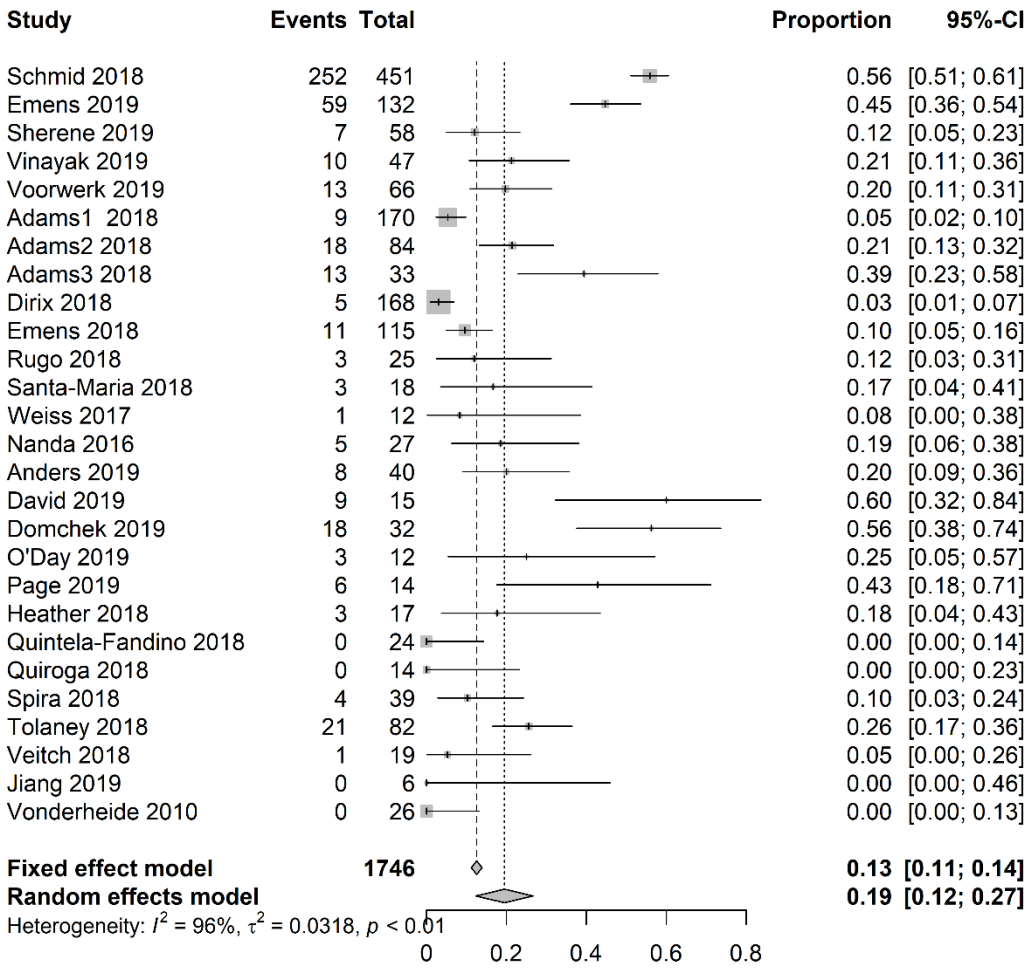
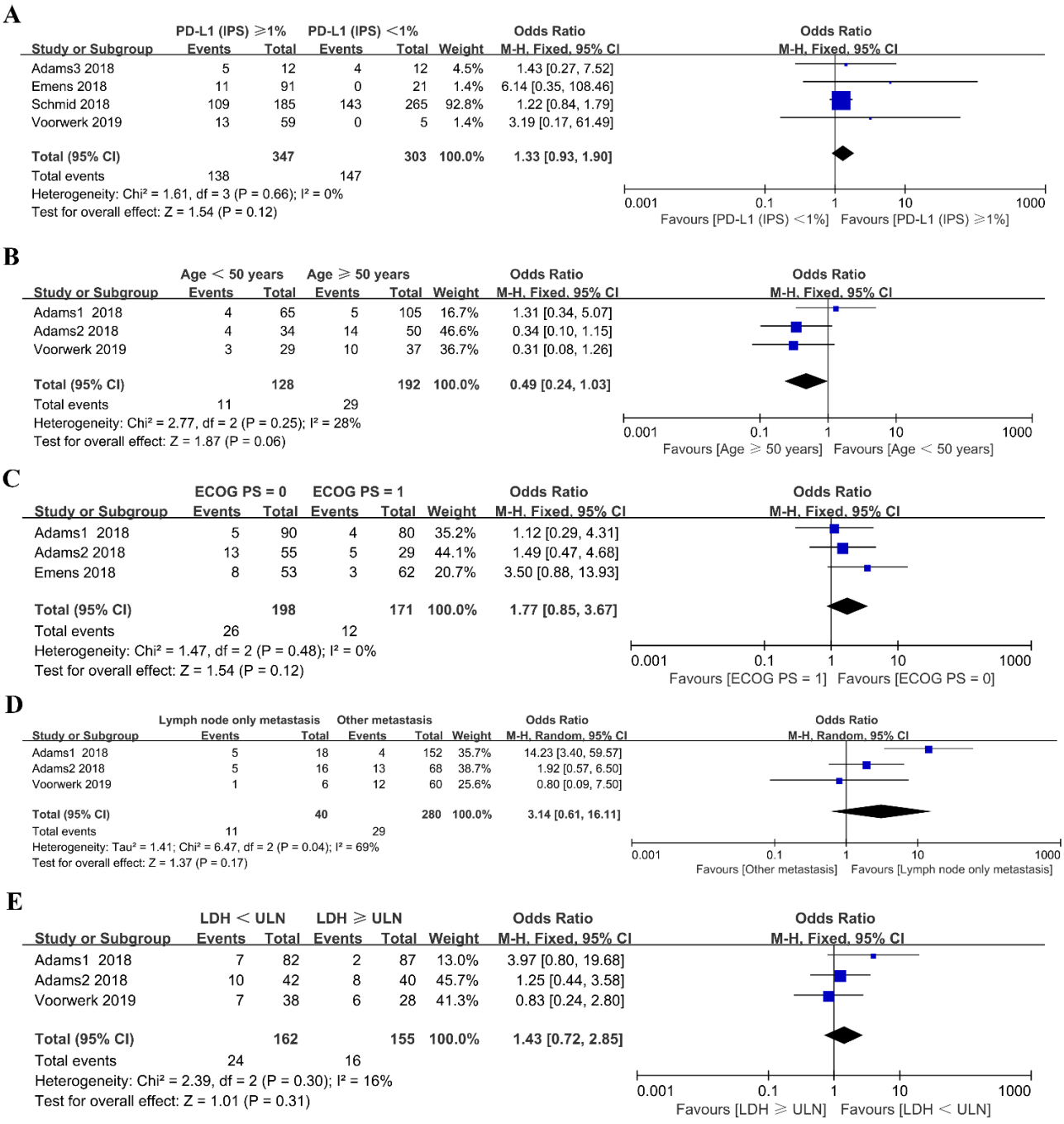
Supplementary Materials



Supplemental Figure 1. Forest plot of ORR of metastatic breast cancer treated with ICI therapy.

