Supplemental Online Content

Turri G, Ostuzzi G, Vita G, et al. Treatment of locally advanced rectal cancer in the era of total neoadjuvant therapy: a systematic review and network meta-analysis. *JAMA Netw Open*. 2024;7(6):e2414702. doi:10.1001/jamanetworkopen.2024.14702

eFigure 1. PRISMA Flowchart eAppendix 1. Search Strategy and Data Extraction eMethods 1. Complete Statistical Methodology eTable 1. List of Studies Included/Excluded/Ongoing/Awaiting Assessment eTable 2. Characteristics of Included Studies eTable 3. Summary of Inclusion and Exclusion Criteria of Included Studies eFigure 2. Risk of Bias of Included Studies eFigure 3. Transitivity Assessment and Meta-Regression eAppendix 2. Primary Outcome: Patients With Pathologic Complete Response eAppendix 3. Primary Outcome: Assessment of Publication Bias eAppendix 4. Primary Outcome: Number Needed to Treat eAppendix 5. Primary Outcome: CINeMA eAppendix 6. Primary Outcome: Sensitivity Analyses eAppendix 7. Tolerability of Treatment eAppendix 8. Toxicity of Treatment eAppendix 9. Dropouts by Any Cause eAppendix 10. Preoperative Treatment Related Deaths eAppendix 11. Rate of Randomized Patients Who Underwent Surgery eAppendix 12. Rate of R0 Resections eAppendix 13. Rate of Negative CRM eAppendix 14. Rate of ypN0 eAppendix 15. Rate of Postoperative Complications Clavien-Dindo 3 or Greater eAppendix 16. Rate of Anastomotic Leak eAppendix 17. Locoregional Recurrence at 3 Years eAppendix 18. Locoregional Recurrence at 5 Years eAppendix 19. Locoregional Failure at 3 Years eAppendix 20. Locoregional Failure at 5 Years eAppendix 21. Distant Recurrence at 3 Years eAppendix 22. Distant Recurrence at 5 Years eAppendix 23. Disease-Free Survival at 3 Years eAppendix 24. Disease-Free Survival at 5 Years eAppendix 25. Overall Survival at 3 Years eAppendix 26. Overall Survival at 5 Years eReferences

This supplemental material has been provided by the authors to give readers additional information about their work.

eFigure 1. PRISMA Flowchart



eAppendix 1. Search Strategy and Data Extraction

We searched without time or language restrictions Medline, EMBASE, Cochrane Central Register of Controlled Trials (CENTRAL), and Web of Science Core Collection electronic databases. Further, we searched ClinicalTrials.gov for unpublished studies. We searched records from database inception to June 8, 2023 (full search strategy attached below). Additionally, we screened references of assessed articles for other potential matches for inclusion.

PubMed (08/06/2023; 561 results)

(("cancer"[Title/Abstract] OR "oncolog*"[Title/Abstract] OR "adenocarcinoma"[Title/Abstract] OR "tumor"[Title/Abstract] OR "neoplas*"[Title/Abstract]) AND ("rectum"[Title/Abstract] OR "rectal"[Title/Abstract]) AND ("neoadjuvant"[Title/Abstract] OR "preoperat*"[Title/Abstract])) AND (randomizedcontrolledtrial[Filter])

Web of Science, Core Collection (08/06/2023; 329 results)

#1 (((((TI=(cancer)) OR TI=(oncolog*)) OR TI=(adenocarcinoma)) OR TI=(tumor)) OR TI=(neoplas*))

#2 (TI=(rectum)) OR TI=(rectal)

#3 (TI=(neoadjuvant)) OR TI=(preoperat*)

((((#1) AND #2) AND #3) AND ALL=(randomized controlled trial)) AND LA=(English)

Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years

CENTRAL (08/06/2023; 837 results)

Neoadjuvant therapy rectal cancer in Title Abstract Keyword AND English in Language - (Word variations have been searched)

ClinicalTrials.gov (08/06/2023, 399 results)

Neoadjuvant therapy; preoperative therapy, Condition: Rectal cancer, Filters: Interventional studies, Adults, Older Adults. Also searched for synonyms: preoperative therapy, induction therapy, treatment, rectal carcinoma, cancer of the rectum.

Two authors (GT, GO) independently assessed titles, abstracts and full texts of potentially relevant articles, and extracted data following recommendations of the Cochrane Handbook for Systematic Reviews of Interventions. Two authors (GT, GV) assessed the methodologic quality of included studies using the Cochrane Risk of Bias version 2 (RoB2) tool. Disagreements were resolved by discussion and consensus with a third senior author (CP, CB).

eMethods 1. Complete Statistical Methodology

We performed a standard pairwise, random-effects meta-analysis for every comparison, and, for each outcome, we also conducted a NMA with a random-effects model in a frequentist framework, using RStudio (version 2023.06.0-421) *netmeta* package and Stata (version 17.0) *mvmeta* package. For dichotomous outcomes, we calculated and pooled relative risks (RRs) with 95% confidence intervals (CIs).

We calculated dichotomous data on a strict intention-to-treat (ITT) basis, considering the total number of randomized participants as denominator. For the primary outcome, where participants had been excluded from the trial before the endpoint, we assumed that they experienced a negative outcome by the end of the trial.

When relevant outcomes were not reported, we asked trial authors to supply the data. In the absence of data from authors, we employed validated statistical methods to impute missing outcomes, with due consideration of the possible bias of these procedures, in accordance with the Cochrane Handbook¹. When standard deviations (SDs) were not reported and not supplied by authors upon request, we calculated them based on the standard error (SE) or t-statistics or P values¹.

For the primary outcome, we calculated the number-needed-to-treat (NNT), defined as the number of individuals needed to be treated with one treatment versus another for one individual to have an additional desirable or undesirable outcome (number-needed-to-treat-to-benefit, NNTB, or to-harm, NNTH, respectively)², using validated methodology³ (see Supplement for details).

For pairwise meta-analyses, we assessed heterogeneity by visual inspection of forest plots, and by the I^2 statistics. For the NMA, common heterogeneity across all comparisons was assumed and estimated in each network^{4,5}.

We assessed global heterogeneity by using the τ^2 (low: $\tau^2 \le 0.010$, moderate: $0.010 < \tau^2 \le 0.242$, high: $\tau^2 > 0.242$) and the I² (low, 0–40%; moderate, 30–60%; substantial, 50–90%; and considerable, 75–100%)⁶. For the NMA, common heterogeneity across all comparisons⁴ was assumed and estimated in each network.

Transitivity assumption is met when effect modifiers are equally distributed across the comparisons. We expected that the inclusion and exclusion criteria would allow to select studies sufficiently similar in terms of characteristics of participants, study design and outcomes, in order for all treatments included in the network to be considered "exchangeable" (as if all of them were part of a large, multi-arm trial). We extracted key study characteristics judged to be potential effect modifiers, namely: study design (open-label or double-blind); number of participants included; definition of LAR; doses and cycles of chemotherapy agents; doses and modality of radiotherapy; months of follow-up; median year of study conduct; proportion of participants discontinuing treatment before study endpoint; percentage of female participants; mean age; percentage of clinical T3-4 (cT4); percentage of participants with clinically suspected nodal metastases (cN+); mean distance from the anal verge (AV); percentage of pathologic T4 after pre-operative treatment (ypT4). By comparing the distribution of these possible effect modifiers across comparisons contributing to the estimation of the treatment effect, we formulated a judgment on whether differences in their distributions were large enough to threaten the validity of the analysis⁷. We considered that distribution differences in specific study characteristics across the different treatment strategies were relevant in case of both significant imbalances according to the Kruskal-Wallis test (continuous variables), the Pearson χ^2 or the Fisher's exact test (categorical variables), and meta-regression analyses showing an actual impact on treatment effect^{8,9}.

We assessed the presence of inconsistency (defined as the statistical disagreement between direct and indirect evidence of a treatment comparison) by comparing direct and indirect evidence within each closed loop¹⁰ and comparing the goodness of fit for an NMA model that assumes consistency with a model that allows for inconsistency in a "design by treatment interaction model" framework¹¹ by using the Stata commands mvmeta¹² and ifplot¹³ and the Stata network suite¹⁴.

For the primary outcome, we calculated the probability of each treatment of being at each possible rank, and produced mean ranks of treatments using the R *gemtc* package.

If ≥ 10 studies were included in the primary outcome, we assessed publication bias by visually inspecting the funnel plot, testing for asymmetry with the Egger's regression test¹⁵, and investigating possible reasons for funnel plot asymmetry.

For the primary outcome, we assessed the confidence of evidence by using the Confidence in Network Meta-Analysis (CINeMA) methodology^{16,17} through its web-based application (<u>http://cinema.ispm.ch</u>).

For the primary outcome, we conducted sensitivity analyses excluding trials with (a) overall high risk of bias according to RoB2; (b) high risk of indirectness; (c) CHT as one of the treatment arms.

eTable 1. List of Studies Included/Excluded/Ongoing/Awaiting Assessment

Study	in/out	Reason	Reason - details	Reference
Aschele 2011	Included			Aschele C, Cionini L, Lonardi S, et al. Primary tumor response to preoperative chemoradiation with or without oxaliplatin in locally advanced rectal cancer: Pathologic results of the STAR-01 randomized phase III trial. <i>J Clin Oncol</i> . 2011;29(20):2773- 2780. doi:10.1200/JCO.2010.34.4911 Aschele C, Lonardi S, Cionini L, et al. Final results of STAR-01: A randomized phase III trial comparing preoperative chemoradiation with or without oxaliplatin in locally advanced rectal cancer. <i>J Clin Oncol</i> . 2016;34(15_suppl):3521-3521.
Bahadoer 2021	Included			doi:10.1200/jco.2016.34.15_suppl.3521 Bahadoer RR, Dijkstra EA, van Etten B, et al. Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial. <i>Lancet Oncol</i> . 2021;22(1):29-42. doi:10.1016/S1470- 2045(20)30555-6 Bahadoer RR, Hospers GAP, Marijnen CAM, et al. Risk and location of distant metastases in patients with locally advanced rectal cancer after total neoadjuvant treatment or chemoradiotherapy in the RAPIDO trial. <i>Eur J Cancer</i> . 2023;185:139-149. doi: 10.1016/j.ejca.2023.02.027 Dijkstra EA, Nilsson PJ, Hospers GAP, et al. Locoregional Failure During and After Short-course Radiotherapy followed by Chemotherapy and Surgery Compared to Long-course Chemoradiotherapy and Surgery – A Five-year Follow-up of the
Bosset 2005	Included			 KAPIDO Trial. Ann Surg. 2023;Published ahead of print(4):766-772. doi: 10.1097/sla.000000000000005799 Bosset JF, Calais G, Mineur L, et al. Enhanced tumorocidal effect of chemotherapy with preoperative radiotherapy for rectal cancer: Preliminary results - EORTC 22921. J Clin Oncol. 2005;23(24):5620-5627. doi:10.1200/JCO.2005.02.113 Bosset JF, Calais G, Mineur L, et al. Fluorouracil-based adjuvant chemotherapy after preoperative chemoradiotherapy in rectal cancer: Long-term results of the EORTC 22921 randomised study. Lancet Oncol. 2014;15(2):184-190. doi:10.1016/S1470-2045(13)70599-0
Bujko 2004	Included			 Bujko K, Nowacki MP, Nasierowska-Guttmejer A, et al. Sphincter preservation following preoperative radiotherapy for rectal cancer: Report of a randomised trial comparing short-term radiotherapy vs. conventionally fractionated radiochemotherapy. In: <i>Radiotherapy and Oncology</i>. Vol 72. Elsevier; 2004:15-24. doi:10.1016/j.radonc.2003.12.006 Bujko K, Nowacki MP, Nasierowska-Guttmejer A, Michalski W, Bebenek M, Kryj M. Long-term results of a randomized trial comparing preoperative short-course radiotherapy with preoperative conventionally fractionated chemoradiation for rectal cancer. <i>Br J Surg</i>. 2006;93(10):1215-1223. doi:10.1002/bjs.5506
Bujko 2013	Included			 Bujko K, Nasierowska-Guttmejer A, Wyrwicz L, et al. Neoadjuvant treatment for unresectable rectal cancer: An interim analysis of a multicentre randomized study. <i>Radiother Oncol.</i> 2013;107(2):171-177. doi:10.1016/j.radonc.2013.03.001 Bujko K, Wyrwicz L, Rutkowski A, et al. Long-course oxaliplatin-based preoperative chemoradiation versus 5 × 5 Gy and consolidation chemotherapy for cT4 or fixed cT3 rectal cancer: Results of a randomized phase III study. <i>Ann Oncol.</i> 2016;27(5):834-842. doi:10.1093/annonc/mdw062
Chakrabarti 2021	Included			Chakrabarti D, Rajan S, Akhtar N, et al. Short-course radiotherapy with consolidation chemotherapy versus conventionally fractionated long-course chemoradiotherapy for locally advanced rectal cancer: randomized clinical trial. <i>Br J Surg.</i> 2021;108(5):511-520. doi:10.1093/bjs/znab020
Conroy 2021	Included			Conroy T, Bosset JF, Etienne PL, et al. Neoadjuvant chemotherapy with FOLFIRINOX and preoperative chemoradiotherapy for patients with locally advanced rectal cancer (UNICANCER-PRODIGE 23): a multicentre, randomised, open-label, phase 3 trial. <i>Lancet Oncol.</i> 2021;22(5):702-715. doi:10.1016/S1470-2045(21)00079-6
Deng 2016	Included			 Deng Y, Chi P, Lan P, et al. Modified FOLFOX6 with or without radiation versus fluorouracil and leucovorin with radiation in neoadjuvant treatment of locally advanced rectal cancer: Initial results of the Chinese FOWARC multicenter, open-label, randomized three-arm phase III trial. <i>J Clin Oncol</i>. 2016;34(27):3300-3307. doi:10.1200/JCO.2016.66.6198 Deng Y, Chi P, Lan P, et al. Neoadjuvant modified folfox6 with or without radiation versus fluorouracil plus radiation for locally advanced rectal cancer: Final results of the Chinese FOWARC trial. <i>J Clin Oncol</i>. 2019;37(34):3223-3233. doi:10.1200/JCO.18.02309

Formándoz Martos		Fernández-Martos C, Garcia-Albeniz X, Pericay C, et al. Chemoradiation, surgery and adjuvant chemotherapy versus induction
2015	Included	chemotherapy followed by chemoradiation and surgery: Long-term results of the Spanish GCR-3 phase II randomized trial.
2013		Ann Oncol. 2015;26(8):1722-1728. doi:10.1093/annonc/mdv223
		Fokas E, Allgäuer M, Polat B, et al. Randomized phase II trial of chemoradiotherapy plus induction or consolidation
		chemotherapy as total neoadjuvant therapy for locally advanced rectal cancer: CAO/ATO/ATO-12. J Clin Oncol.
Fokas 2019	Included	Eokas E. Schlenska J. ange A. Pold B. et al. Chemoradiotherany. Plus Induction or Consolidation Chemotherany as Total
		Neoadiuvant Therany for Patients with Locally Advanced Rectal Cancer Long-term Results of the CAO/ARO/AIO-12
		Randomized Clinical Trial. <i>JAMA Darol.</i> 2022;8(1):e215445-e215445. doi:10.1001/jamaoncol.2021.5445
01 10000	T 1 1 1	Gérard JP, Conroy T, Bonnetain F, et al. Preoperative radiotherapy with or without concurrent fluorouracil and leucovorin in
Gerard 2006	Included	T3-4 rectal cancers: Results of FFCD 9203. J Clin Oncol. 2006;24(28):4620-4625. doi:10.1200/JCO.2006.06.7629
		Gérard JP, Azria D, Gourgou-Bourgade S, et al. Comparison of two neoadjuvant chemoradiotherapy regimens for locally
		advanced rectal cancer: Results of the phase III trial accord 12/0405-Prodige 2. J Clin Oncol. 2010;28(10):1638-1644.
		doi:10.1200/JCO.2009.25.8376
Gérard 2010	Included	Gérard JP, Azria D, Gourgou-Bourgade S, et al. Clinical outcome of the ACCORD 12/0405 PRODIGE 2 randomized trial in
		rectal cancer. J Clin Oncol. 2012;30(36):4558-4565. doi:10.1200/JCO.2012.42.87/1
		AZIA D, DOYELJ, Jamer M, et al. Late toxicities and chinical outcome at 5 years of the ACCORD 12/04/05-rKODDE 02 that
		2442. doi:10.1093/anonc/mtx351
11 11 12017	T 1 1 1	Haddad P, Miraie M, Farhan F, et al. Addition of oxaliplatin to neoadjuvant radiochemotherapy in MRI-defined T3, T4 or N+
Haddad 2017	Included	rectal cancer: a randomized clinical trial. Asia Pac J Clin Oncol. 2017;13(6):416-422. doi:10.1111/ajco.12675
Liao 2015	Included	Jiao D, Zhang R, Gong Z, et al. Fluorouracil-based preoperative chemoradiotherapy with or without oxaliplatin for stage ii/iii
J1a0 2013	menudeu	rectal cancer: A 3-year follow-up study. <i>Chinese J Cancer Res.</i> 2015;27(6):588-596. doi:10.3978/j.issn.1000-9604.2015.12.05
		Jin J, Tang Y, Hu C, et al. Multicenter, Randomized, Phase III Trial of Short-Term Radiotherapy Plus Chemotherapy Versus
Jin 2022	Included	Long-Term Chemoradiotherapy in Locally Advanced Rectal Cancer (STELLAR). J Clin Oncol. 2022;40(15):1681-1692.
		doi:10.1200/jco.21.0166/ Kim SV. Loo L. Kim TW. et al. A Deredomized Disconsideration Chamathemany After Decomposition
Kim 2018	Included	Kim S T, Joo J, Kim T W, et al. A Kandomized Phase 2 That of Consolidation Chemoretapy Alter Preoperative
Killi 2010	menudeu	Rediat Oncol Biol Phys. 2018;101(4):889-40;101(10)(6/; iirobn 2018;04:013)
		Latkauskas T. Pauzas H. Gineikiene L et al. Initial results of a randomized controlled trial comparing clinical and pathological
		downstaging of rectal cancer after preoperative short-course radiotherapy or long-term chemoradiotherapy, both with delayed
Latkouckee 2012	Included	surgery. Color Dis. 2012;14(3):294-298. doi:10.1111/j.1463-1318.2011.02815.x
Latkauskas 2012	menudeu	Latkauskas T, Pauzas H, Kairevice L, et al. Preoperative conventional chemoradiotherapy versus short-course radiotherapy
		with delayed surgery for rectal cancer: Results of a randomized controlled trial. <i>BMC Cancer</i> . 2016;16(1):1-7.
		doi:10.1186/s12885-016-2959-9
Manéahal 2012	Included	Marechal R, Vos B, Polus M, et al. Short course chemotherapy followed by concomitant chemoradiotherapy and surgery in leadly durated expected expected expected expected expected by concomitant chemoradiotherapy and surgery in
Marechai 2012	Included	iocally advanced rectal cancer: A randomized muticentric phase II study. Ann Oncol. 2012;25(6):1525-1530.
		Mei W-I Wang X-Z Li Y-F et al Neoadinyant Chemotherany with CAPOX versus Chemoradiation for Locally Advanced
Mei 2023	Included	Rectal Cancer with Uninvolved Mesorectal Fascia (CONVERT): Initial Results of a Phase III Trial. Ann Surg.
		2023;277(4):557-564. doi:10.1097/sla.000000000005780
		Mohiuddin M, Winter K, Mitchell E, et al. Randomized phase II study of neoadjuvant combined-modality chemoradiation for
Mohiuddin 2006	Included	distal rectal cancer: Radiation therapy oncology group trial 0012. J Clin Oncol. 2006;24(4):650-655.
		doi:10.1200/JCO.2005.03.6095
		Moore J, Price T, Carruthers S, et al. Prospective randomized trial of neoadjuvant chemotherapy during the 'wait period'
Moore 2017	Included	following preoperative chemoradiotherapy for rectal cancer: results of the WAIT trial. <i>Color Dis.</i> 2017;19(11):973-979.
		d01:10.1111/c0d1.15/24 Ngan SV Burmeister B. Fisher RL et al. Dandomized trial of short course radiotherany varsus long course chamoradiation
Ngan 2012	Included	regard ST, Burmerster B, Fisher KJ, et al. Kandolinized that of short-course radionerapy versus iolig-course chemoradiation comparing rates of local recurrence in patients with T3 rectal cancer: Trans.Tasman Radiation Opeology Group Trial 01 04 J
1.5uii 2012	mended	<i>Clin Oncol.</i> 2012;30(31):3827-3833. doi:10.1200/JCO.2012.42.9597

				Ansari N, Solomon MJ, Fisher RJ, et al. Acute Adverse Events and Postoperative Complications in a Randomized Trial of Prooperative Short-course Radiotherapy Versus Long-course Chemoradiotherapy for T3 Adenocarcinoma of the Rectum:
				Trans-Tasman Radiation Oncology Group Trial (TROG 01.04). Ann Surg. 2017;265(5):882-888. doi:10.1097/SLA.000000000001987
O'Connell 2014	Included			O'Connell MJ, Colangelo LH, Beart RW, et al. Capecitabine and oxaliplatin in the preoperative multimodality treatment of rectal cancer: Surgical end points from national surgical adjuvant breast and bowel project trial R-04. <i>J Clin Oncol</i> . 2014;32(18):1927-1934. doi:10.1200/JCO.2013.53.7753
Rodel 2015	Included			Rödel C, Graeven U, Fietkau R, et al. Oxaliplatin added to fluorouracil-based preoperative chemoradiotherapy and postoperative chemotherapy of locally advanced rectal cancer (the German CAO/ARO/AIO-04 study): final results of the multicentre, open-label, randomised, phase 3 trial. <i>Lancet Oncol.</i> 2015;16(8):979-989. doi:10.1016/S1470-2045(15)00159-X
Schmoll 2021	Included			Schmoll HJ, Stein A, van Cutsem E, et al. Pre- And postoperative capecitabine without or with oxaliplatin in locally advanced rectal cancer: PETACC 6 trial by EORTC GITCG and ROG, AIO, AGITG, BGDO, and FFCD. <i>J Clin Oncol</i> . 2021;39(1):17-29. doi:10.1200/JCO.20.01740
Schrag 2023	Included			Schrag D, Shi Q, Weiser M et al. Preoperative treatment for locally advanced rectal cancer. <i>N Engl J Med.</i> 2023 Jun 4. doi: 10.1056/NEJMoa2303269. Online ahead of print.
Wang 2019	Included			Wang J, Guan Y, Gu W, et al. Long-course neoadjuvant chemoradiotherapy with versus without a concomitant boost in locally advanced rectal cancer: A randomized, multicenter, phase II trial (FDRT-002). <i>Radiat Oncol</i> . 2019;14(1). doi:10.1186/s13014-019-1420-z
Choi 2023	ONGOING			NCT05673772. Preoperative Sequential Short-course Radiation Therapy and FOLFOX for Locally Advanced Rectal Cancer (SOLAR). https://clinicaltrials.gov/ct2/show/NCT05673772?term=solar&type=Intr&cond=Rectal+Cancer&cntry=KR&draw=2&rank=1
Kim 2018	ONGOING			Kim CW, Kang BM, Kim IY, et al. Korean Society of Coloproctology (KSCP) trial of cONsolidation Chemotherapy for Locally advanced mid or low rectal cancer after neoadjUvant concurrent chemoraDiothErapy: A multicenter, randomized controlled trial (KONCLUDE). <i>BMC Cancer</i> . 2018;18(1):1-8. doi:10.1186/S12885-018-4466-7/TABLES/1
NCT03177382	ONGOING			Total Neoadjuvant Treatment vs. Chemoradiotherapy in Local Advanced Rectal Cancer With High Risk Factors (TNTCRT). ClinicalTrials.gov Identifier: NCT03177382.
Wang 2022	ONGOING			Wang Y, Shen L, Wan J, et al. Short-course radiotherapy combined with CAPOX and Toripalimab for the total neoadjuvant therapy of locally advanced rectal cancer: a randomized, prospective, multicentre, double-arm, phase II trial (TORCH). BMC Cancer. 2022;22(1):1-8. doi:10.1186/S12885-022-09348-Z/FIGURES/2
Aboelnaga 2015	Excluded	WRONG DESIGN	Single arm	Aboelnaga EM, Daoud MA, Eladl EI, Zaid AM. Induction FOLFOX followed by preoperative hyperfractionated radiotherapy plus bolus 5-fluorouracil in locally advanced rectal carcinoma: single arm phase I-II study. <i>Med Oncol.</i> 2015 Apr;32(4):108. doi: 10.1007/s12032-015-0556-4.
Aghili 2018	Excluded	WRONG DESIGN	Single arm	Aghili M, Sotoudeh S, Ghalehtaki R et al. Preoperative short course radiotherapy with concurrent and consolidation chemotherapies followed by delayed surgery in locally advanced rectal cancer: Preliminary results. <i>Radiat Oncol J</i> 2018; 36(1): 17-24. Doi: 10.3857/roj.2017.00185
Borg 2014	Excluded	WRONG INTERVENTION	Immunotherapy	Borg C, André T, Mantion G et al. Pathological response and safety of two neoadjuvant strategies with bevacizumab in MRI- defined locally advanced T3 resectable rectal cancer: a randomized, noncomparative phase II study. <i>Ann Oncol.</i> 2014 Nov;25(11):2205-2210. doi: 10.1093/annonc/mdu377.
Boulis-Wassif 1984	Excluded	WRONG INTERVENTION	Wrong schedule of treatment	Boulis-Wassif S, Gerard A, Loygue J, Camelot D, Buyse M, Duez N. Final results of a randomized trial on the treatment of rectal cancer with preoperative radiotherapy alone or in combination with 5-fluorouracil, followed by radical surgery. Trial of the European Organization on Research and Treatment of Cancer Gastrointestinal Tract Cancer Cooperative Group. <i>Cancer</i> . 1984 May 1;53(9):1811-8. doi: 10.1002/1097-0142(19840501)53:9<1811::aid-cncr2820530902>3.0.co;2-h.
Cedermark 1995	Excluded	WRONG INTERVENTION	No preoperative therapy in control group	Cedermark B, Johansson H, Rutqvist LE, et al. The Stockholm I trial of preoperative short term radiotherapy in operable rectal carcinoma. A prospective randomized trial. Stockholm Colorectal Cancer Study Group. <i>Cancer</i> . 1995;75:2269–2275. doi: 10.1002/1097-0142(19950501)75:9<2269::aid-cncr2820750913>3.0.co;2-i.
Cercek 2012	Excluded	WRONG DESIGN	Single arm	Cercek A, Goodman KA, Hajj C et al. Neoadjuvant chemotherapy first, followed by chemoradiation and then surgery, in the management of locally advanced rectal cancer. <i>J Natl Compr Canc Netw</i> . 2014 Apr;12(4):513-9. doi: 10.6004/jnccn.2014.0056
Chiorean 2012	Excluded	WRONG DESIGN	Single arm	Chiorean EG, Sanghani S, Schiel MA et al. Phase II and gene expression analysis trial of neoadjuvant capecitabine plus irinotecan followed by capecitabine-based chemoradiotherapy for locally advanced rectal cancer: Hoosier Oncology Group GI03-53. Cancer Chemother Pharmacol. 2012 Jul;70(1):25-32. doi: 10.1007/s00280-012-1883-1

Chua 2010	Excluded	WRONG	Single arm	Chua YJ, Barbachano Y, Cunningham D et al. Neoadjuvant capecitabine and oxaliplatin before chemoradiotherapy and total mesoreated excision in MPL defined poor rick rectal capear a phase 2 trial. <i>Langet Orgol</i> 2010 Mar:11(3):211-8 doi:
Chua 2010	Excluded	DESIGN	Single ann	10.1016/S1470-2045(09)70381-X.
Cotte 2015	Excluded	WRONG	Wrong schedule of	Cotte E, Passot G, Decullier E et al. Pathologic Response, When Increased by Longer Interval, Is a Marker but Not the Cause of Good Prognosis in Pactal Cancer 17 year Follow up of the Lyon P90 01 Pandomized Trial. Int L Padiat Organ Biol Phys.
Colle 2015	Excluded	INTERVENTION	treatment	2016 Mar 1;94(3):544-53. doi: 10.1016/j.ijrobp.2015.10.061.
De Felice 2021	Excluded	WRONG	Single arm	De Felice F, D'Ambrosio G, Iafrate F et al. Intensified Total Neoadjuvant Therapy in Patients With Locally Advanced Rectal
		WEONC		Dewdney A, Cunningham D, Tabernero J et al. Multicenter randomized phase II clinical trial comparing neoadjuvant
Dewdney 2012	Excluded	INTERVENTION	Immunotherapy	oxaliplatin, capecitabine, and preoperative radiotherapy with or without cetuximab followed by total mesorectal excision in
				patients with high-risk rectal cancer (EXPERT-C). J Clin Oncol. 2012 May 10;30(14):1620-7. doi: 10.1200/JCO.2011.39.6036.
Fernández-Martos	Excluded	WRONG	Immunotherapy	Standard Chemoradiotherapy and Surgery in Patients With High-Risk Rectal Adenocarcinoma: The GEMCAD 1402
2019		INTERVENTION	1.7	Randomized Clinical Trial. JAMA Oncol. 2019 Nov 1;5(11):1566-1573. doi: 10.1001/jamaoncol.2019.2294.
Fisher 1988	Excluded	WRONG INTERVENTION	Adjuvant radiotherapy	Fisher B, Wolmark N, Rockette H, et al. Postoperative adjuvant chemotherapy or radiation therapy for rectal cancer: results from NSABP protocol R-01. <i>J Natl Cancer Inst.</i> 1988;80:21–29. doi: 10.1093/jnci/80.1.21.
Folkesson 2005	Excluded	WRONG	No preoperative therapy in	Folkesson J, Birgisson H, Pahlman L, et al. Swedish Rectal Cancer Trial: long lasting benefits from radiotherapy on survival
		WRONG	control group	Gao YH. An X. Sun WJ et al. Evaluation of capecitabine and oxaliplatin administered prior to and then concomitant to
Gao 2014	Excluded	DESIGN	Single arm	radiotherapy in high risk locally advanced rectal cancer. J Surg Oncol. 2014 Apr;109(5):478-82. doi: 10.1002/jso.23516.
		WRONG		Gao YH, Lin JZ, An X et al. Neoadjuvant sandwich treatment with oxaliplatin and capecitabine administered prior to,
Gao 2014	Excluded	DESIGN	Single arm	concurrently with, and following radiation therapy in locally advanced rectal cancer: a prospective phase 2 trial. Int J Radiat Oncol Biol Phys. 2014 Dec 1:90(5):1153-60. doi: 10.1016/j.jirobp.2014.07.021.
Garaia Aquilar 2015	Evoluded	WRONG	Non randomized	Garcia-Aguilar J, Chow OS, Smith DD, et al. Effect of adding mFOLFOX6 after neoadjuvant chemoradiation in locally
Garcia-Aguitai 2013	Excluded	DESIGN	INOILTAIIGOIIIIZEG	advanced rectal cancer: A multicentre, phase 2 trial. Lancet Oncol. 2015;16(8):957-966. doi:10.1016/S1470-2045(15)00004-2
Garcia-Aguilar 2022	Excluded	WRONG	Non-operative management	Garcia-Aguilar J, Patil S, Gollub MJ et al. Organ Preservation in Patients With Rectal Adenocarcinoma Treated With Total
		INTERVENTION		Gérard A Buyse M Nordlinger M et al. Preoperative radiotherapy as adjuvant treatment in rectal cancer. Final results of a
Gerard 1988	Excluded	WRONG	No preoperative therapy in	randomized study of the European Organization for Research and Treatment of Cancer (EORTC). Ann Surg. 1988
		INTERVENTION	control group	Nov;208(5):606-14. doi: 10.1097/00000658-198811000-00011.
		WRONG	Wrong schedule of	Gerard JP, Barbet N, Schiappa R et al. Neoadjuvant chemoradiotherapy with radiation dose escalation with contact x-ray
Gerard 2023	Excluded	WRONG	treatment, non-operative	(OPERA): a phase 3, randomised controlled trial. <i>Lancet Gastroenterol Hepatol.</i> 2023 Apr:8(4):356-367. doi: 10.1016/S2468-
		COMPARISON	management	1253(22)00392-2
~		WRONG	No preoperative therapy in	Goldberg PA, Nicholls RJ, Porter NH, et al. Long-term results of a randomised trial of short-course low-dose adjuvant pre-
Goldberg 1994	Excluded	INTERVENTION	control group	operative radiotherapy for rectal cancer: reduction in local treatment failure. <i>Eur J Cancer</i> . 1994;30A:1602–1606. doi: 10.1016/0959-8049(94)00312-s.
		WRONC		Golo D, But-Hadzic J, Anderluh F et al. Induction chemotherapy, chemoradiotherapy and consolidation chemotherapy in
Golo 2018	Excluded	DESIGN	Single arm	preoperative treatment of rectal cancer - long-term results of phase II OIGIT-01 Trial. Radiol Oncol. 2018 Sep 11;52(3):267-
				274. doi: 10.2478/raon-2018-0028.
Guillem 2005	Excluded	WRONG	Single arm	and total mesorectal excision of locally advanced rectal cancer, Ann Surg. 2005 May:241(5):829-36: discussion 836-8. doi:
		DESIGN		10.1097/01.sla.0000161980.46459.96.
Hartley 2005	Excluded	WRONG	Pooled analysis	Hartley A, Ho KF, McConkey C, Geh JI. Pathological complete response following pre-operative chemoradiotherapy in rectal
		UDSION		Helbling D, Bodoky G, Gautschi O et al. Neoadjuvant chemoradiotherapy with or without panitumumab in patients with wild-
Helbling 2012	Excluded	WRONG	Monoclonal antibody	type KRAS, locally advanced rectal cancer (LARC): a randomized, multicenter, phase II trial SAKK 41/07. Ann Oncol. 2013
		INTERVENTION		Mar;24(3):718-25. doi: 10.1093/annonc/mds519.

