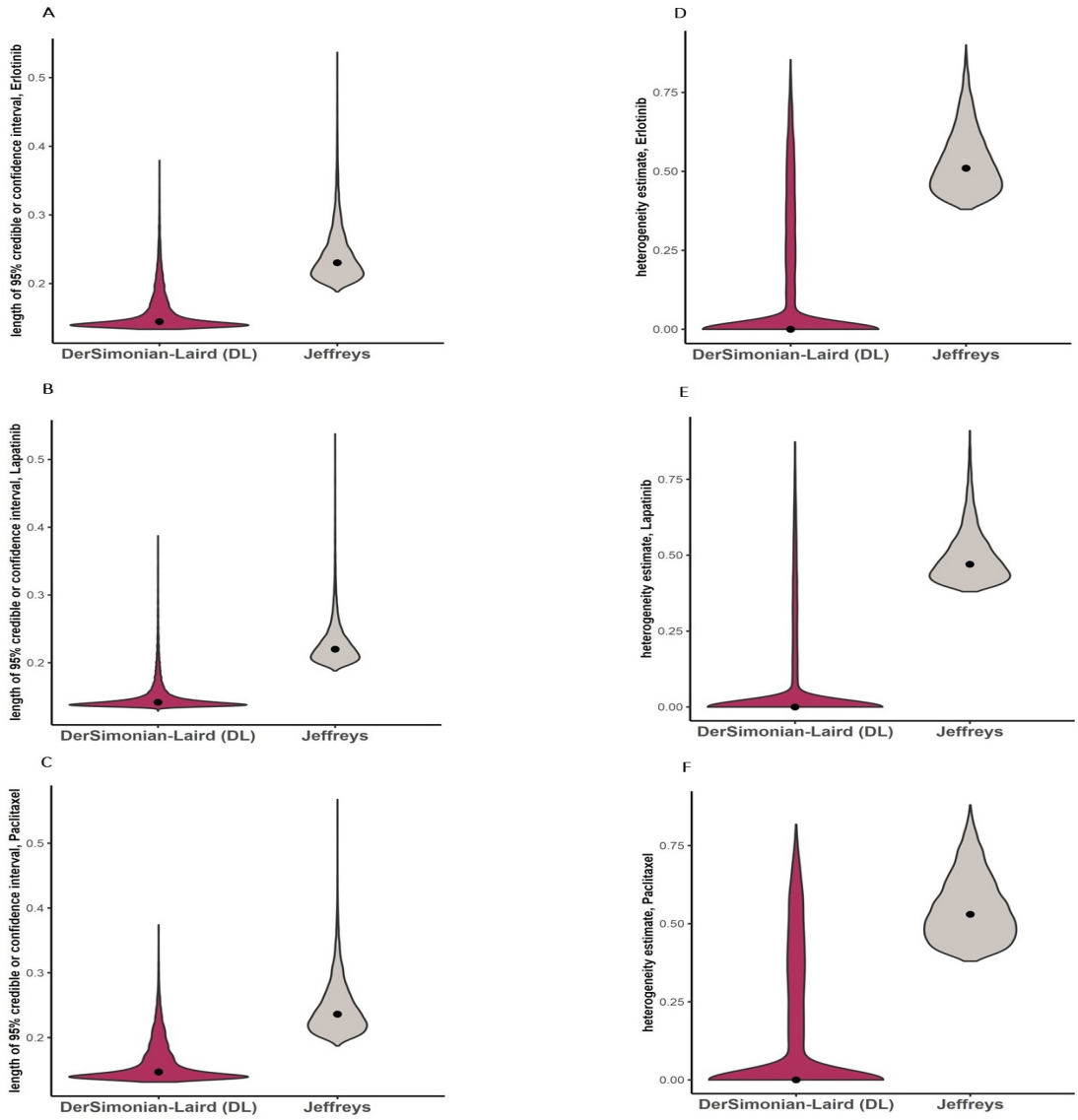


Figure S9. Bayesian versus classical independent to assess gene-drug association using pan-cancer data. Violin plots illustrate (**A-C**) length of 95% confidence and credible intervals and (**D-F**) heterogeneity estimate I^2 using DL and Bayesian Jeffreys procedures. Black dot represents the median.