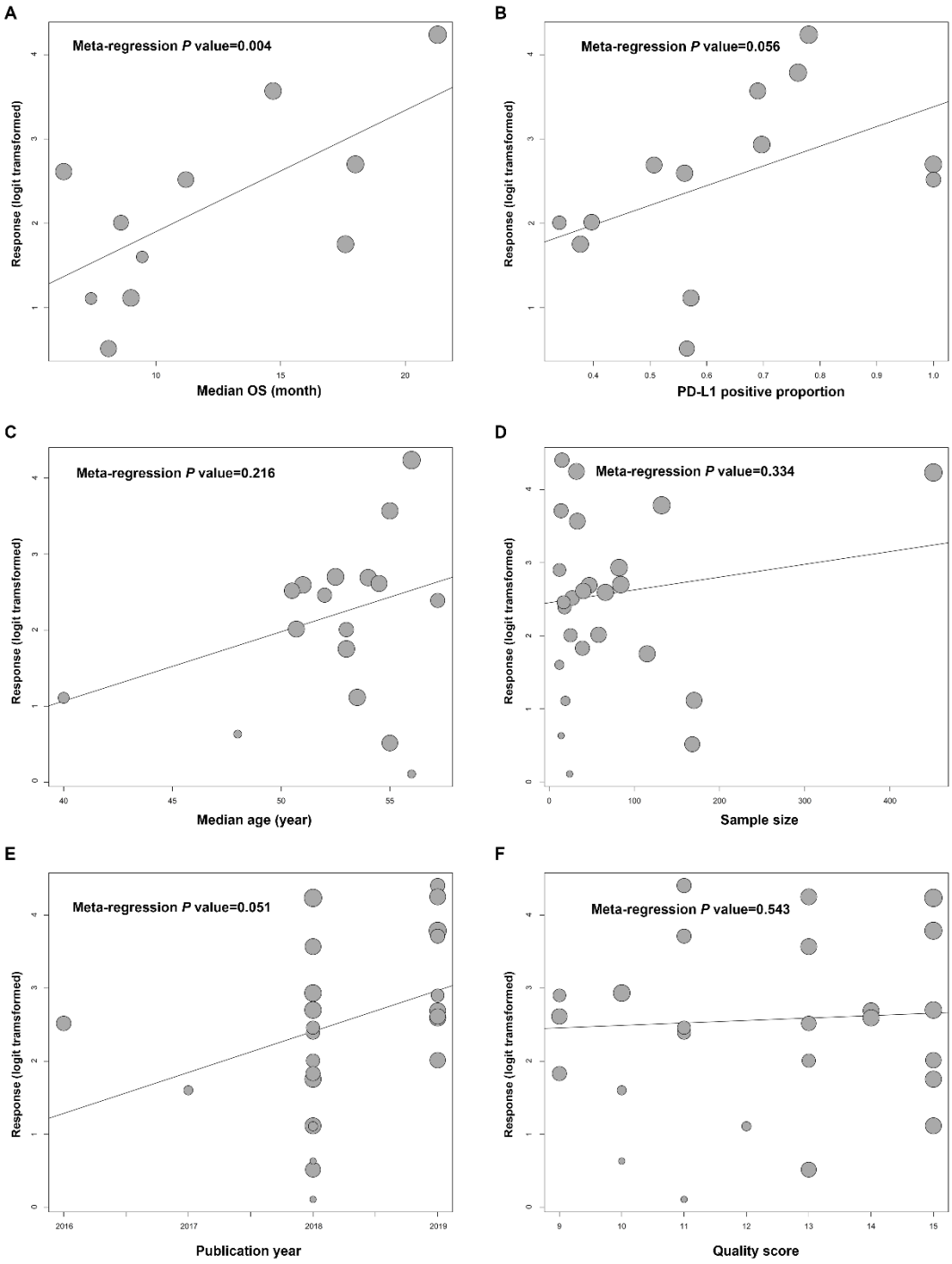
Abbreviations: ORR, objective response rate; ICI, immune checkpoint inhibitor.



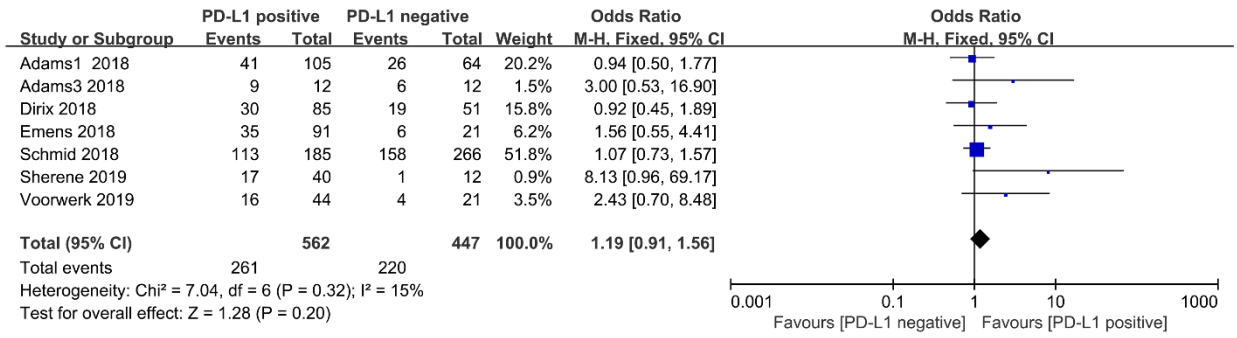
Supplemental Figure 2. Forest plots of comparison of ORR based on (A) PD-L1 (IPS) expression, (B) age, (C) performance status (ECOG), (D) lymph node metastasis and (F) LDH level.

Odds ratio for each study is presented, and horizontal lines indicate the 95% CI.

Abbreviations: ORR, objective response rate; IPS, immune cell proportion score, ECOG, Eastern Cooperative Oncology Group; PS, performance status; LDH, lactate dehydrogenase; ULN, upper limit of normal.



