Langenbeck's Archives of Surgery (2022) 407:1381–1397
https://doi.org/10.1007/s00423-022-02454-3

SYSTEMATIC REVIEWS AND META-ANALYSES

Comparison of estimated treatment effects between randomized controlled trials, case-matched, and cohort studies on laparoscopic versus open distal gastrectomy for advanced gastric cancer: a systematic review and meta-analysis

Ryota Otsuka[1] · Hideki Hayashi[1] · Masaya Uesato[1] · Koichi Hayano[1] · Kentaro Murakami[1] · Masayuki Kano[1] · Takeshi Toyozumi[1] · Hiroshi Suito[1] · Yasunori Matsumoto[1] · Tetsuro Isozaki[1] · Yoshihiro Kurata[1] · Hisahiro Matsubara[1]

Received: 18 September 2021 / Accepted: 24 January 2022 / Published online: 3 February 2022

Abstract

Purpose In actual surgical research, case-matched studies are frequently conducted as an alternative to randomized controlled trials (RCTs). However, it is still unclear what differences there are between RCTs and case-matched studies in upper gastrointestinal surgery, and clarifying them is a very important clinical issue. Thus, the purpose of this study was to investigate estimated treatment effects between RCTs, case-matched studies, and cohort studies regarding laparoscopic distal gastrectomy (LDG) for advanced gastric cancer (AGC).

Methods We searched the PubMed, Cochrane Central Register of Controlled Trials, and Web of Science databases for studies that compared LDG versus open distal gastrectomy for AGC published from the inception of the databases until July 2021. A meta-analysis was performed using the Review Manager version 5.3 software program from the Cochrane Collaboration, and six short-term outcomes and three long-term outcomes were assessed.

Results Twenty-three studies with 13698 patients were included. There was no difference in estimated treatment effects between RCTs and case-matched studies for all outcomes except for the number of retrieved lymph nodes and postoperative complications. In terms of intraoperative blood loss, postoperative hospital stay, number of retrieved lymph nodes, and recurrence, observational studies tended to overestimate the treatment effects.

Conclusion The estimated treatment effects of LDG for AGC in the case-matched study were almost the same as in the RCTs. However, to assess the true magnitude of the treatment effect, the design and actual implementation of the analysis must be critically evaluated.

Keywords Case-matched study · Randomized controlled trial · Laparoscopic distal gastrectomy · Open distal gastrectomy · Advanced gastric cancer

Introduction

It is generally accepted that the gold standard for evaluating the efficacy of therapeutic interventions is randomized controlled trials (RCTs). In RCTs, random assignment of participants to treatment and control groups virtually eliminates distortion of results due to differences in patient characteristics between study groups. However, in most surgical studies, randomization is difficult for ethical and practical reasons [1]. In addition, RCTs are costly and inefficient because they require many resources, including subjects, time, and the cooperation of diverse experts, to estimate treatment effects with sufficient accuracy [2]. Therefore, as a practical alternative, many observational studies have been performed in actual clinical settings.

Observational studies are susceptible to biases such as confounding, selection, and differential ascertainment bias because they lack randomization and other elements of RCT design [3]. Some reports have suggested that both randomized and observational studies may produce very similar

✉ Ryota Otsuka
ryotaotsuka@yahoo.co.jp

[1] Department of Frontier Surgery, Graduate School of Medicine, Chiba University, 1-8-1, Inohana, Chuo-ku, Chiba-shi, Chiba 260-8670, Japan

results [4, 5], while others have reported conflicting results [6]. However, the topics covered in these previous reports are very limited, and more empirical and quantitative evidence is needed to clarify the accuracy of and differences in each study design [7]. In recent years, case-matched studies have been frequently conducted in surgical research for appropriate confounder adjustment in observational studies, and the most common technique is propensity score matching [1]. The propensity score, proposed as a potential solution to the problem of confounding associations between treatment and outcome, represents the probability of being treated with an intervention based on variables measured during or before treatment [8]. Although there are methodological differences between case-matched studies and RCTs, such as patient selection and adjustment for confounders [9, 10], only one report, concerning rectal cancer, has investigated the similarities and differences between different study designs in the field of gastrointestinal surgery [2]. Therefore, it is still unclear what differences there are between RCTs, case-matched studies, and cohort studies in other gastrointestinal surgeries, and clarifying them is a very important clinical issue.

Thus, the purpose of this study was to investigate estimated treatment effects between RCTs, case-matched studies, and cohort studies regarding upper gastrointestinal surgery areas. As a clinical topic, we selected the comparison of laparoscopic distal gastrectomy (LDG) versus open distal gastrectomy (ODG) for advanced gastric cancer (AGC), which is one of the most discussed and interested issues among gastrointestinal surgeons. While there have been several meta-analysis studies that evaluated the efficacy of LDG in AGC [11–14], none of them focused on the differences in study design.

Therefore, in the present study, we evaluated the differences in study designs by addressing this clinical topic for which sufficient evidence has been accumulated.