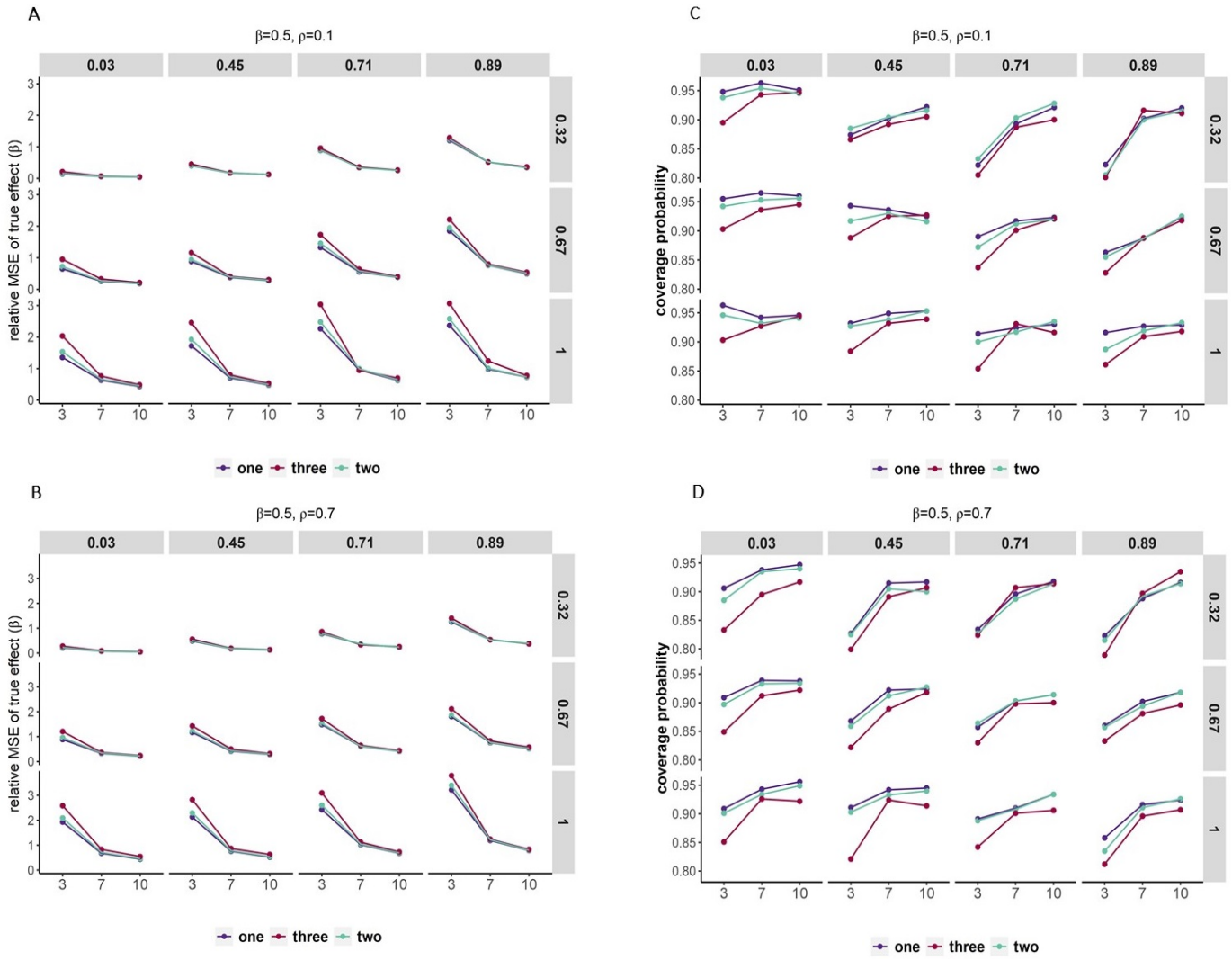


Figure S10. Mean squared error and coverage probability of overall effect estimates. Scenarios containing various within-study variances (row) and heterogeneity across studies (column). The x-axis represents the number of studies and the y-axis shows the relative MSE (**A-B**) and coverage probability of 95% confidence intervals (**C-D**). Set overall effect β = 0.5. Different colors represent a number of duplication or non-independent effects across studies.