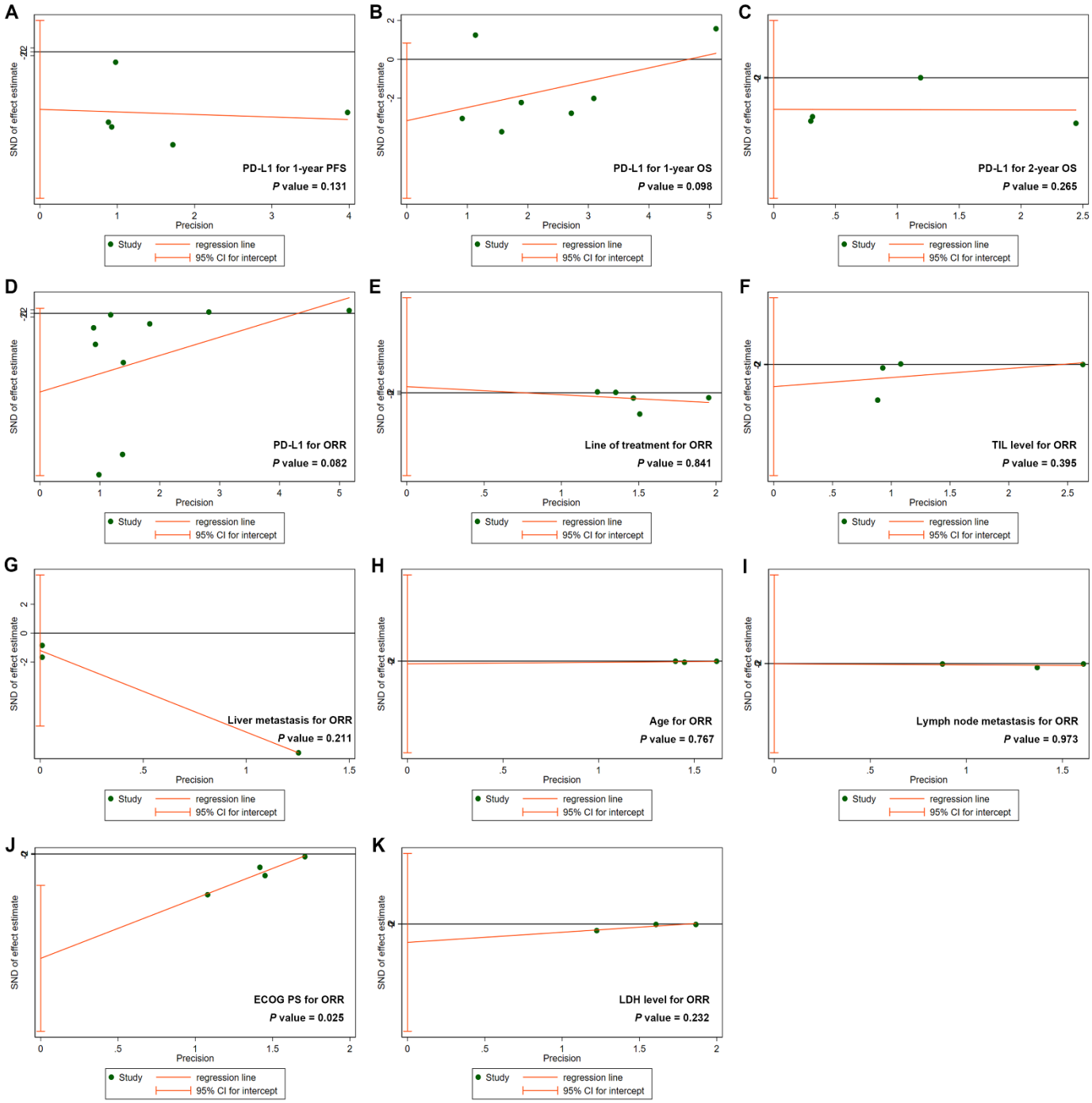
Supplemental Figure 3. Meta-regression analysis for ORR in patients treated with immune checkpoint inhibitor. Bubble plot with fitted meta-regression line of the -log proportion for (A) median overall survival, (B) proportion of PD-L1 positive patients, (C) median age, (D) sample size, (E) year of publication and (F) quality score. Circles size is proportional to the weight of each study in the fitted random-effects meta-regression.



Supplemental Figure 4. Forest plots of comparison of OS rate at the first year based on PD-L1 expression level after receiving ICI treatment.

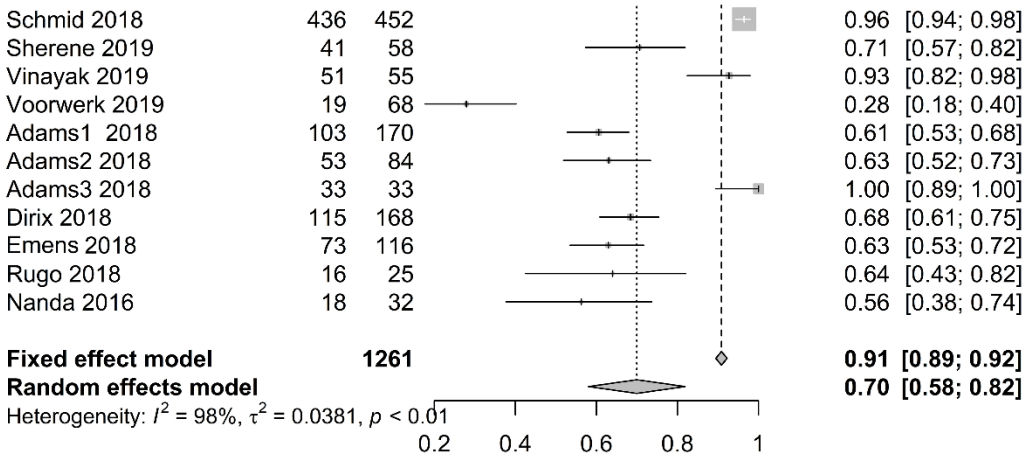
Odds ratio for each study is presented, and horizontal lines indicate the 95% CI.

Abbreviations: OS, overall survival; ICI, immune checkpoint inhibitor; CI, confidence interval.