Hofheinz 2012	Excluded	WRONG INTERVENTION	Comparison of two chemotherapy regimens	Hofheinz RD, Wenz F, Post S et al. Chemoradiotherapy with capecitabine versus fluorouracil for locally advanced rectal cancer: a randomised, multicentre, non-inferiority, phase 3 trial. <i>Lancet Oncol.</i> 2012 Jun;13(6):579-88. doi: 10.1016/S1470-2045(12)70116-X.
Jung 2015	Excluded	WRONG INTERVENTION	Comparison of two chemotherapy regimens	Jung M, Shin SJ, Koom WS et al. A Randomized Phase 2 Study of Neoadjuvant Chemoradiaton Therapy With 5- Fluorouracil/Leucovorin or Irinotecan/S-1 in Patients With Locally Advanced Rectal Cancer. <i>Int J Radiat Oncol Biol Phys</i> . 2015 Dec 1;93(5):1015-22. doi: 10.1016/j.ijrobp.2015.08.037.
Kapiteijn 2001	Excluded	WRONG COMPARISON	No preoperative therapy in control group	Kapiteijn E, Marijnen CA, Nagtegaal ID, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. N Engl J Med. 2001;345:638–646. doi: 10.1056/NEJMoa010580.
Kayal 2014	Excluded	WRONG INTERVENTION	Comparison of two chemotherapy regimens	Kayal PK, Saha A, Dastidar AG, Mahata A, Das A, Sarkar R. A randomized comparative study between neoadjuvant 5- fluorouracil and leukovorin versus 5-fluorouracil and cisplatin along with concurrent radiation in locally advanced carcinoma rectum. <i>Clin Cancer Investig J</i> 2014;3:32-7.
Lefevre 2016	Excluded	WRONG INTERVENTION	Wrong schedule of treatment	Lefevre JH, Mineur L, Kotti S et al. Effect of Interval (7 or 11 weeks) Between Neoadjuvant Radiochemotherapy and Surgery on Complete Pathologic Response in Rectal Cancer: A Multicenter, Randomized, Controlled Trial (GRECCAR-6). <i>J Clin</i> <i>Oncol.</i> 2016 Nov 1;34(31):3773-3780. doi: 10.1200/JCO.2016.67.6049.
Marco 2018	Excluded	WRONG DESIGN	Not randomized	Marco MR, Zhou L, Patil S, et al. Consolidation mFOLFOX6 chemotherapy after chemoradiotherapy improves survival in patients with locally advanced rectal cancer: Final results of a multicenter phase II trial. <i>Dis Colon Rectum</i> . 2018;61(10):1146-1155. doi:10.1097/DCR.00000000001207
Markovina 2017	Excluded	WRONG DESIGN	Not randomized	Markovina S, Youssef F, Roy A, et al. Improved Metastasis- and Disease-Free Survival With Preoperative Sequential Short- Course Radiation Therapy and FOLFOX Chemotherapy for Rectal Cancer Compared With Neoadjuvant Long-Course Chemoradiotherapy: Results of a Matched Pair Analysis. In: <i>International Journal of Radiation Oncology Biology Physics</i> . Vol 99. Int J Radiat Oncol Biol Phys; 2017:417-426. doi:10.1016/j.ijrobp.2017.05.048
Masi 2019	Excluded	WRONG DESIGN	Single arm	Masi G, Vivaldi C, Fornaro L et al. Total neoadjuvant approach with FOLFOXIRI plus bevacizumab followed by chemoradiotherapy plus bevacizumab in locally advanced rectal cancer: the TRUST trial. <i>Eur J Cancer</i> . 2019 Mar;110:32-41. doi: 10.1016/j.ejca.2019.01.006.
NCT02514278	Excluded	WRONG INTERVENTION	Non-operative management, ongoing	Optimisation of Response for Organ Preservation in Rectal Cancer : Neoadjuvant Chemotherapy and Radiochemotherapy vs. Radiochemotherapy (GRECCAR12). ClinicalTrials.gov Identifier: NCT02514278.
NCT02945566	Excluded	WRONG INTERVENTION , WRONG POPULATION	Non-operative management, early rectal cancer	Can the Rectum be Saved by Watchful Waiting or TransAnal Surgery Following (Chemo)Radiotherapy Versus Total Mesorectal Excision for Early REctal Cancer? (STAR-TREC). ClinicalTrials.gov Identifier: NCT02945566.
NCT04246684	Excluded	WRONG INTERVENTION	Non-operative management, ongoing	Short RT Versus RCT,Followed by Chemo.and Organ Preservation for Interm and High-risk Rectal Cancer Patients. ClinicalTrials.gov Identifier: NCT04246684
Peeters 2007	Excluded	WRONG DESIGN	No preoperative therapy in control group	Peeters KCMJ, Marijnen CAM, Nagtegaal ID et al. The TME trial after a median follow-up of 6 years: increased local control but no survival benefit in irradiated patients with resectable rectal carcinoma. <i>Ann Surg</i> . 2007 Nov;246(5):693-701. doi: 10.1097/01.sla.0000257358.56863.ce.
Pettersson 2015	Excluded	WRONG POPULATION	Exclusion of locally- advanced rectal cancer	Pettersson D, Lörinc E, Holm T et al. Tumour regression in the randomized Stockholm III Trial of radiotherapy regimens for rectal cancer. <i>Br J Surg</i> . 2015 Jul;102(8):972-8; discussion 978. doi: 10.1002/bjs.9811.
Rahma 2021	Excluded	WRONG INTERVENTION	Immunotherapy	Rahma OE, Yothers G, Hong TS et al. Use of Total Neoadjuvant Therapy for Locally Advanced Rectal Cancer: Initial Results From the Pembrolizumab Arm of a Phase 2 Randomized Clinical Trial. <i>JAMA Oncol.</i> 2021 Aug 1;7(8):1225-1230. doi: 10.1001/jamaoncol.2021.1683.
Saha 2015	Excluded	WRONG DESIGN	Pilot study	Saha A, Ghosh SK, Roy C et al. A randomized controlled pilot study to compare capecitabine-oxaliplatin with 5-FU- leucovorin as neoadjuvant concurrent chemoradiation in locally advanced adenocarcinoma of rectum. <i>J Cancer Res Ther</i> . 2015 Jan-Mar;11(1):88-93. doi: 10.4103/0973-1482.150341.
Sauer 2012	Excluded	WRONG INTERVENTION	Post-operative chemioradiotherapy	Sauer R, Liersch T, Merkel S et al. Preoperative versus postoperative chemoradiotherapy for locally advanced rectal cancer: results of the German CAO/ARO/AIO-94 randomized phase III trial after a median follow-up of 11 years. <i>J Clin Oncol</i> . 2012 Jun 1;30(16):1926-33. doi: 10.1200/JCO.2011.40.1836.
Siegel 2009	Excluded	NO FINAL DATA REPORTED	No final data reported, only protocol published	Siegel R, Burock S, Wernecke KD et al. Preoperative short-course radiotherapy versus combined radiochemotherapy in locally advanced rectal cancer: a multi-centre prospectively randomised study of the Berlin Cancer Society. <i>BMC Cancer</i> . 2009 Feb 6;9:50. doi: 10.1186/1471-2407-9-50.

Smith 2015	Excluded	WRONG INTERVENTION , WRONG COMPARISON	Not randomized, non- operative management	Smith JJ, Chow OS, Gollub MJ et al. Organ Preservation in Rectal Adenocarcinoma: a phase II randomized controlled trial evaluating 3-year disease-free survival in patients with locally advanced rectal cancer treated with chemoradiation plus induction or consolidation chemotherapy, and total mesorectal excision or nonoperative management. <i>BMC Cancer</i> .2015 Oct 23;15:767. doi: 10.1186/s12885-015-1632-z.
Tveit 1997	Excluded	WRONG INTERVENTION	Post-operative chemioradiotherapy	Tveit KM, Guldvog I, Hagen S, et al. Randomized controlled trial of postoperative radiotherapy and short-term time-scheduled 5-fluorouracil against surgery alone in the treatment of Dukes B and C rectal cancer. Norwegian Adjuvant Rectal Cancer Project Group. <i>Br J Surg.</i> 1997;84:1130–1135.
Wang 2018	Excluded	WRONG INTERVENTION	No preoperative therapy in control group	Wang F, Fan W, Peng J et al. Total mesorectal excision with or without preoperative chemoradiotherapy for resectable mid/low rectal cancer: a long-term analysis of a prospective, single-center, randomized trial. <i>Cancer Commun (Lond)</i> . 2018 Dec 20;38(1):73. doi: 10.1186/s40880-018-0342-8.
Wang 2018	Excluded	WRONG DESIGN	Single arm	Wang X, Yu Y, Meng W et al. Total neoadjuvant treatment (CAPOX plus radiotherapy) for patients with locally advanced rectal cancer with high risk factors: A phase 2 trial. <i>Radiother Oncol.</i> 2018 Nov;129(2):300-305. doi: 10.1016/j.radonc.2018.08.027.
Wong 2012	Excluded	WRONG INTERVENTION	Comparison of two chemotherapy regimens	Wong SJ, Winter K, Meropol NJ et al. Radiation Therapy Oncology Group 0247: a randomized Phase II study of neoadjuvant capecitabine and irinotecan or capecitabine and oxaliplatin with concurrent radiotherapy for patients with locally advanced rectal cancer. Int J Radiat Oncol Biol Phys. 2012 Mar 15;82(4):1367-75. doi: 10.1016/j.ijrobp.2011.05.027.
Wiśniowska 2016	Excluded	WRONG DESIGN	Subgroup analysis of RCT	Wiśniowska K, Nasierowska-Guttmejer A, Polkowski W et al. Does the addition of oxaliplatin to preoperative chemoradiation benefit cT4 or fixed cT3 rectal cancer treatment? A subgroup analysis from a prospective study. <i>Eur J Surg Oncol</i> . 2016 Dec;42(12):1859-1865. doi: 10.1016/j.ejso.2016.08.001.
Wolmark 2000	Excluded	WRONG INTERVENTION	Post-operative chemioradiotherapy	Wolmark N, Wieand HS, Hyams DM, et al. Randomized trial of postoperative adjuvant chemotherapy with or without radiotherapy for carcinoma of the rectum: National Surgical Adjuvant Breast and Bowel Project Protocol R-02. <i>J Natl Cancer</i> <i>Inst.</i> 2000;92:388–396.
Zhu 2013	Excluded	WRONG DESIGN	Single arm	Zhu J, Gu W, Lian P et al. A phase II trial of neoadjuvant IMRT-based chemoradiotherapy followed by one cycle of capecitabine for stage II/III rectal adenocarcinoma. <i>Radiat Oncol.</i> 2013 May 29;8:130. doi: 10.1186/1748-717X-8-130.

eTable 2. Characteristics of Included Studies

Legend: pCR=pathologic complete response; AV distance=distance from the anal verge; L-CRT1=long-course chemoradiotherapy with single-agent fluoropyrimidine-based chemotherapy; L-CRT2=long-course chemoradiotherapy with duplex chemotherapy drug (fluoropyrimidine plus oxaliplatin); FU=fluorouracil; OXA=oxaliplatin; Cape=capecitabine; S-RT=short-course radiotherapy; L-RT=long-course radiotherapy; FU-LV=fluorouracil plus leucovorin; NR=not reported; XELOX/CapOX=capecitabine plus oxaliplatin; FOLFIRONOX=oxaliplatin, irinotecan, FU-LV; mFOLFOX6= leucovorin 400 mg/m2 intravenously followed by fluorouracil 400 mg/m2 intravenously and fluorouracil 2.4 g/m2 by 48-h continuous infravenous infusion plus oxaliplatin 85 mg/m² on day 1 of each chemotherapy cycle; CHT=chemotherapy. Percentages are computed on the intention-to-treat population.

First author	Year	Number of patients	Treatment control arm	Treatment experimental arm	Women (%)	Mean age	Mean AV distance (cm) / Range (%)	Clinical stage exp. arm (%)	pCR exp. arm (%)	R0 exp. arm (%)	ypN0 exp. arm (%)	ypT1-2 exp. arm (%)	Mean follow- up (months)
Aschele	2011	739	L-RT (50.4 Gy) + FU (L-CRT1)	L-RT (50.4 Gy) + OXA + FU (L-CRT2)	33.0	62.5	Max 12 cm <8 cm: 76.9	cT3-4: 95.1 cN+: 66.8	16.0	91.0	67.1	35.1	105.6
Bahadoer	2021	912	L-RT (50.4 Gy) + Cape (L-CRT1)	S-RT (25 Gy) → CAPOX x 6 / FOLFOX4 x 9 (S-RT + consolidation)	32.6	61.0	Max 16 cm <10 cm: 61.5	cT3-4: 97.0 cN+: 90.9	26.0	82.7	68.6	22.3	74.4
Bosset	2005	1011	L-RT (45 Gy)	L-RT (45 Gy) + FU- LV (L-CRT1)	29.6	63.1	Max 15 cm <10 cm: 94.3	cT3-4: 100 cN+: NR	12.8	NR	67.2	40.5	124.8
Bujko	2004	312	S-RT (25 Gy) \rightarrow TME (7 days) (S-RTearly)	L-RT (50.4 Gy) + FU- LV (L-CRT1)	35.0	60.0	5.7	cT3-4: NR cN+: NR	14.0	82.8	68.0	39.5	48.0
Bujko	2013	515	L-RT (50.4 Gy) + OXA + FU-LV (L- CRT2)	S-RT (25 Gy) → FOLFOX4 x 3 (S-RT + consolidation)	31.5	60.0	Max 15 cm <10 cm: 97.3	cT3-4: 96.9 cN+: NR	16.0	77.3	57.5	19.1	35.0
Chakrabarti	2021	140	L-RT (50.4 Gy) + Cape (L-CRT1)	S-RT (25 Gy) \rightarrow XELOX x 2 (S-RT + consolidation)	33.6	42.0	NR	cT3-4: 100 cN+: 68.1	11.6	86.9	75.3	66.7	NR
Conroy	2021	461	L-RT (50 Gy) + Cape (L-CRT1)	FOLFIRINOX x 6 → L-RT + Cape (Induction + L-CRT)	33.6	61.5	Max 15 cm <10 cm: 87%	cT3-4: 96.1 cN+: 89.6	25.5	87.0	75.8	30.7	47.9
Deng	2016	475	L-RT (46-50.4 Gy) +	L-RT (46-50.4 Gy) + mFOLFOX6 (L- CRT2)	34.3	54.0	5.4	cT3-4: 98.2 cN+:81.8	25.4	81.2	50.9	NR	45.1
			FU-LV (L-CKII)	mFOLFOX6 x 4-6 (CHT)	34.3	54.0	6	cT3-4: 99.4 cN+:72.1	6.1	82.4	32.7	NR	45.1
Fernández- Martos	2015	103	L-RT + Cape + OXA (L-CRT2)	$\begin{array}{l} \text{XELOX x } 4 \rightarrow \text{L-RT} \\ + \text{Cape} + \text{OXA} \\ \text{(Induction} + \text{L-CRT)} \end{array}$	34.0	61.0	Max 12 cm	cT3-4: NR cN+: NR	14.3	85.7	67.9	NR	69.0
Fokas	2019	306		FOLFOX x $3 \rightarrow L$ - RT (50.4 Gy) + FU + OXA (Induction + L- CRT)	32	62.0	Max 12 cm <10 cm: 84.0	cT3-4: 96.2 cN+: 85.9	17.3	83.3	67.3	37.2	43.0
FOKAS	2019	306	-	$\begin{array}{l} \text{L-RT} (50.4 \text{ Gy}) + \text{FU} \\ + \text{OXA} \rightarrow \text{FOLFOX x} \\ 3 \text{ (L-CRT +} \\ \text{consolidation)} \end{array}$	33	61.0	Max 12 cm <10 cm: 90.0	cT3-4: 96.7 cN+: 90.0	25.3	85.3	74.0	27.3	43.0
Gérard	2006	733	L-RT (45 Gy)	L-RT (45 Gy) + FU- LV (L-CRT1)	33.5	63.5	NR	cT3-4: 98.4 cN+: NR	10.9	90.1	67.3	29.9	81.0

Gérard	2010	598	L-RT (45 Gy) + Cape (L-CRT1)	L-RT (45 Gy) + CapOX (L-CRT2)	46.0	62.0	NR	cT3-4: 92.7 cN+: 72.5	18.9	44.0 (44.0% missing)	69.1	30.5	60.2
Haddad	2017	63	L-RT (50-50.4 Gy) + Cape (L-CRT1)	L-RT (50-50.4 Gy) + CapOX (L-CRT2)	31.7	57.0	6.2	cT3-4: 97.0 cN+: 93.5	34.0	NR	NR	NR	NR
Jiao	2015	206	L-RT (50 Gy) + Cape (L-CRT1)	L-RT (50 Gy) + XELOX (L-CRT2)	44.2	55.8	Max 12 cm <8 cm: 79.6	cT3-4: 98.1 cN+: 78.6	23.3	97.1	71.8	39.8	48.7
Jin	2022	599	L-RT (50 Gy) + Cape (L-CRT1)	S-RT (25 Gy) → CapOX x 4 (S-RT + consolidation)	29.0	55.5	Max 10 cm <5 cm: 49.3	cT3-4: 98.9 cN+: 86.9	13.1	72.1	56.0	27.5	35.0
Kim	2018	110	L-RT (50.4 Gy) + Cape (L-CRT1)	$L-RT (50.4 \text{ Gy}) + Cape \rightarrow CapOX x 2 (L-CRT + consolidation)$	24.0	55.5	5.5	cT3-4: 100 cN+: 92.5	11.3	90.6	52.8	24.5	NR
Latkauskas	2012	150	S-RT (25 Gy) → TME (6 weeks) (S- RTdelayed)	L-RT (50 Gy) + FU- LV (L-CRT1)	33.6	64.0	Max 15 cm <10 cm: 89.3	cT3-4: NR cN+: 76.0	10.6	85.3	72.0	32	39.7
Maréchal	2012	57	L-RT (45 Gy) + FU (L-CRT1)	mFOLFOX6 x $2 \rightarrow L$ - RT (45 Gy) + FU (Induction + L-CRT)	35.0	62.0	NR	cT3-4: 96.4 cN+: NR	25.0	96.4	46.4	21.4	NR
Mei	2023	589	L-RT (50 Gy) + Cape (L-CRT1)	CapOX x 4 (CHT)	38.0	60.0	Max 12 cm <10 cm: 96.6	cT3-4: 94.7 cN+: 69.3	10.0	90.3	66.7	29.3	NR
Mohiuddin	2006	103	L-RT (45.6 Gy) + FU (L-CRT1)	L-RT (45 Gy) + Irinotecan + FU (L- CRT2)	34.0	57.0	Max 9 cm	cT3-4: 100 cN+: NR	26.4	NR	NR	NR	NR
Moore	2017	49	L-RT (50.4 Gy) + FU (L-CRT1)	L-RT (50.4 Gy) + FU \rightarrow FU x 3 (L-CRT + consolidation)	26.5	60.0	6.3	cT3-4: 100 cN+: 100	16.0	92.0	64.0	24.0	NR
Ngan	2012	323	S-RT (25 Gy) \rightarrow TME (7 days) (S-RTearly)	L-RT (50.4 Gy) + FU (L-CRT1)	27.0	63.5	6.6	cT3-4: 100 cN+: 37.9	14.9	93.7	63.3	29.2	70.8
O'Connell	2014	1276	L-RT (50.4 Gy) + FU or Cape (L-CRT1)	L-RT (50.4 Gy) + FU + Oxa or CapOX (L- CRT2)	32.0	NR	Max 12 cm	cT3-4: NR cN+: NR	19.5	NR	NR	NR	54
Rodel	2015	1232	L-RT (50.4 Gy) + FU (L-CRT1)	L-RT (50.4 Gy) + FU + OXA (L-CRT2)	29.0	62.0	Max 12 cm <10 cm: 89.8	cT3-4: 96.2 cN+: 73.7	17.0	92.3	67.9	32.9	49.6
Schmoll	2021	1068	L-RT (45-50.4 Gy) + Cape (L-CRT1)	L-RT (45-50.4 Gy) + CapOX (L-CRT2)	29.2	62	Max 12 cm ≤5 cm: 45.1	cT3-4: 97.5 cN+: 74.1	14.0	95.8	NR	NR	64.0
Schrag	2023	1128	L-RT (50.4 Gy) + FU or Cape (L-CRT1)	mFOLFOX6 x 6 (CHT) (+ selective L- CRT in 53 patients)	34.5	57.2	8.6	cT3-4: 89.0 cN+: 60.2	20.0	90.4	68.4	40.8	58.0
Wang	2019	120	L-RT (50 Gy) + CapOX (L-CRT2)	L-RT (50 Gy + 5 Gy boost) + CapOX → XELOX (L-CRT + consolidation)	30.0	NR	Max 12 cm ≤5 cm: 66.7	cT3-4: 100 cN+: 76.7	23.3	NR	65.0	40.0	42.0

First author	Year	cT2	cT3	cT4	cN+	Threathened/involved MRF	EMVI +	Recurrent cancers	Distance from AV
Aschele	2011						Not specified		12 cm
Bahadoer	2021								16 cm
Bosset	2005				Not specified	Not specified	Not specified		15 cm
Bujko	2004				Not specified	Not specified	Not specified		Inferior margin palpable on DRE
Bujko	2013				Not specified		Not specified		15 cm
Chakrabarti	2021					Not specified	Not specified		Mid-lower rectum
Conroy	2021					Not specified	Not specified		15 cm
Deng	2016					Not specified	Not specified		12 cm
Fernández-Martos	2015					Not specified	Not specified		12 cm
Fokas	2019					Not specified			12 cm
Gérard	2006				Not specified	Not specified	Not specified		Accessible to DRE
Gérard	2010				Not specified	Not specified	Not specified		Accessible to DRE
Haddad	2017					Not specified	Not specified		15 cm
Jiao	2015					Not specified	Not specified		12 cm
Jin	2022					Not specified	Not specified		Mid-lower rectum
Kim	2018				Not specified	Not specified	Not specified		12 cm
Latkauskas	2012					Not specified	Not specified		15 cm
Maréchal	2012					Not specified	Not specified		Not specified
Mei	2023						Not specified		12 cm
Mohiuddin	2006				Not specified	Not specified	Not specified		9 cm
Moore	2017								12 cm
Ngan	2012					Not specified	Not specified		12 cm
O'Connell	2014					Not specified	Not specified		12 cm
Rodel	2015					Not specified	Not specified		12 cm
Schmoll	2021					Not specified	Not specified		12 cm
Schrag	2023						Not specified		12 cm
Wang	2019					Not specified	Not specified		12 cm

eTable 3. Summary of Inclusion and Exclusion Criteria of Included Studies

eFigure 2. Risk of bias of included studies

Risk of Bias Summary

Total number of studies = 27



	Randomization process	Deviations from intended interventions	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall Bias
Low risk	66,7%	92,5%	92,5%	74,1%	63%	33,3%
Some concerns	33,3%	0%	3,7%	25,9%	7%	55,6%
High risk	0%	7,5%	3,7%	0%	0%	11,1%

Risk of Bias Graph

Unique ID	Study ID	Experimental	Comparator	Outcome	Weight	<u>D1</u>	<u>D2</u>	<u>D3</u>	<u>D4</u>	<u>D5</u>	Overall		
Aschele 2011	1	NA	NA	NA	NA	•	•	•	•	•	!	+	Low risk
Bahadoer 202	12	NA	NA	NA	NA	•	•	•	•	•	•	•	Some concerns
Bosset 2005	3	NA	NA	NA	NA	•	•	•	•	•	•	•	High risk
Bujko 2004	4	NA	NA	NA	NA	•	•	•	!	!	<u> </u>		
Bujko (2) 2013	5	NA	NA	NA	NA	•	•	•	!	•	<u> </u>	D1	Randomisation process
Chakrabarty 2	(6	NA	NA	NA	NA	•	•	•	•	!	!	D2	Deviations from the intended interventions
Conroy 2021	7	NA	NA	NA	NA	+	+	•	•	•	+	D3	Missing outcome data
Deng 2016	8	NA	NA	NA	NA	•	+	•	!	•	!	D4	Measurement of the outcome
Fernandez-Ma	a 9	NA	NA	NA	NA	•	+	•	!	•	!	D5	Selection of the reported result
Fokas 2019	10	NA	NA	NA	NA	•	+	•	•	•	+		
Gerard 2006	11	NA	NA	NA	NA	•	+	•	•	!	!		
Gerard (2) 201	112	NA	NA	NA	NA	•	•	•	•	!	!		
Haddad 2017	13	NA	NA	NA	NA	•	•	•	•	•	•		
Jiao 2015	14	NA	NA	NA	NA	+	+	•	•	!	!		
Jin 2022	15	NA	NA	NA	NA	+	+	!	•	•	!		
Kim 2018	16	NA	NA	NA	NA	•	+	•	•	•	!		
Latkauskas 20	117	NA	NA	NA	NA	!	•	•	!	!	•		
Marechal 2012	2 18	NA	NA	NA	NA	•	+	•	•	!	!		
Mei 2023	19	NA	NA	NA	NA	•	+	•	•	•	+		
Mohiuddin 20	(20	NA	NA	NA	NA	!	+	•	•	!	<u> </u>		
Moore 2017	21	NA	NA	NA	NA	•	•	•	•	•	+		
Ngan 2012	22	NA	NA	NA	NA	•	•	•	!	•	!		
O'Connell 201	423	NA	NA	NA	NA	•	•	•	•	•	!		
Rodel 2015	24	NA	NA	NA	NA	•	•	•	•	•	+		
Schmoll 2021	25	NA	NA	NA	NA	•	+	•	•	•	+		
Schrag 2023	27	NA	NA	NA	NA	!	•	•	!	•	•		
Wang 2019	26	NA	NA	NA	NA	•	•	•	•	•	+		

eFigure 3. Transitivity Assessment and Meta-Regression

Continuous variables

We represented the distribution of the variable within each treatment strategy as a boxplots and performed the Kruskal-Wallis equality-of-populations rank test. The results of the meta-regression analyses are also reported.

Potential effect modifiers	Boxplot	Kruskal-Wallis equality-of- populations rank test	Meta-regression analysis
Year of publication	CHT Induction + L-CRT L-CRT + consolidation L-CRT1 L-CRT2 L-RT S-RT + consolidation S-RTdelayed S-RTearly 2,005 2,010 2,010 2,015 2,020 2,025	chi-squared with ties = 15.938 with 8 d.f. p = 0.0433	Coeff. = $-$ 0.1293549 SE = 0.093929 p = 0.168
Mean follow-up length (weeks)	CHT Induction + L-CRT L-CRT + consolidation L-CRT1 L-CRT2 L-RT S-RT + consolidation S-RTdelayed S-RTearly 40 60 80 100 120	chi-squared with ties = 11.771 with 8 d.f. p = 0.1617	Coeff. = -0.034571 SE = 0.7435485 p = 0.963





eAppendix 2. Primary Outcome: Patients With Pathologic Complete Response (PCR)

Characteristics of the network

Number of treatments:

9

Number of studies:

27

Number of individuals included:

13413

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized	
1	CHT	1050	
2	Induction + L-CRT	469	
3	L-CRT + consolidation	288	
4	L-CRT1	6084	
5	L-CRT2	3169	
6	L-RT	872	
7	S-RT + consolidation	1090	
8	S-RTdelayed	75	
9	S-RTearly	316	

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE(logRR)	Risk	Ratio	RR	9	5%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 V Rodel 2015 Aschele 2011 Gerard (2) 2010 Deng 2016 O'Connell 2014 Schmoll 2021 Haddad 2017 Jiao 2015 Mohiuddin 2006 Common effects model Random effects model Heterogeneity: $J^2 = 4\%$, $\tau^2 = 4\%$	/S L-CRT1 0.2661 -0.0090 0.3253 0.7178 0.0947 0.2007 0.6614 0.1823 0.0158 = 0, p = 0.40	0.1368 0.1663 0.1906 0.2495 0.1171 0.1588 0.5581 0.2688 0.3309	-		1.30 0.99 1.38 2.05 1.10 1.22 1.94 1.20 1.02 1.22 1.22	[1.00; [0.72; [0.95; [1.26; [0.87; [0.90; [0.65; [0.71; [0.53; [1.08;]	1.71] 1.37] 2.01] 3.34] 1.38] 1.67] 5.79] 2.03] 1.94] 1.37]	9.5% 6.4% 4.9% 2.8% 12.9% 7.0% 0.6% 2.5% 1.6% 48.1%	4.5% 4.4% 4.2% 4.5% 4.4% 2.4% 3.9% 3.5%
comparison = L-CRT2 \ Deng 2016	/S CHT 1.4110	0.3351			4.10	[2 .13;	7.91]	1.6%	3.5%
comparison = CHT VS I Deng 2016	- CRT1 -0.6931	0.3713	-+		0.50	[0.24;	1.04]	1.3%	3.3%
comparison = L-RTVS Bosset 2005 Gerard 2006 Common effect model Random effects model Heterogeneity: $J^2 = 0\%$, $\tau^2 =$	L-CRT1 -0.9535 -1.1271 = 0, p = 0.65	0.2268 0.3097	+ + + + + + + + + + + + + + + + + + + +		0.39 0.32 0.36 0.36	[0.25; [0.18; [0.25; [0.25;	0.60] 0.59] 0.52] 0.52]	3.4% 1.8% 5.3%	4.1% 3.6% 7.7%
comparison = L-CRT1 V Bujko 2004 Ngan 2012 Common effect model Random effects model Heterogeneity: $J^2 = 0\%$, $\tau^2 =$	/S S-RTearly 3.0782 2.4787 = 0, <i>p</i> = 0.63	1.0162 0.7275			21.72 11.93 14.61 14.61	[2.96; 1 [2.87; [4.58; [4.58;	59.16] 49.63] 46.58] 46.58]	0.2% 0.3% 0.5%	1.1% 1.8% 2.9%
comparison = L-CRT1 \ Latkauskas 2012	/S S-RTdela 0.9808	yed 0.6570			2.67	[0.74;	9.67]	0.4%	2.0%
comparison = Induction Fernandez-Martos 2015	1 + L-CRT V 0.0364	0.4785	_		1.04	[0.41;	<mark>2.6</mark> 5]	0.8%	2.7%
comparison = Inductior Fokas 2019	+ L-CRT V -0.3810	SL-CRT + consol 0.2242	lidatio	-	0.68	[0.44;	1.06]	3.5%	4.1%
$\begin{array}{l} \mbox{comparison} = \$-RT + cc\\ \mbox{Bahadoer} 2021\\ \mbox{Jin} 2022\\ \mbox{Chakrabarti} 2021\\ \mbox{Common effect model}\\ \mbox{Random effects model}\\ \mbox{Heterogeneity: } I^2 = 22\%, \tau^2 \end{array}$	0.7181 0.3508 0.1621 = 0, <i>p</i> = 0.28	VS L-CRT1 0.1466 0.2365 0.4892	_	■	2.05 1.42 1.18 1.80 1.80	[1.54; [0.89; [0.45; [1.42; [1.42;	2.73] 2.26] 3.07] 2.28] 2.28]	8.2% 3.2% 0.7% 12.1%	4.4% 4.0% 2.7% 11.2%
comparison = L-CRT1 V Kim 2018 Moore 2017 Common effect model Random effects model Heterogeneity: I^2 = 42%, τ^2	/S L-CRT + (-0.7302 0.4463 = 0, <i>p</i> = 0.19	consolidation 0.6804 0.5788	+ - v v		0.48 1.56 0.95 0.95	[0.13; [0.50; [0.40; [0.40;	1.83] 4.86] 2.26] 2.26]	0.4% 0.5% 0.9%	1.9% 2.3% 4.2%
comparison = L-CRT1 V Marechal 2012 Conroy 2021 Common effect model Random effects model Heterogeneity: $J^2 = 71\%$, τ^2	/S Induction 0.0984 -0.8151 = 0, <i>p</i> = 0.06	+ L-CRT 0.4446 0.2162	- + ¢ ¢		1.10 0.44 0.53 0.53	[0.46; [0.29; [0.36; [0.36;	2.64] 0.68] 0.77] 0.77]	0.9% 3.8% 4.7%	2.9% 4.1% 7.0%
comparison = S-RT + co Bujko (2) 2013	onsolidation 0.3783	VS L-CRT2 0.2483		-	1.46	[0.90;	2.38]	2.9%	4.0%
comparison = L-CRT2 \ Wang 2019	/SL-CRT+(-0.5596	0.4039	-+		0.57	[0.26;	1.26]	1.1%	3.1%
comparison = L-CRT1 V Mei 2023 Schrag 2023 Common effect model Random effects model Heterogeneity: $J^2 = 0\%$, τ^2 :	/S CHT 0.2197 0.1326 = 0, <i>p</i> = 0.74	0.2331 0.1143		and the second	1.25 1.14 1.16 1.16	[0.79; [0.91; [0.95; [0.95;	1.97] 1.43] 1.42] 1.42]	3.3% 13.6% 16.8% 	4.0% 4.5% 8.6%
Common effect model Random effects model		0.01 0.	1	1 10 100	1.14 1.16	[1.05; [0.91;	1.24] 1.47]	100.0% 	 100.0%

Heterogeneity: $J^2 = 79\%$, $\tau^2 = 0.3161$, p < 0.01Residual heterogeneity: $J^2 = 10\%$, $\tau^2 = 0$, p = 0.34Test for subgroup differences (common effect): $\chi^2_{13} = 119.67$, df = 13 (p < 0.01) Test for subgroup differences (random effects): $\chi^2_{13} = 119.67$, df = 13 (p < 0.01)

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V 6	V7	V8	V 9
1	CHT			0.80 (0.61 to 1.05)	0.24 (0.12 to 0.50)				
2	0.48 (0.30 to 0.75)	Induction + L-CRT	0.68 (0.41 to 1.15)	1.83 (1.17 to 2.84)	1.04 (0.39 to 2.75)				
3	0.38 (0.23 to 0.65)	0.80 (0.53 to 1.22)	L-CRT + consolidation	1.05 (0.43 to 2.55)	1.75 (0.76 to 4.04)				
4	0.75 (0.57 to 0.98)	1.57 (1.09 to 2.25)	1.95 (1.25 to 3.06)	L-CRT1	0.81 (0.69 to 0.94)	2.77 (1.84 to 4.18)	0.58 (0.43 to 0.78)	2.67 (0.71 to 9.95)	14.65 (4.51 to 47.52)
5	0.59 (0.44 to 0.80)	1.24 (0.85 to 1.80)	1.54 (0.98 to 2.43)	0.79 (0.68 to 0.92)	L-CRT2		0.69 (0.39 to 1.20)		
6	2.07 (1.27 to 3.39)	4.34 (2.51 to 7.50)	5.41 (2.94 to 9.94)	2.77 (1.84 to 4.18)	3.52 (2.27 to 5.44)	L-RT			
7	0.42 (0.29 to 0.62)	0.89 (0.57 to 1.39)	1.11 (0.66 to 1.87)	0.57 (0.43 to 0.74)	0.72 (0.54 to 0.96)	0.21 (0.13 to 0.34)	S-RT + consolidation		
8	1.99 (0.52 to 7.64)	4.18 (1.07 to 16.36)	5.20 (1.30 to 20.91)	2.67 (0.71 to 9.95)	3.38 (0.90 to 12.72)	0.96 (0.24 to 3.82)	4.69 (1.22 to 17.97)	S-RTdelayed	
9	10.94 (3.27 to 36.61)	22.94 (6.70 to 78.58)	28.58 (8.11 to 100.74)	14.65 (4.51 to 47.52)	18.57 (5.67 to 60.84)	5.28 (1.52 to 18.38)	25.76 (7.70 to 86.14)	5.49 (0.94 to 32.11)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).



Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² ≤ 0.010 , moderate with $0.010 < tau² <math>\leq 0.242$, high with tau²> 0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

tau^2= 0.0194 ; tau= 0.1393

 $I^2 = 26.44 \% (0\% to 56.84\%)$

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 27.19 20
 0.1300

 Within designs
 12.01 14
 0.6058

 Between designs
 15.18 6
 0.0189

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:CHT 0.11 1 0.7373 L-CRT1:Induction + L-CRT 3.41 1 0.0646 L-CRT1:L-CRT + consolidation 1.73 1 0.1878 L-CRT1:L-CRT2 3.76 7 0.8073 L-CRT1:L-RT 0.20 1 0.6512 L-CRT1:S-RT + consolidation 2.55 2 0.2791 L-CRT1:S-RTearly 0.23 1 0.6314

Between-designs Q statistic after detaching of single designs

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 15.18 6 0.0189 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method Random effects model:

comparison k prop nma direct indir. RoR z p-value L-CRT1:CHT 3 0.97 1.34 1.25 9.82 0.13 -2.61 0.0089 CHT:L-CRT2 1 0.18 0.59 0.24 0.72 0.34 -2.68 0.0073 Induction + L-CRT:L-CRT + consolidation 1 0.67 0.80 0.68 1.11 0.62 -1.06 0.2903 L-CRT1:Induction + L-CRT 2 0.67 0.64 0.55 0.87 0.63 -1.18 0.2374

Induction + L-CRT:L-CRT2 1 0.15 1.24 1.04 1.27 0.81 -0.38 0.7037 L-CRT1:L-CRT + consolidation 2 0.26 0.51 0.95 0.41 2.29 1.58 0.1135 L-CRT + consolidation:L-CRT2 1 0.30 1.54 1.75 1.46 1.20 0.36 0.7196 L-CRT1:L-CRT2 9 0.89 0.79 0.81 0.67 1.21 0.78 0.4332 L-CRT1:S-RT + consolidation 3 0.78 0.57 0.58 0.54 1.07 0.21 0.8325 L-CRT2:S-RT + consolidation 1 0.27 0.72 0.69 0.74 0.93 -0.21 0.8325

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect)

p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
L-CRT + consolidation	0.9319
S-RT + consolidation	0.8769
Induction + L-CRT	0.7870
L-CRT2	0.6431
L-CRT1	0.4901
CHT	0.3574
S-RTdelayed	0.2250
L-RT	0.1844
S-RTearly	0.0042

eAppendix 3. Primary Outcome: Assessment of publication bias

We assessed the risk of publication bias only for the comparisons L-CRT2 vs L-CRT1, which included 9 RCTs.

For the remaining comparisons, too few RCTs to assess publication bias were included.

Funnel plot



Egger's test

Number of studies $=$ 9)	Root MSE = $.9934$					
Std_Eff Coef.	Std. Err.	t P>	t [95%	Conf. Inte	rval]		
slope 0121893 bias 8727976	.0258203 .7336948	-0.47 -1.19	0.651 0.273	0732446 -2.60771	.048866 .862115		

Test of H0: no small-study effects P = 0.273

eAppendix 4. Primary Outcome: Number-needed-to-treat

We calculated the control event rate (CER) by calculating the mean proportion of PCR in individuals receiving L-CRT1, as indicated by Veroniki et al. (J Clin Epidemiol 2019;111:11-22) \rightarrow 22 RCT; mean 0.14; SD 0.056; range 0.05 to 0.27

Therefore, we calculate the number-needed-to-treat-to-benefit or to-harm (NNTB/NNTH) by applying the formula NNT = 1/((1-RR)*CER). Negative values were interpreted as NNTB and positive values as NNTH:

- L-CRT + consolidation vs. L-CRT1: RR 1.95 [1.25; 3.06] → NNTB 7.5 [28.5; 3.5]
- S-RT + consolidation vs. L-CRT1: RR 1.76 [1.34; 2.30] → NNTB 9.4 [21.0; 5.5]
- Induction + L-CRT vs. L-CRT1: RR 1.57 [1.09; 2.25] → NNTB 12.5 [79.4; 5.7]
- L-CRT2 vs. L-CRT1: RR 1.27 [1.09;1.48] → NNTB 26.4 [79.4; 14.8]
- CHT vs. L-CRT1: RR 0.75 [0.57; 0.98] → NNTH 28.6 [16.6; 357.1]
- S-RTdelayed vs. L-CRT1: RR 0.38 [0.10; 1.40] → NNTH 11.5 [NNTH 7.9; ∞; NNTB 17.8]
- L-RT vs. L-CRT1: RR 0.36 [0.24; 0.54] → NNTH 11.2 [9.4; 15.5]
- S-RTearly vs. L-CRT1: RR 0.07 [0.02; 0.22] → NNTH 7.7 [7.3; 9.1]

eAppendix 5. Primary Outcome: CINeMA

The analysis of the certainty of the evidence was performed with the online application CINeMA, which follows the principles of the GRADE methodology. The following criteria were applied:

- Within-study bias: the "overall" risk of bias of each study was calculated as follows: (a) LOW risk if all domains of the Cochrane RoB2 were at low risk; (b) HIGH risk if at least one domain was at high risk; (c) UNCLEAR RISK if for at least one domain there were "some concerns". For each comparison, the histogram was interpreted according to a "Average risk of bias" rule;
- Reporting bias: a "low risk" was considered for all the included studies;
- Indirectness: the histogram was interpreted according to a "Average risk of bias" rule;
- Imprecision, Heterogeneity, Incoherence: Relative effect estimates below 0.67 and above 1.50 are considered clinically important.

Risk of bias contributions

The bar chart shows the contributions of each piece of study to the network estimate



© 2024 Turri G et al. JAMA Network Open.

Indirectness contributions

The bar chart shows the contributions of each study to the network estimate



CINeMA report

Comparison	Number of	Within-	Reportin	Indirectnes	Imprecisio	Heterogeneit	Incoheren	Confidenc	Reason(s) for
	studies	study bias	g bias	S	n	У	ce	e rating	downgrading
CHT vs. L-CRT1	3	Some concerns	Low risk	No concerns	No concerns	Some concerns	Some concerns	Very low	bias","Heterogeneity" ,"Incoherence"]
CHT vs. L-CRT2	1	Some concerns	Low risk	No concerns	No concerns	No concerns	Some concerns	Low	["Within-study bias","Incoherence"]
Induction + L- CRT vs. L-CRT + consolidation	1	No concerns	Low risk	No concerns	Some concerns	No concerns	No concerns	Moderate	["Imprecision"]
Induction + L- CRT vs. L-CRT1	2	No concerns	Low risk	No concerns	No concerns	Some concerns	No concerns	Moderate	["Heterogeneity"]
Induction + L- CRT vs. L-CRT2	1	No concerns	Low risk	No concerns	Some concerns	No concerns	No concerns	Moderate	["Imprecision"]
L-CRT1 vs. L- CRT + consolidation	2	No concerns	Low risk	No concerns	No concerns	No concerns	No concerns	High	0
L-CRT2 vs. L- CRT + consolidation	1	No concerns	Low risk	No concerns	Some concerns	No concerns	No concerns	Moderate	["Imprecision"]
L-CRT1 vs. L- CRT2	9	Some concerns	Low risk	No concerns	No concerns	Some concerns	No concerns	Low	["Within-study bias","Heterogeneity"]
L-CRT1 vs. L- RT	2	No concerns	Low risk	No concerns	No concerns	No concerns	Major concerns	Low	["Incoherence"]
L-CRT1 vs. S- RT + consolidation	3	Some concerns	Low risk	No concerns	No concerns	No concerns	No concerns	Moderate	["Within-study bias"]
L-CRT1 vs. S- RTdelayed	1	Major concerns	Low risk	No concerns	Some concerns	Some concerns	Major concerns	Very low	["Within-study bias","Imprecision"," Heterogeneity","Inco herence"]
L-CRT1 vs. S- RTearly	2	Some concerns	Low risk	Some concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Indirectness"," Incoherence"]
L-CRT2 vs. S- RT + consolidation	1	Some concerns	Low risk	No concerns	No concerns	Some concerns	No concerns	Low	["Within-study bias","Heterogeneity"]
CHT vs. Induction + L- CRT	0	Some concerns	Low risk	No concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Incoherence"]
CHT vs. L-CRT + consolidation	0	Some concerns	Low risk	No concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Incoherence"]
CHT vs. L-RT	0	Some concerns	Low risk	No concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Incoherence"]
CHT vs. S-RT + consolidation	0	Some concerns	Low risk	No concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Incoherence"]
CHT vs. S- RTdelayed	0	Major concerns	Low risk	No concerns	Major concerns	No concerns	Major concerns	Very low	["Within-study bias","Imprecision"," Incoherence"]
CHT vs. S- RTearly	0	Some concerns	Low risk	Some concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Indirectness"," Incoherence"]
Induction + L- CRT vs. L-RT	0	No concerns	Low risk	No concerns	No concerns	No concerns	Major concerns	Very low	["Incoherence"]
Induction + L- CRT vs. S-RT + consolidation	0	No concerns	Low risk	No concerns	Some concerns	Some concerns	Major concerns	Very low	["Imprecision","Heter ogeneity","Incoheren ce"]
Induction + L- CRT vs. S- RTdelayed	0	Some concerns	Low risk	No concerns	No concerns	Some concerns	Major concerns	Very low	["Within-study bias","Heterogeneity" ,"Incoherence"]
Induction + L- CRT vs. S- RTearly	0	Some concerns	Low risk	Some concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Indirectness"," Incoherence"]
L-CRT + consolidation vs. L-RT	0	No concerns	Low risk	No concerns	No concerns	No concerns	Major concerns	Low	["Incoherence"]
L-CRT + consolidation vs.	0	No concerns	Low risk	No concerns	Major concerns	No concerns	Major concerns	Very low	["Imprecision","Inco herence"]

S-RT + consolidation									
L-CRT + consolidation vs. S-RTdelayed	0	Some concerns	Low risk	No concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Incoherence"]
L-CRT + consolidation vs. S-RTearly	0	Some concerns	Low risk	Some concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Indirectness"," Incoherence"]
L-CRT2 vs. L- RT	0	Some concerns	Low risk	No concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Incoherence"]
L-CRT2 vs. S- RTdelayed	0	Some concerns	Low risk	No concerns	Some concerns	No concerns	Major concerns	Very low	["Within-study bias","Imprecision"," Incoherence"]
L-CRT2 vs. S- RTearly	0	Some concerns	Low risk	Some concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Indirectness"," Incoherence"]
L-RT vs. S-RT + consolidation	0	No concerns	Low risk	No concerns	No concerns	No concerns	Major concerns	Low	["Incoherence"]
L-RT vs. S- RTdelayed	0	Some concerns	Low risk	No concerns	Major concerns	No concerns	Major concerns	Very low	["Within-study bias","Imprecision"," Incoherence"]
L-RT vs. S- RTearly	0	Some concerns	Low risk	Some concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Indirectness"," Incoherence"]
S-RT + consolidation vs. S-RTdelayed	0	Some concerns	Low risk	No concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Incoherence"]
S-RT + consolidation vs. S-RTearly	0	Some concerns	Low risk	Some concerns	No concerns	No concerns	Major concerns	Very low	["Within-study bias","Indirectness"," Incoherence"]
S-RTdelayed vs. S-RTearly	0	Major concerns	Low risk	Some concerns	Some concerns	No concerns	Major concerns	Very low	["Within-study bias","Indirectness"," Imprecision","Incohe rence"]

eAppendix 6. Primary Outcome: Sensitivity analyses for the primary outcome "PCR"

Sensitivity 1 - excluding studies with high risk of bias

Characteristics of the network

Number of treatments:

8

Number of studies:

24

Number of individuals included:

12072

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized	
1	CHT	465	
2	Induction + L-CRT	469	
3	L-CRT + consolidation	288	
4	L-CRT1	5435	
5	L-CRT2	3137	
6	L-RT	872	
7	S-RT + consolidation	1090	
8	S-RTearly	316	

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	(logRR)	Risk	Ratio	RR	9	5%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 V Rodel 2015 Aschele 2011 Gerard (2) 2010 Deng 2016 O'Connell 2014 Schmoll 2021 Jiao 2015 Mohiuddin 2006 Common effect model Random effects model Heterogeneity: $J^2 = 9\%$, $\tau^2 = 3\%$	/S L-CRT1 0.2661 -0.0990 0.3253 0.7178 0.0947 0.2007 0.1823 0.0158 = 0.0030, <i>p</i> =	0.1368 0.1663 0.1906 0.2495 0.1171 0.1588 0.2688 0.3309			1.30 0.99 1.38 2.05 1.10 1.22 1.20 1.02 1.21 1.22	[1.00; [0.72; [0.95; [1.26; [0.87; [0.90; [0.71; [0.53; [1.08; [1.07;]	1.71] 1.37] 2.01] 3.34] 1.38] 1.67] 2.03] 1.94] 1.37] 1.38]	11.1% 7.5% 5.7% 3.3% 15.1% 8.2% 2.9% 1.9% 55.7%	4.8% 4.7% 4.6% 4.3% 4.9% 4.8% 4.2% 3.9%
comparison = L-CRT2 \ Deng 2016	/S CHT 1.4110	0.3351			4.10	[2.13;	7.91]	1.8%	3.9%
comparison = CHT VS I Deng 2016	- CRT1 -0.6931	0.3713		-	0.50	[0.24;	1.04]	1.5%	3.7%
$\label{eq:comparison} \begin{array}{l} \text{comparison} = \text{L-RTVS} \\ \text{Bosset 2005} \\ \text{Gerard 2006} \\ \text{Common effect model} \\ \text{Random effects model} \\ \text{Heterogeneity:} \ J^2 = 0\%, \ \tau^2 \end{array}$	L-CRT1 -0.9535 -1.1271 = 0.0030, <i>p</i> =	0.2268 0.3097 0.65	+ + + + + + + + + + + + + + + + + + + +		0.39 0.32 0.36 0.36	[0.25; [0.18; [0.25; [0.25;	0.60] 0.59] 0.52] 0.52]	4.0% 2.2% 6.2%	4.4% 4.0% 8.4%
$\label{eq:comparison} \begin{array}{l} \text{comparison} = \text{L-CRT1} \ \text{N}\\ \text{Bujko 2004}\\ \text{Ngan 2012}\\ \text{Common effect model}\\ \textbf{Random effects model}\\ \text{Heterogeneity:} \ l^2 = 0\%, \ \tau^2 = 0 \ t^2 = 0\%, \ \tau^2 = 0\%, \ \tau^2$	/S S-RTearl 3.0782 2.4787 = 0.0030, <i>p</i> =	y 1.0162 0.7275 0.63			21.72 11.93 14.61 14.62	[2.96; 1 [2.87; [4.58; [4.57;	59.16] 49.63] 46.58] 46.73]	0.2% 0.4% 0.6%	1.3% 2.0% 3.3%
comparison = Induction Fernandez-Martos 2015	0.0364	S L-CRT2 0.4785	_		1.04	[0.41;	2.65]	0.9%	3.1%
comparison = Inductior Fokas 2019	-0.3810	S L-CRT + consol 0.2242	idatio		0.68	[0.44;	1.06]	4.1%	4.5%
$\begin{array}{l} \mbox{comparison} = \mbox{S-RT} + \mbox{comparison} \\ \mbox{Bahadoer} \ 2021 \\ \mbox{Jin} \ 2022 \\ \mbox{Chakrabarti} \ 2021 \\ \mbox{Chakrabarti} \ 2021 \\ \mbox{Common effect model} \\ \mbox{Random effects model} \\ \mbox{Heterogeneity:} \ \ ^2 = 22\%, \ \tau^2 \end{array}$	0.7181 0.3508 0.1621 = 0.0030, p =	VSL-CRT1 0.1466 0.2365 0.4892	_	₽ ₽ ♦ ♦	2.05 1.42 1.18 1.80 1.79	[1.54; [0.89; [0.45; [1.42; [1.39;	2.73] 2.26] 3.07] 2.28] 2.29]	9.6% 3.7% 0.9% 14.2%	4.8% 4.4% 3.0% 12.2%
comparison = L-CRT1 V Kim 2018 Moore 2017 Common effect model Random effects model Heterogeneity: $J^2 = 42\%$, τ^2	/S L-CRT + -0.7302 0.4463	consolidation 0.6804 0.5788	+ - v v		0.48 1.56 0.95 0.95	[0.13; [0.50; [0.40; [0.40;	1.83] 4.86] 2.26] 2.27]	0.4% 0.6% 1.1%	2.2% 2.6% 4.8%
$\label{eq:comparison} \begin{array}{l} \text{comparison} = \text{L-CRT1} \\ \text{Marechal 2012} \\ \text{Conroy 2021} \\ \text{Common effect model} \\ \text{Random effects model} \\ \text{Heterogeneity:} \ l^2 = 71\%, \ \tau^2 \end{array}$	/S Induction 0.0984 -0.8151 = 0.0030, p =	0.4446 0.2162	+ 0 0		1.10 0.44 0.53 0.53	[0.46; [0.29; [0.36; [0.36;	2.64] 0.68] 0.77] 0.78]	1.0% 4.4% 5.5%	3.3% 4.5% 7.7%
comparison = S-RT + co Bujko (2) 2013	onsolidation 0.3783	0.2483		<u></u>	1.46	[0.90;	2.38]	3.4%	4.3%
comparison = L-CRT2 \ Wang 2019	/SL-CRT+ -0.5596	consolidation 0.4039	-+		0.57	[0. 2 6;	1.26]	1.3%	3.5%
comparison = L-CRT1 \ Mei 2023	/S CHT 0.2197	0.2331		-	1.25	[0. 7 9;	1.97]	3.8%	4.4%
Common effect model Random effects model		[]		¢ 	1.13 1.13	[1.03; [0.87;	1.24] 1.46]	100.0% 	 100.0%
11 J 12 CIC 2	0.0400	0.01 0.	1	1 10 10	00				

Heterogeneity: $l^2 = 81\%$, $\tau^2 = 0.3488$, p < 0.01Residual heterogeneity: $l^2 = 18\%$, $\tau^2 = 0.0030$, p = 0.26Test for subgroup differences (common effect): $\chi^2_{12} = 117.86$, df = 12 (p < 0.01) Test for subgroup differences (random effects): $\chi^2_{12} = 112.17$, df = 12 (p < 0.01)

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8
1	CHT			0.69 (0.45 to 1.07)	0.24 (0.12 to 0.49)			
2	0.38 (0.22 to 0.66)	Induction + L-CRT	0.68 (0.41 to 1.14)	1.83 (1.19 to 2.83)	1.04 (0.39 to 2.74)	100 C	100 C	
3	0.31 (0.17 to 0.56)	0.80 (0.53 to 1.22)	L-CRT + consolidation	1.05 (0.44 to 2.55)	1.75 (0.76 to 4.02)			
4	0.60 (0.39 to 0.90)	1.57 (1.10 to 2.24)	1.95 (1.25 to 3.04)	L-CRT1	0.81 (0.70 to 0.95)	2.77 (1.85 to 4.16)	0.58 (0.43 to 0.78)	14.64 (4.52 to 47.42)
5	0.48 (0.31 to 0.73)	1.25 (0.86 to 1.82)	1.56 (0.99 to 2.45)	0.80 (0.69 to 0.93)	L-CRT2		0.69 (0.39 to 1.19)	
6	1.65 (0.92 to 2.95)	4.34 (2.53 to 7.45)	5.40 (2.96 to 9.86)	2.77 (1.85 to 4.16)	3.46 (2.25 to 5.33)	L-RT		
7	0.34 (0.21 to 0.55)	0.89 (0.57 to 1.39)	1.11 (0.66 to 1.86)	0.57 (0.44 to 0.74)	0.71 (0.53 to 0.95)	0.21 (0.13 to 0.33)	S-RT + consolidation	
8	8.72 (2.51 to 30.34)	22.94 (6.72 to 78.33)	28.55 (8.13 to 100.30)	14.64 (4.52 to 47.42)	18.28 (5.59 to 59.76)	5.29 (1.52 to 18.32)	25.72 (7.71 to 85.76)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).



Assessment of heterogeneity and consistency

Global heterogeneity

^{© 2024} Turri G et al. JAMA Network Open.

We interpreted tau² as follows: heterogeneity low with tau² ≤ 0.010 , moderate with $0.010 < tau² <math>\leq 0.242$, high with tau²> 0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

tau^2= 0.0173 ; tau= 0.1316

I^2= 24.42 % (0 % to 56.69 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 23.82 18 0.1611

 Within designs
 11.10 12 0.5207

 Between designs
 12.72 6 0.0477

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:Induction + L-CRT 3.41 1 0.0646 L-CRT1:L-CRT + consolidation 1.73 1 0.1878 L-CRT1:L-CRT2 2.96 6 0.8139 L-CRT1:L-RT 0.20 1 0.6512 L-CRT1:S-RT + consolidation 2.55 2 0.2791 L-CRT1:S-RTearly 0.23 1 0.6314

Between-designs Q statistic after detaching of single designs

 $\begin{array}{ccccccc} Detached \ design & Q \ df \ p-value \\ Induction + L-CRT:L-CRT + consolidation 11.10 & 5 & 0.0493 \\ Induction + L-CRT:L-CRT2 12.44 & 5 & 0.0293 \\ L-CRT + consolidation:L-CRT2 12.69 & 5 & 0.0264 \\ & L-CRT1:CHT & 8.38 & 5 & 0.1366 \\ L-CRT1:Induction + L-CRT 10.59 & 5 & 0.0601 \\ L-CRT1:L-CRT + consolidation 9.80 & 5 & 0.0810 \\ & L-CRT1:L-CRT + consolidation 9.80 & 5 & 0.0476 \\ L-CRT1:S-RT + consolidation 12.72 & 5 & 0.0261 \\ L-CRT2:S-RT + consolidation 12.72 & 5 & 0.0261 \\ L-CRT1:CHT:L-CRT2 & 3.76 & 4 & 0.4393 \\ \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 12.72 6 0.0477 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value L-CRT1:CHT 2 0.92 1.68 1.45 10.20 0.14 -2.44 0.0147 CHT:L-CRT2 1 0.37 0.48 0.24 0.71 0.34 -2.35 0.0188 Induction + L-CRT:L-CRT + consolidation 1 0.67 0.80 0.68 1.12 0.61 -1.09 0.2773 L-CRT1:Induction + L-CRT 2 0.67 0.64 0.55 0.88 0.62 -1.23 0.2184 Induction + L-CRT:L-CRT2 1 0.15 1.25 1.04 1.30 0.80 -0.42 0.6778 L-CRT1:L-CRT + consolidation 2 0.25 0.51 0.95 0.42 2.28 1.58 0.1136

L-CRT + consolidation:L-CRT2 1 0.29 1.56 1.75 1.49 1.18 0.32 0.7492 L-CRT1:L-CRT2 8 0.89 0.80 0.81 0.70 1.16 0.59 0.5524 L-CRT1:S-RT + consolidation 3 0.79 0.57 0.58 0.55 1.05 0.15 0.8772 L-CRT2:S-RT + consolidation 1 0.27 0.71 0.69 0.72 0.95 -0.15 0.8772

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
L-CRT + consolidation	0.9249
S-RT + consolidation	0.8614
Induction + L-CRT	0.7620
L-CRT2	0.5930
L-CRT1	0.4290
CHT	0.2803
L-RT	0.1488
S-RTearly	0.0007

Sensitivity 2 - excluding studies with high indirectness

Number of treatments:

9

Number of studies:

23

Number of individuals included:

10811

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized	
1	CHT	465	
2	Induction + L-CRT	469	
3	L-CRT + consolidation	288	
4	L-CRT1	4802	
5	L-CRT2	3169	
6	L-RT	367	
7	S-RT + consolidation	1021	
8	S-RTdelayed	75	
9	S-RTearly	155	

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SI	E(logRR)	Risk F	Ratio	RR	95	%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 V Rodel 2015 Aschele 2011 Gerard (2) 2010 Deng 2016 O'Connell 2014 Schmoll 2021 Haddad 2017 Jiao 2015 Mohiuddin 2006 Common effect model Random effects model Heterogeneity: $J^2 = 4\%$, τ^2 :	/S L-CRT1 0.2661 -0.0090 0.3253 0.7178 0.0947 0.2007 0.6614 0.1823 0.0158 = 0.0049, p =	0.1368 0.1663 0.1906 0.2495 0.1171 0.1588 0.5581 0.2688 0.3309			1.30 0.99 1.38 2.05 1.10 1.22 1.94 1.20 1.02 1.22 1.23	[1.00; [0.72; [0.95; [1.26; [0.87; [0.90; [0.65; [0.71; [0.53; [1.08;]	1.71] 1.37] 2.01] 3.34] 1.38] 1.67] 5.79] 2.03] 1.94] 1.37] 1.40]	11.5% 7.8% 5.9% 3.5% 15.8% 8.6% 0.7% 3.0% 2.0% 58.8%	5.4% 5.2% 5.0% 4.6% 5.5% 2.6% 4.5% 4.0%
comparison = L-CRT2 \ Deng 2016	/S CHT 1.4110	0.3351		_ #	4.10	[2.1 3;	7.91]	1.9%	4.0%
comparison = CHT VS I Deng 2016	- CRT1 -0.6931	0.3713			0.50	[0.24;	1.04]	1.6%	3.7%
comparison = L-RT VS Gerard 2006	L-CRT1 -1.1271	0.3097			0.32	[0. 1 8;	0.59]	2.3%	4.2%
comparison = L-CRT1 \ Bujko 2004	/S S-RTea 3.0782	rly 1.0162		+	21.72	[2.96; 15	9.16]	0.2%	1.1%
comparison = L-CRT1 \ Latkauskas 2012	/S S-RTde 0.9808	layed 0.6570	-		2.67	[0.74;	9.67]	0.5%	2.1%
comparison = Inductior Fernandez-Martos 2015	0.0364	/SL-CRT2 0.4785	_		1.04	[0.41;	2 .65]	0.9%	3.0%
comparison = Inductior Fokas 2019	+ L-CRT -0.3810	/S L-CRT + cons 0.2242	olidation		0.68	[0.44;	1.06]	4.3%	4.8%
$\begin{array}{l} \mbox{comparison} = \mbox{S-RT} + \mbox{cal}\\ \mbox{Bahadoer} \ 2021\\ \mbox{Jin} \ 2022\\ \mbox{Common effect model}\\ \mbox{Random effects model}\\ \mbox{Heterogeneity:} \ l^2 = 43\%, \ \tau^2 \end{array}$	0.7181 0.3508 = 0.0049, <i>p</i>	on VS L-CRT1 0.1466 0.2365 = 0.19	-	₩ ◆ ◆	2.05 1.42 1.85 1.83	[1.54; [0.89; [1.45; [1.41;	2.73] 2.26] 2.36] 2.39]	10.1% 3.9% 13.9% 	5.3% 4.7% 10.0%
comparison = L-CRT1 V Kim 2018 Moore 2017 Common effect model Random effects model Heterogeneity: J^2 = 42%, τ^2	/S L-CRT + -0.7302 0.4463	<pre>consolidation</pre>		- - A A	0.48 1.56 0.95 0.95	[0.13; [0.50; [0.40; [0.40;	1.83] 4.86] 2.26] 2.27]	0.5% 0.6% 1.1% 	2.0% 2.5% 4.5%
comparison = L-CRT1 V Marechal 2012 Conroy 2021 Common effect model Random effects model Heterogeneity: I^2 = 71%, τ^2	/S Induction 0.0984 -0.8151 = 0.0049, p	on + L-CRT 0.4446 0.2162 = 0.06	_ +		1.10 0.44 0.53 0.53	[0.46; [0.29; [0.36; [0.36;	2.64] 0.68] 0.77] 0.79]	1.1% 4.6% 5.7%	3.2% 4.8% 8.1%
comparison = S-RT + co Bujko (2) 2013	onsolidatio 0.3783	on VS L-CRT2 0.2483	-	. –	1.46	[0.90;	2.38]	3.5%	4.6%
comparison = L-CRT2 \ Wang 2019	/S L-CRT + -0.5596	consolidation 0.4039		-	0.57	[0.26;	1.26]	1.3%	3.5%
comparison = L-CRT1 \ Mei 2023	/S CHT 0.2197	0.2331	-	-	1.25	[0.79;	1.97]	4.0%	4.7%
Common effect model Random effects model				>	1.18 1.16	[1.08; [0.92;	1.29] 1.47]	100.0% 	 100.0%
Heterogeneity: $l^2 = 77\% \tau^2$	= 0 2485 n	0.01	0.1 1	10	100				

Heterogeneity: $I^{*} = 77\%$, $\tau^{*} = 0.2485$, p < 0.01Residual heterogeneity: $I^{2} = 28\%$, $\tau^{2} = 0.0049$, p = 0.17Test for subgroup differences (common effect): $\chi_{13}^{2} = 87.16$, df = 13 (p < 0.01) Test for subgroup differences (random effects): $\chi_{13}^{2} = 80.60$, df = 13 (p < 0.01)
Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8	V 9
1	CHT		100 C	0.69 (0.44 to 1.08)	0.24 (0.12 to 0.50)	1.1.1	100 C	100 C	
2	0.38 (0.22 to 0.67)	Induction + L-CRT	0.68 (0.40 to 1.16)	1.82 (1.16 to 2.85)	1.04 (0.39 to 2.77)				
3	0.31 (0.16 to 0.57)	0.80 (0.52 to 1.24)	L-CRT + consolidation	1.05 (0.43 to 2.57)	1.75 (0.75 to 4.07)				
4	0.59 (0.39 to 0.91)	1.56 (1.08 to 2.25)	1.94 (1.23 to 3.05)	L-CRT1	0.80 (0.68 to 0.95)	3.09 (1.57 to 6.06)	0.56 (0.40 to 0.77)	2.67 (0.71 to 9.99)	21.72 (2.90 to 162.6
5	0.47 (0.30 to 0.73)	1.23 (0.84 to 1.81)	1.53 (0.96 to 2.43)	0.79 (0.68 to 0.92)	L-CRT2		0.69 (0.39 to 1.21)		
6	1.83 (0.82 to 4.08)	4.80 (2.23 to 10.36)	5.97 (2.65 to 13.49)	3.09 (1.57 to 6.06)	3.90 (1.95 to 7.80)	L-RT			1. Sec. 1. Sec
7	0.33 (0.20 to 0.55)	0.86 (0.54 to 1.37)	1.07 (0.63 to 1.83)	0.55 (0.41 to 0.74)	0.70 (0.51 to 0.95)	0.18 (0.09 to 0.37)	S-RT + consolidation		
8	1.58 (0.39 to 6.35)	4.15 (1.05 to 16.35)	5.16 (1.28 to 20.88)	2.67 (0.71 to 9.99)	3.37 (0.89 to 12.75)	0.86 (0.20 to 3.81)	4.82 (1.25 to 18.64)	S-RTdelayed	
9	12.90 (1.65 to 101.11)	33.80 (4.37 to 261.71)	42.04 (5.34 to 331.28)	21.72 (2.90 to 162.66)	27.45 (3.64 to 206.81)	7.04 (0.84 to 58.83)	39.28 (5.14 to 300.29)	8.14 (0.73 to 90.52)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: other vs 'L-((Random Effects Mode	CRT1' el)	RR	95%-CI
L-CRT + consolidation S-RT + consolidation Induction + L-CRT L-CRT2 CHT S-RTdelayed L-RT S-RTearly			1.94 1.81 1.56 1.26 0.59 0.38 0.32 0.05	[1.23; 3.05] [1.36; 2.41] [1.08; 2.25] [1.08; 1.48] [0.39; 0.91] [0.10; 1.41] [0.16; 0.64] [0.01; 0.34]
0	.01 0.5 1	2 4		

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau² \geq 0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

tau^2= 0.0227 ; tau= 0.1505

I^2= 31.12 % (0 % to 61.64 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 23.23
 16
 0.1077

 Within designs
 10.65
 10
 0.3854

 Between designs
 12.58
 6
 0.0502

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:Induction + L-CRT 3.41 1 0.0646 L-CRT1:L-CRT + consolidation 1.73 1 0.1878 L-CRT1:L-CRT2 3.76 7 0.8073 L-CRT1:S-RT + consolidation 1.74 1 0.1867

Between-designs Q statistic after detaching of single designs

 $\begin{array}{c|c} Detached \ design & Q \ df \ p-value \\ Induction + L-CRT:L-CRT + consolidation 10.98 \ 5 \ 0.0518 \\ Induction + L-CRT:L-CRT2 \ 12.31 \ 5 \ 0.0308 \\ L-CRT + consolidation:L-CRT2 \ 12.55 \ 5 \ 0.0280 \\ L-CRT1:CHT \ 8.28 \ 5 \ 0.1414 \\ L-CRT1:Induction + L-CRT \ 10.48 \ 5 \ 0.0626 \\ L-CRT1:L-CRT + consolidation \ 9.64 \ 5 \ 0.0860 \\ L-CRT1:L-CRT2 \ 7.93 \ 5 \ 0.1604 \\ L-CRT1:S-RT + consolidation \ 12.57 \ 5 \ 0.0277 \\ L-CRT1:CHT:L-CRT2 \ 3.78 \ 4 \ 0.4368 \\ \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 12.02 6 0.0615 0.0515 0.0027

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value L-CRT1:CHT 2 0.92 1.68 1.45 9.60 0.15 -2.31 0.0206 CHT:L-CRT2 1 0.38 0.47 0.24 0.70 0.35 -2.27 0.0234 Induction + L-CRT:L-CRT + consolidation 1 0.66 0.80 0.68 1.10 0.62 -1.04 0.3004 L-CRT1:Induction + L-CRT 2 0.67 0.64 0.55 0.87 0.63 -1.16 0.2468 Induction + L-CRT:L-CRT2 1 0.15 1.23 1.04 1.27 0.82 -0.37 0.7098 L-CRT1:L-CRT + consolidation 2 0.26 0.52 0.95 0.42 2.28 1.56 0.1190 L-CRT + consolidation:L-CRT2 1 0.30 1.53 1.75 1.45 1.21 0.37 0.7117 L-CRT1:L-CRT2 9 0.89 0.79 0.80 0.69 1.16 0.58 0.5634

L-CRT1:S-RT + consolidation 2 0.76 0.55 0.56 0.54 1.03 0.08 0.9353 L-CRT2:S-RT + consolidation 1 0.29 0.70 0.69 0.70 0.97 -0.08 0.9353 Legend: comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) - z-value of test for disagreement (direct versus indirect) Z

p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
L-CRT + consolidation	n 0.9237
S-RT + consolidation	0.8895
Induction + L-CRT	0.7810
L-CRT2	0.6439
L-CRT1	0.4912
CHT	0.3344
S-RTdelayed	0.2429
L-RT	0.1821
S-RTearly	0.0113

Sensitivity 3 - excluding CHT arm

Characteristics of the network

Number of treatments:

8

Number of studies:

24

Number of individuals included:

11201

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized
1	Induction + L-CRT	469
2	L-CRT + consolidation	288
3	L-CRT1	5087
4	L-CRT2	3004
5	L-RT	872
6	S-RT + consolidation	1090
7	S-RTdelayed	75
8	S-RTearly	316