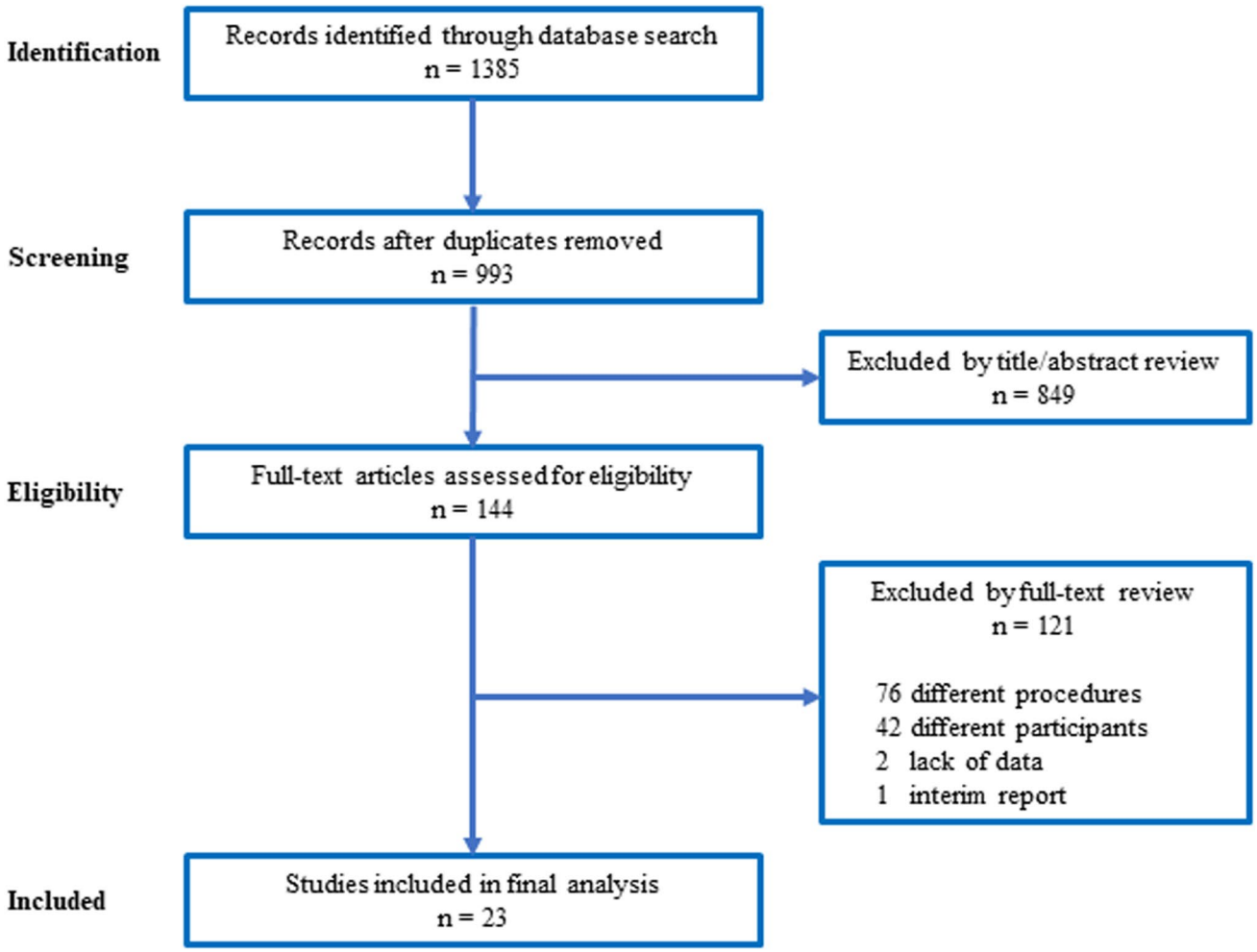

Fig. 1 Flow diagram of study selection

Table 1 Study characteristics

	Year	Country	Institution	Type of study	Participants, *n*		Gender, M:F		Mean age, years		Mean BMI, kg/m^2		Clinical stage, I/II/III/IV		Pathological stage, I/II/III/IV		Operative time, min	
					Lap	Open	Lap	Open	Lap	Open	Lap	Open	Lap	Open	Lap	Open	Lap	Open
RCTs																		
Hu [31]	2016	China	Multi	Prospective	519	520	380:139	346:174	56.5 ± 10.4	55.8 ± 11.1	22.7 ± 3.2	22.7 ± 3.2	NR	NR	151/137/219/11	152/138/221/8	217.3 ± 60.3	186 ± 53.3
Park [33]	2018	Korea	Multi	Prospective	100	96	69:31	65:31	58.6	60.1	23.7 ± 3.0	23.3 ± 3.1	23/52/25/0	22/46/28/0	42/29/28/1	36/33/23/4	257.4 ± 86.7	183 ± 52.5
Wang [37]	2019	China	Multi	Prospective	222	220	144:78	133:87	59.4 ± 12.4	60.6 ± 10.2	23.1 ± 3.1	23.5 ± 3.3	NR	NR	75/63/80/4	68/63/83/6	242.5 ± 63.5	209.9 ± 53.6
Yu [38]	2019	China	Multi	Prospective	519	520	380:139	346:174	56.5 ± 10.4	55.8 ± 11.1	22.7 ± 3.2	22.7 ± 3.2	64/248/207/0	88/247/185/0	151/137/219/11	152/138/221/8	NR	NR
Hyung [40]	2020	Korea	Multi	Prospective	492	482	351:141	335:147	59.8 ± 11.0	59.4 ± 11.5	23.5 ± 2.9	23.7 ± 3.3	134/159/199/0	116/169/197/0	178/148/166/0	165/167/150/0	227.0 ± 67.9	164.4 ± 45.8
Case-matched studies																		
Scatizzi [24]	2011	Italy	Single	Retrospective	30	30	16:14	14:16	70.0 ± 12.0	69.0 ± 10.8	22.0 ± 1.0	24.0 ± 6.5	NR	NR	0/12/18/0	0/10/20/0	240.0 ± 32.5	180.0 ± 30.0
Shuang [25]	2011	China	Single	Retrospective	35	35	30:5	30:5	58.0 ± 10.5	59.0 ± 13.5	21.0 ± 3.0	23.0 ± 3.0	NR	NR	10/15/10/0	9/13/13/0	320.0 ± 77.5	210.0 ± 40.5
Zhang [30]	2015	China	Single	Retrospective	86	86	57:29	61:25	62.0 ± 4.8	61.0 ± 4.0	NR	NR	9/66/11/0	10/67/9/0	5/52/29/0	4/56/26/0	210.0 ± 15.0	180.0 ± 11.7
Yoshida [34]	2018	Japan	Multi	Retrospective	3738	3738	2450:1288	2444:1294	71.0 ± 2.7	70.0 ± 2.5	22.2 ± 3.4	22.2 ± 3.3	NR	NR	0/2318/1220/200	0/2320/1224/194	NR	NR
Kim [35]	2019	Korea	Single	Retrospective	60	60	38:22	41:19	62.5 ± 14.2	62.4 ± 10.4	24.3 ± 4.3	24.1 ± 3.5	22/29/9/0	17/33/10/0	0/37/23/0	0/38/22/0	234.8 ± 46.9	217.2 ± 48.0
Garbarino [39]	2020	Italy	Single	Retrospective	34	34	23:11	21:13	70.9 ± 10.7	71.1 ± 9.1	24.2 ± 4.1	24.2 ± 3.2	NR	NR	2/18/14/0	2/10/22/0	257.2 ± 46.3	197.2 ± 66.4
Wang [41]	2020	China	Single	Retrospective	190	190	134:56	131:59	57.7 ± 10.7	58.3 ± 10.2	21.6 ± 3.1	21.7 ± 2.8	12/87/91/0	14/71/105/0	44/44/102/0	35/37/118/0	195.0 ± 52.2	203.3 ± 63.8
Huang [42]	2021	China	Multi	Retrospective	461	461	135/326	137/324	NR	NR	21.7 ± 3.5	21.0 ± 3.4	NR	NR	NR	NR	169.0 ± 62.8	169.8 ± 47.1
Cohort studies																		
Hur [20]	2008	Korea	Single	Retrospective	26	25	19:7	19:6	NR	NR	NR	NR	NR	NR	NR	NR	255.0 ± 42.5	183.8 ± 31.3
Du [21]	2009	China	Single	Retrospective	78	90	55:23	62:28	56.0 ± 6.0	60.0 ± 8.0	NR	NR	NR	NR	8/27/37/5	9/30/40/11	245.0 ± 35.0	220.0 ± 20.0
Hwang [22]	2009	Korea	Single	Retrospective	45	83	25:20	58:25	55.8 ± 12.5	59.2 ± 12.6	NR	NR	NR	NR	22/10/13/0	34/21/28/0	255.5 ± 58.1	208.3 ± 36.6
Huang [23]	2010	China	Single	Retrospective	66	69	40:26	39:30	55.8 ± 9.2	56.4 ± 10.6	NR	NR	NR	NR	18/21/26/1	16/21/30/2	266.1 ± 55.1	223.8 ± 26.8
Zhao [26]	2011	China	Single	Retrospective	346	313	248:98	221:92	51.4 ± 11.5	52.6 ± 12.4	NR	NR	NR	NR	42/99/199/6	37/87/181/8	211.0 ± 56.0	204.0 ± 41.0
Chun [27]	2012	Korea	Single	Retrospective	52	67	30:22	48:19	61.1 ± 12.6	60.8 ± 11.1	22.8 ± 2.8	22.9 ± 3.0	NR	NR	NR	NR	207.7 ± 40.6	159.9 ± 39.0
Gordon [28]	2013	Japan	Single	Retrospective	66	135	48:18	93:39	64.0 ± 12.1	67.0 ± 11.9	22.8 ± 3.5	22.1 ± 3.3	NR	NR	22/29/15/0	11/50/71/0	291.0 ± 47.8	235.0 ± 55.8
Hosoda [29]	2015	Japan	Single	Retrospective	32	44	21:11	33:11	60.6 ± 11.1	64.6 ± 9.3	23.2 ± 3.4	23.0 ± 3.0	NR	NR	22/5/4/1	31/7/4/2	297.0 ± 12.0	226.0 ± 10.0
Matsuda [32]	2018	Japan	Single	Retrospective	61	67	48:13	36:31	69.5 ± 11.6	70.2 ± 11.1	23.8 ± 4.9	20.9 ± 4.0	27/20/14/0	8/33/26/0	12/34/15/0	9/29/29/0	321.0 ± 77.5	245.0 ± 50.8
Shibuya [36]	2019	Japan	Single	Retrospective	87	27	61:26	15:12	67.5 ± 1.3	69.4 ± 2.3	23.1 ± 0.3	22.5 ± 0.6	36/34/17/0	4/12/11/0	21/33/33/0	2/9/16/0	237.5 ± 50.4	197.6 ± 44.9

Intraoperative blood loss, ml		Postoperative hospital stay, days		Retrieved lymph nodes, *n*		Postoperative complications, *n*		Recurrence, *n*		3-year DFS, %		3-year OS, %	
Lap	Open	Lap	Open	Lap	Open	Lap	Open	Lap	Open	Lap	Open	Lap	Open
RCTs													
105.5 ± 88.6	117.3 ± 84.5	10.8 ± 5.9	11.3 ± 7.6	36.1 ± 16.7	36.9 ± 16.1	79	67	NR	NR	NR	NR	NR	NR
NR	NR	9.8 ± 7.0	9.1 ± 5.5	37.0 ± 13.4	39.7 ± 13.3	17	18	NR	NR	80.1	81.9	NR	NR
91.4 ± 90.9	117.5 ± 103.5	9.9 ± 3.7	10.9 ± 5.2	29.5 ± 10.4	31.4 ± 12.3	29	39	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	36.1 ± 16.7	36.9 ± 16.1	NR	NR	95	85	76.5	77.8	83.1	85.2
152.4 ± 260.5	225.0 ± 211.5	8.0 ± 6.3	9.1 ± 6.3	46.8 ± 18.0	47.2 ± 16.2	77	113	76	72	80.3	81.3	90.6	90.3

Table 1 (continued)

Intraoperative blood loss, ml		Postoperative hospital stay, days		Retrieved lymph nodes, *n*		Postoperative complications, *n*		Recurrence, *n*		3-year DFS, %		3-year OS, %	
Lap	Open	Lap	Open	Lap	Open	Lap	Open	Lap	Open	Lap	Open	Lap	Open
Case-matched studies													
NR	NR	7.0 ± 11.0	9.0 ± 4.3	31.0 ± 11.0	37.0 ± 20.3	2	8	NR	NR	NR	NR	70.9	56.8
200.0 ± 125.0	300.0 ± 250.0	12.0 ± 7.8	17.0 ± 9.3	35.0 ± 14.0	38.0 ± 15.0	2	3	NR	NR	NR	NR	NR	NR
200.0 ± 40.0	260.0 ± 25.0	8.0 ± 1.7	12.0 ± 2.5	20.0 ± 1.2	21.0 ± 1.0	9	13	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	397	451	NR	NR	NR	NR	NR	NR
NR	NR	9.6 ± 4.3	11.5 ± 5.1	30.5 ± 15.5	32.8 ± 16.9	8	8	NR	NR	86.3	75.3	98	86.9
140.8 ± 170.9	180.3 ± 165.3	11.8 ± 8.3	15.8 ± 13.7	26.0 ± 10.6	26.1 ± 12.3	10	12	NR	NR	NR	NR	NR	NR
92.0 ± 79.1	117.0 ± 102.9	11.5 ± 4.5	13.1 ± 5.1	30.2 ± 10.8	28.1 ± 10.1	24	32	49	47	64	60.3	68.7	66.6
78.7 ± 98.9	121.3 ± 115.5	13.0 ± 8.3	14.1 ± 8.6	31.3 ± 10.8	30.5 ± 12.1	68	80	140	137	65.1	63.2	76.1	74.2
Cohort studies													
160.0 ± 87.5	241.3 ± 91.3	7.0 ± 4.0	10.3 ± 3.3	30.5 ± 16.5	41.3 ± 17.8	4	4	8	6	53.4	71.4	88.2	77.2
110.0 ± 25.0	196.0 ± 30.0	8.6 ± 1.2	12.1 ± 2.5	23.5 ± 6.0	21.0 ± 7.5	6	10	22	31	NR	NR	81.2	74.7
333.3 ± 89.2	440.6 ± 156.7	9.8 ± 6.9	11.1 ± 5.6	35.6 ± 14.2	38.3 ± 11.4	7	10	6	17	NR	NR	NR	NR
131.9 ± 88.7	342.3 ± 178.7	9.2 ± 3.4	11.4 ± 4.6	25.8 ± 12.5	24.5 ± 10.3	4	11	NR	NR	NR	NR	NR	NR
128.0 ± 85.0	301.0 ± 156.0	7.9 ± 3.6	10.7 ± 5.8	33.2 ± 12.5	32.8 ± 15.6	24	41	147	141	NR	NR	57.2	54.1
NR	NR	7.0 ± 16.1	7.0 ± 1.6	39.1 ± 15.2	39.3 ± 11.2	5	6	NR	NR	NR	NR	NR	NR
107.0 ± 97.3	495.0 ± 432.5	8.4 ± 4.1	18.1 ± 16.1	35.9 ± 12.6	36.6 ± 14.5	9	33	NR	NR	NR	NR	NR	NR
90.0 ± 27.0	314.0 ± 23.0	9.7 ± 0.5	10.3 ± 0.4	44.0 ± 2.0	39.0 ± 2.0	1	4	3	2	NR	NR	NR	NR
50.0 ± 142.5	333.0 ± 478.5	7.0 ± 13.3	9.0 ± 9.8	29.0 ± 12.5	33.0 ± 12.8	3	5	9	16	NR	NR	NR	NR
34.1 ± 56.4	157.1 ± 129.5	12.3 ± 5.7	14.5 ± 6.8	44.8 ± 14.3	41.7 ± 18.7	4	1	NR	NR	NR	NR	NR	NR

RCTs randomized controlled trials, *Lap* laparoscopy, *CI* confidence interval, *HR* hazard ratio, *DFS* disease-free survival, *OS* overall survival, *NR* not reported

Table 2 Quality assessment of the included RCTs based on the revised Cochrane risk-of-boas tool

Study	Randomization process	Intended interventions	Missing outcome data	Measurement of outcome	Reported result	Overall risk-of-bias judgement
Hu [31]	Low	Low	Low	Low	Low	Low
Park [33]	Low	Low	Low	Low	Low	Low
Wang [37]	Low	Low	Low	Low	Low	Low
Yu [38]	Low	Low	Low	Low	Low	Low
Hyung [40]	Low	Low	Low	Low	Low	Low

RCTs randomized controlled trials

Materials and methods

We performed a systematic review and meta-analysis in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [15].