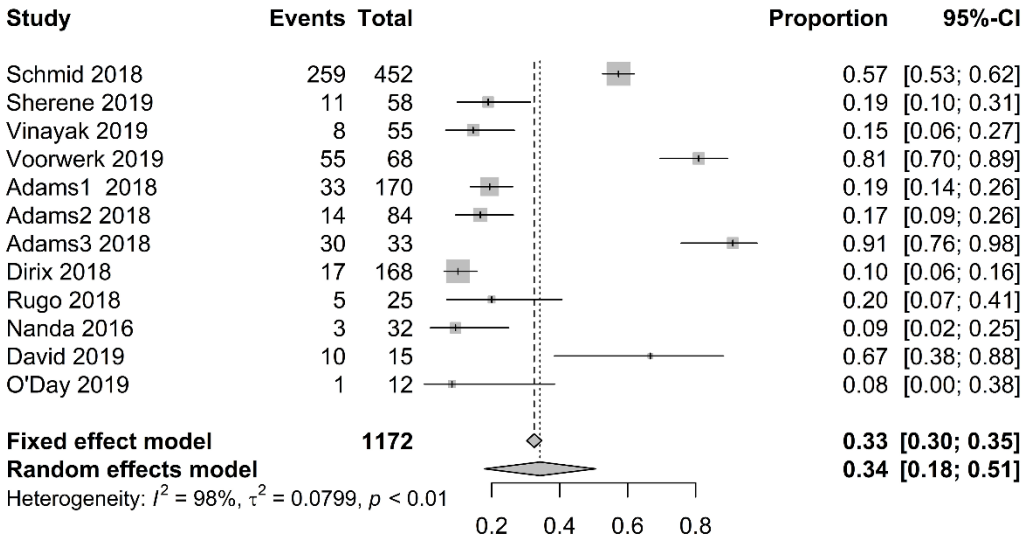


Supplemental Figure 5. Graph of Egger' s test evaluating publication bias of the included studies. *P* value<0.05 indicates publication bias.

A

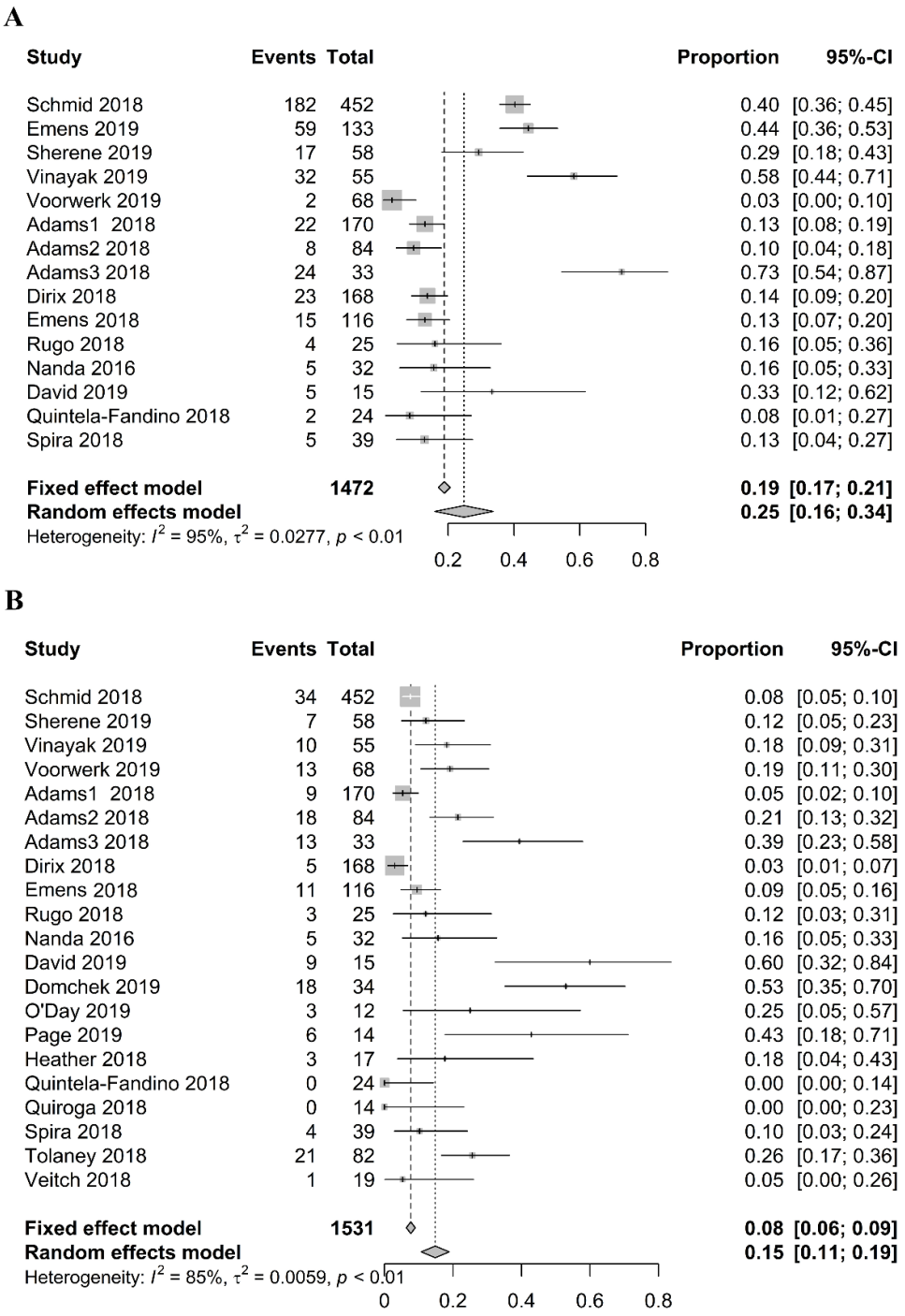


B



Supplemental Figure 6. Forest plots of incidence of any grade of (A) trAE and (B) irAE in metastatic breast cancer with PD-1/PD-L1 inhibitors therapy

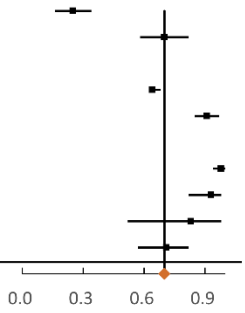
Abbreviations: trAE, treatment-related adverse events; irAE, immune-related adverse events; PD-1, programmed death-1; PD-L1, programmed death ligand 1.



Supplemental Figure 7. Forest plots of incidence⩾3 grade of (A) trAE and (B) irAE in advanced or metastatic breast cancer with PD-1/PD-L1 inhibitors therapy. Abbreviations: trAE, treatment-related adverse events; irAE, immune-related adverse events; PD-1, programmed death-1; PD-L1, programmed death ligand 1.