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE(logRR)	Risk F	Ratio	RR	9	5%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 V Rodel 2015 Aschele 2011 Gerard (2) 2010 O'Connell 2014 Schmoll 2021 Haddad 2017 Jiao 2015 Mohiuddin 2006 Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0\%$	S L-CRT1 0.2661 -0.0090 0.3253 0.0947 0.2007 0.6614 0.1823 0.0158	0.1368 0.1663 0.1906 0.1171 0.1588 0.5581 0.2688 0.3309		► ► ► ► ► ►	1.30 0.99 1.38 1.10 1.22 1.94 1.20 1.02 1.18 1.18	[1.00; [0.72; [0.95; [0.87; [0.87; [0.65; [0.71; [0.73; [1.04;]	1.71] 1.37] 2.01] 1.38] 1.67] 5.79] 2.03] 1.94] 1.33]	12.2% 8.3% 6.3% 9.1% 0.7% 3.2% 2.1% 58.5%	5.6% 5.5% 5.3% 5.7% 2.9% 4.8% 4.3%
comparison = L-RT VS I Bosset 2005 Gerard 2006 Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 =$	L-CRT1 -0.9535 -1.1271	0.2268 0.3097	+ + + + + + + + + + + + + + + + + + + +		0.39 0.32 0.36 0.36	[0.25; [0.18; [0.25; [0.25;	0.60] 0.59] 0.52] 0.52]	4.4% 2.4% 6.8%	5.1% 4.5% 9.6%
comparison = L-CRT1 V Bujko 2004 Ngan 2012 Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 =$	S S-RTearly 3.0782 2.4787	1.0162 0.7275			21.72 11.93 14.61 14.61	[2.96; 1 [2.87; [4.58; [4.58;	59.16] 49.63] 46.58] 46.58]	0.2% 0.4% 0.7%	1.3% 2.1%
comparison = L-CRT1 V Latkauskas 2012	/S S-RTdela 0.9808	yed 0.6570	-		2.67	[0.74;	9.67]	0.5%	2.4%
comparison = Induction Fernandez-Martos 2015	+ L-CRT V 0.0364	0.4785	_	_	1.04	[0.41;	2 .65]	1.0%	3.3%
comparison = Induction Fokas 2019	+ L-CRT V -0.3810	S L-CRT + consol 0.2242	idation =		0.68	[0.44;	1.06]	4.5%	5.1%
comparison = S-RT + cc Bahadoer 2021 Jin 2022 Chakrabarti 2021 Common effect model Random effects model Heterogeneity: $J^2 = 22\%$, τ^2	0.7181 0.3508 0.1621 = 0, <i>p</i> = 0.28	VS L-CRT1 0.1466 0.2365 0.4892		*	2.05 1.42 1.18 1.80 1.80	[1.54; [0.89; [0.45; [1.42; [1.42;	2.73] 2.26] 3.07] 2.28] 2.28]	10.6% 4.1% 1.0% 15.7%	5.6% 5.0% 3.3% 13.9%
comparison = L-CRT1 V Kim 2018 Moore 2017 Common effect model Random effects model Heterogeneity: $I^2 = 42\%$, τ^2	S L-CRT + 0 -0.7302 0.4463 = 0, <i>p</i> = 0.19	consolidation 0.6804 0.5788		- + ^ > >	0.48 1.56 0.95 0.95	[0.13; [0.50; [0.40; [0.40;	1.83] 4.86] 2.26] 2.26]	0.5% 0.7% 1.2% 	2.3% 2.8% 5.1%
comparison = L-CRTI V Marechal 2012 Conroy 2021 Common effect model Random effects model Heterogeneity: $l^2 = 71\%$, τ^2	S Induction 0.0984 -0.8151 = 0, <i>p</i> = 0.06	+ L-CRT 0.4446 0.2162	+		1.10 0.44 0.53 0.53	[0.46; [0.29; [0.36; [0.36;	2.64] 0.68] 0.77] 0.77]	1.2% 4.9% 6.0%	3.6% 5.2% 8.7%
comparison = S-RT + co Bujko (2) 2013	0.3783	VS L-CRT2 0.2483		.	1.46	[0.90;	2.38]	3.7%	4.9%
comparison = L-CRT2 V Wang 2019	/ <mark>S L-CRT +</mark> (-0.5596	0.4039			0.57	[0.26;	1.26]	1.4%	3.8%
Common effect model Random effects model		[]		> 	1.10 1.10	[1.00; [0.85;	1.20] 1.42]	100.0% 	 100.0%
Heterogeneity: $I^2 = 79\%$, τ^2	= 0.2865, p <	0.01 0.	1 1	10 100)				

Heterogeneity: $l^2 = /9\%$, $\tau^2 = 0.2865$, p < 0.01Residual heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, p = 0.54Test for subgroup differences (common effect): $\chi^2_{10} = 98.54$, df = 10 (p < 0.01) Test for subgroup differences (random effects): $\chi^2_{10} = 98.54$, df = 10 (p < 0.01)

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8
1	Induction + L-CRT	0.68 (0.44 to 1.06)	1.90 (1.30 to 2.78)	1.04 (0.41 to 2.65)				
2	0.80 (0.55 to 1.16)	L-CRT + consolidation	1.05 (0.44 to 2.49)	1.75 (0.79 to 3.86)				
3	1.61 (1.17 to 2.22)	2.01 (1.34 to 3.02)	L-CRT1	0.85 (0.75 to 0.96)	2.76 (1.93 to 3.95)	0.56 (0.44 to 0.70)	2.67 (0.74 to 9.67)	14.61 (4.58 to 46.58)
4	1.36 (0.97 to 1.89)	1.69 (1.12 to 2.56)	0.84 (0.75 to 0.95)	L-CRT2		0.69 (0.42 to 1.11)		
5	4.44 (2.75 to 7.18)	5.55 (3.22 to 9.55)	2.76 (1.93 to 3.95)	3.27 (2.24 to 4.78)	L-RT			
6	0.90 (0.61 to 1.32)	1.13 (0.71 to 1.78)	0.56 (0.45 to 0.69)	0.66 (0.53 to 0.84)	0.20 (0.13 to 0.31)	S-RT + consolidation		
7	4.30 (1.14 to 16.19)	5.37 (1.39 to 20.71)	2.67 (0.74 to 9.67)	3.17 (0.87 to 11.54)	0.97 (0.25 to 3.68)	4.77 (1.29 to 17.59)	S-RTdelayed	
8	23.54 (7.07 to 78.35)	29.40 (8.60 to 100.48)	14.61 (4.58 to 46.58)	17.35 (5.41 to 55.65)	5.30 (1.57 to 17.84)	26.13 (8.04 to 84.95)	5.48 (0.97 to 30.99)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).



Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² ≤ 0.010 , moderate with $0.010 < tau² <math>\leq 0.242$, high with tau²> 0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0$; tau = 0

 $I^2 = 0 \% (0 \% \text{ to } 49.97 \%)$

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

Q df p-valueTotal15.64 17 0.5498Within designs11.89 13 0.5364Between designs3.74 4 0.4421

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:Induction + L-CRT 3.41 1 0.0646 L-CRT1:L-CRT + consolidation 1.73 1 0.1878 L-CRT1:L-CRT2 3.76 7 0.8073 L-CRT1:L-RT 0.20 1 0.6512 L-CRT1:S-RT + consolidation 2.55 2 0.2791 L-CRT1:S-RTearly 0.23 1 0.6314

Between-designs Q statistic after detaching of single designs

 $\label{eq:constraint} \begin{array}{cccc} Detached \ design & Q \ df \ p-value \\ Induction + L-CRT:L-CRT + \ consolidation \ 2.03 & 3 & 0.5661 \\ Induction + L-CRT:L-CRT2 \ 3.38 & 3 & 0.3364 \\ L-CRT + \ consolidation:L-CRT2 \ 3.73 & 3 & 0.2918 \\ L-CRT1:Induction + L-CRT \ 1.38 & 3 & 0.7104 \\ L-CRT1:L-CRT + \ consolidation \ 0.93 & 3 & 0.8174 \\ L-CRT1:L-CRT2 \ 3.63 & 3 & 0.3046 \\ L-CRT1:S-RT + \ consolidation \ 3.72 & 3 & 0.2931 \\ L-CRT2:S-RT + \ consolidation \ 3.72 & 3 & 0.2931 \\ \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 3.74 4 0.4421 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value Induction + L-CRT:L-CRT + consolidation 1 0.71 0.80 0.68 1.18 0.58 -1.31 0.1908 L-CRT1:Induction + L-CRT 2 0.70 0.62 0.53 0.91 0.58 -1.54 0.1243 Induction + L-CRT:L-CRT2 1 0.12 1.36 1.04 1.41 0.74 -0.60 0.5485 L-CRT1:L-CRT + consolidation 2 0.22 0.50 0.95 0.41 2.31 1.68 0.0938 L-CRT + consolidation:L-CRT2 1 0.27 1.69 1.75 1.67 1.05 0.09 0.9254 L-CRT1:L-CRT2 8 0.93 0.84 0.85 0.78 1.08 0.34 0.7350 L-CRT1:S-RT + consolidation 3 0.82 0.56 0.56 0.58 0.96 -0.14 0.8869 L-CRT2:S-RT + consolidation 1 0.23 0.66 0.69 0.66 1.04 0.14 0.8869

Legend: comparison - Treatment comparison - Number of studies providing direct evidence k prop - Direct evidence proportion - Estimated treatment effect (RR) in network meta-analysis nma direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
L-CRT + consolidation	0.9372
S-RT + consolidation	0.8572
Induction + L-CRT	0.7661
L-CRT2	0.5714
L-CRT1	0.4195
S-RTdelayed	0.2333
L-RT	0.2110
S-RTearly	0.0044

Sensitivity 4 - PCR on the Per Protocol Populations

Characteristics of the network

Number of treatments:

9

Number of studies:

27

Number of individuals included:

12558

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized
1	CHT	959
2	Induction + L-CRT	437
3	L-CRT + consolidation	269
4	L-CRT1	5707
5	L-CRT2	3009
6	L-RT	836
7	S-RT + consolidation	965
8	S-RTdelayed	68
9	S-RTearly	308

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	(logRR)	Risk	Ratio	RR	9	5%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 V Rodel 2015 Aschele 2011 Gerard (2) 2010 Deng 2016 O'Connell 2014 Schmoll 2021 Haddad 2017 Jiao 2015 Mohiuddin 2006 Common effect model Random effects model Heterogeneity: $J^2 = 12\%$, τ^2	/S L-CRT1 0.2813 -0.0187 0.3185 0.6837 0.0912 0.2517 0.8473 0.1823 -0.0093	0.1366 0.1657 0.1903 0.2464 0.1166 0.1579 0.5455 0.2688 0.3265	-		1.32 0.98 1.38 1.98 1.10 1.29 2.33 1.20 0.99 1.23 1.23	[1.01; [0.71; [0.95; [1.22; [0.87; [0.80; [0.71; [0.80; [1.09; [1.09;]	1.73] 1.36] 2.00] 3.21] 1.38] 1.75] 6.80] 2.03] 1.88] 1.38] 1.39]	9.4% 6.4% 4.8% 2.9% 12.9% 7.0% 0.6% 1.6% 47.9%	4.4% 4.3% 4.0% 4.5% 2.5% 3.8% 3.5% 35.6%
comparison = L-CRT2 \ Deng 2016	/S CHT 1.4309	0.3333			4.18	[2.18;	8.04]	1.6%	3.5%
comparison = CHT VS I Deng 2016	L-CRT1 -0.7472	0.3694			0.47	[0.23;	0.98]	1.3%	3.3%
$\begin{array}{l} \mbox{comparison} = \mbox{L-RTVS}\\ \mbox{Bosset} 2005\\ \mbox{Gerard} 2006\\ \mbox{Common effect model}\\ \mbox{Random effects model}\\ \mbox{Heterogeneity:} \ l^2 = 0\%, \ \tau^2 = 0.0\%, \ \tau^2 = 0.0$	L-CRT1 -0.9618 -1.1514 = 0.0018, p =	0.2262 0.3094	+ + + + +		0.38 0.32 0.36 0.36	[0.25; [0.17; [0.25; [0.25;	0.60] 0.58] 0.51] 0.51]	3.4% 1.8% 5.2%	4.1% 3.6% 7.7%
$\label{eq:comparison} \begin{array}{l} \text{comparison} = \text{L-CRT1} \ \text{W} \\ \text{Bujko 2004} \\ \text{Ngan 2012} \\ \text{Common effect model} \\ \text{Random effects model} \\ \text{Heterogeneity:} \ l^2 = 0\%, \ \tau^2 = 0 \\ \end{array}$	/S S-RTearl 3.1672 2.4913 = 0.0018, p =	y 1.0157 0.7273			23.74 12.08 15.19 15.19	[3.24; 1 [2.90; [4.77; [4.76;	73.80] 50.24] 48.39] 48.48]	0.2% 0.3% 0.5%	1.1% 1.8% 2.9%
comparison = L-CRT1 \ Latkauskas 2012	/S S-RTdela 0.9237	0.6555	-		2.52	[0.70;	9.10]	0.4%	2.0%
comparison = Induction Fernandez-Martos 2015	+ L-CRT V -0.0268	S L-CRT2 0.4771	_		0.97	[0.38;	2.48]	0.8%	2.8%
comparison = Inductior Fokas 2019	+ L-CRT V -0.3417	S L-CRT + consol 0.2222	idatio		0.71	[0.46;	1.10]	3.5%	4.1%
comparison = S-RT + c Bahadoer 2021 Jin 2022 Chakrabarti 2021 Common effect model Random effects model Heterogeneity: $I^2 = 5\%$, $\tau^2 = 5\%$	0.6815 0.3462 0.1981	VSL-CRT1 0.1450 0.2325 0.4854		*	1.98 1.41 1.22 1.76 1.75	[1.49; [0.90; [0.47; [1.39; [1.37;	2.63] 2.23] 3.16] 2.22] 2.23]	8.3% 3.2% 0.7% 12.3%	4.4% 4.0% 2.7% 11.1%
comparison = L-CRT1 V Kim 2018 Moore 2017 Common effect model Random effects model Heterogeneity: $I^2 = 54\%$, τ^2	/S L-CRT + -0.8602 0.4463	consolidation 0.6768 0.5788	- V V		0.42 1.56 0.90 0.90	[0.11; [0.50; [0.38; [0.38;	1.59] 4.86] 2.13] 2.13]	0.4% 0.5% 0.9%	2.0% 2.3% 4.3%
comparison = L-CRT1 V Marechal 2012 Conroy 2021 Common effect model Random effects model Heterogeneity: I ² = 73%, τ ²	/S Induction 0.0972 -0.8426 = 0.0018, p =	0.4417 0.2148	+ \$ \$		1.10 0.43 0.52 0.52	[0.46; [0.28; [0.35; [0.35;	2.62] 0.66] 0.75] 0.76]	0.9% 3.8% 4.7%	2.9% 4.1% 7.0%
comparison = S-RT + co Bujko (2) 2013	onsolidatior 0.3679	0.2472		-	1.44	[0.89;	2.35]	2.9%	3.9%
comparison = L-CRT2 \ Wang 2019	/S L-CRT + -0.4869	consolidation 0.4000	-+		0.61	[0.28;	1.35]	1.1%	3.2%
comparison = L-CRTI V Mei 2023 Schrag 2023 Common effects model Random effects model Heterogeneity: $J^2 = 0\%$, τ^2 :	/S CHT 0.2236 0.1060	0.2315 0.1131 0.65	ļ		1.25 1.11 1.14 1.14	[0.79; [0.89; [0.93; [0.92;	1.97] 1.39] 1.39] 1.41]	3.3% 13.7% 16.9% 	4.0% 4.5% 8.5%
Common effect model Random effects model		0.01 0	1		1.14 1.16	[1.05; [0.91;	1.23] 1.48]	100.0% 	 100.0%

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8	V 9
1	CHT			0.80 (0.60 to 1.06)	0.24 (0.12 to 0.49)				
2	0.47 (0.30 to 0.75)	Induction + L-CRT	0.71 (0.42 to 1.21)	1.85 (1.18 to 2.90)	0.97 (0.36 to 2.60)				
3	0.39 (0.23 to 0.67)	0.83 (0.54 to 1.27)	L-CRT + consolidation	1.12 (0.46 to 2.72)	1.63 (0.70 to 3.77)				
4	0.75 (0.56 to 0.99)	1.58 (1.09 to 2.28)	1.91 (1.21 to 3.00)	L-CRT1	0.80 (0.68 to 0.94)	2.81 (1.85 to 4.29)	0.59 (0.43 to 0.81)	2.52 (0.67 to 9.43)	15.23 (4.68 to 49.60)
5	0.58 (0.42 to 0.80)	1.23 (0.84 to 1.80)	1.48 (0.93 to 2.35)	0.78 (0.67 to 0.91)	L-CRT2		0.69 (0.39 to 1.22)		
6	2.10 (1.27 to 3.49)	4.44 (2.54 to 7.76)	5.36 (2.89 to 9.97)	2.81 (1.85 to 4.29)	3.62 (2.31 to 5.66)	L-RT			
7	0.43 (0.29 to 0.64)	0.91 (0.58 to 1.44)	1.10 (0.65 to 1.87)	0.58 (0.44 to 0.76)	0.74 (0.55 to 1.01)	0.21 (0.12 to 0.34)	S-RT + consolidation		
8	1.88 (0.49 to 7.26)	3.97 (1.01 to 15.64)	4.80 (1.19 to 19.39)	2.52 (0.67 to 9.43)	3.24 (0.86 to 12.23)	0.89 (0.22 to 3.58)	4.35 (1.13 to 16.74)	S-RTdelayed	
9	11.39 (3.38 to 38.34)	24.03 (6.98 to 82.74)	29.03 (8.19 to 102.87)	15.23 (4.68 to 49.60)	19.58 (5.95 to 64.40)	5.41 (1.55 to 18.95)	26.29 (7.82 to 88.38)	6.05 (1.03 to 35.54)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: other vs 'L-CRT (Random Effects Model)	'1' RR	95%-CI
L-CRT + consolidation S-RT + consolidation Induction + L-CRT L-CRT2 CHT S-RTdelayed L-RT S-RTearly	*	1.91 [1.73 [1.58 [1.29 [0.75 [0.40 [0.36 [0.07 [1.21; 3.00] 1.31; 2.28] 1.09; 2.28] 1.10; 1.50] 0.56; 0.99] 0.11; 1.49] 0.23; 0.54] 0.02; 0.21]
C	0.01 0.5 1 2	4	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau² \geq 0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $## tau^2 = 0.0237$; tau = 0.154

I^2= 30.79 % (0 % to 59.32 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 28.90 20
 0.0898

 Within designs
 13.79 14
 0.4653

 Between designs
 15.10
 6
 0.0195

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:CHT 0.21 1 0.6480 L-CRT1:Induction + L-CRT 3.66 1 0.0557 L-CRT1:L-CRT + consolidation 2.15 1 0.1424 L-CRT1:L-CRT2 5.13 7 0.6440 L-CRT1:L-RT 0.24 1 0.6209 L-CRT1:S-RT + consolidation 2.10 2 0.3496 L-CRT1:S-RTearly 0.29 1 0.5884

Between-designs Q statistic after detaching of single designs

 $\begin{array}{c|c} Detached \ design & Q \ df \ p-value \\ Induction + L-CRT:L-CRT + consolidation 13.62 & 5 & 0.0182 \\ Induction + L-CRT:L-CRT2 14.65 & 5 & 0.0120 \\ L-CRT + consolidation:L-CRT2 15.10 & 5 & 0.0099 \\ L-CRT1:CHT & 7.39 & 5 & 0.1934 \\ L-CRT1:Induction + L-CRT 12.83 & 5 & 0.0251 \\ L-CRT1:L-CRT + consolidation 12.77 & 5 & 0.0256 \\ L-CRT1:L-CRT2 & 9.66 & 5 & 0.0855 \\ L-CRT1:S-RT + consolidation 15.09 & 5 & 0.0100 \\ L-CRT1:CHT:L-CRT2 & 3.48 & 4 & 0.4810 \\ \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 15.10 6 0.0195 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method Random effects model:

comparison k prop nma direct indir. RoR z p-value L-CRT1:CHT 3 0.97 1.34 1.25 8.95 0.14 -2.46 0.0139 CHT:L-CRT2 1 0.19 0.58 0.24 0.72 0.33 -2.69 0.0071 Induction + L-CRT:L-CRT + consolidation 1 0.66 0.83 0.71 1.11 0.64 -0.97 0.3331 L-CRT1:Induction + L-CRT 2 0.66 0.63 0.54 0.87 0.62 -1.20 0.2317 Induction + L-CRT:L-CRT2 1 0.15 1.23 0.97 1.28 0.76 -0.50 0.6155 L-CRT1:L-CRT + consolidation 2 0.26 0.52 0.89 0.43 2.06 1.37 0.1708

L-CRT + consolidation:L-CRT2 1 0.30 1.48 1.63 1.42 1.14 0.26 0.7954 L-CRT1:L-CRT2 9 0.88 0.78 0.80 0.65 1.23 0.85 0.3973 L-CRT1:S-RT + consolidation 3 0.78 0.58 0.59 0.54 1.11 0.30 0.7677 L-CRT2:S-RT + consolidation 1 0.28 0.74 0.69 0.77 0.90 -0.30 0.7677 Legend: comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion - Estimated treatment effect (RR) in network meta-analysis nma direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence - Ratio of Ratios (direct versus indirect) RoR - z-value of test for disagreement (direct versus indirect) z p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

 P-score

 L-CRT + consolidation
 0.9231

 S-RT + consolidation
 0.8703

 Induction + L-CRT
 0.7956

 L-CRT2
 0.6474

 L-CRT1
 0.4880

 CHT
 0.3553

 S-RTdelayed
 0.2374

 L-RT
 0.1794

eAppendix 7. Tolerability of treatment

Characteristics of the network

Number of treatments:

8

Number of studies:

25

Number of individuals included:

11987

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized	
1	CHT	1050	
2	Induction + L-CRT	469	
3	L-CRT + consolidation	288	
4	L-CRT1	5373	
5	L-CRT2	2529	
6	L-RT	872	
7	S-RT + consolidation	1090	
8	S-RTearly	316	

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR	SE(logRR)	Risk Ratio	RR	95%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 V Rodel 2015 Aschele 2011 Gerard (2) 2010 Deng 2016 Schmoll 2021 Haddad 2017 Jiao 2015 Mohiuddin 2006 Common effect model Random effects model Heterogeneity: $I^2 = 93\%$, τ^2	/S L-CR 0.0579 -0.2124 -0.0621 0.0650 -0.3001 -0.3831 -0.1054 0.0063	0.0272 0.0363 0.0205 0.0483 0.0483 0.0343 0.1617 0.0636 0.0647	* * *	1.06 0.81 0.94 1.07 0.74 0.68 0.90 1.01 0.92 0.91	[1.00; 1.12] [0.75; 0.87] [0.90; 0.98] [0.97; 1.17] [0.69; 0.79] [0.50; 0.94] [0.79; 1.02] [0.89; 1.14] [0.90; 0.95] [0.84; 0.99]	5.1% 2.9% 9.0% 1.6% 3.2% 0.1% 0.9% 0.9% 23.6%	4.2% 4.0% 4.3% 3.8% 4.1% 1.6% 3.4% 3.4% 28.8%
comparison = L-CRT2 Deng 2016	/S CHT -0.0741	0.0369		0.93	[0.86; 1.00]	2.7%	4.0%
comparison = CHTVS Deng 2016	L-CRT1 0.1391	0.0428		1.15	[1.06; 1.25]	2.0%	3.9%
comparison = L-RT VS Bosset 2005 Gerard 2006 Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, τ^2 :	L-CRT1 0.2087 0.2107 = 0.0110,	0.0228 0.0291 p = 0.96	* + *	1.23 1.23 1.23 1.23	[1.18; 1.29] [1.17; 1.31] [1.19; 1.28] [1.06; 1.43]	7.2% 4.4% 11.6%	4.3% 4.2% 8.4%
Bujko 2004 Ngan 2012	-0.0463	0.0237	*	0.95	[0.91; 1.00]	6.7% 0.0%	4.2% 0.0%
comparison = Induction Fernandez-Martos 2015	n + L-CR 0.0679	0.0919	2	1.07	[0.89; 1.28]	0.4%	2.8%
comparison = Induction Fokas 2019	n + L-CR 0.0597	CT VS L-CRT 0.0466	+ consolidation	1.06	[0.97; 1.16]	1.7%	3.8%
$\begin{array}{l} \mbox{comparison = S-RT + c} \\ \mbox{Bahadoer 2021} \\ \mbox{Jin 2022} \\ \mbox{Chakrabarti 2021} \\ \mbox{Common effect model} \\ \mbox{Random effects model} \\ \mbox{Heterogeneity: } \end{tabular}^2 = 88\%, \end{tabular}^2 \end{array}$	-0.0567 -0.2192 -0.0022	ation VS L-CI 0.0260 0.0371 0.0469	₹T1 	0.94 0.80 1.00 0.91 0.91	[0.90; 0.99] [0.75; 0.86] [0.91; 1.09] [0.88; 0.95] [0.80; 1.03]	5.5% 2.7% 1.7% 10.0%	4.2% 4.0% 3.8% 12.1%
comparison = L-CRT1 V Kim 2018 Moore 2017 Common effect model Random effects model Heterogeneity: I^2 = 0%, τ^2 :	/S L-CR 0.0430 0.0400 = 0.0110,	r + consolid 0.0549 0.0400 p = 0.96	ation	1.04 1.04 1.04 1.04	[0.94; 1.16] [0.96; 1.13] [0.98; 1.11] [0.89; 1.22]	1.2% 2.3% 3.6%	3.6% 4.0% 7.6%
$\label{eq:comparison} \begin{array}{l} \text{comparison} = \text{L-CRT1} \\ \text{Marechal 2012} \\ \text{Conroy 2021} \\ \text{Common effect model} \\ \text{Random effects model} \\ \text{Heterogeneity:} \ l^2 = 50\%, \ \tau^2 \end{array}$	/S Indue 0.0357 0.0966 = 0.0110	ction + L-CR 0.0357 0.0237	T ++ ◆	1.04 1.10 1.08 1.07	[0.97; 1.11] [1.05; 1.15] [1.04; 1.12] [0.92; 1.24]	2.9% 6.7% 9.6%	4.1% 4.2% 8.3%
comparison = S-RT + c Bujko (2) 2013	onsolida -0.0513	ation VS L-CI 0.0654	RT2	0.95	[0.84; 1.08]	0.9%	3.4%
comparison = L-CRT2 Wang 2019	/S L-CR 0.0000	T + consolid 0.0339	ation	1.00	[0.94; 1.07]	3.3%	4.1%
$\label{eq:comparison} \begin{array}{l} \mbox{comparison} = \mbox{L-CRT1} \\ \mbox{Mei 2023} \\ \mbox{Schrag 2023} \\ \mbox{Common effect model} \\ \mbox{Random effects model} \\ \mbox{Heterogeneity:} \ \mbox{I}^2 = \mbox{62\%, } \ \mbox{C}^2 \\ \mbox{Heterogeneity:} \ \mbox{I}^2 = \mbox{62\%, } \ \mbox{C}^2 \\ \mbox{Heterogeneity:} \ \mbox{I}^2 = \mbox{62\%, } \ \mbox{C}^2 \\ \mbox{Heterogeneity:} \ \mbox{I}^2 = \mbox{62\%, } \ \mbox{C}^2 \\ \mbox{Heterogeneity:} \ \mbox{I}^2 = \mbox{62\%, } \ \mbox{C}^2 \\ \mbox{Heterogeneity:} \ \mbox{I}^2 = \mbox{62\%, } \ \mbox{C}^2 \\ \mbox{Heterogeneity:} \ \mbox{I}^2 = \mbox{62\%, } \mbox{62\%, } \mbox{I}^2 = \mbox{62\%, } $	/S CHT 0.0527 -0.0003	0.0295 0.0139 0, <i>p</i> = 0.10	+	1.05 1.00 1.01 1.03	[0.99; 1.12] [0.97; 1.03] [0.98; 1.03] [0.88; 1.19]	4.3% 19.5% 23.8% 	4.2% 4.3% 8.5%
Common effect model Random effects model				1.01 0.99	[0.99; 1.02] [0.94; 1.04]	100.0% 	 100.0%
		0.	.5 1	2			

 $\begin{array}{ccc} 0.5 & 1 \\ \text{Heterogeneity: } l^2 = 93\%, \ \tau^2 = 0.0145, \ p < 0.01 \\ \text{Residual heterogeneity: } l^2 = 89\%, \ \tau^2 = 0.0110, \ p < 0.01 \\ \text{Test for subgroup differences (common effect): } \ \chi^2_{\frac{1}{42}} = 235.61, \ \text{df} = 12 \ (p < 0.01) \\ \text{Test for subgroup differences (random effect): } \ \chi^2_{\frac{1}{12}} = 18.71, \ \text{df} = 12 \ (p = 0.10) \end{array}$

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8
1	CHT			1.02 (0.92 to 1.14)	1.08 (0.90 to 1.29)			
2	1.05 (0.92 to 1.21)	Induction + L-CRT	1.06 (0.88 to 1.28)	0.93 (0.82 to 1.06)	1.07 (0.84 to 1.37)			
3	1.08 (0.94 to 1.24)	1.02 (0.91 to 1.16)	L-CRT + consolidation	0.96 (0.84 to 1.10)	1.00 (0.83 to 1.20)			
4	1.01 (0.91 to 1.11)	0.95 (0.86 to 1.06)	0.93 (0.84 to 1.03)	L-CRT1	1.10 (1.02 to 1.18)	0.81 (0.72 to 0.92)	1.10 (0.99 to 1.22)	0.95 (0.80 to 1.14)
5	1.10 (0.99 to 1.23)	1.05 (0.94 to 1.17)	1.02 (0.92 to 1.14)	1.10 (1.03 to 1.17)	L-CRT2		1.05 (0.85 to 1.30)	
6	0.82 (0.70 to 0.95)	0.77 (0.66 to 0.91)	0.76 (0.64 to 0.89)	0.81 (0.72 to 0.92)	0.74 (0.64 to 0.85)	L-RT		
7	1.12 (0.97 to 1.28)	1.06 (0.92 to 1.22)	1.03 (0.90 to 1.18)	1.11 (1.01 to 1.22)	1.01 (0.91 to 1.13)	1.37 (1.17 to 1.60)	S-RT + consolidation	
8	0.96 (0.79 to 1.17)	0.91 (0.75 to 1.11)	0.89 (0.73 to 1.09)	0.95 (0.80 to 1.14)	0.87 (0.72 to 1.05)	1.18 (0.95 to 1.46)	0.86 (0.71 to 1.05)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).



Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau² \geq 0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

tau^2= 0.0073 ; tau= 0.0856

I^2= 85.47 % (78.66 % to 90.11 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 123.92 18 < 0.0001</td>

 Within designs
 107.82 12 < 0.0001</td>

 Between designs
 16.11 6
 0.0132

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:CHT 2.64 1 0.1040 L-CRT1:Induction + L-CRT 2.02 1 0.1555 L-CRT1:L-CRT + consolidation 0.00 1 0.9648 L-CRT1:L-CRT2 85.91 6 < 0.0001 L-CRT1:L-RT 0.00 1 0.9569 L-CRT1:S-RT + consolidation 17.25 2 0.0002

Between-designs Q statistic after detaching of single designs

 $\begin{array}{ccccccc} Detached \ design & Q \ df \ p-value \\ Induction + L-CRT:L-CRT + consolidation 14.30 \ 5 \ 0.0138 \\ Induction + L-CRT:L-CRT2 \ 15.77 \ 5 \ 0.0075 \\ L-CRT + consolidation:L-CRT2 \ 16.00 \ 5 \ 0.0069 \\ L-CRT1:CHT \ 12.84 \ 5 \ 0.0249 \\ L-CRT1:Induction + L-CRT \ 13.83 \ 5 \ 0.0167 \\ L-CRT1:L-CRT + consolidation \ 14.23 \ 5 \ 0.0142 \\ L-CRT1:L-CRT + consolidation \ 14.23 \ 5 \ 0.0177 \\ L-CRT1:S-RT + consolidation \ 15.72 \ 5 \ 0.0077 \\ L-CRT1:CHT \ 12.72 \ 5 \ 0.0077 \\ L-CRT1:CHT \ 12.72 \ 3.54 \ 4 \ 0.4717 \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 3.73 6 0.7133 0.0943 0.0089

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value L-CRT1:CHT 3 0.90 0.99 0.98 1.19 0.82 -1.15 0.2506 CHT:L-CRT2 1 0.35 1.10 1.08 1.12 0.96 -0.32 0.7468 Induction + L-CRT:L-CRT + consolidation 1 0.40 1.02 1.06 1.00 1.06 0.46 0.6419 L-CRT1:Induction + L-CRT 2 0.64 1.05 1.07 1.01 1.06 0.55 0.5857 Induction + L-CRT:L-CRT2 1 0.20 1.05 1.07 1.04 1.03 0.19 0.8461 L-CRT1:L-CRT + consolidation 2 0.54 1.07 1.04 1.11 0.94 -0.63 0.5260

L-CRT + consolidation:L-CRT2 1 0.34 1.02 1.00 1.03 0.97 -0.29 0.7735 L-CRT1:L-CRT2 8 0.78 1.10 1.10 1.09 1.00 0.04 0.9708 L-CRT1:S-RT + consolidation 3 0.81 1.11 1.10 1.16 0.95 -0.43 0.6689 L-CRT2:S-RT + consolidation 1 0.25 1.01 1.05 1.00 1.05 0.43 0.6689 Legend: comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence

indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect)

z - z-value of test for disagreement (direct versus indirect)

p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
L-RT	0.9894
S-RTearly	0.7106
L-CRT1	0.6395
CHT	0.6333
Induction + L-CRT	0.4059
L-CRT + consolidation	0.2893
L-CRT2	0.1778
S-RT + consolidation	0.1542

eAppendix 8. Toxicity of treatment

Characteristics of the network

Number of treatments:

8

Number of studies:

22

Number of individuals included:

11568

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized	
1	CHT	885	
2	Induction + L-CRT	469	
3	L-CRT + consolidation	263	
4	L-CRT1	5243	
5	L-CRT2	3004	
6	L-RT	367	
7	S-RT + consolidation	1021	
8	S-RTearly	316	

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study			Dick	Patio	DD	95% CI	Weight	Weight
olddy	logitit t		Nak	i auto		3070-01	(common)	(random)
comparison = L-CRT2 \	/SL-CRT	1		E E			10.00	5.000
Rodel 2015	0.1340	0.1073			1.14	[0.93; 1.41]	10.2%	5.3%
Aschele 2011	1.1100	0.2021		· · · ·	3.05	[2.05; 4.54]	2.9%	5.0%
O'Connoll 2014	0.0402	0.1947		i i	2.33	[1.09, 3.41]	3.1% 10.70/	0.1% 5.20/
Schmoll 2021	0.0490	0.1187		\$ 	2 4 9	[1.21, 1.00]	8.4%	5.2%
Haddad 2017	2 0477	1 0303			7 75	[1.03:58.38]	0.1%	2.0%
Jiao 2015	0.6931	0.3420		<u><u></u></u>	2.00	[1.02: 3.91]	1.0%	4.5%
Mohiuddin 2006	0.2041	0.2227	-	<u> </u>	1.23	[0.79; 1.90]	2.4%	5.0%
Common effect model				Ê\$	1.63	[1.48; 1.80]	46.8%	
Random effects model					1.84	[1.42; 2.39]		37.4%
Heterogeneity: $I^2 = 84\%$, τ^2	= 0.0926,	p < 0.01						
comparison = L-RTVS	L-CRT1							
Gerard 2006	-1.6832	0.3358			0.19	[0.10; 0.36]	1.0%	4.6%
comparison = L-CRT1 \	/S S-RTe	arly						
Bujko 2004	1.7099	0.4721		§ 	5.53	[2.19; 13.95]	0.5%	4.0%
Ngan 2012	2.6794	0.5862		÷	14.58	[4.62; 45.99]	0.3%	3.5%
Common effect model				\sim	8.10	[3.94; 16.64]	0.9%	
Random effects model				····	8.30	[3.59; 19.21]		7.5%
Heterogeneity: $I^2 = 40\%$, τ^2	= 0.0926,	p = 0.20		с с с				
comparison = Induction	+L-CRT	VSL-CRT2						
Fernandez-Martos 2015	-0.3203	0.3333	-+		0.73	[0.38; 1.40]	1.1%	4.6%
comparison = Induction	+L-CRT	VSL-CRT+co	nsolidatio	c c c c c c c c c c c c c c c c c c c				
Fokas 2019	0 2726	0 1708	Teenaaro		1 31	[0 94 [.] 1 84]	4 0%	5 1%
1 01100 2010	0.2720	0.1100		e ç		[0.01, 1.01]		0.170
comparison = S-RT + co	onsolidati	ion VS L-CRT1		6				
Bahadoer 2021	0.6714	0.0967		4 ·	1.96	[1.62; 2.37]	12.6%	5.3%
Jin 2022	0.7416	0.1814			2.10	[1.47; 3.00]	3.6%	5.1%
Common effect model				\diamond	1.99	[1.68; 2.35]	16.2%	
Random effects model		0.70			2.02	[1.27; 3.21]		10.4%
Heterogeneity: $I^- = 0\%$, $\tau^- =$	= 0.0926, p	= 0.73		i u				
comparison = L-CRT1 \	/SL-CRT	+ consolidation	n	6				
Kim 2018	-0 9533	0 8142	·		0.39	[0.08 [.] 1.90]	0.2%	2.6%
				6				
comparison = L-CRT1 \	/S Induct	ion + L-CRT						
Marechal 2012	-1.1337	0.7723	+		0.32	[0.07; 1.46]	0.2%	2.8%
Conroy 2021	-0.2647	0.1145	+	6	0.77	[0.61; 0.96]	9.0%	5.3%
Common effect model					0.75	[0.60; 0.94]	9.2%	
Kandom effects model	- 0.0000	n = 0.07	\sim	Ta	0.68	[0.38; 1.24]		8.0%
neterogeneity. 7 – 19%, 1	- 0.0920,	p = 0.27						
comparison = S-RT + co	onsolidati	ion VS L-CRT2		u 1				
Bujko (2) 2013	0.0969	0.1666	-	12 A	1.10	[0.79; 1.53]	4.2%	5.1%
, , ,						. , ,		
comparison = L-CRT2 \	/SL-CRT	+ consolidation	n	c c				
Wang 2019	-0.3102	0.3525	-+		0.73	[0.37; 1.46]	0.9%	4.5%
comparison = L CDT()								
Comparison = L-CRTTN Moi 2023	0 3055	0.2488	-+		0.67	IO 41: 1 101	1 0%	1 0%
Schrag 2023	-0.5859	0.0932		u u	0.56	[0.46: 0.67]	13.6%	5.3%
Common effect model	5.0000	0.0002	\$	i i	0.57	[0.48: 0.68]	15.5%	0.070
Random effects model			\diamond		0.60	[0.37; 0.97]		10.2%
Heterogeneity: $I^2 = 0\%$, $\tau^2 =$	= 0.0926, <i>p</i>	= 0.47		6		- / .		
o					4.00	14 40. 4 0	400.007	
Common effect model				\$:	1.26	[1.18; 1.35]	100.0%	400.00/
Random effects model				ř.	1.33	[0.92; 1.91]		100.0%
			01 05	1 2 10				
Heterogeneity: $I^2 = 92\% \tau^2$	= 0.6426	p < 0.01	0.1 0.0	12 10				
Residual heterogeneity: I^2 =	= 77%, τ ² =	0.0926, p < 0.01						

Test for subgroup differences (common effect): $\chi_{10}^2 = 224.03$, df = 10 (p < 0.01) Test for subgroup differences (random effect): $\chi_{10}^2 = 67.19$, df = 10 (p < 0.01)

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8
1	CHT			1.67 (1.06 to 2.64)				
2	1.02 (0.54 to 1.93)	Induction + L-CRT	1.31 (0.69 to 2.51)	1.45 (0.83 to 2.54)	0.73 (0.31 to 1.71)		100 C	
3	1.01 (0.49 to 2.11)	0.99 (0.58 to 1.68)	L-CRT + consolidation	2.59 (0.48 to 14.06)	1.36 (0.56 to 3.31)			
4	1.67 (1.06 to 2.64)	1.63 (1.06 to 2.52)	1.65 (0.93 to 2.93)	L-CRT1	0.55 (0.43 to 0.70)	5.38 (2.27 to 12.74)	0.50 (0.32 to 0.77)	8.28 (3.63 to 18.89)
5	0.92 (0.55 to 1.54)	0.90 (0.58 to 1.41)	0.91 (0.51 to 1.61)	0.55 (0.44 to 0.69)	L-CRT2		0.91 (0.48 to 1.73)	
6	8.99 (3.39 to 23.86)	8.78 (3.34 to 23.07)	8.88 (3.15 to 25.01)	5.38 (2.27 to 12.74)	9.75 (4.00 to 23.76)	L-RT		
7	0.83 (0.46 to 1.50)	0.81 (0.46 to 1.42)	0.82 (0.42 to 1.60)	0.50 (0.34 to 0.72)	0.90 (0.61 to 1.34)	0.09 (0.04 to 0.24)	S-RT + consolidation	
8	13.83 (5.38 to 35.52)	13.52 (5.32 to 34.34)	13.66 (5.00 to 37.31)	8.28 (3.63 to 18.89)	15.00 (6.38 to 35.27)	1.54 (0.47 to 5.07)	16.66 (6.75 to 41.12)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: other vs 'L-CRT1 (Random Effects Model)	RR	95%-CI
S-RT + consolidation L-CRT2 CHT L-CRT + consolidation Induction + L-CRT L-RT S-RTearly	*	2.01 1.81 1.67 1.65 1.63 0.19 0.12	[1.39; 2.91] [1.44; 2.27] [1.06; 2.64] [0.93; 2.93] [1.06; 2.52] [0.08; 0.44] [0.05; 0.28]
0	.01 0.5 1 2 4	4	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0.0805$; tau = 0.2837

I^2= 71.54 % (52.87 % to 82.82 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 52.71 15 < 0.0001</td>

 Within designs
 47.61 11 < 0.0001</td>

 Between designs
 5.11 4
 0.2765

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:CHT 0.51 1 0.4737 L-CRT1:Induction + L-CRT 1.24 1 0.2657 L-CRT1:L-CRT2 44.08 7 < 0.0001 L-CRT1:S-RT + consolidation 0.12 1 0.7328 L-CRT1:S-RTearly 1.66 1 0.1978

Between-designs Q statistic after detaching of single designs

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 2.00 4 0.7349 0.3105 0.0964

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value Induction + L-CRT:L-CRT + consolidation 1 0.67 0.99 1.31 0.56 2.35 1.48 0.1375 L-CRT1:Induction + L-CRT 2 0.60 0.61 0.69 0.51 1.35 0.66 0.5094 Induction + L-CRT:L-CRT2 1 0.27 0.90 0.73 0.98 0.74 -0.58 0.5623 L-CRT1:L-CRT + consolidation 1 0.12 0.61 0.39 0.64 0.60 -0.56 0.5764 L-CRT + consolidation:L-CRT2 1 0.41 0.91 1.36 0.69 1.99 1.17 0.2440 L-CRT1:L-CRT 2 8 0.84 0.55 0.55 0.59 0.93 -0.25 0.8064 L-CRT1:S-RT + consolidation 2 0.71 0.50 0.50 0.50 0.99 -0.03 0.9757 L-CRT2:S-RT + consolidation 1 0.37 0.90 0.91 0.90 1.01 0.03 0.9757

k

Legend: comparison - Treatment comparison - Number of studies providing direct evidence prop - Direct evidence proportion

nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
S-RT + consolidation	0.8454
L-CRT2	0.7465
CHT	0.6675
L-CRT + consolidation	0.6590
Induction + L-CRT	0.6428
L-CRT1	0.2960
L-RT	0.1087
S-RTearly	0.0342

eAppendix 9. Dropouts by any cause

Characteristics of the network

Number of treatments:

9

Number of studies:

27

Number of individuals included:

13383

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized
1	CHT	1050
2	Induction + L-CRT	469
3	L-CRT + consolidation	288
4	L-CRT1	6075
5	L-CRT2	3147
6	L-RT	872
7	S-RT + consolidation	1090
8	S-RTdelayed	75
9	S-RTearly	317

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR	SE(logRR)	Risk Ratio	RR	95%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 Rodel 2015	/S L-CR 1.1084	.T1 0.5132		3.03	[1.11; 8.28]	1.7%	2.0%
Aschele 2011	-0.1958	0.3375		0.82	[0.42; 1.59]	3.8%	4.2%
Gerard (2) 2010	-0.3116	0.4572		0.73	[0.30; 1.79]	2.1%	2.4%
Deng 2016 O'Connoll 2014	-0.2719	0.3130		0.76	[0.41; 1.41]	4.4%	4.8%
Schmoll 2021	0.3338	0.2886	1-	1 40	[0.32, 1.01]	5.2%	5.4%
Haddad 2017	0.9491	0.6284		2.58	[0.75; 8.85]	1.1%	1.3%
Jiao 2015	0.0000	1.4073		1.00	[0.06; 15.77]	0.2%	0.3%
Mohiuddin 2006	-0.3460	0.7379		0.71	[0.17; 3.00]	0.8%	1.0%
Common effect model Random effects model	- 0.0154	n = 0.26	-	1.05 1.06	[0.81; 1.37] [0.80; 1.40]	24.6% 	26.8%
comparison = L CDT2		r, p = 0.20					
Deng 2016	0.2076	0.3568		1.23	[0.61; 2.48]	3.4%	3.8%
comparison = CHT VS Deng 2016	L-CRT1 -0.4796	0.3353		0.62	[0.32; 1.19]	3.9%	4.2%
comparison = L-RTVS	L-CRT1	0.0505	_				0.00/
Bosset 2005	-0.3234	0.3585		0.72	[0.36; 1.46]	3.4%	3.8%
Common effect model	-0.8051	0.4472		0.45	[0.19, 1.07]	Z.2%	2.5%
Random effects model Heterogeneity: $I^2 = 0\%$, $\tau^2 =$	= 0.0154,	p = 0.40	Ä	0.60	[0.34; 1.06]		6.3%
comparison = L-CRT1	/S S-RT	early					
Bujko 2004	0.9857	0.4274		2.68	[1.16; 6.19]	2.4%	2.8%
Ngan 2012	0.0062	0.6983		1.01	[0.26; 3.95]	0.9%	1.1%
Common effects model				2.05	[1.00; 4.19]	3.3%	3 0%
Heterogeneity: $I^2 = 30\%$, τ^2	= 0.0154	l, p = 0.23		2.05	[0.97, 4.25]		3.970
comparison = L-CRT1 V Latkauskas 2012	/S S-RT -0.8473	delayed 0.6705		0.43	[0. 12 ; 1.59]	1.0%	1.2%
comparison = Induction	1 + L-CR	T VS L-CRT2		0.20	10.02: 2.041	0.2%	0.49/
Fernandez-Martos 2015	-1.1958	1.1377 —		0.30	[0.03; 2.81]	0.3%	0.4%
comparison = Induction Fokas 2019	1 + L-CR 0.5204	0.4282	consolidation *	1.68	[0.73; 3.89]	2.4%	2.7%
comparison = S-RT + c	onsolida	ation VS L-CR	F1				
Bahadoer 2021	-0.3548	0.2083		0.70	[0.47; 1.05]	10.0%	9.0%
JIN 2022 Chakrabarti 2021	-0.0169	0.1580		0.98	[0.72; 1.34]	17.5%	12.8%
Common effect model	0.2799	0.4740	4	0.89	[0.32, 3.33]	29.4%	2.3%
Random effects model			4	0.89	[0.67; 1.19]		24.1%
Heterogeneity: $I^2 = 17\%$, τ^2	= 0.0154	l, p = 0.30			- / -		
comparison = L-CRT1 \	/SL-CR	T + consolidat	ion				
Kim 2018	-1.1357	0.6383		0.32	[0.09; 1.12]	1.1%	1.3%
Moore 2017	0.0408	1.3850		1.04	[0.07; 15.73]	0.2%	0.3%
Common effect model				0.39	[0.13; 1.23]	1.3%	4 69/
Heterogeneity: $I^2 = 0\%$, τ^2 :	= 0.0154,	p = 0.44		0.40	[0.13; 1.26]		1.0%
comparison = L-CRT1 \	/S Indu	tion + L-CRT					
Marechal 2012	-0.0351	1.3892		0.97	[0.06; 14.70]	0.2%	0.3%
Conroy 2021	-0.4011	0.3608		0.67	[0.33; 1.36]	3.3%	3.7%
Common effect model				0.69	[0.35; 1.36]	3.6%	
Heterogeneity: $I^2 = 0\%$, τ^2 :	= 0.0154,	p = 0.80		0.69	[0.33; 1.41]		4.0%
comparison = S-RT + c	onsolida	tion VS L-CR	Г2				
Bujko (2) 2013	-0.1384	0.3220		0.87	[0.46; 1.64]	4.2%	4.5%
comparison = L-CRT2 V Wang 2019	/S L-CR 0.8473	T + consolidat 0.6655	ion	2.33	[0.63; 8.60]	1.0%	1.2%
comparison = L CDT4							
Mei 2023	0.0374	0.2542		1.04	[0.63: 1.71]	6.7%	6.7%
Schrag 2023	-0.3410	0.2162		0.71	[0.47, 1.09]	9.3%	8.5%
Common effect model			4	0.83	[0.60; 1.15]	16.1%	
Random effects model Heterogeneity: $I^2 = 22\%$, τ^2	= 0.0154	l, p = 0.26	1	0.84	[0.58; 1.21]		15.2%
Common offerst made			l l	0.04	TO 00- 4 0 47	100.007	
Random effects model			, <u> </u>	0.91	[0.80; 1.04]		 100.0%
			01 051 2 10				

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8	V 9
1	CHT			1.02 (0.72 to 1.46)	0.81 (0.37 to 1.77)		100 C		1.00
2	0.74 (0.36 to 1.50)	Induction + L-CRT	1.68 (0.68 to 4.17)	1.45 (0.68 to 3.09)	0.30 (0.03 to 2.89)				
3	1.02 (0.46 to 2.23)	1.38 (0.68 to 2.81)	L-CRT + consolidation	2.51 (0.78 to 8.10)	0.43 (0.11 to 1.65)				
4	1.05 (0.74 to 1.48)	1.42 (0.76 to 2.66)	1.03 (0.51 to 2.09)	L-CRT1	0.94 (0.70 to 1.26)	1.68 (0.92 to 3.07)	1.12 (0.81 to 1.56)	2.01 (0.94 to 4.31)	0.43 (0.11 to 1.67)
5	0.91 (0.60 to 1.39)	1.24 (0.64 to 2.41)	0.90 (0.44 to 1.86)	0.87 (0.67 to 1.14)	L-CRT2		1.15 (0.56 to 2.35)		
6	1.76 (0.88 to 3.53)	2.39 (1.00 to 5.70)	1.73 (0.68 to 4.38)	1.68 (0.92 to 3.07)	1.93 (1.00 to 3.72)	L-RT			1.00
7	1.15 (0.73 to 1.82)	1.56 (0.78 to 3.12)	1.13 (0.53 to 2.43)	1.10 (0.81 to 1.49)	1.26 (0.87 to 1.82)	0.65 (0.33 to 1.28)	S-RT + consolidation		
8	2.11 (0.91 to 4.87)	2.86 (1.07 to 7.68)	2.07 (0.73 to 5.86)	2.01 (0.94 to 4.31)	2.31 (1.03 to 5.17)	1.20 (0.45 to 3.16)	1.83 (0.81 to 4.16)	S-RTearly	
9	0.45 (0.11 to 1.82)	0.61 (0.14 to 2.72)	0.44 (0.10 to 2.04)	0.43 (0.11 to 1.67)	0.49 (0.12 to 1.96)	0.25 (0.06 to 1.12)	0.39 (0.10 to 1.57)	0.21 (0.04 to 1.01)	S-RTdelayed

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: other vs 'L-CRT1 (Random Effects Model)	RR	95%-CI
S-RTdelayed Induction + L-CRT L-CRT2 CHT L-CRT + consolidation S-RT + consolidation L-RT S-RTearly		2.33 1.42 1.14 1.05 1.03 0.91 0.59 0.50	[0.60; 9.07] [0.76; 2.66] [0.88; 1.50] [0.74; 1.48] [0.51; 2.09] [0.67; 1.23] [0.33; 1.08] [0.23; 1.06]
0	.01 0.5 1 2 4	4	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau² \geq 0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

tau^2= 0.0306 ; tau= 0.1748

I^2= 17.86 % (0 % to 51.45 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 24.35 20
 0.2274

 Within designs
 15.20 14
 0.3648

 Between designs
 9.15 6
 0.1652

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:CHT 1.29 1 0.2569 L-CRT1:Induction + L-CRT 0.07 1 0.7987 L-CRT1:L-CRT + consolidation 0.60 1 0.4404 L-CRT1:L-CRT2 8.71 7 0.2740 L-CRT1:L-RT 0.71 1 0.4007 L-CRT1:S-RT + consolidation 2.40 2 0.3009 L-CRT1:S-RTearly 1.43 1 0.2315

Between-designs Q statistic after detaching of single designs

 $\begin{array}{c|c} & Detached \ design & Q \ df \ p-value \\ Induction + L-CRT:L-CRT + \ consolidation \ 8.67 \ 5 \ 0.1228 \\ Induction + L-CRT:L-CRT2 \ 7.42 \ 5 \ 0.1912 \\ L-CRT + \ consolidation:L-CRT2 \ 7.38 \ 5 \ 0.1942 \\ L-CRT1:CHT \ 7.09 \ 5 \ 0.2143 \\ L-CRT1:Induction + L-CRT \ 9.15 \ 5 \ 0.1034 \\ L-CRT1:L-CRT + \ consolidation \ 5.41 \ 5 \ 0.3680 \\ L-CRT1:L-CRT2 \ 9.15 \ 5 \ 0.1033 \\ L-CRT1:S-RT + \ consolidation \ 9.07 \ 5 \ 0.1063 \\ L-CRT1:CHT2:S-RT + \ consolidation \ 9.07 \ 5 \ 0.1063 \\ L-CRT1:CHT:L-CRT2 \ 5.49 \ 4 \ 0.2406 \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 8.56 6 0.1999 0.1085 0.0118

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value L-CRT1:CHT 3 0.96 0.95 0.98 0.52 1.87 0.68 0.4946 CHT:L-CRT2 1 0.29 0.91 0.81 0.96 0.85 -0.35 0.7229 Induction + L-CRT:L-CRT + consolidation 1 0.61 1.38 1.68 1.01 1.67 0.69 0.4920 L-CRT1:Induction + L-CRT 2 0.69 0.70 0.69 0.74 0.93 -0.10 0.9220 Induction + L-CRT:L-CRT2 1 0.09 1.24 0.30 1.42 0.21 -1.28 0.1993

L-CRT1:L-CRT + consolidation 2 0.36 0.97 0.40 1.62 0.25 -1.87 0.0619 L-CRT + consolidation:L-CRT2 1 0.29 0.90 0.43 1.22 0.35 -1.28 0.2009 L-CRT1:L-CRT2 9 0.82 0.87 0.94 0.62 1.52 1.17 0.2438 L-CRT1:S-RT + consolidation 3 0.85 1.10 1.12 0.99 1.13 0.29 0.7720 L-CRT2:S-RT + consolidation 1 0.27 1.26 1.15 1.30 0.88 -0.29 0.7720

Legend:

D

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
S-RTdelayed	0.8802
Induction + L-CRT	0.7914
L-CRT2	0.6712
CHT	0.5525
L-CRT + consolidation	0.5192
L-CRT1	0.4899
S-RT + consolidation	0.3812
L-RT	0.1322
S-RTearly	0.0822

eAppendix 10. Pre-operative treatment related deaths

Characteristics of the network

Number of treatments:

8

Number of studies:

24

Number of individuals included:

11963

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized	
1	CHT	1050	
2	Induction + L-CRT	469	
3	L-CRT + consolidation	288	
4	L-CRT1	5359	
5	L-CRT2	2519	
6	L-RT	872	
7	S-RT + consolidation	1090	
8	S-RTearly	316	

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	E(logRR)	Risk Ratio	RR	95%-CI	Weight (common)	Weight (random)
comparison = L-CRT2	/S L-CRT1	0.9640		2.02	10 27: 11 061	6.0%	6.0%
Aschelo 2011	0.7095	1 2225		- 2.03	[0.37, 11.00]	3.4%	3.4%
Gerard (2) 2010	0.0068	1.4118		1.01	[0.06: 16.02]	2.6%	2.6%
Deng 2016	0.0000	1.4099		1.00	[0.06; 15.85]	2.6%	2.6%
O'Connell 2014	0.1761	0.6029		1.19	[0.37; 3.89]	14.1%	14.1%
Haddad 2017	-0.0317	1.3916		0.97	[0.06; 14.82]	2.7%	2.7%
Mohiuddin 2006	-0.0583	1.4004		0.94	[0.06; 14.68]	2.6%	2.6%
Common effect model			\sim	1.32	[0.62; 2.80]	34.9%	
Heterogeneity: $I^2 = 0\%$, $\tau^2 =$	= 0, <i>p</i> = 1.00			1.32	[0.62; 2.80]		34.9%
comparison = L-CRT2 V Deng 2016	/S CHT 0.0000	1.4099		1.00	[0.06; 15.85]	2.6%	2.6%
comparison = CHT VS I Deng 2016	L-CRT1 0.0000	1.4099		1.00	[0.06; 15.85]	2.6%	2.6%
comparison = L-RTVS	L-CRT1						
Bosset 2005	0.0020	1.4128		1.00	[0.06: 15.98]	2.6%	2.6%
Gerard 2006	0.0216	1.4123		1.02	[0.06; 16.28]	2.6%	2.6%
Common effect model				1.01	[0.14; 7.17]	5.1%	
Random effects model				1.01	[0.14; 7.17]		5.1%
Heterogeneity: $I^- = 0\%$, τ^- :	= 0, p = 0.99						
comparison = L-CRT1 \	/S S-RTear	lv					
Buiko 2004	0.6803	1.2195		- 1.97	[0.18: 21.55]	3.5%	3.5%
Ngan 2012	-0.0062	1.4098		0.99	[0.06; 15.75]	2.6%	2.6%
Common effect model				1.47	[0.24; 8.97]	6.0%	
Random effects model				1.47	[0.24; 8.97]		6.0%
Heterogeneity: $I^2 = 0\%$, τ^2 :	= 0, <i>p</i> = 0.71						
comparison = Induction Fernandez-Martos 2015	1 + L-CRT \ -0.7903	/S L-CRT2 1.2087 -		0.45	[0.04; 4.85]	3.5%	3.5%
comparison = Induction	1 + L-CRT \	/SL-CRT+	consolidation	0.06	10.06: 15.221	2.6%	2.6%
FUNdS 2019	-0.0392	1.4090	ſ	0.90	[0.00, 15.25]	2.0 /0	2.0 %
comparison = S-RT + c	onsolidatio	n VS L-CR	F1				
Bahadoer 2021	-1.1249	1.1528 -		0.32	[0.03; 3.11]	3.9%	3.9%
Jin 2022	-0.0169	1.4118		0.98	[0.06; 15.65]	2.6%	2.6%
Chakrabarti 2021	0.7217	1.2130		- 2.06	[0.19; 22.18]	3.5%	3.5%
Common effect model				0.83	[0.20; 3.39]	9.9%	0.0%
Heterogeneity: $I^2 = 0\% \tau^2$:	$= 0 \ p = 0.54$			0.85	[0.20; 3.39]		9.9%
neterogeneity. r = 070, t	- 0, p - 0.04						
comparison = L-CRT1 \	/SL-CRT+	- consolidat	ion				
Kim 2018	-0.0370	1.4011		0.96	[0.06; 15.01]	2.6%	2.6%
Moore 2017	0.0408	1.3850	Î	1.04	[0.07; 15.73]	2.7%	2.7%
Common effect model				1.00	[0.15; 6.91]	5.3%	
Heterogeneity: $l^2 = 0\% \tau^2$:	$= 0 \ n = 0.97$			1.00	[0.15; 6.91]		5.3%
rieterogeneity. r = 070, t	- 0, p - 0.31						
comparison = L-CRT1 \	/S Inductio	n + L-CRT					
Marechal 2012	-0.0351	1.3892		0.97	[0.06; 14.70]	2.7%	2.7%
Conroy 2021	-1.0943	1.1509 -	*	0.33	[0.04; 3.19]	3.9%	3.9%
Common effect model				0.52	[0.09; 2.93]	6.5%	
Random effects model	- 0 0 - 6			0.52	[0.09; 2.93]		6.5%
Helefogeneity. $I = 0\%$, $\tau =$	- 0, p - 0.50						
comparison = S-RT + c	onsolidatio	n VS L-CR	ſ2				
Bujko (2) 2013	-0.8745	0.6844		0.42	[0.11; 1.60]	11.0%	11.0%
comparison = L-CRT2	/SL-CRT+	consolidat	ion				
Wang 2019	0.0000	1.4024		1.00	[0.06; 15.62]	2.6%	2.6%
-							
comparison = L-CRT1 \	/S CHT	1 0000	_	0.50	10.05.5.00	0.407	0.407
Mel 2023 Schrag 2022	-0.0058	1.2220		0.52	[0.05; 5.69]	3.4%	3.4%
Schrag 2023	-1.0241	1.1532 -		0.30	[0.04, 3.44]	3.9%	3.9%
Random effects model				0.43	[0.08. 2.21]	1.3%	7.3%
Heterogeneity: $I^2 = 0\%$, τ^2 :	= 0, <i>p</i> = 0.83			0.40	[0.00, 2.21]		
			ļ				
Common effect model				0.88	[0.56; 1.37]	100.0%	400.0%
Random effects model				0.88	[0.00; 1.37]		100.0%
			0.1 0.5 1 2 10				

Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, p = 1.00Residual heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, p = 1.00Test for subgroup differences (common effect): $\chi^2_{12} = 4.10$, df = 12 (p = 0.98) Test for subgroup differences (random effects): $\chi^2_{12} = 4.10$, df = 12 (p = 0.98)

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	1/2						
		٧Z	V3	V4	V5	V6	V7	V8
1 C	нт			1.87 (0.46 to 7.70)	1.00 (0.06 to 15.85)			
2 1.46 (0.2	23 to 9.50)	Induction + L-CRT	0.96 (0.06 to 15.23)	1.94 (0.34 to 11.03)	0.45 (0.04 to 4.85)			
3 1.60 (0.2	3 to 11.33)	1.09 (0.20 to 5.85)	L-CRT + consolidation	1.00 (0.14 to 6.88)	1.00 (0.06 to 15.62)			
4 1.90 (0.4	19 to 7.39)	1.30 (0.35 to 4.77)	1.19 (0.29 to 4.93)	L-CRT1	0.76 (0.36 to 1.61)	0.99 (0.14 to 7.00)	1.21 (0.29 to 4.94)	1.47 (0.24 to 8.97
5 1.28 (0.3	30 to 5.55)	0.88 (0.23 to 3.39)	0.80 (0.18 to 3.50)	0.68 (0.35 to 1.31)	L-CRT2		2.40 (0.63 to 9.17)	
6 1.88 (0.1	7 to 20.33)	1.28 (0.12 to 13.45)	1.17 (0.10 to 13.19)	0.99 (0.14 to 7.00)	1.46 (0.19 to 11.57)	L-RT		
7 2.67 (0.5	0 to 14.41)	1.83 (0.36 to 9.16)	1.67 (0.30 to 9.28)	1.41 (0.50 to 3.95)	2.09 (0.75 to 5.79)	1.43 (0.16 to 13.03)	S-RT + consolidation	
8 2.79 (0.2	9 to 26.82)	1.91 (0.21 to 17.71)	1.74 (0.17 to 17.43)	1.47 (0.24 to 8.97)	2.18 (0.32 to 14.95)	1.49 (0.10 to 21.39)	1.05 (0.13 to 8.38)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: other vs 'L-CR' (Random Effects Model)	T1' RR	95%-Cl
CHT L-CRT2 Induction + L-CRT L-CRT + consolidation L-RT S-RT + consolidation S-RTearly		→ 1.90 - 1.48 → 1.30 → 1.19 → 1.01 0.71 → 0.68	[0.49; 7.39] [0.76; 2.88] [0.35; 4.77] [0.29; 4.93] [0.14; 7.17] [0.25; 1.99] [0.11; 4.14]
0	01 0.5 1 2	4	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² ≤ 0.010 , moderate with $0.010 < tau² <math>\leq 0.242$, high with tau²> 0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0$; tau = 0

 $I^2 = 0 \% (0 \% \text{ to } 48.92 \%)$

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 3.31 18
 0.9999

 Within designs
 2.33 12
 0.9987

 Between designs
 0.98 6
 0.9863

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:CHT 0.05 1 0.8265 L-CRT1:Induction + L-CRT 0.34 1 0.5571 L-CRT1:L-CRT + consolidation 0.00 1 0.9685 L-CRT1:L-CRT2 0.56 5 0.9898 L-CRT1:L-RT 0.00 1 0.9922 L-CRT1:S-RT + consolidation 1.24 2 0.5386 L-CRT1:S-RTearly 0.14 1 0.7127

Between-designs Q statistic after detaching of single designs

 $\begin{array}{c|c} Detached \ design & Q \ df \ p-value \\ Induction + L-CRT:L-CRT + \ consolidation \ 0.97 \ 5 \ 0.9651 \\ Induction + L-CRT:L-CRT2 \ 0.54 \ 5 \ 0.9905 \\ L-CRT + \ consolidation:L-CRT2 \ 0.95 \ 5 \ 0.9668 \\ L-CRT1:CHT \ 0.78 \ 5 \ 0.9780 \\ L-CRT1:Induction + L-CRT \ 0.51 \ 5 \ 0.9919 \\ L-CRT1:L-CRT + \ consolidation \ 0.91 \ 5 \ 0.9692 \\ L-CRT1:L-CRT2 \ 0.79 \ 5 \ 0.9780 \\ L-CRT1:L-CRT2 \ 0.79 \ 5 \ 0.9780 \\ L-CRT1:S-RT + \ consolidation \ 0.88 \ 5 \ 0.9714 \\ L-CRT2:S-RT + \ consolidation \ 0.88 \ 5 \ 0.9714 \\ L-CRT1:CHT:L-CRT2 \ 0.71 \ 4 \ 0.9502 \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 0.98 6 0.9863 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value L-CRT1:CHT 3 0.93 0.53 0.53 0.45 1.18 0.06 0.9490 CHT:L-CRT2 1 0.28 1.28 1.00 1.41 0.71 -0.21 0.8355 Induction + L-CRT:L-CRT + consolidation 1 0.37 1.09 0.96 1.18 0.82 -0.11 0.9088 L-CRT1:Induction + L-CRT 2 0.56 0.77 0.52 1.30 0.40 -0.69 0.4909 Induction + L-CRT:L-CRT2 1 0.33 0.88 0.45 1.20 0.38 -0.66 0.5078

L-CRT1:L-CRT + consolidation 2 0.54 0.84 1.00 0.69 1.46 0.26 0.7953 L-CRT + consolidation:L-CRT2 1 0.29 0.80 1.00 0.73 1.37 0.19 0.8509 L-CRT1:L-CRT2 7 0.78 0.68 0.76 0.45 1.69 0.64 0.5215 L-CRT1:S-RT + consolidation 3 0.54 1.41 1.21 1.68 0.72 -0.31 0.7532 L-CRT2:S-RT + consolidation 1 0.58 2.09 2.40 1.72 1.39 0.31 0.7532

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

P-score
0.7391
0.6839
0.5751
0.5297
0.4674
0.4224
0.3156
0.2669

eAppendix 11. Rate of randomized patients who underwent surgery

Characteristics of the network

Characteristics of the network

Number of treatments:

9

Number of studies:

27

Number of individuals included:

13413

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized	
1	CHT	1050	
2	Induction + L-CRT	469	
3	L-CRT + consolidation	288	
4	L-CRT1	6084	
5	L-CRT2	3169	
6	L-RT	872	
7	S-RT + consolidation	1090	
8	S-RTdelayed	75	
9	S-RTearly	316	