Literature search strategy

We searched the PubMed, Cochrane Central Register of Controlled Trials, and Web of Science databases for studies in which LDG was compared with ODG for AGC published from inception until July 2021. The search terms used were "laparoscopy" OR "laparoscopic" AND "stomach neoplasms" OR "gastric cancer" OR "stomach cancer" AND "open gastrectomy" AND "distal gastrectomy" (Appendix S1). The reference lists of all relevant articles were evaluated to identify other related papers. The study title, study authors, year of publication, and study characteristics were checked, and duplicates were removed. Two authors (R.O. and Y.M.) independently reviewed the title and abstract of articles after eliminating duplicates. The same authors then evaluated the full text according to the study eligibility criteria described below. In cases of disagreement, the authors discussed or consulted a third author until agreement was reached.

Eligibility

The inclusion criteria were as follows: (1) RCTs, case-matched studies, or cohort studies; (2) studies that compared LDG versus ODG for AGC; (3) studies that provided available outcome data; and (4) articles written in English.

The exclusion criteria were as follows: (1) studies without appropriate data; (2) laboratory or animal studies; and (3) papers identified as letters, comments, correspondence, editorials, or reviews.

Table 3 Quality assessment of observational studies

Studies	Selection				Comparability	Outcome			Total star
	1	2	3	4	5	6	7	8	
Scatizzi [24]	*	*	*	*	**	*		*	8
Shuang [25]	*	*	*	*	**	*	*	*	9
Zhang [30]	*	*	*	*	**	*	*	*	9
Yoshida [34]	*	*	*	*	**	*	*	*	9
Kim [35]	*	*	*	*	**	*	*	*	9
Garbarino [39]	*	*	*	*	**	*	*	*	9
Wang [41]	*	*	*	*	**	*	*	*	9
Huang [42]	*	*	*	*	**	*	*	*	9
Hur [20]	*	*	*	*	*	*	*	*	8
Du [21]	*	*	*	*	*	*	*	*	8
Hwang [22]	*	*	*	*	*	*	*	*	8
Huang [23]	*	*	*	*	*	*	*	*	8
Zhao [26]	*	*	*	*	*	*	*	*	8
Chun [27]	*	*	*	*	*	*	*	*	8
Gordon [28]	*	*	*	*	*	*	*	*	8
Hosoda [29]	*	*	*	*	*	*	*	*	8
Matsuda [32]	*	*	*	*	*	*	*	*	8
Shibuya [36]	*	*	*	*	*	*	*	*	8

Data extraction and outcome parameters

Two authors (R.O. and Y.M.) collected the data independently. The following data were extracted: population characteristics (year of publication, study design, country in which the study was performed, number of patients), short-term outcome parameters (operative time, intraoperative blood loss, postoperative hospital stay, retrieved lymph nodes, postoperative complications), and long-term outcome parameters (recurrence, 3-year disease-free survival (DFS), 3-year overall survival (OS)). The collected data were double-checked by each author, and any discrepancies were resolved by rechecking and discussion.

Assessment of study quality and risk of bias

RCTs were assessed using the revised Cochrane risk-of-bias tool [16]. For observational studies, the Newcastle–Ottawa quality assessment scale (NOS) was used to assess the quality of the included studies [17]. The score ranged from 0 to 9 stars, and studies with a score of ≥ 6 were considered to be of a high quality. For each outcome, a funnel plot was used to examine the publication bias among the included studies.

Statistical analyses

All statistics analyses were carried out using Review Manager version 5.3 software (The Cochrane Collaboration, Oxford, UK). The random effects model were used. Heterogeneity was assessed using the I^2 statistic. Odds ratio (OR) with corresponding 95% confidence interval (CI) was evaluated for categorical variables. The mean difference (MD) with corresponding CI was assessed for continuous variables. The mean with standard deviation (SD) was estimated from the median, the range, and the size of a sample using the method of Hozo et al. [18]. Survival outcome was analyzed according to the pooled hazard ratio (HR) and 95% CI. If the HR was not provided directly, an estimated HR was calculated from Kaplan–Meier curves according to the method of Tierney et al. [19]. The P value of < 0.05 was defined statistically significant.

Results

Study characteristics

The comprehensive electronic literature search detected 1385 articles. In total, 392 articles were removed due to duplication. According to the eligible criteria, 849 were excluded by title/abstract screening. The remaining 144 articles were evaluated by full-text review. Ultimately, 23 studies with 13698 patients were included (Fig. 1) [20–42]. Although two RCTs were from the same trial (CLASS-01 trial, NCT01609309) [31, 38], one reported the short-term outcomes [31] and the other was a follow-up that reported the long-term outcomes [38], so both were included in this study to analyze the results of each. The included studies were 5 RCTs, 8 case-matched studies, and 10 cohort studies. The characteristics of the included studies are summarized in Table 1.

The study quality and risk of bias

The risk of bias assessed using the revised Cochrane risk-of-bias tool is shown in Table 2. For overall risk-of-bias judgement, all included RCTs were rated as low risk of bias. The quality of the included observational studies was assessed using the NOS, and all studies were graded as a high quality (Table 3). In addition, we conducted a funnel plot analysis to assess the possibility of a publication bias (Fig. 2). The spread of the distribution of the effect sizes of the studies in the funnel plot was more pronounced in observational studies than in others.

Short-term outcomes

Operative time

A total of 21 studies with 6222 patients (4 RCTs with 2651 patients, 7 case-matched studies with 1792 patients, and 10 cohort studies with 1779 patients) reported operative time (Table 4). The meta-analysis showed that the operative time of the LDG group was significantly longer than that in the ODG group in RCTs (MD: 49.2, 95% CI: 29.38 to 69.02, $P < 0.00001$), case-matched studies (MD: 32.25, 95% CI: 15.2 to 55.3, $P = 0.0006$), and cohort studies (MD: 47.85, 95% CI: 29.37 to 66.33, $P < 0.00001$) (Fig. 3).

Intraoperative blood loss

In total, 17 studies with 4831 patients (3 RCTs with 1562 patients, 5 case-matched studies with 1612 patients, and 9 cohort studies with 1657 patients) revealed intraoperative blood loss (Table 4). The LDG group showed significantly less intraoperative blood loss than the ODG group in RCTs (MD: − 35.91, 95% CI: − 67.54 to − 4.28, $P = 0.03$), case-matched studies (MD: − 44.89, 95% CI: − 64.65 to − 25.12, $P < 0.00001$), and cohort studies (MD: − 179.3, 95% CI: − 235.81 to − 122.8, $P < 0.00001$) (Fig. 4).

Postoperative hospital stay

Twenty-one studies with 6222 patients (4 RCTs with 2651 patients, 7 case-matched studies with 1792 patients,

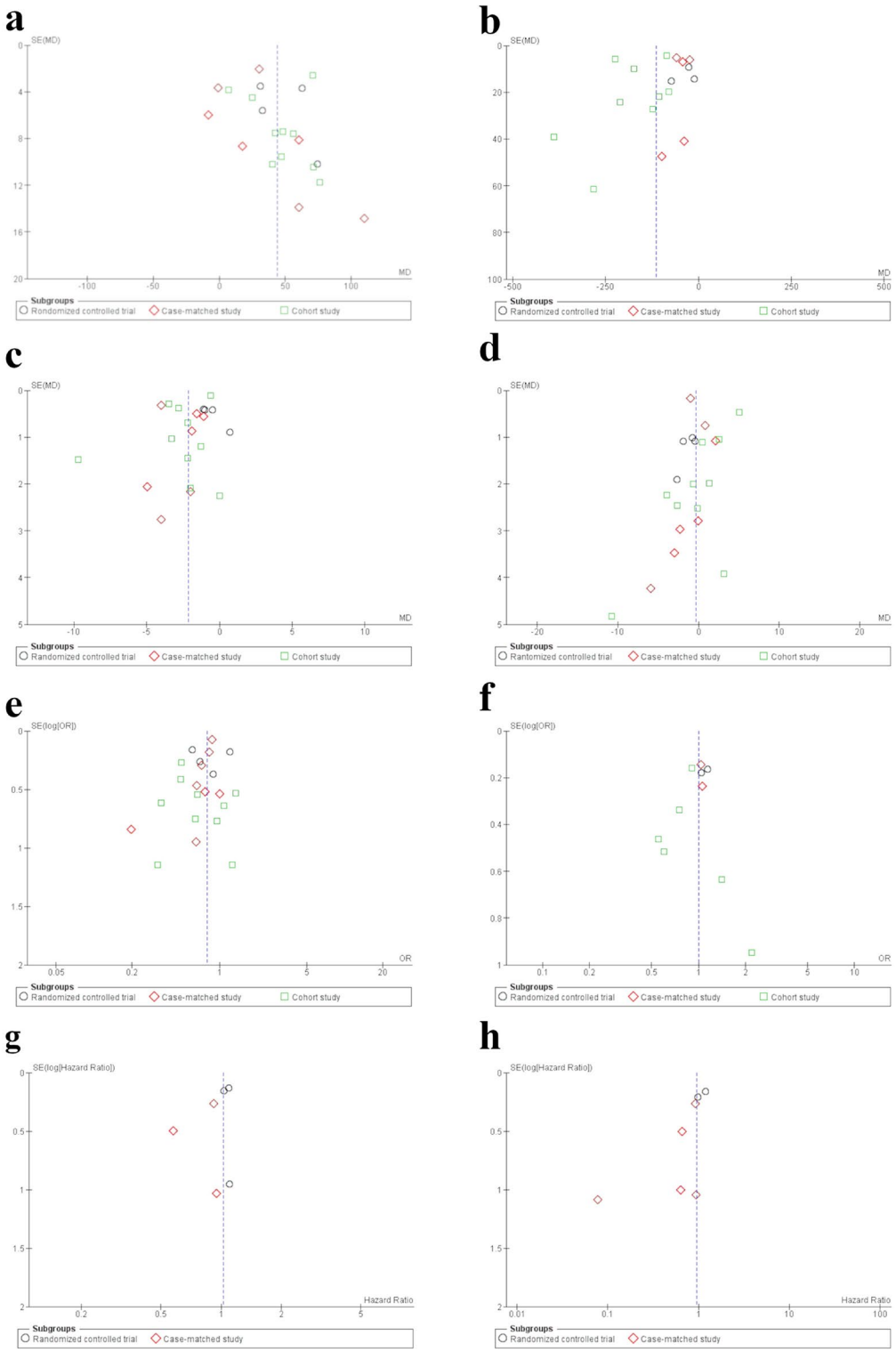

Fig. 2 Funnel plot of publication bias. **a** Operative time. **b** Intraoperative blood loss. **c** Postoperative hospital stay. **d** Retrieved lymph nodes. **e** Postoperative complications. **f** Recurrence. **g** The 3-year disease-free survival. **h** The 3-year overall survival

and 10 cohort studies with 1779 patients) showed postoperative hospital stay (Table 4). The LDG group had significantly less postoperative hospital stay than the ODG group in RCTs (MD: − 0.73, 95% CI: − 1.28 to − 0.19, $P = 0.009$), case-matched studies (MD: − 2.49, 95% CI: − 3.84 to − 1.13, $P = 0.0003$), and cohort studies (MD: − 2.75, 95% CI: − 4.1 to − 1.41, $P < 0.00001$) (Fig. 5).