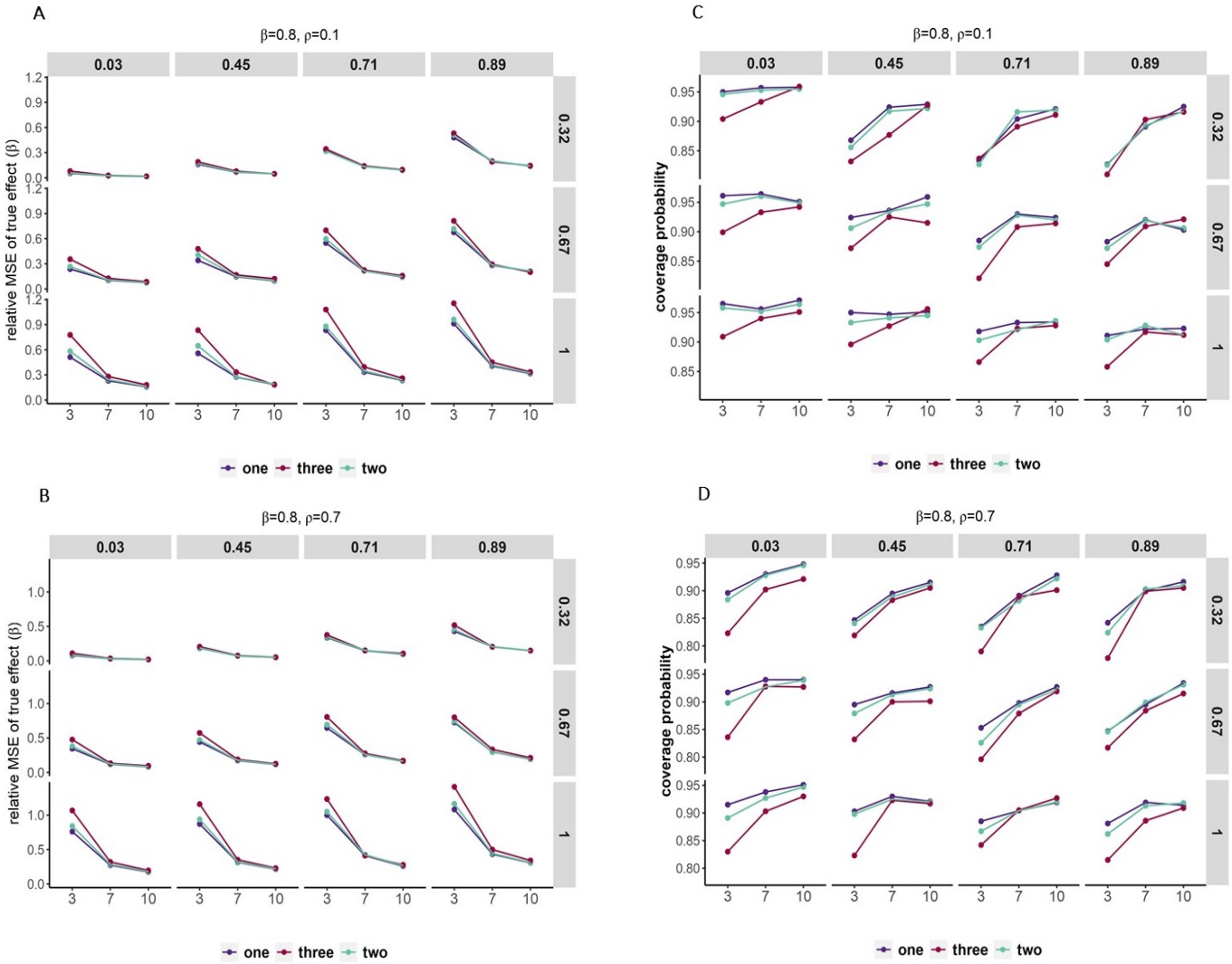


Figure S11. Mean squared error and coverage probability of overall effect estimates. Scenarios containing various within-study variances (row) and heterogeneity across studies (column). The x-axis represents the number of studies and the y-axis shows the relative MSE (**A-B**) and coverage probability of 95% confidence intervals (**C-D**). Set overall effect β = 0.8. Different colors represent a number of duplication or non-independent effects across studies.

