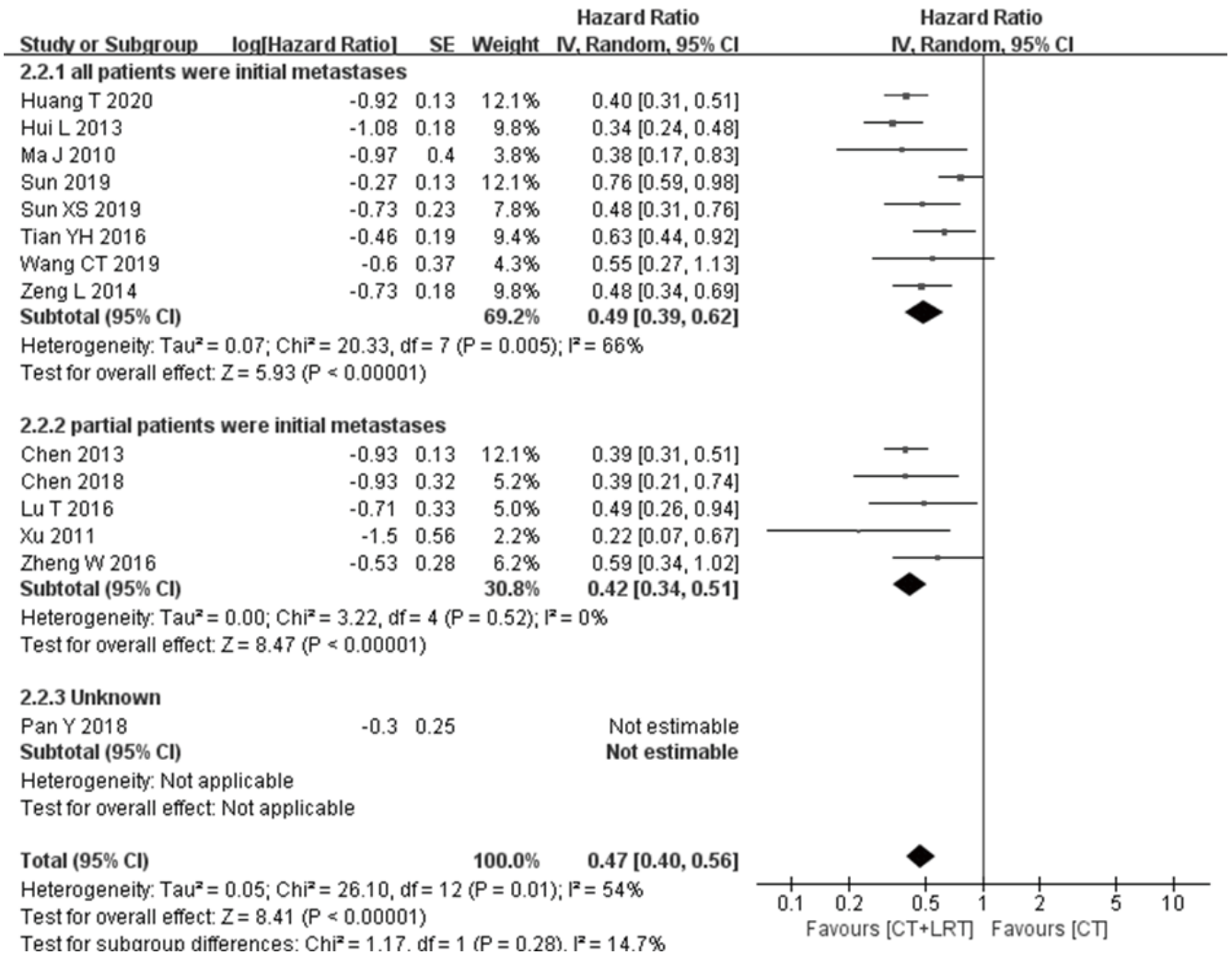

Figure S1 Forest plots of subgroup analysis according to initial metastases.