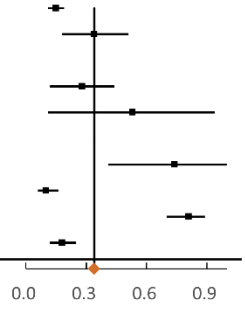
A

	No. of Study	No. of Patients	trAE	95%CI
trAE (3-4 Grade)	15	1472	0.25	0.16-0.34
trAE (total)	11	1261	0.70	0.58-0.82
Therapy				
Monotherapy trAE	7	607	0.64	0.64-0.68
Combination therapy trAE	5	652	0.91	0.85-0.97
Combination therapy				
Nab-paclitaxel or Paclitaxel trAE	2	485	0.98	0.94-1.00
PARP Inhibitors trAE	1	55	0.93	0.82-0.98
Radiotherapy trAE	1	12	0.83	0.52-0.98
Trastuzumab or T-DM1 trAE	1	58	0.71	0.57-0.82



B

	No. of Study	No. of Patients	irAE	95%CI
irAE (3-4 Grade)	21	1531	0.15	0.11-0.19
irAE (total)	12	1172	0.34	0.18-0.51
Immune Checkpoint Inhibitor				
PD-1 Inhibitor irAE	9	519	0.28	0.12-0.44
PD-L1 Inhibitor irAE	3	653	0.53	0.11-0.94
Antibody				
Atezolizumab irAE	2	485	0.74	0.41-1.00
Avelumab irAE	1	168	0.10	0.06-0.16
Nivolumab irAE	1	68	0.81	0.70-0.89
Pembrolizumab irAE	8	451	0.18	0.12-0.25



Supplemental Figure 8. Forest plots of incidence of (A) trAE in different therapy and combination therapy and (B) irAE with different immune checkpoint inhibitors and antibodies.

Abbreviations: trAE, treatment-related adverse events; irAE, immune-related adverse events.

Supplemental Table 1. Methodological quality of included randomized controlled trials for prognosis based on Cochrane risk of bias tool.

Scoring items: ① Random sequence generation； ② Allocation concealment; ③ Blinding of participants and personnel; ④ Blinding of outcome assessment; ⑤ Incomplete outcome data; ⑥Selective reporting; ⑦Other sources of bias.

Study	Selection bias		Performance bias	Detection bias	Attrition bias	Reporting bias	Other bias
	①	②	③	④	⑤	⑥	⑦
Schmid 2018	low-risky	low-risky	low-risky	low-risky	low-risky	low-risky	unclear
Emens 2019	low-risky	low-risky	low-risky	unclear	low-risky	low-risky	unclear

Supplemental Table 2. Methodological quality of included non-randomized studies for prognosis based on the methodological index for non-randomized studies (MINORS).

Scoring items: ①A clearly stated aim； ②Inclusion of consecutive patients; ③ Prospective collection of data; ④Endpoints appropriate to the aim of the study; ⑤ Unbiased assessment of the study endpoint; ⑥Follow-up period appropriate to the aim of the study; ⑦Loss to follow up less than 5%; ⑧Prospective calculation of the study size.

Study	①	②	③	④	⑤	⑥	⑦	⑧	Score
Sherene 2019	2	2	2	2	1	2	2	2	15
Vinayak 2019	2	2	2	2	1	2	1	2	14
Voorwerk 2019	2	2	2	2	1	2	1	2	14
Adams1 2018	2	2	2	2	1	1	1	2	13
Adams2 2018	2	2	2	2	1	2	2	2	15
Adams3 2018	2	2	2	2	1	2	2	2	15
Dirix 2018	2	2	2	2	1	1	1	2	13
Emens 2018	2	2	2	2	1	2	2	2	15
Rugo 2018	2	2	2	2	1	1	1	2	13
Santa-Maria 2018	2	2	2	2	1	0	1	1	11
Weiss 2017	2	2	2	2	1	0	1	0	10
Nanda 2016	2	2	2	2	1	1	1	2	13
Anders 2019	2	2	2	2	1	0	0	0	9
David 2019	2	2	2	2	1	2	0	0	11
Domchek 2019	2	2	2	2	1	2	2	0	13
O'Day 2019	2	2	2	2	1	0	0	0	9
Page 2019	2	2	2	2	1	0	2	0	11
Heather 2018	2	2	2	2	1	0	2	0	11
Quintela-Fandino 2018	2	2	2	2	1	0	2	0	11
Quiroga 2018	2	2	2	2	1	0	1	0	10
Spira 2018	2	2	2	2	1	0	0	0	9
Tolaney 2018	2	2	2	2	1	0	1	0	10
Veitch 2018	2	2	2	2	1	1	2	0	12

Supplemental Table 3. Risk of bias assessment of included studies for prognostic analysis based on Quality In Prognosis Studies (QUIPS) tool.

Scoring items: ①Study participation; ②Study Attrition; ③Prognostic Factor ; ④ Outcome Measurement; ⑤Study Confounding; ⑥Statistical Analysis and Reporting.

H: High risk of bias; M: Moderate risk of bias; L: Low risk of bias.

Study	①	②	③	④	⑤	⑥
Sherene 2019	M	L	L	L	H	L
Vinayak 2019	M	M	L	L	H	L
Voorwerk 2019	M	L	L	L	M	L
Adams1 2018	L	L	L	L	M	L
Adams2 2018	L	L	L	L	M	L
Adams3 2018	M	L	L	L	M	L
Dirix 2018	L	L	L	L	H	L
Rugo 2018	M	L	L	L	H	L
Domchek 2019	H	M	H	M	H	M
Tolaney 2018	H	H	H	M	M	M
Emens 2019	M	L	H	M	M	M
Emens 2018	L	L	L	L	M	L
Schmid 2018	L	L	L	L	M	L

Supplemental Table 4. GRADE quality of evidence summary.