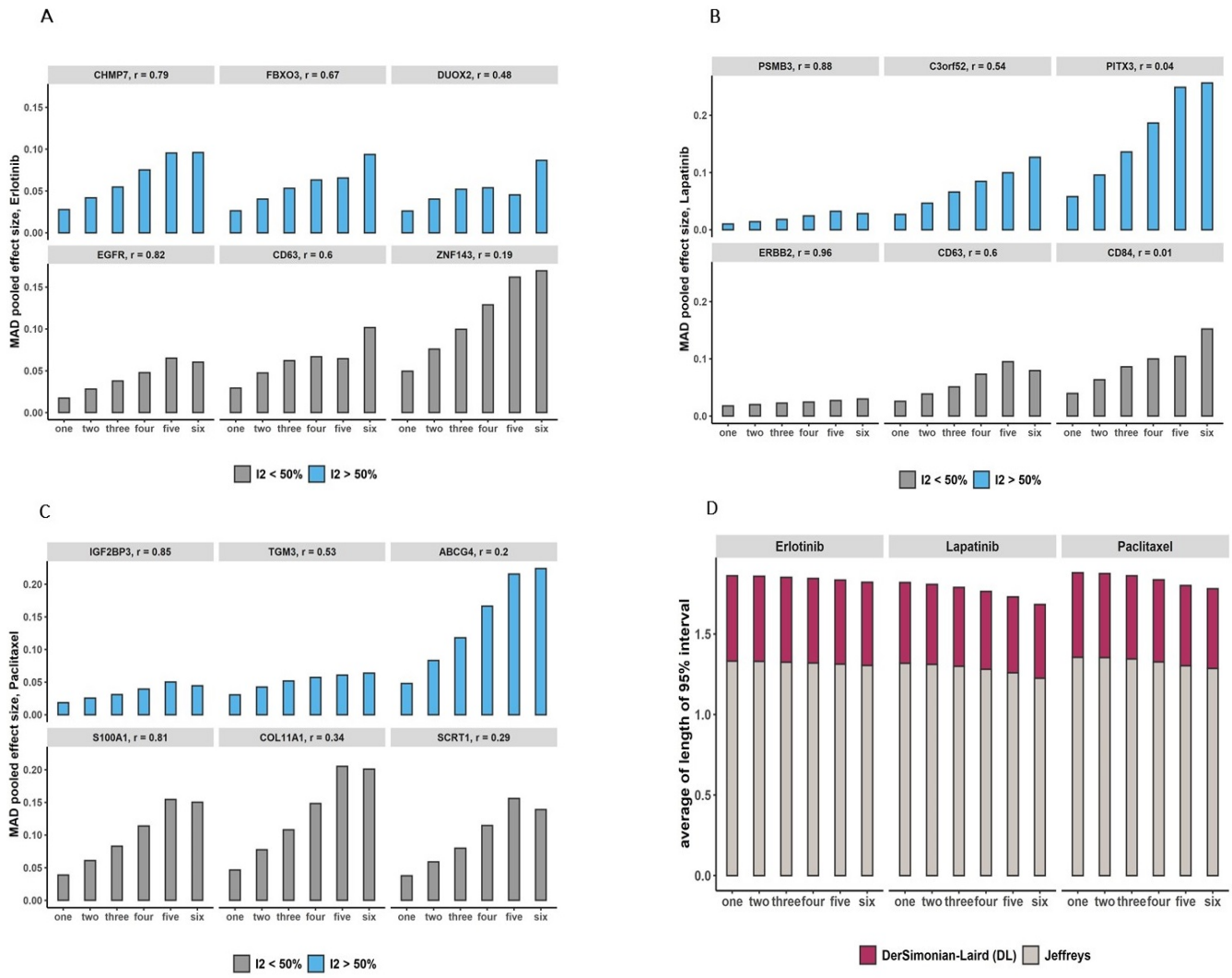


Figure S12. Breast cancer non-independent Jeffreys Bayesian meta-analyses. (**A-C**) Bar plots present increases in the number of duplications and its impact on the estimated overall effect using MAD metric across drugs and selected genes with substantial (blue) and non-substantial (gray) heterogeneity estimation. Note that x-axis presents the number of duplicate study effects. (**D**) Bar plots illustrate an average of 95% confidence or credible intervals for specific genes over the number of duplications using DL and Bayesian Jeffreys procedures.