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logPD	SE(logPP)	Dick Datio	DD	95% CI	Weight	Weight (random)
Study	IOGKK	SE(IOGKK)	RISK RAUO	ĸĸ	90%-01	(common)	(random)
comparison = L-CRT2	/SL-CR	T1					
Rodel 2015	-0.0152	0.0082		0.98	[0.97; 1.00]	17.9%	8.6%
Aschele 2011	0.0097	0.0140	主	1.01	[0.98; 1.04]	6.1%	6.2%
Gerard (2) 2010	0.0068	0.0109	1.	1.01	[0.99; 1.03]	10.1%	7.5%
Deng 2016 O'Connoll 2014	0.0341	0.0392	<u> </u>	1.03	[0.96; 1.12]	0.8%	1.6%
C Connell 2014 Schmoll 2021	0.0035	0.0108		1.00	[0.98, 1.03]	10.3%	7.0%
Haddad 2017	-0.0510	0.0104		0.95	[0.92, 0.96]	4.5%	0.4%
Jiao 2015	0.0000	0.0097	÷	1 00	[0.00, 1.00]	12.9%	8.0%
Mohiuddin 2006	0.0251	0.0536	_ 	1.03	[0.92; 1.14]	0.4%	0.9%
Common effect model			¢	1.00	[0.99; 1.00]	63.1%	
Random effects model			4	1.00	[0.98; 1.01]		46.0%
Heterogeneity: $I^2 = 51\%$, τ^2	= 0.0002	p = 0.04					
comparison - L CDT21							
Deng 2016	-0 0199	0.0342		0.98	[0 92 [.] 1 05]	1.0%	2.0%
20192010	0.0100	0.0012		0.00	[0.02, 1.00]		2.070
comparison = CHTVS	L-CRT1						
Deng 2016	0.0541	0.0374	+	1.06	[0.98; 1.14]	0.9%	1.8%
comparison = L-RTVS	L-CRT1	0.0404	L	4.04	10.00.4.041	4 70/	E E0/
Bossel 2005 Corard 2006	0.0083	0.0101	L	1.01	[0.98, 1.04]	4.7%	0.0%
Common effect model	0.0245	0.0131		1.02	[1.00, 1.03]	11 7%	0.0 %
Random effects model			Ę.	1.02	[0.99:1.05]		12 1%
Heterogeneity: $I^2 = 0\%$. τ^2 :	= 0.0002.	p = 0.44			[0.00, 1.00]		12.170
comparison = L-CRT1	/S S-RT	early					
Bujko 2004	-0.0890	0.0322		0.91	[0.86; 0.97]	1.2%	2.2%
Ngan 2012	-0.0125	0.0177	-	0.99	[0.95; 1.02]	3.8%	5.0%
Common effect model				0.97	[0.94; 1.00]	5.0%	
Random effects model	0.0000	0.04	\sim	0.97	[0.93; 1.00]		7.2%
Heterogeneity: $I^{-} = 17\%$, τ^{-}	= 0.0002	p = 0.04					
comparison = L-CRT1	/9 9 PT	bevelab					
Latkauskas 2012	0.0572	0 0439		1.06	[0 97· 1 15]	0.6%	1.3%
Eunauonuo 2012	0.0012	0.0100		1.00	[0.07, 1.10]	0.070	1.070
comparison = Induction	1 + L-CR	TVSL-CRT2					
Fernandez-Martos 2015	0.0625	0.0361	+	1.06	[0.99; 1.14]	0.9%	1.9%
comparison = Induction	1 + L-CR	TVSL-CRT	⊦ consolidation				
Fokas 2019	-0.0392	0.0302		0.96	[0.91; 1.02]	1.3%	2.5%
comparison = S-RT+c	oneolida	tion VSL-CE	71				
Bahadoer 2021	0.0367	0 0215		1.04	IN 99- 1 081	2.6%	4.0%
Jin 2022	0.0046	0.0428		1.04	[0.92 1.09]	0.7%	1.4%
Chakrabarti 2021	-0.0360	0.0609		0.96	[0.86; 1.09]	0.3%	0.7%
Common effect model			♦	1.02	[0.99; 1.06]	3.6%	
Random effects model			⇒	1.02	[0.98; 1.06]		6.1%
Heterogeneity: $I^2 = 0\%$, $\tau^2 =$	= 0.0002,	p = 0.47					
comparison = L-CRT1	/SL-CR	T + consolida	ation		10.00 4.041	0.00/	0.00/
Kim 2018	0.1300	0.0701		1.14	[0.99; 1.31]	0.2%	0.6%
Noore 2017	0.0000	0.0404		1.00	[0.92, 1.08]	0.7%	1.5%
Random effects model			5	1.03	10.96.1.11	1.0 %	2 1%
Heterogeneity: $I^2 = 61\%$. τ^2	= 0.0002	p = 0.11		1.00	[0.30, 1.11]		2.170
fieldrogeneitj. F		., p					
comparison = L-CRT1	/S Induc	tion + L-CRT					
Marechal 2012	0.0013	0.0505		1.00	[0.91; 1.11]	0.5%	1.0%
Conroy 2021	0.0275	0.0246	+	1.03	[0.98; 1.08]	2.0%	3.3%
Common effect model				1.02	[0.98; 1.07]	2.5%	
Random effects model	0.0000	0.01	A 1	1.02	[0.97; 1.07]		4.4%
Heterogeneity: $I^- = 0\%$, $\tau^- = 0\%$	= 0.0002,	p = 0.64					
comparison = S-RT+c	onsolida	tion VSL-CE	772				
Buiko (2) 2013	0 0104	0 0242		1 01	[0.96 [,] 1.06]	2 1%	34%
20110 (2) 2010	0.0101	0.0212			[0.00, 1.00]	2.170	0.170
comparison = L-CRT2	/SL-CR	T + consolida	ation				
Wang 2019	-0.0728	0.0555	—• +	0.93	[0.83; 1.04]	0.4%	0.9%
comparison = L-CRT1	/S CHT	0.0007		4.00	10.05.1.05		0.00
Mel 2023 Schrog 2022	-0.0039	0.0267	1_	1.00	[0.95; 1.05]	1.7%	3.0%
Common effect model	0.0200	0.0107	Ľ	1.03	[0.99, 1.00]	4.3% 6.0%	0.3%
Random effects model			Ę.	1.02	[0.98; 1.05]	0.0 %	8.3%
Heterogeneity: $I^2 = 0\%$. τ^2 :	= 0.0002.	p = 0.33	ſ	1.02	[5.670
5	,						
Common effect model			¢	1.00	[0.99; 1.01]	100.0%	
Random effects model			· · · · · · · · · · · · · · · · · · ·	1.00	[0.99; 1.01]		100.0%
			0.75 1	1.5			
			0.10	1.0			

Heterogeneity: $l^2 = 45\%$, $\tau^2 = 0.0003$, p < 0.01Residual heterogeneity: $l^2 = 43\%$, $\tau^2 = 0.0002$, p = 0.03Test for subgroup differences (common effect): $\chi^2_{\frac{1}{2}3} = 24.16$, df = 13 (p = 0.03) Test for subgroup differences (random effects): $\chi^2_{\frac{1}{1}3} = 17.94$, df = 13 (p = 0.16)

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8	V9
1	CHT			1.00 (0.96 to 1.03)	1.02 (0.95 to 1.10)				
2	1.00 (0.95 to 1.05)	Induction + L-CRT	0.96 (0.90 to 1.03)	0.98 (0.93 to 1.03)	1.06 (0.98 to 1.15)				
3	0.98 (0.93 to 1.04)	0.98 (0.93 to 1.03)	L-CRT + consolidation	0.97 (0.90 to 1.04)	1.08 (0.96 to 1.20)				
4	1.00 (0.96 to 1.03)	0.99 (0.96 to 1.03)	1.01 (0.96 to 1.06)	L-CRT1	1.00 (0.99 to 1.02)	0.98 (0.95 to 1.01)	0.98 (0.94 to 1.02)	1.06 (0.97 to 1.16)	0.96 (0.93 to 1.00
5	1.00 (0.97 to 1.04)	1.00 (0.96 to 1.04)	1.02 (0.97 to 1.07)	1.01 (0.99 to 1.02)	L-CRT2		0.99 (0.93 to 1.05)		
6	0.98 (0.94 to 1.02)	0.98 (0.93 to 1.03)	0.99 (0.94 to 1.05)	0.98 (0.95 to 1.01)	0.97 (0.94 to 1.01)	L-RT			
7	0.98 (0.94 to 1.03)	0.98 (0.93 to 1.03)	1.00 (0.94 to 1.06)	0.99 (0.95 to 1.02)	0.98 (0.94 to 1.01)	1.00 (0.96 to 1.05)	S-RT + consolidation		
8	1.05 (0.96 to 1.16)	1.05 (0.95 to 1.16)	1.07 (0.97 to 1.19)	1.06 (0.97 to 1.16)	1.05 (0.96 to 1.15)	1.08 (0.98 to 1.19)	1.07 (0.97 to 1.18)	S-RTdelayed	
9	0.96 (0.91 to 1.01)	0.96 (0.91 to 1.01)	0.98 (0.92 to 1.04)	0.96 (0.93 to 1.00)	0.96 (0.92 to 1.00)	0.98 (0.93 to 1.03)	0.98 (0.93 to 1.03)	0.91 (0.82 to 1.01)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

	Comparison: other vs 'L-CRT1'		
Treatment	(Random Effects Model)	RR	95%-CI
S-RTearly		1.04	[1.00; 1.08]
L-RT		1.02	[0.99; 1.05]
S-RT + consolidation		1.01	[0.98; 1.05]
L-CRT + consolidation		1.01	[0.96; 1.06]
CHT		1.00	[0.96; 1.03]
Induction + L-CRT		0.99	[0.96; 1.03]
L-CRT2		0.99	[0.98; 1.01]
S-RTdelayed	÷	0.94	[0.86; 1.04]
0	.01 0.5 1 2 4		

Assessment of heterogeneity and consistency

Global heterogeneity
We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau² \geq 0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

tau^2= 3e-04 ; tau= 0.0164

I^2= 44.04 % (6.35 % to 66.56 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 35.74 20
 0.0165

 Within designs
 25.44 14
 0.0304

 Between designs
 10.30 6
 0.1127

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:CHT 0.94 1 0.3320 L-CRT1:Induction + L-CRT 0.22 1 0.6403 L-CRT1:L-CRT + consolidation 2.58 1 0.1080 L-CRT1:L-CRT2 15.27 7 0.0327 L-CRT1:L-RT 0.60 1 0.4393 L-CRT1:S-RT + consolidation 1.52 2 0.4674 L-CRT1:S-RTearly 4.31 1 0.0378

Between-designs Q statistic after detaching of single designs

 $\begin{array}{c|c} & \text{Detached design} & \mathbb{Q} \ \text{df } p\text{-value} \\ & \text{Induction} + \text{L-CRT:L-CRT} + \text{consolidation} 9.45 & 5 & 0.0923 \\ & \text{Induction} + \text{L-CRT:L-CRT2} & 6.15 & 5 & 0.2915 \\ & \text{L-CRT} + \text{consolidation:L-CRT2} & 9.18 & 5 & 0.1020 \\ & \text{L-CRT1:CHT} & 8.11 & 5 & 0.1505 \\ & \text{L-CRT1:Induction} + \text{L-CRT} & 9.21 & 5 & 0.1010 \\ & \text{L-CRT1:L-CRT} + \text{consolidation} & 7.36 & 5 & 0.1953 \\ & \text{L-CRT1:L-CRT2} & 9.57 & 5 & 0.0882 \\ & \text{L-CRT1:S-RT} + \text{consolidation} & 9.84 & 5 & 0.0798 \\ & \text{L-CRT2:S-RT} + \text{consolidation} & 9.84 & 5 & 0.0798 \\ & \text{L-CRT1:CHT:L-CRT2} & 7.05 & 4 & 0.1333 \\ \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 8.50 6 0.2037 0.0148 0.0002

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value L-CRT1:CHT 3 0.92 1.00 1.00 1.02 0.99 -0.22 0.8271 CHT:L-CRT2 1 0.22 1.00 1.02 1.00 1.02 0.47 0.6363 Induction + L-CRT:L-CRT + consolidation 1 0.55 0.98 0.96 1.01 0.95 -0.90 0.3681 L-CRT1:Induction + L-CRT 2 0.58 1.01 1.02 0.99 1.03 0.85 0.3946 Induction + L-CRT:L-CRT2 1 0.26 1.00 1.06 0.98 1.09 1.79 0.0732

L-CRT1:L-CRT + consolidation 2 0.45 0.99 1.04 0.95 1.09 1.67 0.0952 L-CRT + consolidation:L-CRT2 1 0.19 1.02 1.08 1.01 1.07 1.01 0.3107 L-CRT1:L-CRT2 9 0.91 1.01 1.00 1.05 0.96 -1.51 0.1305 L-CRT1:S-RT + consolidation 3 0.66 0.99 0.98 1.00 0.98 -0.53 0.5961 L-CRT2:S-RT + consolidation 1 0.39 0.98 0.99 0.97 1.02 0.53 0.5961

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

	P-score
S-RTearly	0.8927
L-RT	0.7185
S-RT + consolidation	0.6686
L-CRT + consolidation	0.6181
L-CRT1	0.4658
CHT	0.3896
Induction + L-CRT	0.3578
L-CRT2	0.2828
S-RTdelayed	0.1062

eAppendix 12. Rate of R0 resections

Characteristics of the network

Number of treatments:

8

Number of studies:

18

Number of individuals included:

9145

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized	
1	CHT	1050	
2	Induction + L-CRT	441	
3	L-CRT + consolidation	228	
4	L-CRT1	4007	
5	L-CRT2	2093	
6	S-RT + consolidation	1090	
7	S-RTdelayed	75	
8	S-RTearly	161	

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

						Weight	Weight
Study	logRR	SE(logRR)	Risk Ratio	RR	95%-CI	(common)	(random)
		74	t				
comparison = L-CR12 \	/SL-CR	0.0455	. t	0.00	10.00.4.001	40.00/	44.50/
Acchelo 2011	-0.0134	0.0100	8	1.04	[0.90, 1.02]	10.9%	14.3%
Aschele 2011	0.0400	0.0230		1.04	[0.99, 1.09]	0.170	0.1%
Schmoll 2021	0.0450	0.0302		0.08	[0.94, 1.17]	21.6%	2.0%
Jiao 2015	0.0202	0.0280	_ i	1 02	[0.93, 1.00]	5.7%	6.7%
Common effect model	0.0202	0.0200	4	0.99	[0.98: 1.01]	55.8%	
Random effects model			\$	1.00	[0.98: 1.03]		47.4%
Heterogeneity: $I^2 = 47\%$, τ^2	= 0.0003	, p = 0.11			• / •		
comparison = L-CRT2 \	/S CHT	0.0540		0.00	10.00 4.001	4 70/	0.40/
Deng 2016	-0.0148	0.0519		0.99	[0.89; 1.09]	1.7%	2.4%
comparison = CHT VS I	L-CRT1						
Deng 2016	0.0606	0.0552	<u> </u>	1.06	[0.95; 1.18]	1.5%	2.1%
comparison = L-CRT1 \	/S S-RT	early					
Ngan 2012	0.0005	0.0301		1.00	[0.94; 1.06]	5.0%	6.0%
		de la const					
comparison = L-CR I1 \	0 1 1 E O	delayed		1 1 2	10.06-1.221	0.7%	1 00/
Laikauskas 2012	0.1156	0.0600		1.12	[0.90, 1.32]	0.7%	1.0%
comparison = Induction	+L-CR	TVSL-CRT2					
Fernandez-Martos 2015	-0.0326	0.0643		0.97	[0.85: 1.10]	1.1%	1.6%
					[]		
comparison = Induction	1 + L-CR	TVSL-CRT+	- consolidation				
Fokas 2019	-0.0237	0.0493		0.98	[0.89; 1.08]	1.9%	2.6%
comparison = S-RT + co	onsolida	tion VS L-CR	.T1	4 00	10.07.4.401	4.50/	5 50/
banadoer 2021	0.0330	0.0318		1.03	[0.97; 1.10]	4.5%	0.0% 0.0%
JIII 2022 Chakrabarti 2021	0.0454	0.0552		0.06	[0.94, 1.10]	1.0%	2.3%
Common effect model	-0.0500	0.0005		1 02	[0.00, 1.09]	7.3%	1.0 /0
Random effects model			5	1.02	[0.97 1 08]	1.070	9.6%
Heterogeneity: $I^2 = 0\%$, $\tau^2 =$	= 0.0003.	p = 0.55		1.02	[0.07, 1.00]		0.070
, , , , , , , , , , , , , , , , , , ,							
comparison = L-CRT1 \	/SL-CR	T + consolida	tion				
Kim 2018	0.2506	0.0884	· · · · · · · · · · · · · · · · · · ·	- 1.28	[1.08; 1.53]	0.6%	0.9%
Moore 2017	-0.0036	0.0852		1.00	[0.84; 1.18]	0.6%	0.9%
Common effect model				1.13	[1.00; 1.27]	1.2%	
Random effects model	0.0000	0.04		1.13	[1.00; 1.27]		1.8%
Heterogeneity: $I^{-} = II\%$, τ^{-}	= 0.0003	p = 0.04					
comparison = L-CRT1	/S Induc	tion + L-CRT					
Conrov 2021	0 0093	0 0353		1 01	[0 94 [.] 1 08]	3.6%	47%
0011072021	0.0000	0.0000			[0.01, 1.00]	0.070	
comparison = S-RT + co	onsolida	tion VS L-CR	T2				
Bujko (2) 2013	0.0993	0.0529		1.10	[1.00; 1.23]	1.6%	2.3%
comparison = L-CR I1 \	/S CHI	0.0070		1 00	10 04: 4 0F1	C 10/	7 40/
Mei 2023	-0.0041	0.0273	¥	1.00	[0.94, 1.05]	0.1%	1.1%
Common effect model	0.0001	0.0190	, i i i i i i i i i i i i i i i i i i i	1.01	[0.97, 1.03]	12.0%	11.0 /0
Random effects model			4	1.00	[0.96 1.04]	10.7 /0	18.6%
Heterogeneity: $I^2 = 0\%$. $\tau^2 =$	= 0.0003.	p = 0.71		1.00	[0.00, 1.04]		10.070
· · · · · · · · · · · · · · · · · · ·	,						
Common effect model			A	1.00	[0.99; 1.02]	100.0%	
Random effects model			P	1.01	[0.99; 1.03]		100.0%
			0.75 4	5			
Heterogeneity: $I^2 = 2504$ $-^2$	- 0.0002	n = 0.15	0.75 1	C.1			
Residual heterogeneity: $I = 20\%$, τ	= 0.0002 = 40% τ^2	= 0.0003 n = 0.003 n) 10				
Test for subgroup differenc	es (com	non effect): γ^2 .	= 12.20, df = 11 (p = 0.35)				
Test for subgroup differenc	, es (rando	m effects): χ_{11}^{21}	= 9.89, df $= 11 (p = 0.54)$				

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8
1	CHT			1.00 (0.97 to 1.05)	1.01 (0.91 to 1.13)			
2	1.04 (0.96 to 1.12)	Induction + L-CRT	0.98 (0.88 to 1.08)	0.99 (0.91 to 1.07)	0.97 (0.85 to 1.10)	100 C		
3	1.06 (0.96 to 1.17)	1.02 (0.94 to 1.11)	L-CRT + consolidation	0.89 (0.78 to 1.00)			-	
4	1.00 (0.96 to 1.04)	0.97 (0.91 to 1.03)	0.95 (0.87 to 1.03)	L-CRT1	1.00 (0.97 to 1.03)	0.98 (0.92 to 1.03)	1.12 (0.95 to 1.32)	1.00 (0.93 to 1.07)
5	1.01 (0.96 to 1.05)	0.97 (0.91 to 1.03)	0.95 (0.87 to 1.04)	1.00 (0.98 to 1.03)	L-CRT2	0.91 (0.81 to 1.01)		
6	0.97 (0.91 to 1.03)	0.93 (0.86 to 1.01)	0.91 (0.82 to 1.01)	0.96 (0.92 to 1.01)	0.96 (0.91 to 1.01)	S-RT + consolidation		
7	1.13 (0.95 to 1.33)	1.08 (0.91 to 1.29)	1.06 (0.88 to 1.28)	1.12 (0.95 to 1.32)	1.12 (0.95 to 1.32)	1.17 (0.98 to 1.38)	S-RTdelayed	100 A
8	1.00 (0.93 to 1.09)	0.97 (0.88 to 1.06)	0.95 (0.84 to 1.06)	1.00 (0.93 to 1.07)	1.00 (0.92 to 1.08)	1.04 (0.95 to 1.13)	0.89 (0.75 to 1.06)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: other vs 'L-CF (Random Effects Model)	RT1' RR	95%-CI
S-RT + consolidation	+	1.04	[0.99; 1.09]
CHT	+	1.00	[0.96; 1.04]
S-RTearly		1.00	[0.93; 1.07]
L-CRT2		1.00	[0.97; 1.02]
Induction + L-CRT		0.97	[0.91; 1.03]
L-CRT + consolidation	L 🖻	0.95	[0.87; 1.03]
S-RTdelayed	r	0.89	[0.76; 1.05]
C	0.01 0.5 1 2	4	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² ≤ 0.010 , moderate with $0.010 < tau² <math>\leq 0.242$, high with tau²> 0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

tau^2= 4e-04 ; tau= 0.0204

I^2= 31.7 % (0 % to 64.7 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 17.57 12
 0.1294

 Within designs
 12.39 7
 0.0885

 Between designs
 5.18 5
 0.3941

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:CHT 0.13 1 0.7143 L-CRT1:L-CRT + consolidation 4.29 1 0.0384 L-CRT1:L-CRT2 6.76 3 0.0800 L-CRT1:S-RT + consolidation 1.21 2 0.5459

Between-designs Q statistic after detaching of single designs

 $\begin{array}{cccccccc} & Detached \ design & Q \ df \ p-value \\ Induction + L-CRT:L-CRT + \ consolidation \ 2.96 \ 4 \ 0.5652 \\ Induction + L-CRT:L-CRT2 \ 5.17 \ 4 \ 0.2704 \\ & L-CRT1:CHT \ 4.64 \ 4 \ 0.3264 \\ & L-CRT1:Induction + L-CRT \ 3.79 \ 4 \ 0.4345 \\ & L-CRT1:L-CRT + \ consolidation \ 2.96 \ 4 \ 0.5652 \\ & L-CRT1:L-CRT \ 2.507 \ 4 \ 0.2803 \\ & L-CRT1:S-RT + \ consolidation \ 3.80 \ 4 \ 0.4331 \\ & L-CRT2:S-RT + \ consolidation \ 3.80 \ 4 \ 0.4331 \\ & L-CRT1:CHT:L-CRT2 \ 3.75 \ 3 \ 0.2902 \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 4.55 5 0.4738 0.0220 0.0005

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value
L-CRT1:CHT 3 0.95 1.00 1.00 1.03 0.97 -0.34 0.7338
CHT:L-CRT2 1 0.17 1.01 1.01 1.00 1.01 0.18 0.8568Induction + L-CRT:L-CRT + consolidation 1 0.65 1.02 0.98 1.11 0.88 -1.41 0.1581
L-CRT1:Induction + L-CRT 1 0.63 1.04 1.01 1.08 0.93 -1.05 0.2942
Induction + L-CRT:L-CRT2 1 0.25 0.97 0.97 1.00 0.00 0.9989
L-CRT1:L-CRT + consolidation 2 0.51 1.06 1.13 0.99 1.14 1.41 0.1581
L-CRT1:L-CRT + consolidation 3 0.81 0.96 0.98 0.90 1.08 1.19 0.2339
L-CRT2:S-RT + consolidation 1 0.23 0.96 0.91 0.98 0.93 -1.19 0.2339

k

Legend: comparison - Treatment comparison - Number of studies providing direct evidence

prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

	P-score
S-RT + consolidation	0.9156
CHT	0.6396
L-CRT1	0.6114
S-RTearly	0.5972
L-CRT2	0.5698
Induction + L-CRT	0.3192
L-CRT + consolidation	0.2265
S-RTdelayed	0.1206

eAppendix 13. Rate of negative CRM

Characteristics of the network

Number of treatments:

8

Number of studies:

11

Number of individuals included:

4963

Number of individuals randomized to each treatment:

Treatment name	N. individuals randomized	
CHT	585	
Induction + L-CRT	415	
L-CRT + consolidation	203	
L-CRT1	2312	
L-CRT2	756	
S-RT + consolidation	462	
S-RTdelayed	75	
S-RTearly	155	
	Treatment name CHT Induction + L-CRT L-CRT + consolidation L-CRT1 L-CRT2 S-RT + consolidation S-RTdelayed S-RTearly	Treatment nameN. individuals randomizedCHT585Induction + L-CRT415L-CRT + consolidation203L-CRT12312L-CRT2756S-RT + consolidation462S-RTdelayed75S-RTearly155

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	E(logRR)	Risk Ratio	RR	95%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 V Aschele 2011 Gerard (2) 2010 Jiao 2015 Common effect model Random effects model Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0\%$	/S L-CRT1 0.0357 0.0068 0.0202	0.0668 0.0917 0.0280		1.04 1.01 1.02 1.02 1.02	[0.91; 1.18] [0.84; 1.21] [0.97; 1.08] [0.97; 1.07] [0.97; 1.07]	3.9% 2.1% 22.1% 28.1% 	5.9% 3.4% 18.1% 27.4%
comparison = L-CRT1 \ Bujko 2004	/S S-R Tear 0.0589	ly 0.0560		1.06	[0.95; 1.18]	5.6%	7.8%
comparison = L-CRT1 \ Latkauskas 2012	/S S-RTdel 0.0645	ayed 0.0750		1.07	[0.92; 1.24]	3.1%	4.9%
comparison = Inductior Fokas 2019	• + L-CRT \ -0.0700	/SL-CRT+ 0.0481	consolidation	0.93	[0.85; 1.02]	7.5%	9.8%
comparison = S-RT + co Bahadoer 2021	onsolidatio 0.0325	n VS L-CR 0.0311	T1	1.03	[0.97; 1.10]	18.0%	16.4%
comparison = L-CRT1 \ Kim 2018	/S L-CRT + 0.2232	consolida 0.1040	tion	- 1.25	[1.02; 1.53]	1.6%	2.7%
$\label{eq:comparison} \begin{array}{l} \text{comparison} = \text{L-CRT1} \\ \text{Marechal 2012} \\ \text{Conroy 2021} \\ \text{Common effect model} \\ \text{Random effects model} \\ \text{Heterogeneity: } I^2 = 0\%, \tau^2 = 0 \\ \end{array}$	/S Inductio -0.1121 -0.1450 = 0, p = 0.75	n + L-CRT 0.0827 0.0618	+	0.89 0.87 0.88 0.88	[0.76; 1.05] [0.77; 0.98] [0.79; 0.96] [0.79; 0.96]	2.5% 4.6% 7.1%	4.1% 6.7% 10.8%
comparison = L-CRT1 V Schrag 2023	/S CHT -0.0105	0.0245	+	0.99	[0.94; 1.04]	28.9%	20.2%
Common effect model Random effects model			÷	1.00 1.00	[0.98; 1.03] [0.97; 1.04]	100.0% 	 100.0%
Heterogeneity: $I^2 = 44\%$, τ^2 Residual heterogeneity: $I^2 =$ Test for subgroup differenc Test for subgroup differenc	= 0.0010, <i>p</i> = 0%, τ ² = 0, es (common es (random e	= 0.06 p = 0.98 effect): χ_{2}^{2} = effects): χ_{7}^{2}	0.75 1 1 17.79, df = 7 (p = 0.01) 17.79, df = 7 (p = 0.01)	.5			

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8
1	CHT			1.01 (0.84 to 1.22)				
2	0.96 (0.76 to 1.22)	Induction + L-CRT	0.93 (0.76 to 1.14)	1.14 (0.97 to 1.34)				
3	1.01 (0.78 to 1.32)	1.06 (0.89 to 1.25)	L-CRT + consolidation	0.80 (0.61 to 1.05)			-	
4	1.01 (0.84 to 1.22)	1.05 (0.91 to 1.22)	1.00 (0.83 to 1.20)	L-CRT1	0.98 (0.86 to 1.11)	0.97 (0.80 to 1.17)	1.07 (0.85 to 1.35)	1.06 (0.86 to 1.31)
5	0.99 (0.79 to 1.24)	1.03 (0.85 to 1.25)	0.98 (0.78 to 1.22)	0.98 (0.86 to 1.11)	L-CRT2			
6	0.98 (0.75 to 1.28)	1.02 (0.80 to 1.30)	0.97 (0.74 to 1.26)	0.97 (0.80 to 1.17)	0.99 (0.79 to 1.24)	S-RT + consolidation		
7	1.08 (0.80 to 1.45)	1.12 (0.85 to 1.48)	1.06 (0.79 to 1.44)	1.07 (0.85 to 1.35)	1.09 (0.84 to 1.42)	1.10 (0.82 to 1.49)	S-RTdelayed	
8	1.07 (0.81 to 1.42)	1.12 (0.86 to 1.44)	1.06 (0.80 to 1.41)	1.06 (0.86 to 1.31)	1.08 (0.85 to 1.39)	1.10 (0.82 to 1.46)	0.99 (0.73 to 1.36)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: other vs 'L-C (Random Effects Mode	RT1' I) RR	95%-CI
Induction + L-CRT S-RT + consolidation L-CRT2 CHT L-CRT + consolidation S-RTearly S-RTdelayed		1.05 1.03 1.02 1.01 1.00 0.94 0.94	[0.91; 1.22] [0.85; 1.25] [0.90; 1.16] [0.84; 1.22] [0.83; 1.20] [0.76; 1.16] [0.74; 1.18]
0	.01 0.5 1	2 4	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

tau^2= 0.0085 ; tau= 0.0922

I^2= 66.25 % (12.07 % to 87.05 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 11.85
 4
 0.0185

 Within designs
 0.17
 3
 0.9816

 Between designs
 11.68
 1
 0.0006

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:Induction + L-CRT 0.10 1 0.7495 L-CRT1:L-CRT2 0.07 2 0.9644

Between-designs Q statistic after detaching of single designs

Detached design Q df p-value

- Induction + L-CRT:L-CRT + consolidation 0.00 0 ---
 - L-CRT1:Induction + L-CRT 0.00 0 --
 - L-CRT1:L-CRT + consolidation 0.00 0 --

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 11.68 1 0.0006 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value Induction + L-CRT:L-CRT + consolidation 1 0.71 1.06 0.93 1.42 0.65 -2.20 0.0276 L-CRT1:Induction + L-CRT 2 0.82 0.95 0.88 1.34 0.65 -2.20 0.0276 L-CRT1:L-CRT + consolidation 1 0.48 1.00 1.25 0.82 1.53 2.20 0.0276

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion - Estimated treatment effect (RR) in network meta-analysis nma - Estimated treatment effect (RR) derived from direct evidence direct - Estimated treatment effect (RR) derived from indirect evidence indir. - Ratio of Ratios (direct versus indirect) RoR - z-value of test for disagreement (direct versus indirect) z p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

 P-score

 Induction + L-CRT
 0.7006

 S-RT + consolidation
 0.6056

 L-CRT2
 0.5824

СНТ	0.5310
L-CRT1	0.4795
L-CRT + consolidation	0.4794
S-RTearly	0.3130
S-RTdelayed	0.3085

eAppendix 14. Rate of ypN0

Characteristics of the network

Number of treatments:

9

Number of studies:

21

Number of individuals included:

10070

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized	
1	CHT	885	
2	Induction + L-CRT	469	
3	L-CRT + consolidation	288	
4	L-CRT1	4498	
5	L-CRT2	1732	
6	L-RT	872	
7	S-RT + consolidation	1090	
8	S-RTdelayed	75	
9	S-RTearly	161	

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR	SE(logRR)	Risk Ratio	RR	95%-CI	Weight (common)	Weight (random)		
comparison = L-CRT2 V Rodel 2015 Aschele 2011 Gerard (2) 2010 Jiao 2015 Common effect model Random effects model Heterogeneity: $J^2 = 0\%$, $\tau^2 = 0\%$	/S L-CRT -0.0005 -0.0260 0.0320 0.1297 = 0.0011, <i>p</i>	1 0.0391 0.0492 0.0568 0.0974		1.00 0.97 1.03 1.14 1.01 1.01	[0.93; 1.08] [0.88; 1.07] [0.92; 1.15] [0.94; 1.38] [0.96; 1.06] [0.95; 1.08]	13.4% 8.5% 6.3% 2.2% 30.3%	7.8% 7.0% 6.4% 3.8% 25.0%		
comparison = L-RTVS Bosset 2005 Gerard 2006 Common effect model Random effects model Heterogeneity: l^2 = 80%, τ^2	L-CRT1 -0.1640 0.0004 = 0.0011,	0.0496 0.0554 <i>p</i> = 0.03	* *	0.85 1.00 0.91 0.92	[0.77; 0.94] [0.90; 1.12] [0.85; 0.98] [0.84; 1.00]	8.3% 6.7% 15.0% 	7.0% 6.5% 13.5%		
comparison = L-CRT1 V Ngan 2012	/S S-RTe 0.0649	arly 0.0891		1.07	[0.90; 1.27]	2.6%	4.2%		
comparison = L-CRT1 V Latkauskas 2012	/S S-RTd 0.2278	elayed 0.1229	*	1.26	[0.99; 1.60]	1.4%	2.8%		
comparison = Induction Fernandez-Martos 2015	1 + L-CRT -0.0705	VS L-CRT2 0.1201		0.93	[0.74; 1.18]	1.4%	2.9%		
comparison = Induction Fokas 2019	1 + L-CRT -0.0948	VS L-CRT + coi 0.0739	nsolidation	0.91	[0.79; 1.05]	3.8%	5.1%		
$\begin{array}{l} \mbox{comparison} = \mbox{S-RT} + \mbox{cm}\\ \mbox{Bahadoer} \ 2021\\ \mbox{Jin} \ 2022\\ \mbox{Chakrabarti} \ 2021\\ \mbox{Common effect model}\\ \mbox{Random effects model}\\ \mbox{Heterogeneity:} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	0.1231 0.0385 0.0286 = 0.0011, <i>p</i>	ion VS L-CRT1 0.0493 0.0745 0.0994 = 0.52	* * * *	1.13 1.04 1.03 1.09 1.09	[1.03; 1.25] [0.90; 1.20] [0.85; 1.25] [1.01; 1.18] [1.00; 1.18]	8.4% 3.7% 2.1% 14.2%	7.0% 5.1% 3.7% 15.8%		
$\begin{array}{l} \mbox{comparison} = \mbox{L-CRT1} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	/S L-CRT -0.0734 0.2127 = 0.0011,	<pre>+ consolidation 0.1889 0.1829 p = 0.28</pre>		0.93 1.24 1.08 1.08	[0.64; 1.35] [0.86; 1.77] [0.83; 1.39] [0.83; 1.40]	0.6% 0.6% 1.2%	1.4% 1.5% 2.9%		
comparison = L-CRT1 V Marechal 2012 Conroy 2021 Common effect model Random effects model Heterogeneity: $J^2 = 42\%$, τ^2	/S Induct 0.1725 -0.1837 = 0.0011,	ion + L-CRT 0.2631 0.0627 p = 0.19	+ \$	- 1.19 0.83 0.85 0.85	[0.71; 1.99] [0.74; 0.94] [0.75; 0.96] [0.75; 0.97]	0.3% 5.2% 5.5%	0.8% 5.9% 6.7%		
comparison = S-RT + c Bujko (2) 2013	onsolidat 0.0708	ion VS L-CRT2 0.0791		1.07	[0.92; 1.25]	3.3%	4.8%		
comparison = L-CRT2 V Wang 2019	/S L-CRT -0.2963	+ consolidation 0.1637	*	0.74	[0.54; 1.02]	0.8%	1.8%		
$\begin{array}{l} \mbox{comparison} = \mbox{L-CRT1} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	/S CHT 0.1050 0.0492 = 0.0011, p	0.0537 0.0389 = 0.40	++ • •	1.11 1.05 1.07 1.07	[1.00; 1.23] [0.97; 1.13] [1.01; 1.14] [0.99; 1.16]	7.1% 13.5% 20.6%	6.6% 7.9% 14.5%		
Common effect model Random effects model			0.75 1 1.5	1.01 1.01	[0.98; 1.04] [0.96; 1.06]	100.0% 	 100.0%		
Heterogeneity: $l^2 = 56\%$, $\tau^2 = 0.0059$, $\rho < 0.01$ Residual heterogeneity: $l^2 = 26\%$, $\tau^2 = 0.0011$, $\rho = 0.21$ Test for subgroup differences (common effect): $\chi^2_{41} = 33.51$, df = 11 ($\rho < 0.01$) Test for subgroup differences (random effects): $\chi^2_{11} = 25.99$, df = 11 ($\rho < 0.01$)									