http://dx.doi.org/10.21037/apm-20-1561

Table S1 Result of quality assessment by using the methodological index for non-randomized studies (minors) for involved studies

Study	A clearly stated aim[a]	Inclusion of consecutive patients[b]	Prospective collection of data[c]	Endpoints appropriate to the aim of the study[d]	Unbiased assessment of the study endpoint[e]	Follow-up period appropriate to the aim of the study[f]	Loss to follow up less than 5%[g]	Prospective calculation of the study size[h]	An adequate control group[i]	Contemporary groups	Baseline equivalence of groups[j]	Adequate statistical analyses	Score
Cao *et al.* 2011; (20)	1	2	0	1	1	0	2	0	2	2	1	2	14
Chen *et al.* 2013; (21)	2	2	0	1	1	1	1	0	1	2	1	2	14
Chen *et al.* 2018; (22)	1	2	0	1	1	0	2	0	1	2	1	2	13
Gu *et al.* 2020 (23)	1	2	0	1	1	1	2	0	2	2	1	2	15
Huang *et al.* 2020; (24)	2	2	0	2	1	0	2	0	1	2	1	2	15
Lin *et al.* 2013; (25)	2	2	0	2	1	2	2	0	2	2	1	2	18
Lu *et al.* 2016; (26)	1	2	0	1	1	1	2	0	1	2	1	2	14
Ma *et al.* 2010; (27)	2	2	0	2	1	2	2	0	1	2	1	2	17
Pan *et al.* 2018; (28)	2	2	0	1	1	0	2	0	1	2	1	2	14
Sun *et al.* 2019; (29)	1	2	0	1	1	1	2	0	1	2	1	2	14
Sun *et al.* 2019; (30)	1	2	0	1	1	1	2	0	2	2	1	2	15
Tian *et al.* 2016; (31)	1	2	0	1	1	0	2	0	1	2	1	2	13
Wang *et al.* 2009; (32)	1	2	0	1	1	2	2	0	1	2	0	2	14
Zeng *et al.* 2014; (33)	2	2	0	1	1	1	2	0	2	2	1	2	16
Zeng *et al.* 2016; (34)	2	2	0	1	1	2	2	0	1	2	1	2	16

The items are scored 0 (not reported), 1 (reported but inadequate) or 2 (reported and adequate). [a], if the aim of the study was definite and to compare LRT + CT and CT alone, the item was scored 2; if not, the item was scored 1. [b], if the period of recruiting patients was consecutive, the item was scored 2; if not, the item was scored 1; if not reported, the item was scored 0. [c], if all data were collected proactively, the item was scored 2; if partial data were collected proactively, the item was scored 1. [d], the adequate endpoints meant the study should have two endpoints, including PFS and OS, to evaluate the efficacy. If the study only had one endpoint, the item was scored 1. [e], if author declared explicitly assessment was blind and mutually independent, the item was scored 2. If assessment was only conducted independently between the researchers, the item was scored 1. If the detail of assessment wasn't reported, the item was scored 0. [f], if median follow-up duration was not less than 36 months, the item was scored 2; if median follow-up duration was less than 36 months, the item was scored 1; if the information of follow-up wasn't reported, the item was scored 0. [g], if the rate of lost to follow-up was less than 5%, the item was scored 2; if not, the item was scored 1; if the rate of lost to follow-up wasn't reported, the item was scored 0. [h], if the authors calculated the study size in advance and mentioned in the paper, the item was scored 2; if the authors mentioned the anticipated size, but didn't demonstrated the detail, the item was scored 1. [i], an adequate control group meant the same disease, adequate size, same treatment (except LRT) compare to research group. If the control group didn't meet all requirements, the item was scored 1. If the control group wasn't reported in the study, the item was scored 0. [j], if the baseline of recruited patients between control and research group had no significant difference, the item was scored 2; if not, the item was scored 1; if the baseline wasn't reported, the item was scored 0.