Table 4 Summary of meta-analysis

		Randomized controlled trials				Case-matched studies				Cohort studies			
Outcomes	Measures	Study, *n*	Patients, *n*	Point estimation	95% CI	Study, *n*	Patients, *n*	Point estimation	95% CI	Study, *n*	Patients, *n*	Point estimation	95% CI
Short-term													
Operative time	MD	4	2651	49.2	29.38, 69.02	7	1792	32.25	15.2, 55.3	10	1779	47.85	29.37, 66.33
Intraoperative blood loss	MD	3	1562	− 35.91	− 67.54, − 4.28	5	1612	− 44.89	− 64.65, − 25.12	9	1657	− 179.3	− 235.81, − 122.81
Postoperative hospital stay	MD	4	2651	− 0.73	− 1.28, − 0.19	7	1792	− 2.49	− 3.84, − 1.13	10	1779	− 2.75	− 4.1, − 1.41
Retrieved lymph nodes	MD	4	2651	− 1.19	− 2.23, − 0.04	7	1792	− 0.14	− 1.63, − 1.35	10	1779	0.21	− 2.16, 2.58
Postoperative complications	OR	4	2651	0.82	0.56, 1.20	8	9268	0.84	0.74, 0.95	10	1779	0.6	0.44, 0.84
Long-term													
Recurrence	OR	2	2013	1.1	0.87, 1.39	2	1302	1.04	0.82, 1.32	6	1210	0.85	0.66, 1.09
3-year DFS	HR	3	2209	1.07	0.88, 1.31	3	1422	0.83	0.53, 1.30	NA	NA	NA	NA
3-year OS	HR	2	2013	1.11	0.87, 1.43	5	1552	0.68	0.38, 1.24	NA	NA	NA	NA

CI confidence interval, *MD* mean difference, *OR* odds ratio, *HR* hazard ratio, *DFS* disease-free survival, *OS* overall survival, *NA* not applicable

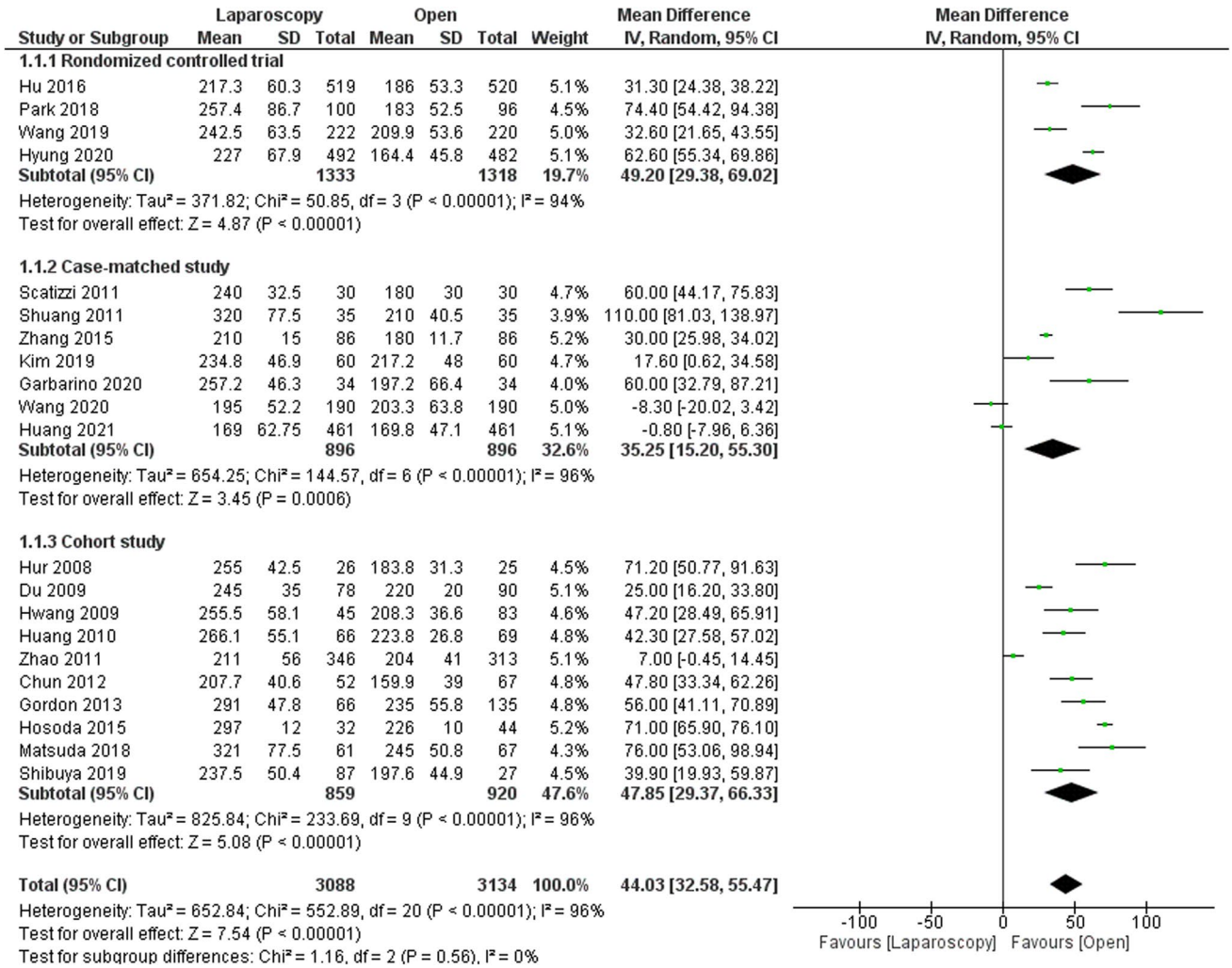

Fig. 3 Results of the meta-analysis of operative time stratified by study design

Number of retrieved lymph nodes

A total of 21 studies with 6222 patients (4 RCTs with 2651 patients, 7 case-matched studies with 1792 patients, and 10 cohort studies with 1779 patients) reported the number of retrieved lymph nodes (Table 4). The number of retrieved lymph nodes was significantly larger in the ODG group than in the LDG group in RCTs (MD: − 1.19, 95% CI: − 2.23 to − 0.04, $P = 0.04$). In contrast, there were no significant differences between the groups in case-matched studies (MD: − 0.14, 95% CI: − 1.63 to 1.35, $P = 0.85$) and cohort studies (MD: 0.21, 95% CI: − 2.16 to 2.58, $P = 0.86$) (Fig. 6).

Postoperative complications

Twenty-two studies with 13,698 patients (4 RCTs with 2651 patients, 8 case-matched studies with 9268 patients, and 10 cohort studies with 1779 patients) revealed the incidence of postoperative complications (Table 4). There were no significant differences between the two groups in RCTs (OR: 0.82, 95% CI: 0.56 to 1.20, $P = 0.30$). Conversely, the LDG group had a significantly lower incidence of postoperative complications than the ODG group in case-matched studies (OR: 0.84, 95% CI: 0.74 to 0.95, $P = 0.006$) and cohort studies (OR: 0.60, 95% CI: 0.44 to 0.84, $P = 0.002$) (Fig. 7).

Results of long-term outcomes

Recurrence

In total, 10 studies with 4525 patients (2 RCTs with 2013 patients, 2 case-matched studies with 1302 patients, and 6 cohort studies with 1210 patients) showed the incidence of recurrence (Table 4). There were no significant differences between the two groups in RCTs (OR: 1.10, 95% CI: 0.87 to 1.39, $P = 0.45$), case-matched studies (OR: 1.04, 95% CI: 0.82 to 1.32, $P = 0.76$), and