No. of study	Design	Risk of bias	Inconsistency	Imprecision	Indirectness	Publication bias	No. of patients	Effect	Quality	Importance
PD-L1 status for 2-year OS										
4	randomized trials	Serious[1]	-	-	-	-	639	2.28 (1.16, 4.48)	MODERATE	CRITICAL
PD-L1 status for 1-year OS										
7	randomized trials	Serious[1]	-	-	-	-	1009	1.19 (0.91, 1.56)	MODERATE	CRITICAL
PD-L1 status for 1-year PFS										
6	randomized trials	Serious[1]	-	-	-	-	873	1.55 (1.02, 2.36)	MODERATE	CRITICAL
PD-L1 status for ORR										
10	randomized trials	Serious[1]	-	-	-	-	1252	1.44 (1.09, 1.91)	MODERATE	IMPORTANT
TIL for ORR										
4	randomized trials	Serious[1]	-	-	-	-	374	2.53 (1.39, 4.61)	MODERATE	IMPORTANT
CD8+ T cell for ORR										
2	randomized trials	Serious[1]	-	Serious[3]	-	Serious[4]	168	4.33 (1.53, 12.22)	VERY LOW	IMPORTANT
Liver metastasis for ORR										
3	randomized trials	Serious[1]	-	-	-	-	369	0.19 (0.06, 0.66)	MODERATE	IMPORTANT
Line of ICI treatment for ORR										
6	randomized trials	Serious[1]	-	-	-	Serious[4]	353	2.00 (1.13, 3.52)	LOW	IMPORTANT
Lymph node only metastasis for ORR										
3	randomized trials	Serious[1]	Serious[2]	Serious[3]	-	-	320	3.14 (0.61, 16.11)	VERY LOW	IMPORTANT
Age for ORR										
3	randomized trials	Serious[1]	-	-	-	Serious[4]	320	0.49 (0.24, 1.03)	LOW	IMPORTANT
ECOG PS for ORR										
3	randomized trials	Serious[1]	-	-	-	Serious[5]	369	1.77 (0.85, 3.67)	LOW	IMPORTANT
LDH level for ORR										
3	randomized trials	Serious[1]	-	-	-	Serious[4]	317	1.43 (0.72, 2.85)	LOW	NOT IMPORTANT

[1] Selective outcome reporting of biomarkers.
[2] Inconsistency index I^2 >50% for the effect.
[3] Wide 95% CI of the effect.
[4] More than 20% of the studies are from abstract.
[5] Egger's test indicates publication bias.

Appendix 1: Search strategy

Pubmed

(("Breast Neoplasms"[Mesh]) OR (Breast Neoplasm[Title/Abstract]) OR (Breast Tumors[Title/Abstract]) OR (Breast Tumor[Title/Abstract]) OR (Breast Cancer[Title/Abstract]) OR (Breast Carcinoma[Title/Abstract]) OR (Breast Carcinomas[Title/Abstract]) OR (Breast Malignant Neoplasm[Title/Abstract]) OR (Breast Malignant Tumor[Title/Abstract]) OR (Mammary Neoplasm, Human[Title/Abstract]) OR (Mammary Neoplasms, Human[Title/Abstract]) OR (Mammary Carcinoma, Human[Title/Abstract]) OR (Mammary Carcinomas, Human[Title/Abstract]) OR (Mammary Cancer, Human[Title/Abstract]) OR (Mammary Cancers, Human[Title/Abstract])) AND ((Ipilimumab[Title/Abstract]) OR (Yervoy[Title/Abstract]) OR (Tremelimumab[Title/Abstract]) OR (Ticilimumab[Title/Abstract]) OR (Nivolumab[Title/Abstract]) OR (Opdivo[Title/Abstract]) OR (BMS-936558[Title/Abstract]) OR (MDX-1106[Title/Abstract]) OR (Pembrolizumab[Title/Abstract]) OR (Keytruda[Title/Abstract]) OR (MK-3475[Title/Abstract]) OR (Lambrolizumab[Title/Abstract]) OR (Atezolizumab[Title/Abstract]) OR (Tecentriq[Title/Abstract]) OR (MPDL3280A[Title/Abstract]) OR (Durvalumab[Title/Abstract]) OR (MEDI4736[Title/Abstract]) OR (BMS-936559[Title/Abstract]) OR (MDX-1105[Title/Abstract]) OR (Avelumab[Title/Abstract]) OR (MSB0010718C[Title/Abstract]) OR (Immune checkpoint inhibitor[Title/Abstract]) OR (PD-1 inhibitor[Title/Abstract]) OR (PD-1 antibody[Title/Abstract]) OR (PD-1[Title/Abstract]) OR (PD-L1 inhibitor[Title/Abstract]) OR (PD-L1 antibody[Title/Abstract]) OR (PD-L1[Title/Abstract]) OR (CTLA-4 inhibitor[Title/Abstract]) OR (CTLA-4 antibody[Title/Abstract]) OR (CTLA-4[Title/Abstract]) OR ("Programmed Cell Death 1 Receptor"[Mesh]))

Embase

#1. 'breast cancer'/exp

#2. 'breast neoplasm':ab,ti

#3. 'breast tumor':ab,ti

#4. 'breast cancer':ab,ti

#5. 'breast carcinoma':ab,ti

#6. 'breast malignant neoplasm':ab,ti

#7. 'breast malignant tumor':ab,ti

#8. 'mammary neoplasm':ab,ti

#9. 'mammary carcinoma':ab,ti

#10. 'mammary cancer':ab,ti

#11. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10

#12. 'ipilimumab':ab,ti

#13. 'yervoy':ab,ti

#14. 'tremelimumab':ab,ti

#15. 'ticilimumab':ab,ti

#16. 'nivolumab':ab,ti

#17. 'opdivo':ab,ti

#18. 'bms-936558':ab,ti

#19. 'mdx-1106':ab,ti

#20. 'pembrolizumab':ab,ti

#21. 'keytruda':ab,ti

#22. 'mk-3475':ab,ti

#23. 'lambrolizumab':ab,ti

#24. 'atezolizumab':ab,ti

#25. 'tecentriq':ab,ti

#26. 'mpdl3280a':ab,ti

#27. 'durvalumab':ab,ti

#28. 'medi4736':ab,ti

#29. 'bms-936559':ab,ti

#30. 'mdx-1105':ab,ti

#31. 'avelumab':ab,ti

#32. 'msb0010718c':ab,ti

#33. 'immune checkpoint inhibitor':ab,ti

#34. 'pd-1 inhibitor':ab,ti

#35. 'pd-1 antibody':ab,ti

#36. 'pd-l1 inhibitor':ab,ti

#37. 'pd-l1 antibody':ab,ti

#38. 'ctla-4 inhibitor':ab,ti

#39. 'ctla-4 antibody':ab,ti

#40. #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39

#41. #11 AND #40

Cochrane library

#1 MeSH descriptor: [Breast Neoplasms] explode all trees

#2 (Breast Neoplasm):ti,ab,kw OR (Breast Tumor):ti,ab,kw OR (Breast Cancer):ti,ab,kw OR (Breast Carcinoma):ti,ab,kw (Word variations have been searched)

#3 (Breast Malignant Neoplasm):ti,ab,kw OR (Breast Malignant Tumor):ti,ab,kw OR (Mammary Neoplasm):ti,ab,kw OR (Mammary Carcinoma):ti,ab,kw OR (Mammary Cancer):ti,ab,kw (Word variations have been searched)

#4 (Ipilimumab):ti,ab,kw OR (Yervoy):ti,ab,kw OR (Tremelimumab):ti,ab,kw OR (Ticilimumab):ti,ab,kw OR (Nivolumab):ti,ab,kw (Word variations have been searched)