http://dx.doi.org/10.21037/apm-20-1561

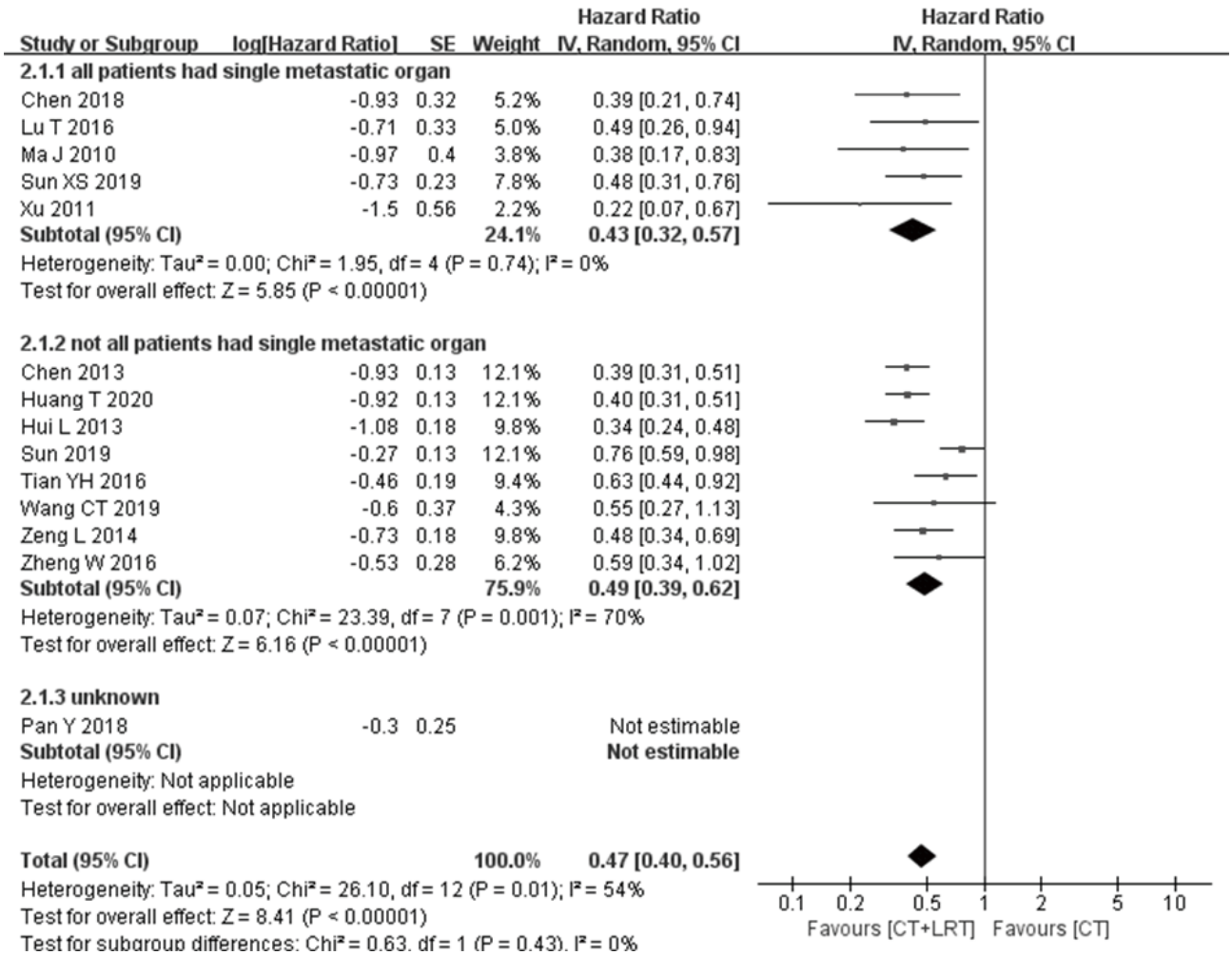

Figure S2 Forest plots of subgroup analysis according to single metastatic organ.