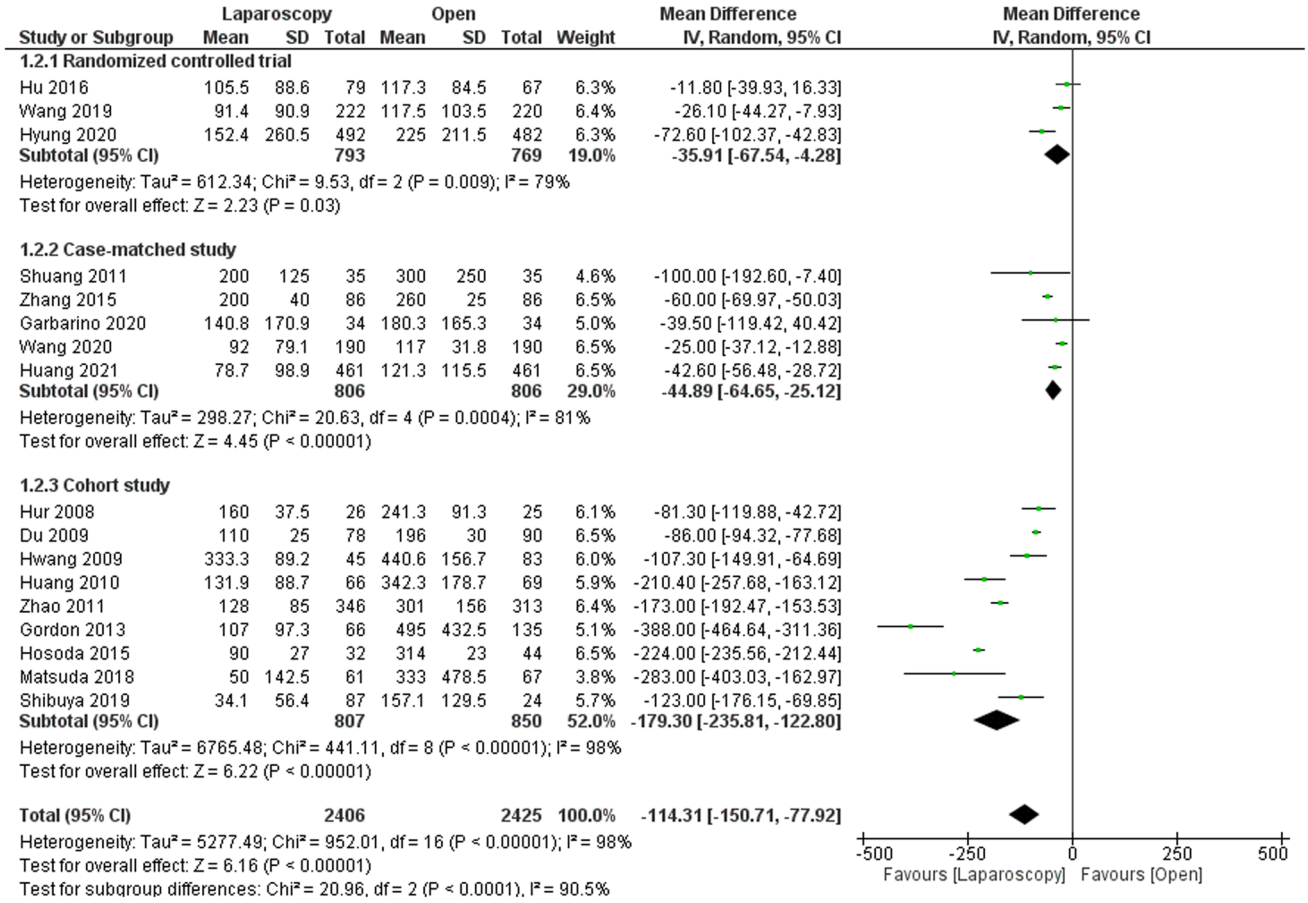

Fig. 4 Results of the meta-analysis of intraoperative blood loss stratified by study design

cohort studies (OR: 0.85, 95% CI: 0.66 to 1.09, $P=0.21$) (Fig. 8).

The 3-year DFS

Six studies with 3631 patients (3 RCTs with 2209 patients and 3 case-matched studies with 1422 patients) reported the 3-year DFS (Table 4). There were no significant differences between the two groups in RCTs (HR: 1.07, 95% CI: 0.88 to 1.31, $P=0.51$) and case-matched studies (HR: 0.83, 95% CI: 0.53 to 1.3, $P=0.42$) (Fig. 9). However, the LDG group tended to be correlated with favorable 3-year DFS in case-matched studies compared to in RCTs.

The 3-year OS

A total of 7 studies with 3565 patients (2 RCTs with 2013 patients and 5 case-matched studies with 1552 patients) showed the 3-year OS (Table 4). There were no significant differences between the two groups in RCTs (HR: 1.11, 95% CI: 0.87 to 1.43, $P=0.40$) and case-matched studies (HR: 0.68, 95% CI: 0.38 to 1.24, $P=0.21$) (Fig. 10). However, the LDG group tended to be associated to favorable 3-year OS in case-matched studies compared to in RCTs.

Discussion

In this study, we performed a meta-analysis including 23 studies for 5 short-term outcomes and 3 long-term outcomes. There was no difference in estimated treatment effects between RCTs and case-matched studies for all outcomes except for the number of retrieved lymph nodes and postoperative complications. For all analyzable items, the results of cohort studies were similar to those of case-matched studies. In terms of short-term outcomes, both RCTs and case-matched studies found significantly longer operative time, less intraoperative blood loss, and shorter postoperative hospital stay in LDG compared to ODG. Postoperative complications were significantly less in case-matched studies but not in RCTs. However, given

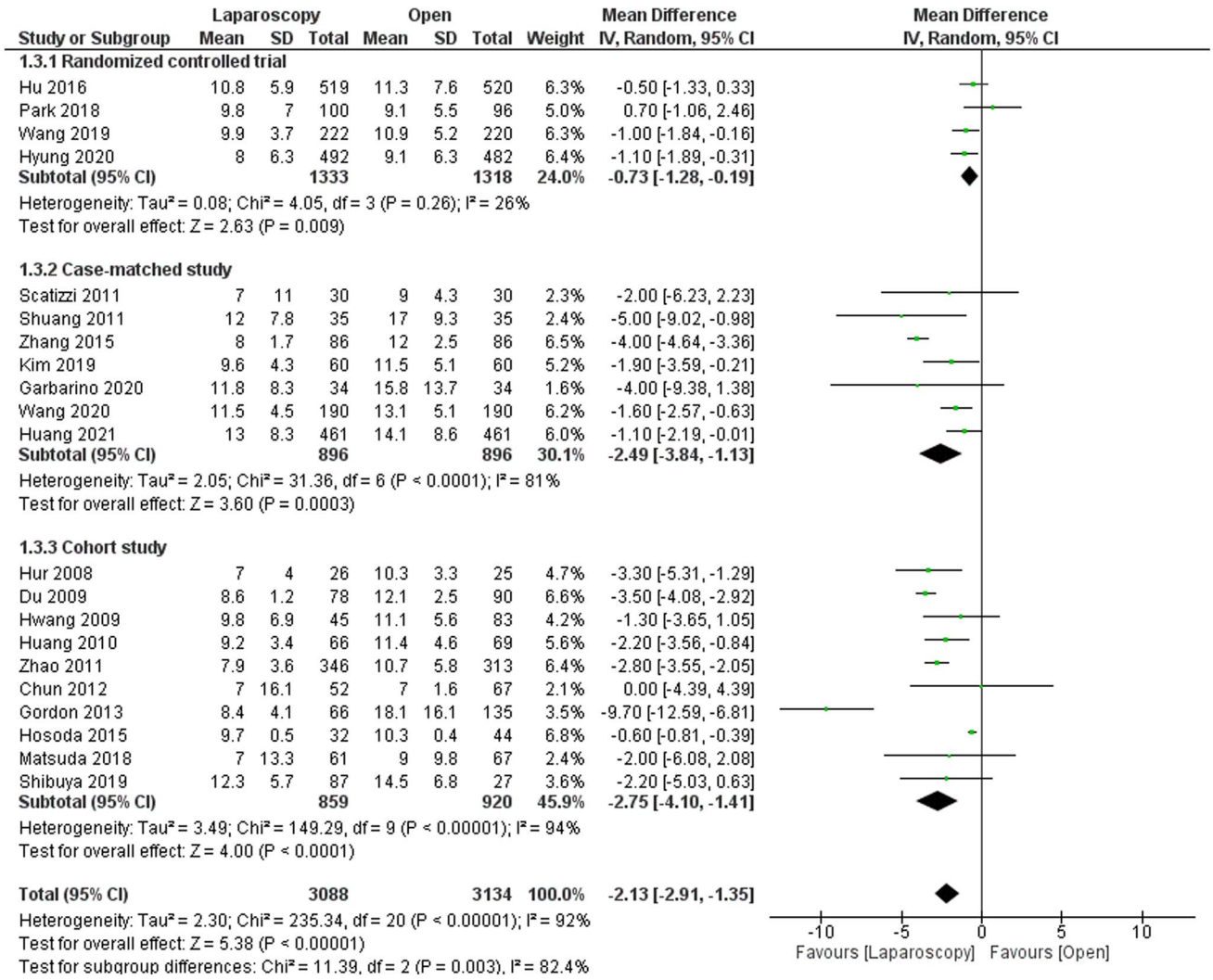

Fig. 5 Results of the meta-analysis of postoperative hospital stay stratified by study design

the distribution of the 95% CIs for postoperative complications, we considered the estimated treatment effect in both studies to be comparable. Regarding long-term outcomes, although LDG had relatively better 3-year DFS and 3-year OS in case–matched studies, there was no significant difference between LDG and ODG in both RCTs and case–matched studies. Thus, the findings of RCTs and case-matched studies were similar for almost all outcomes.

The estimated treatment effects of LDG in case-matched studies were intermediate between RCTs and cohort studies in terms of intraoperative blood loss, postoperative hospital stay, retrieved lymph nodes, and recurrence. RCTs can adjust for all confounders (including unknown ones), whereas propensity score matching, a typical case-matching method, has been shown to potentially miss some confounders [43]. Therefore, such differences in estimated treatment effects among study designs may be due to the different degree of adjustment for covariates in each study design. The amount of intraoperative blood loss, hospitalization period, number of retrieved lymph nodes, and recurrence are outcomes that can be objectively assessed from medical records, surgical records, pathology reports, and imaging findings. Hence, it is suggested that differences in research design may affect even objective endpoints. In addition, it has been reported that observational studies such as case-matched studies and cohort studies may overestimate treatment effects [44]. The nature of the overestimation of observational studies, which was also observed in this study regarding upper gastrointestinal surgery areas, is consistent with the results of a study