#5 (Opdivo):ti,ab,kw OR (BMS-936558):ti,ab,kw OR (MDX-1106):ti,ab,kw OR (Pembrolizumab):ti,ab,kw OR (Keytruda):ti,ab,kw (Word variations have been searched)

#6 (MK-3475):ti,ab,kw OR (Lambrolizumab):ti,ab,kw OR (Atezolizumab):ti,ab,kw OR (Tecentriq):ti,ab,kw OR (MPDL3280A):ti,ab,kw (Word variations have been searched)

#7 (Durvalumab):ti,ab,kw OR (MEDI4736):ti,ab,kw OR (BMS-936559):ti,ab,kw OR (MDX-1105):ti,ab,kw OR (Avelumab):ti,ab,kw (Word variations have been searched)

#8 (MSB0010718C):ti,ab,kw OR (Immune checkpoint inhibitor):ti,ab,kw OR (PD-1 inhibitor):ti,ab,kw OR (PD-1 antibody):ti,ab,kw OR (PD-1):ti,ab,kw (Word variations have been searched)

#9 (PD-L1 inhibitor):ti,ab,kw OR (PD-L1 antibody):ti,ab,kw OR (PD-L1):ti,ab,kw OR (CTLA-4 inhibitor):ti,ab,kw OR (CTLA-4 antibody):ti,ab,kw (Word variations have been searched)

#10 (CTLA-4):ti,ab,kw OR (Programmed Cell Death 1 Receptor):ti,ab,kw (Word variations have been searched)

#11 (#1 OR #2 OR #3) AND (#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10)

Web of Science

((Ipilimumab) OR (Yervoy) OR (Tremelimumab) OR (Ticilimumab) OR (Nivolumab) OR (Opdivo) OR (BMS-936558) OR (MDX-1106) OR (Pembrolizumab) OR (Keytruda) OR (MK-3475) OR (Lambrolizumab) OR (Atezolizumab) OR (Tecentriq) OR (MPDL3280A) OR (Durvalumab) OR (MEDI4736) OR (BMS-936559) OR (MDX-1105) OR (Avelumab) OR (MSB0010718C)) AND ((Breast Neoplasm) OR (Breast Neoplasms) OR (Breast Tumor) OR (breast cancer) OR (Breast Carcinoma) OR (Breast Malignant Neoplasm) OR (Breast Malignant Tumor) OR (Mammary Neoplasm) OR (Mammary Carcinoma) OR (Mammary Cancer))

US National Institutes of Health Ongoing Trials Register

www.clinicaltrials.gov

Appendix 2

Protocol of Systematic Review and Meta-Analysis

Title: Efficacy and Predictive Factors of Immune Checkpoint Inhibitors for Metastatic Breast Cancer: A Systematic Review and Meta-Analysis

Review background

Description of the health condition and context

According to the global statistics, breast cancer is the most common cancer and the second leading cause of cancer-related death among female patients. Approximately 20% of the patients will experience occurrence with distant metastatic disease in the first five years. Patients with recurrence or metastatic breast cancer predict a poor prognosis with a 5-year relative survival rate of 27%. Recently, Immune checkpoint inhibitors (ICI) have shown encouraging prospects in metastatic breast cancer in several clinical trials. Patients with metastatic breast cancer showed an objective response rate (ORR) of 3%~45% after treated with ICI in different reported phase 2 clinical trials. However, response only occurred in a small population. Understanding the efficacy and predictive factors is critical for clinical practice. Additionally, the incidence of adverse events of ICI treatment is needed to analyzed.

Description of the predictive factors

1. The type of ICI agents and combined therapy;
2. Biomarkers that have been reported in other studies: PD-L1 expression, tumor mutation burden (TMB), tumor-infiltrating lymphocytes level (TIL), CD8+ T cell level, microsatellite instability (MSI) and so on;
3. Baseline characteristics of patients: line of ICI treatment, subtype of breast cancer, metastatic site, age, menopausal status, performance status and so on.

Health outcomes

The outcomes will include objective response rate (ORR), treatment-related adverse

events (trAEs), immune-related adverse events (irAEs), progression-free survival (PFS) and overall survival (OS).

Why it is important to do this review

Immune checkpoint inhibitors have shown encouraging prospects in metastatic breast cancer in several clinical trials. However, response only occurred in a small population. Understanding the efficacy and predictive factors is critical for clinical practice.

Objectives

Primary objectives

The PICOTS format consists of the following elements (Riley et al, BMJ 2019):

• Population—Patients with metastatic breast cancer.

• Index prognostic factor—Particular biomarker (PD-L1, TMB, MSI, TIL, ICI regimen, metastatic sites, age, performance status and so on).

• Comparator prognostic factors—Not applicable to this review.

• Outcomes—Objective response rate (ORR), treatment-related adverse events (trAEs), immune-related adverse events (irAEs), progression-free survival (PFS) and overall survival (OS).

• Timing—Biomarker measurement had to be done before ICI treatment and all follow-up information on the outcomes was extracted from the studies.

• Setting—Hospital/treatment center.

Secondary objectives

Secondary objectives were to evaluate the accuracy and consistency of each biomarker if possible. Comparing the performance of each summarized prognostic biomarkers in predicting response and survival.

Investigation of sources of heterogeneity between studies

Subgroup analysis: Determination of PD-L1 status, ICI regimen, metastatic sites, age, menopausal status, performance status and so on. Meta-regression analysis:

Covariation between the outcomes and baseline characteristics of each included study.

Methods

Criteria for considering studies for this review

Only prospective clinical trials of patients with breast cancer treated with an ICI (including anti-PD-1, anti-PD-L1 and anti-CTLA-4 inhibitor) that reported response outcomes and adverse events data will be included in this review. Articles published online "ahead of print" will be included. Meeting abstracts without published full-text original articles will be eligible for this study. Exclusion criteria will be insufficient data, not advanced or metastatic breast cancer, preclinical studies, case reports, letters, commentaries and reviews. Additionally, retrospective studies will be excluded in this review.

Types of studies

Clinical trials will be included in this review. English will be set as restricted language. When duplicate studies from the same trial will be identified, only the most complete and updated data of the study will be included.

Targeted population

Patients with metastatic breast cancer treated with an ICI (including anti-PD-1, anti-PD-L1 and anti-CTLA-4 inhibitor)

Types of prognostic / predictive factor(s) or model(s)

Details of biomarkers and baseline characteristics of patients will be showed as follows (The content in parentheses is the detection method): PD-L1 expression (immunohistochemical staining), tumor mutation burden (next generation sequencing), tumor-infiltrating lymphocytes level (HE staining), CD8+ T cell level (immunohistochemical staining), microsatellite instability (next generation sequencing), line of ICI treatment (describe in text), subtype of breast cancer(describe in text), metastatic site(describe in text), age(describe in text), menopausal

status(describe in text), performance status (ECOG scale) and so on.