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8	V 9
1	CHT			0.93 (0.85 to 1.02)					
2	0.86 (0.74 to 1.00)	Induction + L-CRT	0.91 (0.77 to 1.08)	1.17 (1.01 to 1.35)	0.93 (0.72 to 1.20)				
3	0.81 (0.68 to 0.97)	0.95 (0.82 to 1.09)	L-CRT + consolidation	0.93 (0.71 to 1.21)	1.34 (0.96 to 1.88)				
4	0.93 (0.85 to 1.02)	1.08 (0.96 to 1.22)	1.14 (0.98 to 1.33)	L-CRT1	0.99 (0.92 to 1.06)	1.09 (0.99 to 1.20)	0.92 (0.84 to 1.02)	1.26 (0.97 to 1.63)	1.07 (0.88 to 1.30)
5	0.91 (0.82 to 1.02)	1.07 (0.94 to 1.21)	1.13 (0.96 to 1.31)	0.98 (0.92 to 1.05)	L-CRT2		0.93 (0.78 to 1.12)		
6	1.01 (0.89 to 1.16)	1.18 (1.01 to 1.38)	1.25 (1.04 to 1.49)	1.09 (0.99 to 1.20)	1.11 (0.99 to 1.25)	L-RT	100 C		
7	0.86 (0.76 to 0.97)	1.00 (0.87 to 1.15)	1.06 (0.89 to 1.25)	0.92 (0.85 to 1.00)	0.94 (0.85 to 1.03)	0.85 (0.74 to 0.96)	S-RT + consolidation		
8	1.17 (0.89 to 1.54)	1.36 (1.02 to 1.81)	1.44 (1.07 to 1.94)	1.26 (0.97 to 1.63)	1.28 (0.98 to 1.67)	1.15 (0.87 to 1.52)	1.36 (1.04 to 1.79)	S-RTdelayed	
9	0.99 (0.80 to 1.23)	1.16 (0.92 to 1.45)	1.22 (0.95 to 1.57)	1.07 (0.88 to 1.30)	1.08 (0.88 to 1.34)	0.98 (0.78 to 1.22)	1.16 (0.93 to 1.44)	0.85 (0.61 to 1.18)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: other vs 'L-CRT1 (Random Effects Model)	RR	95%-CI
L-CRT + consolidation S-RT + consolidation Induction + L-CRT L-CRT2 S-RTearly CHT L-RT S-RTdelayed		1.14 1.08 1.08 1.02 0.94 0.93 0.92 0.80	[0.98; 1.33] [1.00; 1.18] [0.96; 1.22] [0.95; 1.09] [0.77; 1.14] [0.85; 1.02] [0.83; 1.01] [0.62; 1.03]
0	01 0.5 1 2 4	1	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0.0023$; tau = 0.0475

I^2= 32.23 % (0 % to 64.22 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 19.18 13
 0.1176

 Within designs
 12.09
 9
 0.2081

 Between designs
 7.09
 4
 0.1312

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:CHT 0.71 1 0.3996 L-CRT1:Induction + L-CRT 1.73 1 0.1878 L-CRT1:L-CRT + consolidation 1.18 1 0.2767 L-CRT1:L-CRT2 2.26 3 0.5195 L-CRT1:L-RT 4.90 1 0.0269 L-CRT1:S-RT + consolidation 1.31 2 0.5204

Between-designs Q statistic after detaching of single designs

 $\label{eq:constraint} \begin{array}{c} \mbox{Detached design} & \mbox{Q df } \mbox{p-value} \\ \mbox{Induction } + \mbox{L-CRT:L-CRT + consolidation } 6.08 & 3 & 0.1076 \\ \mbox{Induction } + \mbox{L-CRT:L-CRT2 } 5.00 & 3 & 0.1719 \\ \mbox{L-CRT + consolidation:L-CRT2 } 5.97 & 3 & 0.1130 \\ \mbox{L-CRT1:Induction } + \mbox{L-CRT 3.30 } 3 & 0.3480 \\ \mbox{L-CRT1:L-CRT + consolidation } 3.00 & 3 & 0.3915 \\ \mbox{L-CRT1:L-CRT + consolidation } 7.09 & 3 & 0.0692 \\ \mbox{L-CRT2:S-RT + consolidation } 7.09 & 3 & 0.0692 \\ \mbox{L-CRT2:S-RT + consolidation } 7.09 & 3 & 0.0692 \\ \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 6.19 4 0.1855 0.0369 0.0014

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

 $\begin{array}{c} {\rm comparison\ k\ prop\ nma\ direct\ indir.\ RoR\ z\ p-value} \\ {\rm Induction\ +\ L-CRT:L-CRT\ +\ consolidation\ 1\ 0.67\ 0.95\ 0.91\ 1.03\ 0.89\ -0.79\ 0.4280} \\ {\rm L-CRT1:Induction\ +\ L-CRT\ 2\ 0.62\ 0.92\ 0.86\ 1.05\ 0.82\ -1.63\ 0.1034} \\ {\rm Induction\ +\ L-CRT:L-CRT\ 2\ 1\ 0.24\ 1.07\ 0.93\ 1.11\ 0.84\ -1.19\ 0.2336} \\ {\rm L-CRT1:L-CRT\ +\ consolidation\ 2\ 0.32\ 0.87\ 1.08\ 0.79\ 1.36\ 1.87\ 0.0620} \\ {\rm L-CRT\ +\ consolidation:L-CRT\ 2\ 1\ 0.22\ 1.13\ 1.34\ 1.07\ 1.26\ 1.18\ 0.2373} \\ {\rm L-CRT1:L-CRT\ +\ consolidation\ 3\ 0.80\ 0.92\ 0.99\ 0.98\ 1.01\ 0.10\ 0.9209} \\ {\rm L-CRT1:S-RT\ +\ consolidation\ 3\ 0.80\ 0.92\ 0.92\ 0.92\ 1.01\ 0.08\ 0.9337} \\ {\rm L-CRT2:S-RT\ +\ consolidation\ 1\ 0.30\ 0.94\ 0.93\ 0.94\ 0.99\ -0.08\ 0.9337} \end{array}$

Legend: comparison - Treatment comparison k - Number of studies providing direct evidence

prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

	P-score
L-CRT + consolidation	0.9141
S-RT + consolidation	0.8155
Induction + L-CRT	0.7881
L-CRT2	0.5824
L-CRT1	0.5085
S-RTearly	0.3354
CHT	0.2597
L-RT	0.2247
S-RTdelayed	0.0715

eAppendix 15. Rate of post-operative complications Clavien-Dindo III or Greater

Characteristics of the network

Number of treatments:

6

Number of studies:

9

Number of individuals included:

3525

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized
1	Induction + L-CRT	356
2	L-CRT + consolidation	212
3	L-CRT1	1629
4	L-CRT2	883
5	S-RT + consolidation	295
6	S-RTearly	150

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	(logRR)	Risk Ratio	RR	95%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 ' Rodel 2015 Gerard (2) 2010 Common effect model Random effects model Heterogeneity: $I^2 = 0\%$, τ^2	VS L-CRT1 0.1984 -0.0274 = 0, <i>p</i> = 0.39	0.1438 0.2187		1.22 0.97 1.14 1.14	[0.92; 1.62] [0.63; 1.49] [0.90; 1.44] [0.90; 1.44]	36.5% 15.8% 52.2% 	36.5% 15.8% 52.2%
comparison = L-CRT1 Bujko 2004	VS S-RTear -0.3293	ly 0.3536		0.72	[0.36; 1.44]	6.0%	6.0%
comparison = Induction Fokas 2019	n + L-CRT V 0.0870	/S L-CRT + co 0.2704	onsolidation	1.09	[0.64; 1.85]	10.3%	10.3%
$\begin{array}{l} \text{comparison} = \text{L-CRT1} \\ \text{Kim 2018} \\ \text{Moore 2017} \\ \text{Common effect model} \\ \text{Random effects model} \\ \text{Heterogeneity: } \textit{J}^2 = 0\%, \tau^2 \end{array}$	VS L-CRT + -1.5533 -0.3646 = 0, <i>p</i> = 0.34	consolidatio 1.0991 0.5788	n	0.21 0.69 0.54 0.54	[0.02; 1.82] [0.22; 2.16] [0.20; 1.46] [0.20; 1.46]	0.6% 2.3% 2.9% 	0.6% 2.3% 2.9%
comparison = L-CRT1 Conroy 2021	VS Inductio 0.2996	n + L-CRT 0.2699		1.35	[0.80; 2.29]	10.4%	10.4%
comparison = S-RT + c Jin 2022 Chakrabarti 2021 Common effect model Random effects model Heterogeneity: $J^2 = 0\%$, τ^2	onsolidatio -0.1085 0.0645 = 0, <i>p</i> = 0.75	n VS L-CRT1 0.2224 0.5034	+	0.90 1.07 0.92 0.92	[0.58; 1.39] [0.40; 2.86] [0.62; 1.38] [0.62; 1.38]	15.2% 3.0% 18.2%	15.2% 3.0% 18.2%
Common effect model Random effects model			0.1 0.5 1 2 10	1.06 1.06	[0.89; 1.25] [0.89; 1.25]	100.0% 	 100.0%

Heterogeneity: $l^2 = 0\%$, $\tau^2 < 0.0001$, p = 0.61Residual heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, p = 0.62Test for subgroup differences (common effect): $\chi_5^2 = 4.60$, df = 5 (p = 0.47) Test for subgroup differences (random effects): $\chi_5 = 4.60$, df = 5 (p = 0.47)

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6
1	Induction + L-CRT	1.09 (0.63 to 1.89)	0.74 (0.43 to 1.28)			
2	0.90 (0.55 to 1.48)	L-CRT + consolidation	1.87 (0.68 to 5.14)	-		
3	0.89 (0.54 to 1.47)	0.99 (0.53 to 1.83)	L-CRT1	0.88 (0.68 to 1.15)	1.08 (0.71 to 1.64)	0.72 (0.35 to 1.46)
4	0.79 (0.45 to 1.38)	0.87 (0.45 to 1.71)	0.88 (0.68 to 1.15)	L-CRT2		
5	0.97 (0.51 to 1.85)	1.07 (0.51 to 2.25)	1.08 (0.71 to 1.64)	1.22 (0.75 to 2.00)	S-RT + consolidation	
6	0.64 (0.27 to 1.53)	0.71 (0.28 to 1.82)	0.72 (0.35 to 1.46)	0.81 (0.38 to 1.73)	0.67 (0.29 to 1.52)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: other (Random Effect	vs 'L-CRT1 s Model)	RR	95%-CI
S-RTearly L-CRT2 L-CRT + consolidation S-RT + consolidation nduction + L-CRT		**	1.39 1.13 0.99 0.92 0.89	[0.68; 2.82] [0.87; 1.47] [0.53; 1.83] [0.61; 1.41] [0.54; 1.47]
0	.01 (0.5 1 2 4	Ļ	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau² \geq 0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

tau^2= 0.0058 ; tau= 0.0764

I^2= 5.98 % (0 % to 80.45 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 4.25
 4
 0.3727

 Within designs
 1.76
 3
 0.6239

 Between designs
 2.50
 1
 0.1142

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:L-CRT + consolidation 0.92 1 0.3386 L-CRT1:L-CRT2 0.74 1 0.3882 L-CRT1:S-RT + consolidation 0.10 1 0.7532

Between-designs Q statistic after detaching of single designs

Detached design Q df p-value

 $Induction + L\text{-}CRT\text{-}L\text{-}CRT + consolidation \ 0.00 \ 0 \qquad \text{--}$

L-CRT1:Induction + L-CRT 0.00 0 ---

L-CRT1:L-CRT + consolidation 0.00 0 ---

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 2.50 1 0.1142 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

 comparison k prop nma direct indir. RoR z p-value

 Induction + L-CRT:L-CRT + consolidation 1 0.81 0.90
 1.09
 0.40 2.75
 1.55
 0.1201

 L-CRT1:Induction + L-CRT 1 0.81 1.12
 1.35
 0.49 2.75
 1.55
 0.1201

 L-CRT1:L-CRT + consolidation 2 0.37 1.01
 0.54
 1.47 0.36 -1.55
 0.1201

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion - Estimated treatment effect (RR) in network meta-analysis nma - Estimated treatment effect (RR) derived from direct evidence direct - Estimated treatment effect (RR) derived from indirect evidence indir. - Ratio of Ratios (direct versus indirect) RoR - z-value of test for disagreement (direct versus indirect) z p-value - p-value of test for disagreement (direct versus indirect)

^{© 2024} Turri G et al. JAMA Network Open.

	P-score
S-RTearly	0.7915
L-CRT2	0.6713
L-CRT + consolidation	0.4601
L-CRT1	0.4368
S-RT + consolidation	0.3408
Induction + L-CRT	0.2994

eAppendix 16. Rate of anastomotic leak

Characteristics of the network

Number of treatments:

9

Number of studies:

17

Number of individuals included:

8333

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized
1	CHT	424
2	Induction + L-CRT	294
3	L-CRT + consolidation	126
4	L-CRT1	3752
5	L-CRT2	2481
6	L-RT	360
7	S-RT + consolidation	670
8	S-RTdelayed	68
9	S-RTearly	158

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study			Risk Ratio	PP	95%-CI	Weight	Weight (random)		
Study	IUGRA 3E	(logkk)	RISK Rauo		90 /0-01	(common)	(ranuoni)		
comparison = L-CRT2 V Rodel 2015 Gerard (2) 2010 Deng 2016 O'Connell 2014 Schmoll 2021 Common effect model Random effects model Heterogeneity: $I^2 = 0\%$, τ^2 :	VS L-CRT1 0.2484 -0.1823 -0.0705 0.1979 0.0510 = 0, <i>p</i> = 0.84	0.2267 0.4200 0.2431 0.3690 0.2307		1.28 0.83 0.93 1.22 1.05 1.08 1.08	[0.82; 2.00] [0.37; 1.90] [0.58; 1.50] [0.59; 2.51] [0.67; 1.65] [0.85; 1.36]	14.8% 4.3% 12.9% 5.6% 14.3% 52.0%	11.9% 5.2% 11.0% 6.3% 11.7% 46.2%		
comparison = L-CRT2 \ Deng 2016	/S CHT 0.8309	0.3272		2.30	[1.21; 4.36]	7.1%	7.6%		
comparison = CHT VS I Deng 2016	L-CRT1 -0.9014	0.3248		0.41	[0.21; 0.77]	7.2%	7.6%		
comparison = L-RT VS Gerard 2006	L-CRT1 -0.0028	0.3705		1.00	[0.48; 2.06]	5.5%	6.3%		
comparison = L-CRT1 V Ngan 2012	/S S-RTear -0.5533	ly 0.6166		0.58	[0.17; 1.93]	2.0%	2.7%		
comparison = L-CRT1 \ Latkauskas 2012	/S S-RTdel 0.1660	ayed 0.6492		1.18	[0.33; 4.21]	1.8%	2.5%		
comparison = Inductior Fernandez-Martos 2015	0.5328	1.2082		1.70	[0.16; 18.19]	0.5%	0.8%		
comparison = S-RT + co Bahadoer 2021	onsolidatio -0.2453	n VS L-CRT1 0.6015		0.78	[0.24; 2.54]	2.1%	2.8%		
$\label{eq:comparison} \begin{array}{l} \mbox{comparison} = \mbox{L-CRT1} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	/S L-CRT + -0.1671 0.0408 = 0, <i>p</i> = 0.92	consolidation 1.3993 1.3850		— 0.85 — 1.04 0.94 0.94	[0.05; 13.14] [0.07; 15.73] [0.14; 6.47] [0.14; 6.47]	0.4% 0.4% 0.8%	0.6% 0.6% 1.2%		
comparison = L-CRT1 V Marechal 2012 Conroy 2021 Common effect model Random effects model Heterogeneity: $J^2 = 0\%$, τ^2 :	/S Inductio -0.0364 0.1512 = 0, <i>p</i> = 0.89	n + L-CRT 1.3883 0.2799	+	— 0.96 1.16 1.15 1.15	[0.06; 14.65] [0.67; 2.01] [0.67; 1.98] [0.67; 1.98]	0.4% 9.7% 10.1%	0.6% 9.3% 9.9%		
comparison = S-RT + co Bujko (2) 2013	onsolidatio 0.4479	n VS L-CRT2 0.4400		1.57	[0.66; 3.71]	3.9%	4.8%		
comparison = L-CRT2 \ Wang 2019	/S L-CRT + 0.0728	consolidation 1.4013		1.08	[0.07; 16.76]	0.4%	0.6%		
comparison = L-CRT1 \ Mei 2023	/S CHT 0.0413	0.3428		1.04	[0.53; 2.04]	6.5%	7.1%		
Common effect model Random effects model				1.06 1.05	[0.89; 1.26] [0.85; 1.30]	100.0% 	 100.0%		
Heterogeneity: $l^2 = 1\%$, $\tau^2 = 0.0449$, $p = 0.45$ Residual heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, $p = 0.96$ Test for subgroup differences (common effect): $\chi_{12}^2 = 16.66$, df = 12 ($p = 0.16$) Test for subgroup differences (random effects): $\chi_{12}^2 = 16.66$, df = 12 ($p = 0.16$)									

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7	V8	V9
1	CHT			0.61 (0.38 to 0.97)	0.44 (0.23 to 0.83)				
2	0.69 (0.34 to 1.36)	Induction + L-CRT		0.87 (0.51 to 1.48)	1.70 (0.16 to 18.19)				
3	0.59 (0.11 to 3.04)	0.86 (0.16 to 4.55)	L-CRT + consolidation	1.06 (0.15 to 7.33)	0.93 (0.06 to 14.49)				
4	0.62 (0.39 to 0.96)	0.90 (0.53 to 1.52)	1.04 (0.21 to 5.07)	L-CRT1	0.93 (0.73 to 1.18)	1.00 (0.49 to 2.07)	1.28 (0.39 to 4.15)	1.18 (0.33 to 4.21)	0.58 (0.17 to 1.93)
5	0.57 (0.36 to 0.92)	0.83 (0.47 to 1.47)	0.97 (0.20 to 4.73)	0.93 (0.74 to 1.17)	L-CRT2		0.64 (0.27 to 1.51)		
6	0.62 (0.26 to 1.45)	0.90 (0.37 to 2.21)	1.05 (0.18 to 5.95)	1.00 (0.49 to 2.07)	1.08 (0.50 to 2.31)	L-RT	100 C		
7	0.48 (0.21 to 1.10)	0.70 (0.29 to 1.68)	0.81 (0.14 to 4.55)	0.78 (0.38 to 1.58)	0.83 (0.41 to 1.68)	0.77 (0.28 to 2.14)	S-RT + consolidation		
8	0.73 (0.19 to 2.80)	1.06 (0.27 to 4.20)	1.23 (0.16 to 9.36)	1.18 (0.33 to 4.21)	1.27 (0.35 to 4.63)	1.18 (0.27 to 5.09)	1.52 (0.35 to 6.54)	S-RTdelayed	
9	0.35 (0.10 to 1.28)	0.52 (0.14 to 1.93)	0.60 (0.08 to 4.38)	0.58 (0.17 to 1.93)	0.62 (0.18 to 2.12)	0.57 (0.14 to 2.35)	0.74 (0.18 to 3.01)	0.49 (0.08 to 2.82)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: othe (Random Effec	r vs 'L-CR' ts Model)	T1' RR	95%-CI
S-RTearly S-RT + consolidation L-CRT2 L-CRT + consolidation L-RT Induction + L-CRT S-RTdelayed CHT		******	→ 1.74 1.29 1.08 → 1.04 1.00 0.90 - 0.85 0.62	[0.52; 5.82] [0.63; 2.63] [0.85; 1.36] [0.21; 5.07] [0.48; 2.06] [0.53; 1.52] [0.24; 3.02] [0.39; 0.96]
0	.01	0.5 1 2	4	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0$; tau = 0

 $I^2 = 0 \% (0 \% \text{ to } 60.23 \%)$

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 5.88 10
 0.8251

 Within designs
 1.00
 5
 0.9624

 Between designs 4.88
 5
 0.4308

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:Induction + L-CRT 0.02 1 0.8947 L-CRT1:L-CRT + consolidation 0.01 1 0.9159 L-CRT1:L-CRT2 0.97 3 0.8077

Between-designs Q statistic after detaching of single designs

Detached design Q df p-value Induction + L-CRT:L-CRT2 4.51 4 0.3414 L-CRT + consolidation:L-CRT2 4.88 4 0.3000 L-CRT1:CHT 1.85 4 0.7641 L-CRT1:Induction + L-CRT 4.51 4 0.3414 L-CRT1:L-CRT + consolidation 4.88 4 0.3000 L-CRT1:L-CRT2 4.48 4 0.3454 L-CRT1:S-RT + consolidation 3.79 4 0.4346 L-CRT2:S-RT + consolidation 3.79 4 0.4346 L-CRT1:CHT:L-CRT2 1.54 3 0.6739

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 4.88 5 0.4308 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value L-CRT1:CHT 2 0.94 1.62 1.64 1.39 1.18 0.16 0.8712 CHT:L-CRT2 1 0.55 0.57 0.44 0.80 0.55 -1.24 0.2159 L-CRT1:Induction + L-CRT 2 0.95 1.11 1.15 0.54 2.13 0.61 0.5431 Induction + L-CRT:L-CRT2 1 0.06 0.83 1.70 0.80 2.13 0.61 0.5431 L-CRT1:L-CRT + consolidation 2 0.67 0.96 0.94 1.00 0.94 -0.04 0.9714 L-CRT + consolidation:L-CRT2 1 0.33 0.97 0.93 0.99 0.94 -0.04 0.9714 L-CRT1:L-CRT + consolidation 1 0.36 0.78 1.28 0.58 2.20 1.04 0.2975 L-CRT2:S-RT + consolidation 1 0.66 0.83 0.64 1.40 0.46 -1.04 0.2975

Legend:

k - Number of studies providing direct evidence prop - Direct evidence proportion

nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

	P-score
S-RTearly	0.7871
S-RT + consolidation	0.6921
L-CRT2	0.5900
L-CRT + consolidation	0.5152
L-RT	0.4965
L-CRT1	0.4893
S-RTdelayed	0.4042
Induction + L-CRT	0.4023
CHT	0.1233

eAppendix 17. Locoregional recurrence at 3 years

Characteristics of the network

Number of treatments:

7

Number of studies:

11

Number of individuals included:

5749

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized	
1	CHT	165	
2	Induction + L-CRT	387	
3	L-CRT + consolidation	210	
4	L-CRT1	2319	
5	L-CRT2	1486	
6	S-RT + consolidation	1021	
7	S-RTearly	161	

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	(logRR)	Risk Ratio	RR	95%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 ¹ Rodel 2015 Gerard (2) 2010 Deng 2016 Jiao 2015 Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, τ^2	VS L-CRT1 -0.4464 -0.3186 -0.1054 -0.1823 = 0.0139, <i>p</i> =	0.3043 - 0.3544 - 0.4461 - 0.5893 -	*	0.64 0.73 0.90 0.83 0.73 0.73	[0.35; 1.16] [0.36; 1.46] [0.38; 2.16] [0.26; 2.64] [0.50; 1.07] [0.49; 1.09]	11.8% 8.7% 5.5% 3.2% 29.3%	10.9% 8.9% 6.3% 3.9% 30.0%
comparison = L-CRT2 Deng 2016	VS CHT -0.2007	0.4358 -		0.82	[0.35; 1.92]	5.8%	6.5%
comparison = CHT VS Deng 2016	L-CRT1 0.0953	0.4228		1.10	[0.48; 2.52]	6.1%	6.8%
comparison = L-CRT1 Ngan 2012	VS S-RTearl -0.5452	y 0.4624 ——		0.58	[0.23; 1.43]	5.1%	5.9%
comparison = Induction Fokas 2019	n + L-CRT V 0.2485	S L-CRT + con 0.5278	solidation	1.28	[0.46; 3.61]	3.9%	4.8%
comparison = S-RT + c Bahadoer 2021 Jin 2022 Common effect model Random effects model Heterogeneity: $I^2 = 67\%$, τ^2	onsolidatior 0.5202 -0.2638 ² = 0.0139, p =	VS L-CRT1 0.3730 0.2539		1.68 0.77 0.98 1.00	[0.81; 3.50] [0.47; 1.26] [0.65; 1.49] [0.64; 1.57]	7.9% 17.0% 24.9% 	8.2% 13.5% 21.8%
comparison = L-CRT1 Conroy 2021	VS Induction 0.2667	1 + L-CRT 0.4102		1.31	[0.58; 2.92]	6.5%	7.1%
comparison = S-RT + c Bujko (2) 2013	onsolidatior 0.6378	0.2763		1.89	[1.10; 3.25]	14.4%	12.3%
comparison = L-CRT2 Wang 2019	VS L-CRT + 0.1542	consolidation 0.5255		1.17	[0.42; 3.27]	4.0%	4.8%
Common effect model Random effects model				1.00 0.99	[0.81; 1.22] [0.78; 1.27]	100.0% 	 100.0%
$\begin{array}{cccc} 0.5 & 1 & 2 \\ \text{Heterogeneity: } l^2 = 13\%, \tau^2 = 0.0514, p = 0.31 \\ \text{Residual heterogeneity: } l^2 = 0\%, \tau^2 = 0.0139, p = 0.48 \\ \text{Test for subgroup differences (common effect): } \gamma_{\pm}^2 = 10.37, \text{ df} = 8 (p = 0.24) \end{array}$							

Test for subgroup differences (common effect): $\chi_8 = 10.37$, df = 8 (p = 0.24) Test for subgroup differences (random effects): $\chi_8^2 = 9.06$, df = 8 (p = 0.34)

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7
1	CHT			1.10 (0.48 to 2.52)	1.22 (0.52 to 2.87)	-	
2	1.26 (0.46 to 3.46)	Induction + L-CRT	1.28 (0.46 to 3.61)	0.77 (0.34 to 1.71)			
3	1.64 (0.55 to 4.85)	1.30 (0.57 to 2.95)	L-CRT + consolidation		0.86 (0.31 to 2.40)		
4	0.96 (0.46 to 2.00)	0.76 (0.37 to 1.54)	0.59 (0.25 to 1.35)	L-CRT1	1.37 (0.94 to 2.00)	1.02 (0.67 to 1.53)	0.58 (0.23 to 1.43)
5	1.42 (0.68 to 2.98)	1.13 (0.53 to 2.38)	0.87 (0.38 to 1.97)	1.48 (1.07 to 2.05)	L-CRT2	0.53 (0.31 to 0.91)	
6	0.89 (0.40 to 1.95)	0.70 (0.32 to 1.53)	0.54 (0.22 to 1.30)	0.92 (0.65 to 1.31)	0.62 (0.42 to 0.92)	S-RT + consolidation	
7	0.56 (0.17 to 1.79)	0.44 (0.14 to 1.39)	0.34 (0.10 to 1.16)	0.58 (0.23 to 1.43)	0.39 (0.15 to 1.02)	0.63 (0.24 to 1.66)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: othe (Random Effe	er vs 'L-CRT1 cts Model)	RR	95%-CI
S-RTearly S-RT + consolidation CHT Induction + L-CRT L-CRT2 L-CRT + consolidation	n		1.72 1.08 0.96 0.76 0.67 0.59	[0.70; 4.27] [0.76; 1.53] [0.46; 2.00] [0.37; 1.54] [0.49; 0.93] [0.25; 1.35]
(0.0 <mark>1</mark>	0.5 1 2 4	Ļ	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0$; tau = 0

I^2= 0 % (0 % to 70.81 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 4.20
 6
 0.6495

 Within designs
 3.20
 3
 0.3613

 Between designs
 1.00
 3
 0.8018

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:L-CRT2 0.18 2 0.9117 L-CRT1:S-RT + consolidation 3.02 1 0.0823

Between-designs Q statistic after detaching of single designs

 $\begin{array}{c|c} & Detached \ design & Q \ df \ p-value \\ Induction + L-CRT:L-CRT + \ consolidation \ 1.00 & 2 & 0.6077 \\ & L-CRT + \ consolidation:L-CRT2 \ 1.00 & 2 & 0.6077 \\ & L-CRT1:Induction + L-CRT \ 1.00 & 2 & 0.6077 \\ & L-CRT1:L-CRT2 \ 0.95 & 2 & 0.6221 \\ & L-CRT1:S-RT + \ consolidation \ 0.28 & 2 & 0.8711 \\ & L-CRT2:S-RT + \ consolidation \ 0.28 & 2 & 0.8711 \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 0.88 3 0.8308 0.0945 0.0089

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

 $\begin{array}{c} {\rm comparison\ k\ prop\ nma\ direct\ indir.\ RoR\ z\ p-value}\\ {\rm L-CRT1:CHT\ 1\ 0.79\ 1.04\ 0.91\ 1.72\ 0.53\ -0.70\ 0.4852}\\ {\rm CHT:L-CRT2\ 1\ 0.75\ 1.42\ 1.22\ 2.24\ 0.55\ -0.70\ 0.4852}\\ {\rm Induction\ +\ L-CRT1:L-CRT\ +\ consolidation\ 1\ 0.63\ 1.30\ 1.28\ 1.33\ 0.97\ -0.04\ 0.9681\\ {\rm L-CRT1:Induction\ +\ L-CRT\ 1\ 0.78\ 1.32\ 1.31\ 1.35\ 0.97\ -0.04\ 0.9681\\ {\rm L-CRT\ +\ consolidation:L-CRT2\ 1\ 0.63\ 0.87\ 0.86\ 0.89\ 0.97\ -0.04\ 0.9681\\ {\rm L-CRT\ 1:L-CRT2\ 4\ 0.73\ 1.48\ 1.37\ 1.84\ 0.74\ -0.79\ 0.4299\\ {\rm L-CRT\ 1:S-RT\ +\ consolidation\ 2\ 0.72\ 0.92\ 1.02\ 0.73\ 1.40\ 0.85\ 0.3956\\ {\rm L-CRT\ 2:S-RT\ +\ consolidation\ 1\ 0.51\ 0.62\ 0.53\ 0.74\ 0.71\ -0.85\ 0.3956\end{array}$

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect)

z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

	P-score
S-RTearly	0.8989
S-RT + consolidation	0.6972
L-CRT1	0.6090
CHT	0.5527
Induction + L-CRT	0.3623
L-CRT2	0.2047
L-CRT + consolidation	0.1751

eAppendix 18. Locoregional recurrence at 5 years

Characteristics of the network

Number of treatments:

7

Number of studies:

7

Number of individuals included:

4886

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized
1	CHT	585
2	Induction + L-CRT	54
3	L-CRT1	2370
4	L-CRT2	887
5	L-RT	367
6	S-RT + consolidation	462
7	S-RTearly	161

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SI	E(logRR)	R	lisk Ratio	R	R 95%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 Gerard (2) 2010 Schmoll 2021 Common effect model Random effects model Heterogeneity: $J^2 = 0\%$, τ^2	VS L-CRT1 -0.1158 -0.3536	0.2740 0.2189		<u> </u>	0.8 0.7 0.7	9 [0.52; 1.52] 0 [0.46; 1.08] 7 [0.55; 1.08] 7 [0.55; 1.08]	16.0% 25.1% 41.1% 	17.1% 19.4% 36.6%
comparison = L-RT VS Gerard 2006	8 L-CRT1 0.7331	0.2232		-	2.0	8 [1.34; 3.22]	24.1%	19.3%
comparison = L-CRT1 Ngan 2012	VS S-RTear -0.2939	ly 0.4267	-		0.7	5 [0.32; 1.72]	6.6%	11.7%
comparison = Induction Fernandez-Martos 2015	on + L-CRT \ 1.0014	/SL-CRT2 1.1377	_		2.7	2 [0.29; 25.31]	0.9%	2.7%
comparison = S-RT + Bahadoer 2021	consolidatio 0.4998	n VS L-CRT 0.2383	1	-	1.6	5 [1.03; 2 .63]	21.2%	18.6%
comparison = L-CRT1 Schrag 2023	VS CHT -0.1262	0.4455	-		0.8	8 [0.37; 2.11]	6.1%	11.1%
Common effect model Random effects mode	1		Γ	- 	1.1 1.1	7 [0.94; 1.45] 3 [0.77; 1.67]	100.0% 	 100.0%
Heterogeneity: / ² = 65%, / Residual heterogeneity: / ² Test for subgroup differen	$t^2 = 0.1537, p$ $t^2 = 0\%, \tau^2 = 0,$	< 0.01 p = 0.50 effect): $\gamma_{r}^{2} = 1$	0.1 (16.77. df =	$0.5 \ 1 \ 2$ 5 (p < 0.01)	10			

Test for subgroup differences (common effects): $\chi_5^2 = 16.77$, di = 5 (p < 0.01) Test for subgroup differences (random effects): $\chi_5^2 = 16.77$, df = 5 (p < 0.01)

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



Netleague table Relative risks (RRs) and 95% confidence intervals (CIs) are reported. Results of the network metaanalysis are reported in the lower left part of the table and results from the pairwise meta-analysis are reported in the upper right part of the table. RRs higher than 1 favor the column-defining treatment.