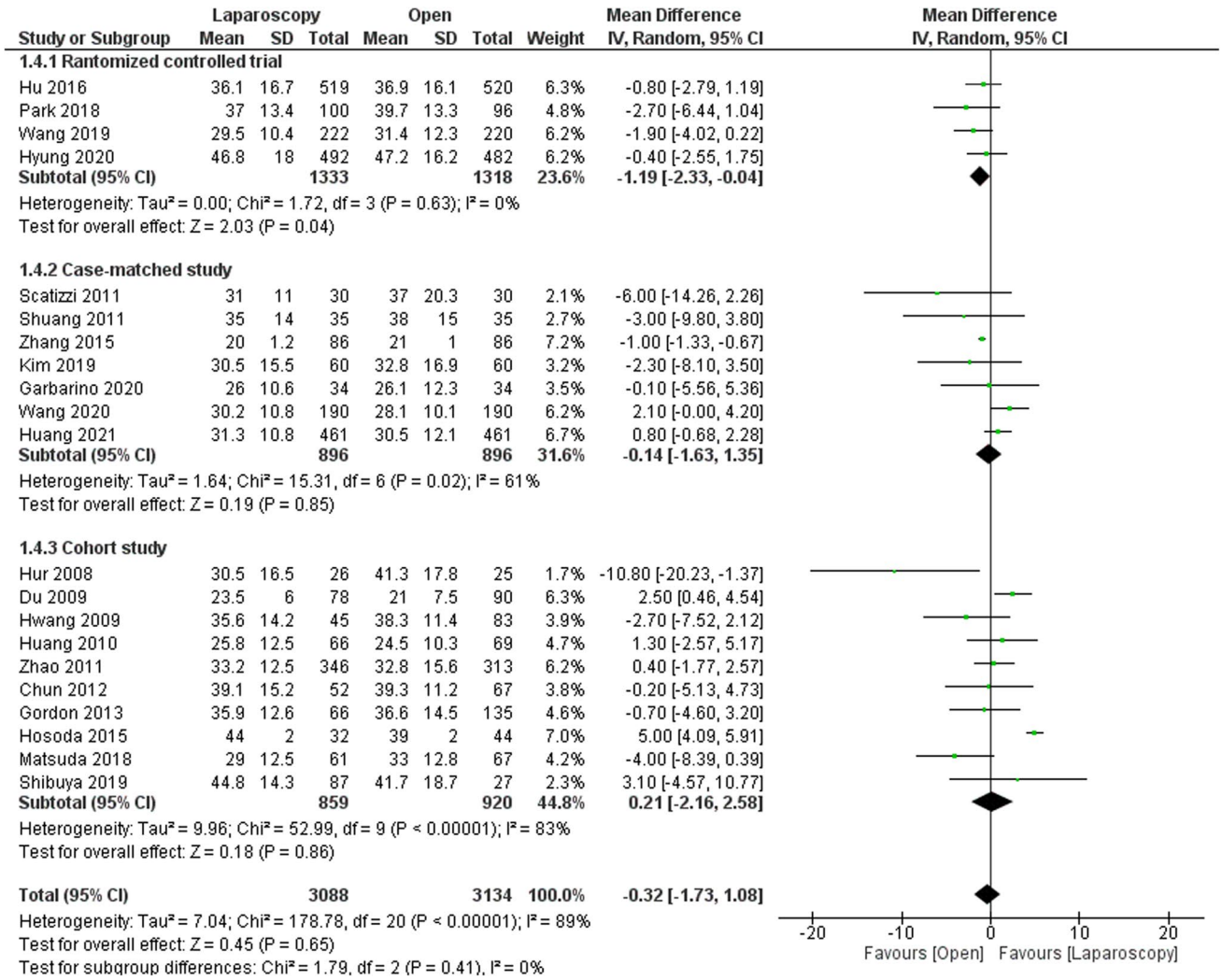

Fig. 6 Results of the meta-analysis of retrieved lymph nodes stratified by study design

conducted in the lower gastrointestinal surgery field [2], indicating that this review is significant in terms of accumulating evidence regarding differences in treatment effects among study designs in the gastrointestinal surgery field. Possible causes of overestimation in observational studies include missing data, possible crossover, publication bias, selective reporting of results, selection bias, outcome ascertainment bias, immortal-time bias, and residual confounding [43, 44]. Therefore, although observational studies are a very useful research tool in real clinical practice, the design and actual implementation of the analysis must be critically evaluated on a case-by-case basis in order to assess the true magnitude of treatment effects.

The strengths of this study are its novelty in the absence of similar studies in the field of upper gastrointestinal surgery and the inclusion of a relatively large number of studies to assess the differences among RCTs, case-matched studies, and cohort studies. However, there are several limitations of the present study. First, cohort studies lacked the long-term outcome data for calculating the HR needed to conduct a meta-analysis. These would have allowed us to examine in more detail the differences in long-term outcomes between the study designs. Second, only published studies were

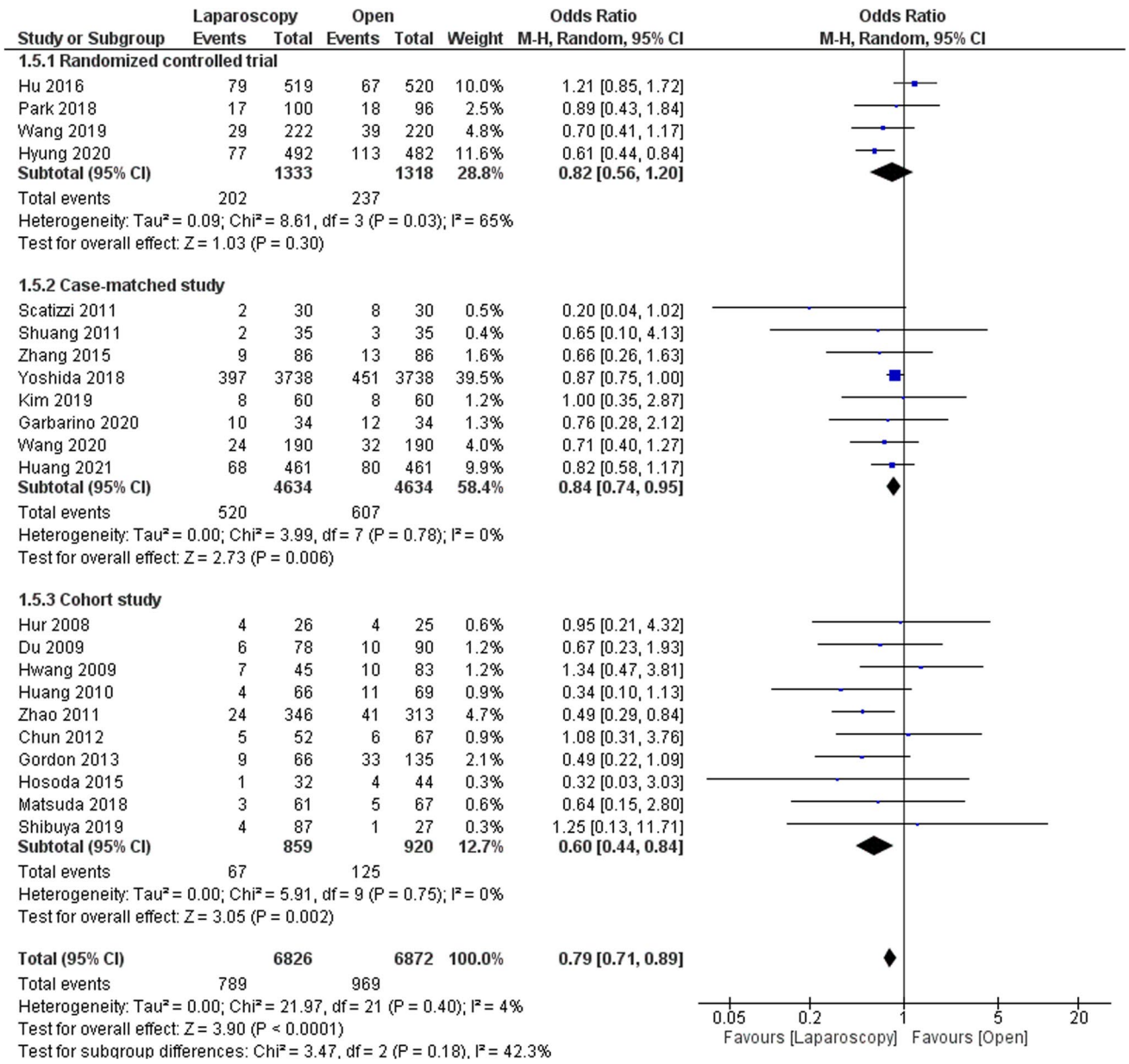

Fig. 7 Results of the meta-analysis of postoperative complications stratified by study design

included in the present study, which made it difficult to eliminate potential publication bias. Finally, most of the articles included in this review were conducted in East Asia. Therefore, more extensive studies should be conducted in other countries and regions to improve the quality of the research and to find general trends in differences among study designs.

In recent years, observational data representative of clinical practice has become available from nationwide clinical databases, such as the National Clinical Database (NCD) and the National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB) [45, 46]. Given this background, case-matched studies using modern design methods such as propensity score matching will become more and more important in the future because it is an efficient way to evaluate the effects of interventions in typical clinical settings [1]. In addition, the results of