Types of outcomes to be predicted

Objective response rate (ORR) will be assessed after the treatment of ICI according to RECIST 1.1 guideline. Treatment-related adverse events (trAEs) and immune-related adverse events (irAEs) will be assessed after the treatment of ICI according to Common Terminology Criteria for Adverse Events (CTCAE). Progression-free survival (PFS) will be defined as the interval between start of ICI therapy and disease progression (either local or distant) or dead of any reason. Overall survival (OS) will be defined as the interval between start of ICI therapy and dead of any reason.

Search methods for identification of studies

Electronic searches

The comprehensive search of online databases will include PubMed, Embase, the Cochrane library, Web of Science online databases and www.clinicaltrials.gov. The retrieval strategy will contain the following keywords: Nivolumab, Pembrolizumab, Atezolizumab, Durvalumab, Avelumab, Ipilimumab, Tremelimumab, immune checkpoint inhibitor, PD-1 inhibitor, PD-L1 inhibitor, CTLA-4 inhibitor, and breast cancer. We also will review abstracts from American Society of Clinical Oncology (ASCO), conferences using the same criteria reported below. The reference lists from these studies will be hand searched for potential eligible articles. All the search strategies will be conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

Detailed retrieval terms are presented in **Appendix 1**.

Searching other resources

List ‘grey’ literature sources, such as reports and conference proceedings. If journals are specifically hand-searched for the review, this should also be noted. List people (for example, researchers, experts) and/or organizations who will be contacted. List any

other sources, which may include, for example, reference lists, the World Wide Web or personal collections of articles.

Data collection

Selection of studies All the search results will be independently inspected by two authors with discrepancies consulted by a third reviewer. Criteria of selection will be applied by reviewers after screening the potentially included studies. Duplicates will be removed using Endnote X9 software or manually.

Data extraction and management

Baseline characteristics of each study (authors, year of publication or conference presentation, line of ICI treatment, type of ICI agents, breast cancer type, number of patients enrolled, combination therapy and median survival) will be recorded by two reviewers independently. Objective response rate (ORR), treatment-related adverse events (trAEs), immune-related adverse events (irAEs), PFS and OS will be extracted from studies. Additionally, we will extract data from different subgroups in the same trial to identify biomarkers (PD-L1 expression, tumor mutation burden (TMB), tumor-infiltrating lymphocytes level (TIL), CD8+ T cell level, microsatellite instability (MSI)) that predict ORR, PFS or OS of ICI treatment. These results will be described by odds ratio (OR) and 95% confidence interval (CI).

Assessment of risk of bias in included studies

The quality of each randomized controlled trial (RCT) and non-randomized trial will be assessed by using the Cochrane risk of bias tool and the methodological index for non-randomized studies (MINORS), respectively. Cochrane risk of bias tool expressed risk of bias as low, high, or unclear risk including the aspects of selection, performance, detection, attrition, reporting and other bias. MINORS is recognized as the most appropriate guideline to evaluate the methodological quality of non-randomized trial which contained eight specific items. Two reviewers will make the assessment with disagreements consulted by a third reviewer. All RCTs and non-randomized trials will

be scored and recorded. Additionally, we will visually use the funnel plots and Egger's test to assess publication bias.

Assessment of heterogeneity

Cochran's Q test (reported with a $\chi 2$ value and P value) will be used to estimate study heterogeneity. Heterogeneity will be indicated if $P<0.1$. I2 statistic was also used with values over 50% suggesting heterogeneity. (Debray et al, BMJ 2017 and Snell et al, J Clin Epidemiol 2016).

Assessment of reporting deficiencies

If any tests or investigations are undertaken to detect reporting biases the methods for statistical tests for detecting asymmetry in hazard ratios from prognostic factor studies should be used (Riley et al, BMJ 2016 for more information on funnel plots).

Data synthesis

Data synthesis and meta-analysis approaches

The ORR, trAE and irAE of the patients will be extracted and pooled using both the fixed- and random-effects model in R software, version 3.5.0. ORs describing PFS and OS were synthesized in Review Manager software (version 5.3, Cochrane Collaboration). We will use Cochran's Q test (reported with a $\chi 2$ value and P value) to estimate study heterogeneity. Heterogeneity was indicated if $P<0.1$. I2 statistic will be also used with values over 50% suggesting heterogeneity. OR and 95% CI for each of comparisons in the subgroup will be pooled using the fixed-effects model (if heterogeneity Cochran's Q test $P>0.1$) and the random-effects model (if heterogeneity Cochran's Q test $P<0.1$) in Review Manager software. (Debray et al, BMJ 2017, Riley et al, Res Synth Methods 2010, Debray et al, Stat Med 2014). No transformation will be necessary for pooling statistic.

Subgroup analysis and investigation of heterogeneity

Subgroup analysis will be used to deal with the potential heterogeneity observed after

data synthesis. Several subgroups are analyzed, such as detection of PD-L1 expression (immunohistochemical staining), tumor mutation burden (next generation sequencing), tumor-infiltrating lymphocytes level (HE staining), CD8+ T cell level (immunohistochemical staining), microsatellite instability (next generation sequencing), line of ICI treatment (describe in text), subtype of breast cancer(describe in text), metastatic site(describe in text), age(describe in text), menopausal status(describe in text), performance status (ECOG scale), etc. Meta-regression, which extends a standard random effects meta-analysis model by including study-level covariates, can be used to formally test for subgroup differences (Berkey et al, Stat Med 1995). Covariates may represent case-mix (participant selection), follow-up duration, predictor and outcome measurement, study design, risk of bias, etc.

Sensitivity analysis

Sensitivity analyses will be done by omitting every study one-by-one from the meta-analysis to determine the influence of individual study exerting on the combined result. (Snell et al, J Clin Epidemiol 2016).

Conclusions and summary of findings

The summary of findings will be displayed as the form of tables and forest plots. For the analysis of prognostic biomarkers, the synthesized outcome of each biomarker is recorded and summarized if there are enough studies. The biomarker that just reported in single study will be summarized in the result finding of text. The GRADE approach will be used and adopted to evaluate the reliability of evidence if necessary. The conclusions will be made strictly following the above standards in method section. The conclusions must focus on which biomarkers or factors can predict the response and survival of metastatic breast cancer after receiving ICI drugs.