http://dx.doi.org/10.21037/apm-20-1561

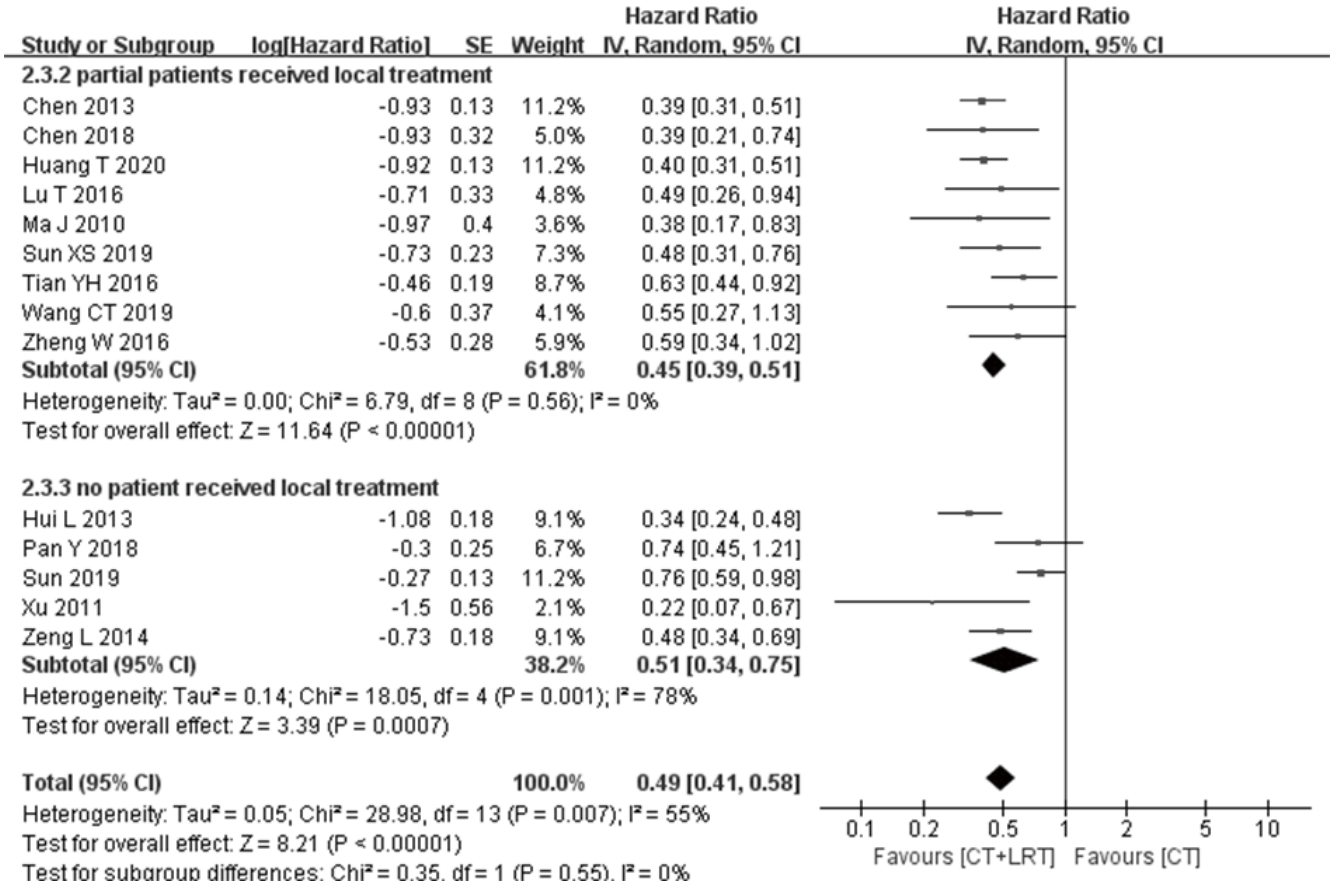

Figure S3 Forest plots of subgroup analysis according to local treatment.