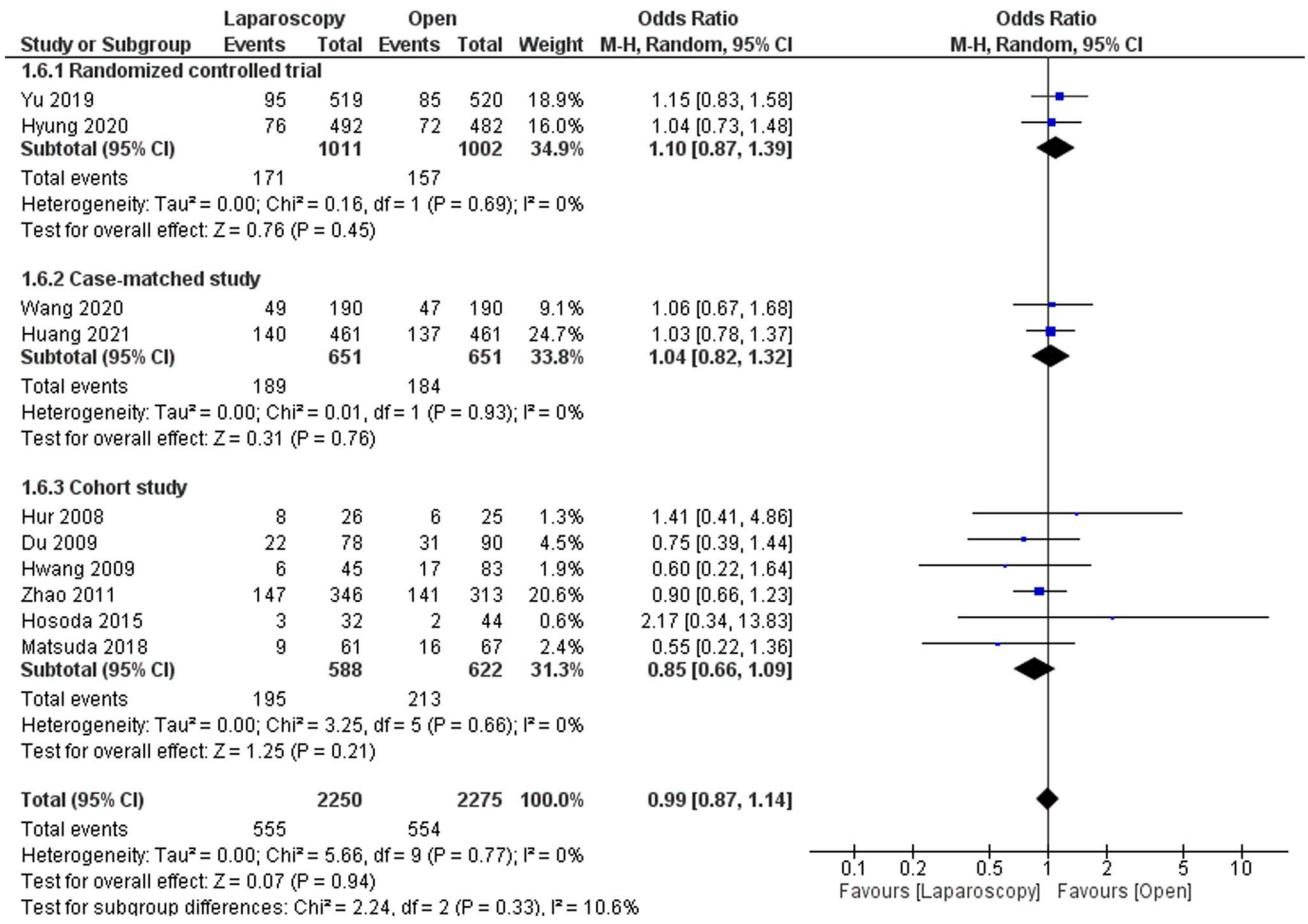

Fig. 8 Results of the meta-analysis of recurrence stratified by study design

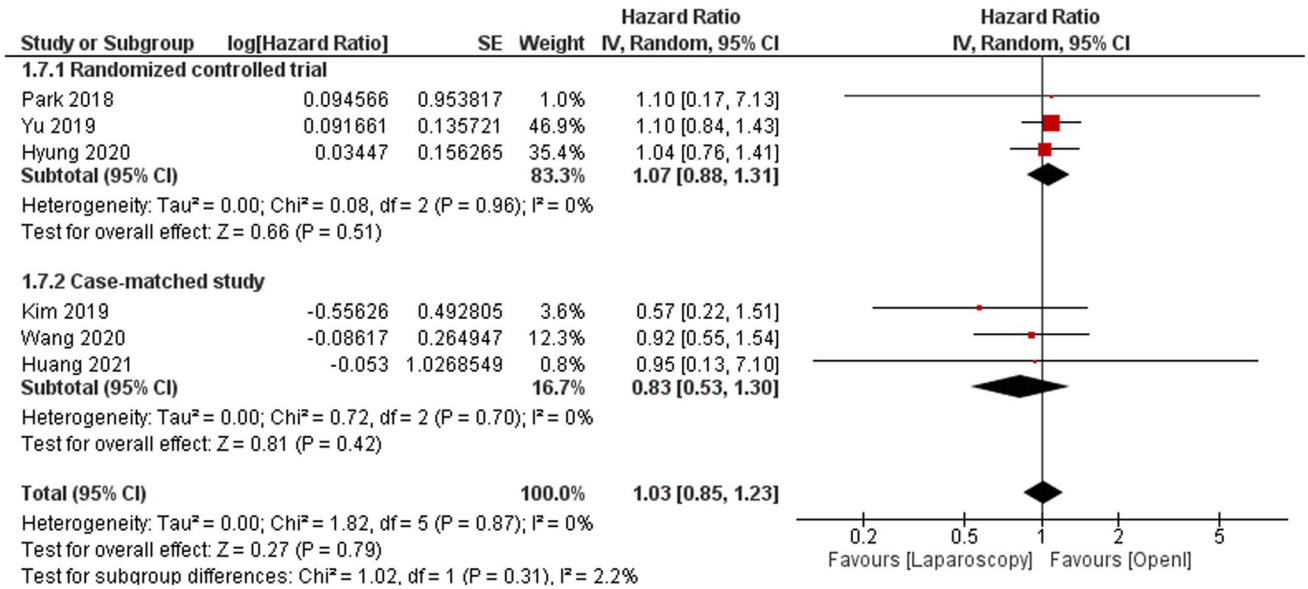

Fig. 9 Results of the meta-analysis of the 3-year disease-free survival stratified by study design

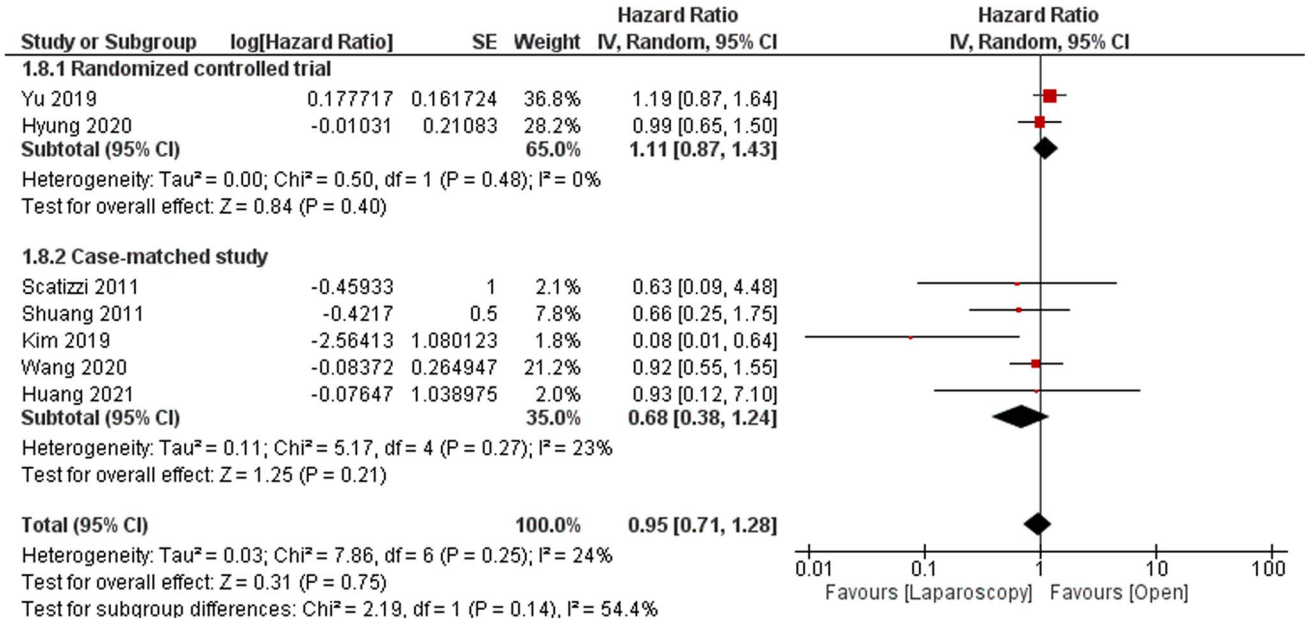

Fig. 10 Results of the meta-analysis of the 3-year overall survival stratified by study design

properly conducted and analyzed observational studies are expected to help prioritize research needs that should be addressed in more resource-intensive RCTs. Therefore, this study, which compares the estimated treatment effects of RCTs and observational studies, has important implications for clinical practice and future research.

Conclusion

Our analysis indicated that the estimated treatment effects of LDG for AGC in the case-matched study were almost the same as in the RCTs. However, to assess the true magnitude of the treatment effect, the design and actual implementation of the analysis must be critically evaluated.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00423-022-02454-3.

Authors' contributions Ryota Otsuka conceived and designed the study. Yasunori Matsumoto and Ryota Otsuka performed the literature search and data acquisition. Takashi Toyozumi, Hiroshi Suito, Tetsuro Isozaki, Yoshihiro Kurata, and Ryota Otsuka contributed to the data analysis and interpretation. The manuscript was prepared by Ryota Otsuka under the supervision of Hideki Hayashi, Masaya Uesato, Koichi Hayano, Kentaro Murakami, Masayuki Kano, and Hisahiro Matsubara. All authors read and approved the final manuscript.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Ethics approval This study was performed in line with the principles of the Declaration of Helsinki.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

References

1. Adamina M, Guller U, Weber WP, Oertli D (2006) Propensity scores and the surgeon. Br J Surg 93(4):389–394. https://doi.org/10.1002/bjs.5265
2. Hoshino N, Fukui Y, Hida K, Obama K (2021) Similarities and differences between study designs in short- and long-term outcomes of laparoscopic versus open low anterior resection for rectal cancer: a systematic review and meta-analysis of randomized, case-matched, and cohort studies. Ann Gastroenterol Surg 5(2):183–193. https://doi.org/10.1002/ags3.12409
3. Byar DP (1980) Why data bases should not replace randomized clinical trials. Biometrics 36(2):337–342
4. Benson K, Hartz AJ (2000) A comparison of observational studies and randomized, controlled trials. N Engl J Med 342(25):1878–1886. https://doi.org/10.1056/nejm200006223422506
5. Concato J, Shah N, Horwitz RI (2000) Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med 342(25):1887–1892. https://doi.org/10.1056/nejm200006223422507
6. Ioannidis JP, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG, Contopoulos-Ioannidis DG, Lau J (2001) Comparison of evidence of treatment effects in randomized and