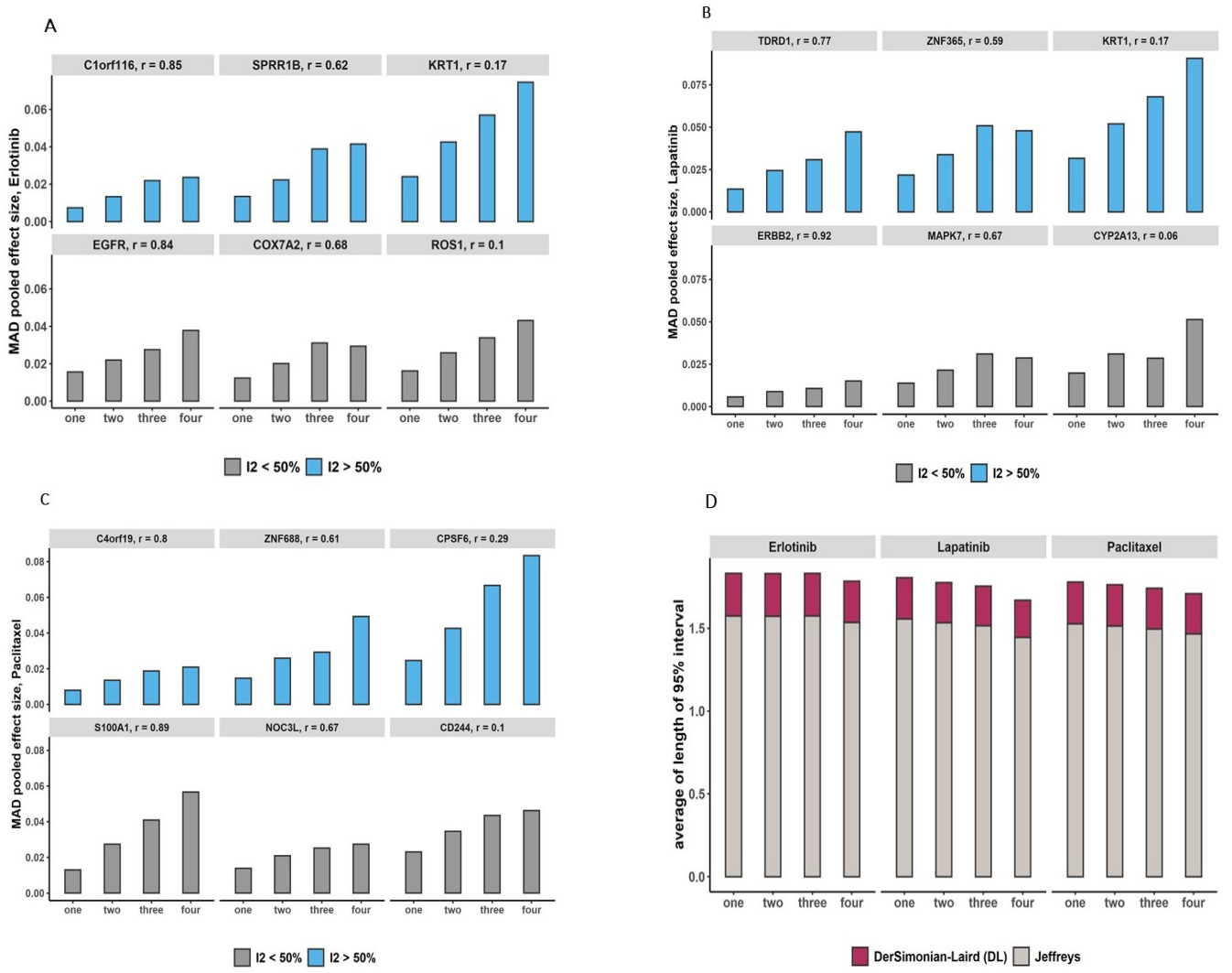


Figure S13. Pan-cancer non-independent Jeffreys Bayesian meta-analyses. (**A-C**) Bar plots present increases in the number of duplications and its impact on the estimated overall effect using MAD metric across drugs and selected genes with substantial (blue) and non-substantial (gray) heterogeneity estimation. Note that x-axis presents the number of duplicate study effects. (**D**) Bar plots illustrate an average of 95% confidence or credible intervals for specific genes over the number of duplications using DL and Bayesian Jeffreys procedures.