http://dx.doi.org/10.21037/apm-20-1561

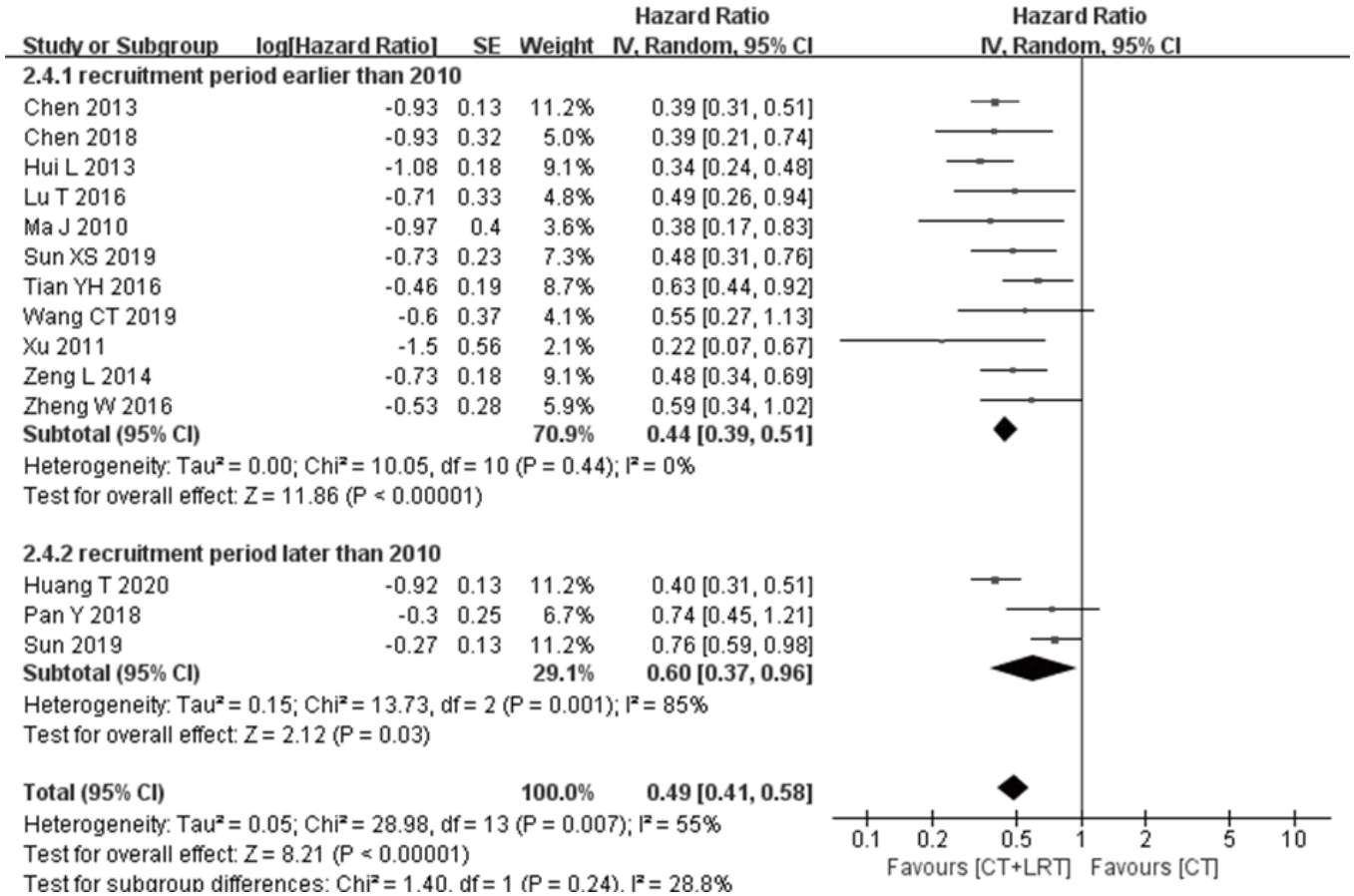

Figure S4 Forest plots of subgroup analysis according to recruitment period.

http://dx.doi.org/10.21037/apm-20-1561