nonrandomized studies. JAMA 286(7):821–830. https://doi.org/10.1001/jama.286.7.821
7. Ioannidis JP, Haidich AB, Lau J (2001) Any casualties in the clash of randomised and observational evidence? BMJ 322(7291):879–880. https://doi.org/10.1136/bmj.322.7291.879
8. Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70(1):41–55
9. D'Agostino RB Jr (1998) Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med 17(19):2265–2281. https://doi.org/10.1002/(sici)1097-0258(19981015)17:19%3c2265::aid-sim918%3e3.0.co;2-b
10. Ali MS, Groenwold RH, Belitser SV, Pestman WR, Hoes AW, Roes KC, Boer A, Klungel OH (2015) Reporting of covariate selection and balance assessment in propensity score analysis is suboptimal: a systematic review. J Clin Epidemiol 68(2):112–121. https://doi.org/10.1016/j.jclinepi.2014.08.011
11. Qiu J, Pankaj P, Jiang H, Zeng Y, Wu H (2013) Laparoscopy versus open distal gastrectomy for advanced gastric cancer: a systematic review and meta-analysis. Surg Laparosc Endosc Percutaneous Tech 23(1):1–7. https://doi.org/10.1097/SLE.0b013e3182747af7
12. Lu C, Zhou S, Peng Z, Chen L (2015) Quality of D2 lymphadenectomy for advanced gastric cancer: is laparoscopic-assisted distal gastrectomy as effective as open distal gastrectomy? Surg Endosc 29(6):1537–1544. https://doi.org/10.1007/s00464-014-3838-6
13. Aurello P, Sagnotta A, Terrenato I, Berardi G, Nigri G, D'Angelo F, Ramacciato G (2016) Oncologic value of laparoscopy-assisted distal gastrectomy for advanced gastric cancer: a systematic review and meta-analysis. J Min Access Surg 12(3):199–208. https://doi.org/10.4103/0972-9941.181283
14. Chen X, Feng X, Wang M, Yao X (2020) Laparoscopic versus open distal gastrectomy for advanced gastric cancer: a meta-analysis of randomized controlled trials and high-quality nonrandomized comparative studies. Eur J Surg Oncol: The journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology 46(11):1998–2010. https://doi.org/10.1016/j.ejso.2020.06.046
15. Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. PLoS Med 6(7):e1000097. https://doi.org/10.1371/journal.pmed.1000097
16. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng HY, Corbett MS, Eldridge SM, Emberson JR, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT (2019) RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ (Clinical research ed) 366:l4898. https://doi.org/10.1136/bmj.l4898
17. Stang A (2010) Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol 25(9):603–605. https://doi.org/10.1007/s10654-010-9491-z
18. Hozo SP, Djulbegovic B, Hozo I (2005) Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol 5:13. https://doi.org/10.1186/1471-2288-5-13
19. Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR (2007) Practical methods for incorporating summary time-to-event data into meta-analysis. Trials 8:16. https://doi.org/10.1186/1745-6215-8-16
20. Hur H, Jeon HM, Kim W (2008) Laparoscopy-assisted distal gastrectomy with D2 lymphadenectomy for T2b advanced gastric cancers: three years' experience. J Surg Oncol 98(7):515–519. https://doi.org/10.1002/jso.21155
21. Du XH, Li R, Chen L, Shen D, Li SY, Guo Q (2009) Laparoscopy-assisted D2 radical distal gastrectomy for advanced gastric cancer: initial experience. Chin Med J 122(12):1404–1407
22. Hwang SI, Kim HO, Yoo CH, Shin JH, Son BH (2009) Laparoscopic-assisted distal gastrectomy versus open distal gastrectomy for advanced gastric cancer. Surg Endosc 23(6):1252–1258. https://doi.org/10.1007/s00464-008-0140-5
23. Huang JL, Wei HB, Zheng ZH, Wei B, Chen TF, Huang Y, Guo WP, Hu B (2010) Laparoscopy-assisted D2 radical distal gastrectomy for advanced gastric cancer. Dig Surg 27(4):291–296. https://doi.org/10.1159/000281818
24. Scatizzi M, Kröning KC, Lenzi E, Moraldi L, Cantafio S, Feroci F (2011) Laparoscopic versus open distal gastrectomy for locally advanced gastric cancer: a case-control study. Updat Surg 63(1):17–23. https://doi.org/10.1007/s13304-011-0043-1
25. Shuang J, Qi S, Zheng J, Zhao Q, Li J, Kang Z, Hua J, Du J (2011) A case-control study of laparoscopy-assisted and open distal gastrectomy for advanced gastric cancer. J Gastrointest Surg: official journal of the Society for Surgery of the Alimentary Tract 15(1):57–62. https://doi.org/10.1007/s11605-010-1361-1
26. Zhao Y, Yu P, Hao Y, Qian F, Tang B, Shi Y, Luo H, Zhang Y (2011) Comparison of outcomes for laparoscopically assisted and open radical distal gastrectomy with lymphadenectomy for advanced gastric cancer. Surg Endosc 25(9):2960–2966. https://doi.org/10.1007/s00464-011-1652-y
27. Chun HT, Kim KH, Kim MC, Jung GJ (2012) Comparative study of laparoscopy-assisted versus open subtotal gastrectomy for pT2 gastric cancer. Yonsei Med J 53(5):952–959. https://doi.org/10.3349/ymj.2012.53.5.952
28. Gordon AC, Kojima K, Inokuchi M, Kato K, Sugihara K (2013) Long-term comparison of laparoscopy-assisted distal gastrectomy and open distal gastrectomy in advanced gastric cancer. Surg Endosc 27(2):462–470. https://doi.org/10.1007/s00464-012-2459-1
29. Hosoda K, Sakuramoto S, Katada N, Yamashita K, Moriya H, Mieno H, Kikuchi S, Watanabe M (2015) Laparoscopic versus open distal gastrectomy with D2 lymph node dissection for cT2 gastric cancer: a retrospective cohort study of short- and long-term outcomes. Int Surg 100(9–10):1315–1322. https://doi.org/10.9738/intsurg-d-15-00027.1
30. Zhang Y, Qi F, Jiang Y, Zhai H, Ji Y (2015) Long-term follow-up after laparoscopic versus open distal gastrectomy for advanced gastric cancer. Int J Clin Exp Med 8(8):13564–13570
31. Hu Y, Huang C, Sun Y, Su X, Cao H, Hu J, Xue Y, Suo J, Tao K, He X, Wei H, Ying M, Hu W, Du X, Chen P, Liu H, Zheng C, Liu F, Yu J, Li Z, Zhao G, Chen X, Wang K, Li P, Xing J, Li G (2016) Morbidity and mortality of laparoscopic versus open D2 distal gastrectomy for advanced gastric cancer: a randomized controlled trial. J Clin Oncol: official journal of the American Society of Clinical Oncology 34(12):1350–1357. https://doi.org/10.1200/jco.2015.63.7215
32. Matsuda S, Booka E, Mori K, Mihara K, Nishiya S, Handa K, Ono S, Ito Y, Shibutani S, Egawa T (2018) Comparison of laparoscopic distal gastrectomy with open distal gastrectomy for patients with advanced gastric cancer: a single-center analysis from a community hospital. Int Surg 103(11–12):585–592. https://doi.org/10.9738/intsurg-d-18-00011.1
33. Park YK, Yoon HM, Kim YW, Park JY, Ryu KW, Lee YJ, Jeong O, Yoon KY, Lee JH, Lee SE, Yu W, Jeong SH, Kim T, Kim S, Nam BH (2018) Laparoscopy-assisted versus open D2 distal gastrectomy for advanced gastric cancer: results from a randomized phase II multicenter clinical trial (COACT 1001). Ann Surg 267(4):638–645. https://doi.org/10.1097/sla.0000000000002168
34. Yoshida K, Honda M, Kumamaru H, Kodera Y, Kakeji Y, Hiki N, Etoh T, Miyata H, Yamashita Y, Seto Y, Kitano S, Konno

H (2018) Surgical outcomes of laparoscopic distal gastrectomy compared to open distal gastrectomy: a retrospective cohort study based on a nationwide registry database in Japan. Ann Gastroenterol Surg 2(1):55–64. https://doi.org/10.1002/ags3.12054

35. Kim SH, Chung Y, Kim YH, Choi SI (2019) Oncologic outcomes after laparoscopic and open distal gastrectomy for advanced gastric cancer: propensity score matching analysis. J Gastric Cancer 19(1):83–91. https://doi.org/10.5230/jgc.2019.19.e4
36. Shibuya K, Kawamura H, Takahashi S, Ohno Y, Ichikawa N, Yoshida T, Homma S, Ishizu H, Takahashi M, Taketomi A (2019) Short-term and long-term outcomes following laparoscopic gastrectomy for advanced gastric cancer compared with open gastrectomy. Surg Laparosc Endosc Percutaneous Tech 29(4):297–303
37. Wang Z, Xing J, Cai J, Zhang Z, Li F, Zhang N, Wu J, Cui M, Liu Y, Chen L, Yang H, Zheng Z, Wang X, Gao C, Wang Z, Fan Q, Zhu Y, Ren S, Zhang C, Liu M, Ji J, Su X (2019) Short-term surgical outcomes of laparoscopy-assisted versus open D2 distal gastrectomy for locally advanced gastric cancer in North China: a multicenter randomized controlled trial. Surg Endosc 33(1):33–45. https://doi.org/10.1007/s00464-018-6391-x
38. Yu J, Huang C, Sun Y, Su X, Cao H, Hu J, Wang K, Suo J, Tao K, He X, Wei H, Ying M, Hu W, Du X, Hu Y, Liu H, Zheng C, Li P, Xie J, Liu F, Li Z, Zhao G, Yang K, Liu C, Li H, Chen P, Ji J, Li G (2019) Effect of laparoscopic vs open distal gastrectomy on 3-year disease-free survival in patients with locally advanced gastric cancer: the CLASS-01 randomized clinical trial. JAMA 321(20):1983–1992. https://doi.org/10.1001/jama.2019.5359
39. Garbarino GM, Costa G, Laracca GG, Castagnola G, Mercantini P, Di Paola M, Vita S, Masoni L (2020) Laparoscopic versus open distal gastrectomy for locally advanced gastric cancer in middle-low-volume centers in Western countries: a propensity score matching analysis. Langenbeck's Arch Surg 405(6):797–807. https://doi.org/10.1007/s00423-020-01951-7
40. Hyung WJ, Yang HK, Park YK, Lee HJ, An JY, Kim W, Kim HI, Kim HH, Ryu SW, Hur H, Kim MC, Kong SH, Cho GS, Kim JJ, Park DJ, Ryu KW, Kim YW, Kim JW, Lee JH, Han SU (2020) Long-term outcomes of laparoscopic distal gastrectomy for locally advanced gastric cancer: the KLASS-02-RCT randomized clinical trial. J Clin Oncol: official journal of the American Society of Clinical Oncology 38(28):3304–3313. https://doi.org/10.1200/jco.20.01210
41. Wang JB, Zhong Q, Chen QY, Lin GT, Liu ZY, Huang XB, Xie JW, Lin JX, Lu J, Cao LL, Lin M, Tu RH, Huang ZN, Lin JL, Zheng HL, Zheng CH, Huang CM, Li P (2020) Well-designed retrospective study versus small-sample prospective study in research based on laparoscopic and open radical distal gastrectomy for advanced gastric cancer. Surg Endosc 34(10):4504–4515. https://doi.org/10.1007/s00464-019-07237-4
42. Huang ZN, Ma Y, Chen QY, Zheng CH, Li P, Xie JW, Wang JB, Lin JX, Lu J, Cao LL, Lin M, Tu RH, Lin JL, Zheng HL, Huang CM (2021) Potential survival benefits of open over laparoscopic radical gastrectomy for gastric cancer patients beyond three years after surgery: result from multicenter in-depth analysis based on propensity matching. Surg Endosc. https://doi.org/10.1007/s00464-021-08430-0
43. Lonjon G, Porcher R, Ergina P, Fouet M, Boutron I (2017) Potential pitfalls of reporting and bias in observational studies with propensity score analysis assessing a surgical procedure: a methodological systematic review. Ann Surg 265(5):901–909. https://doi.org/10.1097/sla.0000000000001797
44. Dahabreh IJ, Sheldrick RC, Paulus JK, Chung M, Varvarigou V, Jafri H, Rassen JA, Trikalinos TA, Kitsios GD (2012) Do observational studies using propensity score methods agree with randomized trials? A systematic comparison of studies on acute coronary syndromes. Eur Heart J 33(15):1893–1901. https://doi.org/10.1093/eurheartj/ehs114
45. Nakayama T, Imanaka Y, Okuno Y, Kato G, Kuroda T, Goto R, Tanaka S, Tamura H, Fukuhara S, Fukuma S, Muto M, Yanagita M, Yamamoto Y (2017) Analysis of the evidence-practice gap to facilitate proper medical care for the elderly: investigation, using databases, of utilization measures for National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB). Environ Health Prev Med 22(1):51. https://doi.org/10.1186/s12199-017-0644-5
46. Kakeji Y, Takahashi A, Hasegawa H, Ueno H, Eguchi S, Endo I, Sasaki A, Takiguchi S, Takeuchi H, Hashimoto M, Horiguchi A, Masaki T, Marubashi S, Yoshida K, Gotoh M, Konno H, Yamamoto H, Miyata H, Seto Y, Kitagawa Y (2020) Surgical outcomes in gastroenterological surgery in Japan: report of the National Clinical Database 2011–2018. Ann Gastroenterol Surg 4(3):250–274. https://doi.org/10.1002/ags3.12324