	V1	V2	V3	V4	V5	V6	V7
1	CHT		1.13 (0.47 to 2.72)				
2	0.54 (0.05 to 6.07)	Induction + L-CRT		2.72 (0.29 to 25.31)			
3	1.13 (0.47 to 2.72)	2.10 (0.22 to 19.99)	L-CRT1	1.30 (0.93 to 1.82)	0.48 (0.31 to 0.74)	0.61 (0.38 to 0.97)	0.75 (0.32 to 1.72)
4	1.47 (0.58 to 3.75)	2.72 (0.29 to 25.31)	1.30 (0.93 to 1.82)	L-CRT2			
5	0.55 (0.21 to 1.45)	1.01 (0.10 to 10.02)	0.48 (0.31 to 0.74)	0.37 (0.21 to 0.64)	L-RT		
6	0.69 (0.26 to 1.85)	1.27 (0.13 to 12.72)	0.61 (0.38 to 0.97)	0.47 (0.26 to 0.83)	1.26 (0.67 to 2.39)	S-RT + consolidation	
7	0.85 (0.25 to 2.83)	1.56 (0.14 to 17.32)	0.75 (0.32 to 1.72)	0.57 (0.23 to 1.41)	1.55 (0.60 to 3.99)	1.23 (0.47 to 3.20)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: oth (Random Eff	ner vs 'L-CRT1 ects Model)	RR	95%-CI
Induction + L-CRT L-RT S-RT + consolidation S-RTearly CHT L-CRT2			2.10 2.08 1.65 1.34 1.13 0.77	[0.22; 19.99] [1.34; 3.22] [1.03; 2.63] [0.58; 3.10] [0.47; 2.72] [0.55; 1.08]
0.	01	0.5 1 2 4	Ļ	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0$; tau = 0

I^2= 0 % (NA % to NA %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

Q df p-valueTotal0.4610.4976Within designs0.4610.4976Between designs0.000--

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:L-CRT2 0.46 1 0.4976

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 0.00 0 -- 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k nma direct indir. RoR z p-value

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

		P-score
	L-RT	0.8278
9	S-RT + consolidation	0.6778
]	Induction + L-CRT	0.6612
	S-RTearly	0.5206
	CHT	0.4077
	L-CRT1	0.3081
	L-CRT2	0.0967

eAppendix 19. Locoregional failure at 3 years

Characteristics of the network

Number of treatments:

7

Number of studies:

10

Number of individuals included:

5234

Number of individuals randomized to each treatment:

Treatment name	N. individuals randomized	
CHT	165	
Induction + L-CRT	387	
L-CRT + consolidation	210	
L-CRT1	2319	
L-CRT2	1232	
S-RT + consolidation	760	
S-RTearly	161	
	Treatment name CHT Induction + L-CRT L-CRT + consolidation L-CRT1 L-CRT2 S-RT + consolidation S-RTearly	Treatment nameN. individuals randomizedCHT165Induction + L-CRT387L-CRT + consolidation210L-CRT12319L-CRT21232S-RT + consolidation760S-RTearly161

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	(logRR)	Risk Ratio	RR	95%-CI	Weight (common)	Weight (random)	
comparison = L-CRT2 V Rodel 2015 Gerard (2) 2010 Deng 2016 Jiao 2015 Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, τ^2	VS L-CRT1 -0.4538 -0.2163 -0.1942 -0.1823	0.2485 0.2884 0.3437 0.5893	***	0.64 0.81 0.82 0.83 0.74 0.74	[0.39; 1.03] [0.46; 1.42] [0.42; 1.62] [0.26; 2.64] [0.54; 1.01] [0.53; 1.02]	13.6% 10.1% 7.1% 2.4% 33.2%	11.7% 9.9% 7.9% 3.4% 32.9%	
comparison = L-CRT2 Deng 2016	/S CHT -0.3054	0.3346		0.74	[0.38; 1.42]	7.5%	8.2%	
comparison = CHT VS Deng 2016	L-CRT1 0.1112	0.3152		1.12	[0.60; 2.07]	8.4%	8.8%	
comparison = L-CRT1 Ngan 2012	/S S-RTear -0.5452	0.4624 —		0.58	[0.23; 1.43]	3.9%	5.1%	
comparison = Induction Fokas 2019	n + L-CRT V 0.2973	S L-CRT + con 0.3979	solidation	1.35	[0.62; 2.94]	5.3%	6.4%	
$\begin{array}{l} \mbox{comparison} = \mbox{S-RT} + \mbox{c}\\ \mbox{Bahadoer} \mbox{2021}\\ \mbox{Jin} \mbox{2022}\\ \mbox{Common effect model}\\ \mbox{Random effects model}\\ \mbox{Heterogeneity:} \mbox{I^2} = \mbox{60\%}, \mbox{τ^2} \end{array}$	0.2774 -0.2638 = 0.0077, p	0.2273 0.2539 = 0.11		1.32 0.77 1.04 1.03	[0.85; 2.06] [0.47; 1.26] [0.74; 1.45] [0.73; 1.47]	16.2% 13.0% 29.2%	12.8% 11.4% 24.2%	
comparison = L-CRT1 Conroy 2021	VS Inductio 0.8253	n + L-CRT 0.3496		2.28	[1.15; 4.53]	6.9%	7.7%	
comparison = L-CRT2 V Wang 2019	VS L-CRT + 0.4418	consolidation 0.3863		1.56	[0.73; 3.32]	5.6%	6.7%	
Common effect model Random effects model				0.97 0.98	[0.81; 1.16] [0.78; 1.23]	100.0% 	 100.0%	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$								

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.


	V1	V2	V3	V4	V5	V6	V7
1	CHT			1.12 (0.60 to 2.07)	1.36 (0.70 to 2.61)	-	
2	2.17 (0.98 to 4.81)	Induction + L-CRT	1.35 (0.62 to 2.94)	0.44 (0.22 to 0.87)			
3	2.57 (1.13 to 5.82)	1.18 (0.63 to 2.22)	L-CRT + consolidation		0.64 (0.30 to 1.37)		
4	1.05 (0.60 to 1.84)	0.48 (0.27 to 0.87)	0.41 (0.22 to 0.78)	L-CRT1	1.36 (0.99 to 1.86)	0.96 (0.69 to 1.34)	0.58 (0.23 to 1.43)
5	1.46 (0.83 to 2.57)	0.67 (0.36 to 1.25)	0.57 (0.31 to 1.06)	1.39 (1.02 to 1.88)	L-CRT2		
6	1.02 (0.53 to 1.94)	0.47 (0.24 to 0.92)	0.40 (0.19 to 0.81)	0.96 (0.69 to 1.34)	0.70 (0.44 to 1.09)	S-RT + consolidation	
7	0.61 (0.21 to 1.77)	0.28 (0.10 to 0.83)	0.24 (0.08 to 0.72)	0.58 (0.23 to 1.43)	0.42 (0.16 to 1.09)	0.60 (0.23 to 1.58)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: o (Random E	ther vs 'L-CRT1 ffects Model)	RR	95%-CI
S-RTearly CHT S-RT + consolidation L-CRT2 Induction + L-CRT L-CRT + consolidation	n		1.72 1.05 1.04 0.72 0.48 0.41	[0.70; 4.27] [0.60; 1.84] [0.74; 1.45] [0.53; 0.98] [0.27; 0.87] [0.22; 0.78]
(0.01	0.5 1 2 4	Ļ	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0$; tau = 0

I^2= 0 % (0 % to 74.62 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 3.43
 5
 0.6341

 Within designs
 2.99
 3
 0.3938

 Between designs
 0.44
 2
 0.8011

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:L-CRT2 0.47 2 0.7925 L-CRT1:S-RT + consolidation 2.52 1 0.1123

Between-designs Q statistic after detaching of single designs

Detached design Q df p-value Induction + L-CRT:L-CRT + consolidation 0.14 1 0.7129 L-CRT + consolidation:L-CRT2 0.14 1 0.7129 L-CRT1:Induction + L-CRT 0.14 1 0.7129 L-CRT1:L-CRT2 0.43 1 0.5110

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 0.44 2 0.8011 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value
L-CRT1:CHT 1 0.81 0.95 0.89 1.22 0.73 -0.43 0.6650
CHT:L-CRT2 1 0.74 1.46 1.36 1.81 0.75 -0.43 0.6650Induction + L-CRT:L-CRT + consolidation 1 0.65 1.18 1.35 0.93 1.45 0.56 0.5788
L-CRT1:Induction + L-CRT 1 0.73 2.06 2.28 1.57 1.45 0.56 0.5788
L-CRT + consolidation:L-CRT2 1 0.67 0.57 0.64 0.44 1.45 0.56 0.5788
L-CRT1:L-CRT2 4 0.94 1.39 1.36 1.98 0.69 -0.56 0.5788

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion - Estimated treatment effect (RR) in network meta-analysis nma direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) - z-value of test for disagreement (direct versus indirect) Z p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
S-RTearly	0.9159
S-RT + consolidation	0.6904
CHT	0.6898
L-CRT1	0.6555
L-CRT2	0.3439
Induction + L-CRT	0.1436
L-CRT + consolidation	0.0610

eAppendix 20. Locoregional failure at 5 years

Characteristics of the network

Number of treatments:

7

Number of studies:

6

Number of individuals included:

3792

Number of individuals randomized to each treatment:

Treatment name	N. individuals randomized
CHT	585
Induction + L-CRT	54
L-CRT1	1823
L-CRT2	340
L-RT	367
S-RT + consolidation	462
S-RTearly	161
	Treatment name CHT Induction + L-CRT L-CRT1 L-CRT2 L-RT S-RT + consolidation S-RTearly

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	(logRR)	Risk Rati	o RR	95%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 Gerard (2) 2010	VS L-CRT1 -0.1630	0.2479		0.85	[0.52; 1.38]	18.3%	20.6%
comparison = L-RTVS Gerard 2006	0.5281	0.1694	•	- 1.70	[1.22; 2.36]	39.1%	27.9%
comparison = L-CRT1 Ngan 2012	VS S-RTear -0.2939	ly 0.4267		0.75	[0.32; 1.72]	6.2%	10.6%
comparison = Inductio Fernandez-Martos 2015	on + L-CRT \ 1.1556	/S L-CRT2 0.7771			[0.69; 14.57]	1.9%	3.9%
comparison = S-RT + o Bahadoer 2021	consolidatio 0.3605	n VS L-CR 0.2056	.71	- 1.43	[0.96; 2.15]	26.6%	24.3%
comparison = L-CRT1 Schrag 2023	VS CHT -0.0686	0.3742		0.93	[0.45; 1.94]	8.0%	12.7%
Common effect model Random effects mode	I		, — , 	1.31 1.23	[1.06; 1.61] [0.90; 1.69]	100.0% 	 100.0%
Heterogeneity: I ² = 47%, 1 Residual heterogeneity: I ² Test for subgroup differen Test for subgroup differen	$c^2 = 0.0648, p$ = NA%, $\tau^2 = 0$ ces (common ces (random 6	= 0.09 0, p = NA effect): $\chi_{\frac{5}{2}}^2$ = effects): χ_5^2 =	0.1 0.5 1 2 = 9.43, df = 5 (p = 0.0 = 9.43, df = 5 (p = 0.0	2 10 9) 9)			

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



	V1	V2	V3	V4	V5	V6	V7
1	CHT		1.07 (0.51 to 2.23)				
2	0.40 (0.07 to 2.30)	Induction + L-CRT		3.18 (0.69 to 14.57)			
3	1.07 (0.51 to 2.23)	2.70 (0.55 to 13.35)	L-CRT1	1.18 (0.72 to 1.91)	0.59 (0.42 to 0.82)	0.70 (0.47 to 1.04)	0.75 (0.32 to 1.72)
4	1.26 (0.52 to 3.04)	3.18 (0.69 to 14.57)	1.18 (0.72 to 1.91)	L-CRT2			
5	0.63 (0.28 to 1.41)	1.59 (0.31 to 8.14)	0.59 (0.42 to 0.82)	0.50 (0.28 to 0.90)	L-RT		
6	0.75 (0.32 to 1.72)	1.88 (0.36 to 9.78)	0.70 (0.47 to 1.04)	0.59 (0.32 to 1.11)	1.18 (0.70 to 1.99)	S-RT + consolidation	
7	0.80 (0.26 to 2.43)	2.01 (0.33 to 12.22)	0.75 (0.32 to 1.72)	0.63 (0.24 to 1.67)	1.26 (0.51 to 3.11)	1.07 (0.42 to 2.70)	S-RTearly

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

C Treatment	Comparison: othe (Random Effec	r vs 'L-CRT1' ts Model)	RR	95%-CI
Induction + L-CRT L-RT S-RT + consolidation S-RTearly CHT L-CRT2			2.70 1.70 1.43 1.34 1.07 0.85	[0.55; 13.35] [1.22; 2.36] [0.96; 2.15] [0.58; 3.10] [0.51; 2.23] [0.52; 1.38]
0.0)1	0.5 1 2 4		

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

Tau^2= NA ; tau= NA

I^2= NA % (NA % to NA %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

Q df p-value

Total 0.00 0 --

Within designs 0.00 0 ---

Between designs 0.00 0 ---

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 0.00 0 -- 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k nma direct indir. RoR z p-value

Legend: comparison – Treatment comparison k - Number of studies providing direct evidence nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. – Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
Induction + L-CRT	0.8215
L-RT	0.7627
S-RT + consolidation	0.6180
S-Rtearly	0.5341
CHT	0.3577
L-CRT1	0.2617
L-CRT2	0.1444

eAppendix 21. Distant recurrence rate at 3 years

Characteristics of the network

Number of treatments:

5

Number of studies:

8

Number of individuals included:

4811

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized
1	Induction + L-CRT	387
2	L-CRT + consolidation	150
3	L-CRT1	1992
4	L-CRT2	1261
5	S-RT + consolidation	1021

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	(logRR)	Risk Ratio	RR	95%-CI	(common)	(random)
comparison = L-CRT2 Rodel 2015 Gerard (2) 2010 Jiao 2015 Common effect model Random effects mode Heterogeneity: I^2 = 1%, τ^2	VS L-CRT1 -0.1981 -0.0940 -0.5341 = < 0.0001, <i>p</i>	0.1129 0.1483 0.2718	the state of the s	0.82 0.91 0.59 0.82 0.82	[0.66; 1.02] [0.68; 1.22] [0.34; 1.00] [0.69; 0.97] [0.69; 0.97]	22.1% 12.8% 3.8% 38.8%	17.2% 13.5% 6.2% 36.9%
comparison = Inductio Fokas 2019	n + L-CRT V 0.1149	S L-CRT + cons 0.2536	olidation	1.12	[0.68; 1.84]	4.4%	6.9%
comparison = S-RT + σ Bahadoer 2021 Jin 2022 Common effect model Random effects model Heterogeneity: $J^2 = 29\%$, σ	-0.3003 -0.0741 -0.0001,	n VS L-CRT1 0.1214 0.1477 p = 0.24	anand and a state and	0.74 0.93 0.81 0.81	[0.58; 0.94] [0.70; 1.24] [0.68; 0.98] [0.68; 0.98]	19.1% 12.9% 32.1%	16.3% 13.6% 29.8%
comparison = L-CRT1 Conroy 2021	VS Inductio 0.2714	n + L-CRT 0.1654		1.31	[0.95; 1.81]	10.3%	12.1%
comparison = S-RT + c Bujko (2) 2013	onsolidatio 0.0954	n VS L-CRT2 0.1398		1.10	[0.84; 1.45]	14.4%	14.3%
Common effect model Random effects mode	I	Г		0.91 0.92	[0.82; 1.01] [0.79; 1.07]	100.0% 	 100.0%
Heterogeneity: $l^2 = 49\%$, τ Residual heterogeneity: l^2 Test for subgroup differen Test for subgroup differen	$r^{2} = 0.0214, p$ = 12%, $\tau^{2} < 0$ ces (common ces (random 6	0.5 = 0.06 .0001, p = 0.33 effect): χ_4^2 = 10.35 effects): χ_4^2 = 10.35	1 2 , df = 4 ($p = 0.03$) , df = 4 ($p = 0.03$)				

Weight Weight

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



	V1	V2	V3	VA	V5
	••	12	••	••	••
1	Induction + L-CRT	1.12 (0.68 to 1.84)	0.76 (0.55 to 1.05)		
2	1.12 (0.68 to 1.84)	L-CRT + consolidation			
3	0.76 (0.55 to 1.05)	0.68 (0.38 to 1.23)	L-CRT1	1.22 (1.03 to 1.44)	1.23 (1.03 to 1.48)
4	0.95 (0.66 to 1.36)	0.85 (0.46 to 1.56)	1.24 (1.07 to 1.44)	L-CRT2	0.91 (0.69 to 1.20)
5	0.91 (0.64 to 1.31)	0.82 (0.44 to 1.51)	1.20 (1.02 to 1.41)	0.96 (0.80 to 1.16)	S-RT + consolidation

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: othe (Random Effec	er vs 'L-0 cts Mode	CRT1 9)	RR	95%-CI
S-RT + consolidation L-CRT2 Induction + L-CRT L-CRT + consolidation	·			0.83 0.80 0.76 0.68	[0.71; 0.98] [0.69; 0.93] [0.55; 1.05] [0.38; 1.23]
0	.01	0.5 1	2 4	1	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0$; tau = 0

I^2= 0 % (0 % to 79.2 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 3.75
 4
 0.4415

 Within designs
 3.42
 3
 0.3313

 Between designs
 0.33
 1
 0.5680

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:L-CRT2 2.02 2 0.3641 L-CRT1:S-RT + consolidation 1.40 1 0.2368

Between-designs Q statistic after detaching of single designs

Detached design Q df p-value L-CRT1:L-CRT2 0.00 0 --L-CRT1:S-RT + consolidation 0.00 0 --

L-CRT2:S-RT + consolidation 0.00 0 --

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 0.24 1 0.6210 0.0566 0.0032

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-valueL-CRT1:L-CRT2 3 0.80 1.241.221.36 0.90 -0.570.5680L-CRT1:S-RT + consolidation 2 0.75 1.201.231.11 1.110.570.5680L-CRT2:S-RT + consolidation 1 0.45 0.960.911.01 0.90 -0.570.5680

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion - Estimated treatment effect (RR) in network meta-analysis nma - Estimated treatment effect (RR) derived from direct evidence direct - Estimated treatment effect (RR) derived from indirect evidence indir. - Ratio of Ratios (direct versus indirect) RoR - z-value of test for disagreement (direct versus indirect) z p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

 L-CRT1
 P-score

 S-RT + consolidation
 0.9585

 L-CRT2
 0.4174

Induction + L-CRT 0.3565 L-CRT + consolidation 0.2449

eAppendix 22. Distant recurrence rate at 5 years

Characteristics of the network

Number of treatments:

5

Number of studies:

5

Number of individuals included:

3016

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized
1	Induction + L-CRT	54
2	L-CRT1	1452
3	L-CRT2	887
4	S-RT + consolidation	462
5	S-RTearly	161

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SI	E(logRR)	Risk Ratio	RR	95%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 Gerard (2) 2010 Schmoll 2021 Common effect model Random effects mode Heterogeneity: $I^2 = 0\%$, τ^2	VS L-CRT1 -0.0646 -0.1082 = 0, <i>p</i> = 0.81	0.1303 0.1201		0.94 0.90 0.92 0.92	[0.73; 1.21] [0.71; 1.14] [0.77; 1.09] [0.77; 1.09]	23.7% 27.9% 51.6%	23.8% 27.8% 51.6%
comparison = L-CRT1 Ngan 2012	VS S-RTea 0.0808	rly 0.1766		1.08	[0.77; 1.53]	12.9%	13.3%
comparison = Inductio Fernandez-Martos 2015	n + L-CRT 0.0699	/SL-CRT2 0.3590 —		— 1.07	[0.53; 2.17]	3.1%	3.3%
comparison = S-RT + c Bahadoer 2021	onsolidatio -0.2682	on VS L-CRT1 0.1116		0.76	[0.61; 0.95]	32.3%	31.8%
Common effect model Random effects mode	I			0.89 0.89	[0.78; 1.00] [0.78; 1.01]	100.0% 	 100.0%
Heterogeneity: $I^2 = 0\%$, τ^2 Residual heterogeneity: I^2 Test for subgroup differen Test for subgroup differen	= 0.0010, p = 0%, τ^2 = 0, ces (commor ces (random	0.5 = 0.47 p = 0.81 n effect): $\chi_2^2 = 3.4$ effects): $\chi_3^2 = 3.4$	7, df = 3 (<i>p</i> = 0.33) 7, df = 3 (<i>p</i> = 0.33)	2			

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



	V1	V2	V3	V4	V 5
1	Induction + L-CRT		1.07 (0.53 to 2.17)		
2	0.98 (0.48 to 2.03)	L-CRT1	1.09 (0.92 to 1.30)	1.08 (0.77 to 1.53)	1.31 (1.05 to 1.63)
3	1.07 (0.53 to 2.17)	1.09 (0.92 to 1.30)	L-CRT2		
4	1.06 (0.48 to 2.38)	1.08 (0.77 to 1.53)	0.99 (0.67 to 1.46)	S-RTearly	1. Sec. 1. Sec
5	1.28 (0.60 to 2.74)	1.31 (1.05 to 1.63)	1.20 (0.91 to 1.58)	1.21 (0.80 to 1.82)	S-RT + consolidation

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: othe (Random Effe	er vs 'L-CR' cts Model)	T1' RR	95%-CI
Induction + L-CRT S-RTearly L-CRT2 S-RT + consolidation	I		0.98 0.92 0.92 0.76	[0.48; 2.03] [0.65; 1.30] [0.77; 1.09] [0.61; 0.95]
0	.01	0.5 1 2	4	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0$; tau = 0

I^2= 0 % (NA % to NA %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 0.06
 1
 0.8056

 Within designs
 0.06
 1
 0.8056

 Between designs
 0.00
 0
 -

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:L-CRT2 0.06 1 0.8056

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 0.00 0 -- 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k nma direct indir. RoR z p-value

Legend: comparison - Treatment comparison k - Number of studies providing direct evidence nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
L-CRT1	0.7573
Induction + L-CRT	0.5898
S-RTearly	0.5232
L-CRT2	0.4910
S-RT + consolidation	0.1387

eAppendix 23. Disease-Free Survival at 3 years

Characteristics of the network

Number of treatments:

7

Number of studies:

13

Number of individuals included:

7409

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized	
1	CHT	165	
2	Induction + L-CRT	387	
3	L-CRT + consolidation	210	
4	L-CRT1	3156	
5	L-CRT2	2395	
6	S-RT + consolidation	1021	
7	S-RTdelayed	75	

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	(logRR)	Risk Ratio	RR	95%-CI	Weight (common)	Weight (random)	
comparison = L-CRT2 V Rodel 2015 Aschele 2011 Gerard (2) 2010 Deng 2016 Schmoll 2021 Jiao 2015 Common effect model Random effects model Heterogeneity: $J^2 = 0\%$, $\tau^2 = 0\%$	/S L-CRT1 0.0624 0.0518 0.0735 0.0567 -0.0072 0.1422	0.0342 0.0454 0.0518 0.0639 0.0338 0.0807 = 0.49		1.06 1.05 1.08 1.06 0.99 1.15 1.05 1.05	[1.00; 1.14] [0.96; 1.15] [0.97; 1.19] [0.93; 1.20] [0.93; 1.06] [0.98; 1.35] [1.01; 1.08] [1.01; 1.09]	16.8% 9.5% 7.3% 4.8% 17.1% 3.0% 58.6%	14.8% 9.6% 7.7% 5.3% 15.0% 3.5%	
comparison = L-CRT2 V Deng 2016	/S CHT 0.0484	0.0634		1.05	[0.93; 1.19]	4.9%	5.4%	
comparison = CHT VS Deng 2016	L-CRT1 0.0083	0.0669	<u> </u>	1.01	[0.88; 1.15]	4.4%	4.9%	
comparison = L-CRT1 V Latkauskas 2012	/S S-RTdel 0.2412	ayed 0.1180			[1.01; 1.60]	1.4%	1.7%	
comparison = Induction Fokas 2019	1 + L-CRT V -0.0035	S L-CRT + c 0.0692	onsolidation	1.00	[0.87; 1.14]	4.1%	4.6%	
comparison = S-RT + c Bahadoer 2021 Jin 2022 Common effect model Random effects model Heterogeneity: $J^2 = 0\%$, τ^2 :	0.0911 0.0366 = < 0.0001, <i>p</i>	n VS L-CRT1 0.0406 0.0627 = 0.47		1.10 1.04 1.08 1.08	[1.01; 1.19] [0.92; 1.17] [1.01; 1.15] [1.01; 1.15]	11.9% 5.0% 16.9% 	11.4% 5.5% 16.9%	
comparison = L-CRT1 V Conroy 2021	/S Inductio -0.0972	n + L-CRT 0.0574	<u> </u>	0.91	[0.81; 1.02]	6.0%	6.4%	
comparison = S-RT + c Bujko (2) 2013	onsolidatio 0.0173	n VS L-CRT2 0.0840		1.02	[0.86; 1 .20]	2.8%	3.2%	
comparison = L-CRT2 V Wang 2019	/S L-CRT + -0.1872	consolidatio 0.1431 —	n	0.83	[0.63; 1.10]	1.0%	1.2%	
Common effect model Random effects model				1.04 1.04	[1.01; 1.07] [1.01; 1.07]	100.0% 	 100.0%	
Heterogeneity: $l^2 = 22\%$, $\tau^2 = 0.0005$, $p = 0.21$ Residual heterogeneity: $l^2 = 0\%$, $\tau^2 < 0.0001$, $p = 0.55$								

Test for subgroup differences (common effect): $\chi_g^2 = 12.98$, df = 8 (p = 0.11) Test for subgroup differences (random effects): $\chi_g^2 = 12.77$, df = 8 (p = 0.12)

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



	V1	V2	V3	V4	V5	V6	V7
1	CHT			1.01 (0.88 to 1.15)	0.95 (0.84 to 1.08)	-	
2	0.90 (0.77 to 1.04)	Induction + L-CRT	1.00 (0.87 to 1.14)	1.10 (0.98 to 1.23)			
3	0.87 (0.72 to 1.05)	0.97 (0.86 to 1.10)	L-CRT + consolidation		1.21 (0.91 to 1.60)		
4	1.00 (0.89 to 1.12)	1.12 (1.01 to 1.24)	1.15 (0.99 to 1.33)	L-CRT1	0.96 (0.92 to 0.99)	0.93 (0.87 to 0.99)	1.27 (1.01 to 1.60)
5	0.96 (0.86 to 1.07)	1.07 (0.96 to 1.20)	1.10 (0.94 to 1.28)	0.96 (0.92 to 0.99)	L-CRT2	0.98 (0.83 to 1.16)	
6	0.93 (0.82 to 1.06)	1.04 (0.92 to 1.18)	1.07 (0.91 to 1.25)	0.93 (0.87 to 0.99)	0.97 (0.91 to 1.04)	S-RT + consolidation	
7	1.27 (0.99 to 1.65)	1.42 (1.10 to 1.84)	1.46 (1.11 to 1.92)	1.27 (1.01 to 1.60)	1.33 (1.05 to 1.68)	1.37 (1.08 to 1.74)	S-RTdelayed

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: other vs (Random Effects	s 'L-CRT' Model)	1' RR	95%-CI
L-CRT + consolidation Induction + L-CRT S-RT + consolidation L-CRT2 CHT S-RTdelayed	r		1.15 1.12 1.08 1.05 1.00 0.79	[0.99; 1.33] [1.01; 1.24] [1.01; 1.14] [1.01; 1.08] [0.89; 1.12] [0.62; 0.99]
0	.01 0.5	1 2	4	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0$; tau = 0

 $I^2 = 0 \% (0 \% \text{ to } 64.8 \%)$

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 5.60
 8
 0.6917

 Within designs
 4.94
 5
 0.4228

 Between designs
 0.66
 3
 0.8829

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:L-CRT2 4.41 4 0.3532 L-CRT1:S-RT + consolidation 0.53 1 0.4654

Between-designs Q statistic after detaching of single designs

 $\begin{array}{c|c} Detached \ design & Q \ df \ p-value \\ Induction + L-CRT:L-CRT + \ consolidation \ 0.05 & 2 & 0.9742 \\ L-CRT + \ consolidation:L-CRT2 \ 0.05 & 2 & 0.9742 \\ L-CRT1:Induction + L-CRT \ 0.05 & 2 & 0.9742 \\ L-CRT1:L-CRT2 \ 0.66 & 2 & 0.7195 \\ L-CRT1:S-RT + \ consolidation \ 0.64 & 2 & 0.7276 \\ L-CRT2:S-RT + \ consolidation \ 0.64 & 2 & 0.7276 \end{array}$

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 0.66 3 0.8829 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

 $\begin{array}{c} {\rm comparison \ k \ prop \ nma \ direct \ indir. \ RoR \ z \ p-value} \\ {\rm L-CRT1:CHT \ l \ 0.74 \ 1.00} & 0.99 & 1.02 \ 0.97 \ -0.20 & 0.8408 \\ {\rm CHT:L-CRT2 \ l \ 0.82 \ 0.96} & 0.95 & 0.98 \ 0.97 \ -0.20 & 0.8408 \\ {\rm Induction \ + \ L-CRT:L-CRT \ + \ consolidation \ l \ 0.83 \ 0.97 & 1.00 & 0.87 \ 1.14 & 0.78 & 0.4362 \\ {\rm L-CRT1:Induction \ + \ L-CRT \ l \ 0.89 \ 0.89 & 0.91 & 0.79 \ 1.14 & 0.78 & 0.4362 \\ {\rm L-CRT1:Induction \ + \ L-CRT2 \ l \ 0.29 \ 1.10 & 1.21 & 1.06 \ 1.14 \ 0.78 & 0.4362 \\ {\rm L-CRT1:L-CRT2 \ 6 \ 0.95 \ 0.96 & 0.96 & 0.97 \ 0.98 \ -0.24 & 0.8087 \\ {\rm L-CRT1:S-RT \ + \ consolidation \ 2 \ 0.86 \ 0.93 & 0.93 & 0.94 \ 0.99 \ -0.15 & 0.8804 \\ {\rm L-CRT2:S-RT \ + \ consolidation \ 1 \ 0.17 \ 0.97 & 0.98 & 0.97 \ 1.01 \ 0.15 & 0.8804 \\ \end{array}$

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect)

z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
L-CRT + consolidatio	n 0.8687
Induction + L-CRT	0.8106
S-RT + consolidation	n 0.6872
L-CRT2	0.5319
CHT	0.3326
L-CRT1	0.2568
S-RTdelayed	0.0121

eAppendix 24. Disease-Free Survival at 5 years

Characteristics of the network

Number of treatments:

6

Number of studies:

7

Number of individuals included:

5302

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized
1	CHT	585
2	Induction + L-CRT	54
3	L-CRT1	2585
4	L-CRT2	1249
5	L-RT	367
6	S-RT + consolidation	462

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	(logRR)	I	Risk Ratio	F	RR 95%	%-CI	Weight (common)	Weight (random)
comparison = L-CRT2 Aschele 2011 Gerard (2) 2010 Schmoll 2021 Common effect model Random effects model Heterogeneity: $I^2 = 0\%$, τ^2	VS L-CRT1 0.0446 0.0534 -0.0103 = 0, <i>p</i> = 0.56	0.0507 0.0591 0.0387			- 1. - 1. 0. 1. 1.	05 [0.95; 05 [0.94; 99 [0.92; 02 [0.97; 02 [0.97;	1.15] 1.18] 1.07] 1.08] 1.08]	12.0% 8.8% 20.5% 41.3%	13.9% 11.0% 19.9% 44.8%
comparison = L-RTVS Gerard 2006	L-CRT1 -0.0675	0.0632			0.	93 [0.83; 1	1.06]	7.7%	9.9%
comparison = Inductio Fernandez-Martos 2015	n + L-CRT V -0.0383	/S L-CRT2 0.1412			— 0.	96 [0.73; ⁻	1.27]	1.5%	2.4%
comparison = S-RT + c Bahadoer 2021	onsolidatio 0.0911	n VS L-CRT1 0.0444	1		- 1.	10 [1.00; 1	1.20]	15.6%	16.7%
comparison = L-CRT1 Schrag 2023	VS CHT -0.0278	0.0301			0.	97 [0.92; ⁻	1.03]	33.9%	26.2%
Common effect model Random effects model	l				1. 1.	01 [0.97; 1 01 [0.97; 1	1.04] 1.06]	100.0% 	 100.0%
Heterogeneity: $l^2 = 23\% \tau^2$	$^{2} = 0.0010 p$	= 0.25	0.8	1	1.25				

Residual heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, p = 0.56Test for subgroup differences (common effect): $\chi_4^2 = 6.61$, df = 4 (p = 0.16) Test for subgroup differences (random effects): $\chi_4^2 = 6.61$, df = 4 (p = 0.16)

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



	V1	V2	V3	V4	V5	V6
1	CHT		1.03 (0.97 to 1.09)			
2	1.05 (0.79 to 1.40)	Induction + L-CRT		0.96 (0.73 to 1.27)		
3	1.03 (0.97 to 1.09)	0.98 (0.74 to 1.30)	L-CRT1	0.98 (0.93 to 1.03)	1.07 (0.95 to 1.21)	0.91 (0.84 to 1.00)
4	1.01 (0.93 to 1.09)	0.96 (0.73 to 1.27)	0.98 (0.93 to 1.03)	L-CRT2		
5	1.10 (0.96 to 1.26)	1.05 (0.77 to 1.43)	1.07 (0.95 to 1.21)	1.09 (0.95 to 1.25)	L-RT	
6	0.94 (0.84 to 1.04)	0.90 (0.67 to 1.20)	0.91 (0.84 to 1.00)	0.93 (0.84 to 1.03)	0.85 (0.73 to 0.99)	S-RT + consolidation

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: othe (Random Effec	r vs 'L- ts Mod	CR1 el)	[1' RF	R 95%-CI
S-RT + consolidation CHT L-CRT2 Induction + L-CRT L-RT	·		1	1.10 1.03 1.02 0.98 	0 [1.00; 1.20] 3 [0.97; 1.09] 2 [0.97; 1.08] 3 [0.74; 1.30] 3 [0.83; 1.06]
C	.01	0.5 1	2	4	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

 $tau^2 = 0$; tau = 0

I^2= 0 % (0 % to 89.6 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

Q df p-valueTotal1.1720.5581Within designs1.1720.5581Between designs0.000--

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:L-CRT2 1.17 2 0.5581

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 0.00 0 -- 0 0

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k nma direct indir. RoR z p-value

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
S-RT + consolidation	0.9049
CHT	0.6128
L-CRT2	0.5523
Induction + L-CRT	0.4135
L-CRT1	0.3698
L-RT	0.1465

eAppendix 25. Overall survival at 3 years

Characteristics of the network

Number of treatments:

7

Number of studies:

11

Number of individuals included:

5576

Number of individuals randomized to each treatment:

Treatment name	N. individuals randomized	
CHT	165	
Induction + L-CRT	387	
L-CRT + consolidation	210	
L-CRT1	2232	
L-CRT2	1486	
S-RT + consolidation	1021	
S-RTdelayed	75	
	Treatment name CHT Induction + L-CRT L-CRT + consolidation L-CRT1 L-CRT2 S-RT + consolidation S-RTdelayed	Treatment nameN. individuals randomizedCHT165Induction + L-CRT387L-CRT + consolidation210L-CRT12232L-CRT21486S-RT + consolidation1021S-RTdelayed75

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	(logRR)	Risk Ratio	RR	95%-CI	Weight (common)	Weight (random)	
comparison = L-CRT2 V Rodel 2015 Gerard (2) 2010 Deng 2016 Jiao 2015 Common effect model Random effects model Heterogeneity: $J^2 = 0\%$, $\tau^2 = 0\%$	/S L-CRT1 0.0089 0.0145 -0.0202 0.0440	0.0207 0.0275 0.0367 0.0507		1.01 1.01 0.98 1.04 1.01 1.01	[0.97; 1.05] [0.96; 1.07] [0.91; 1.05] [0.95; 1.15] [0.98; 1.04] [0.96; 1.06]	20.8% 11.8% 6.6% 3.5% 42.7%	14.2% 11.0% 7.8% 4.9% 38.0%	
comparison = L-CRT2 \ Deng 2016	/S CHT -0.0202	0.0367		0.98	[0.91; 1.05]	6.6%	7.8%	
comparison = CHT VS I Deng 2016	L-CRT1 0.0000	0.0348		1.00	[0.93; 1.07]	7.3%	8.4%	
comparison = L-CRT1 \ Latkauskas 2012	/S S-RTdela 0.0667	o.0819		1.07	[0.91; 1.26]	1.3%	2.2%	
comparison = Inductior Fokas 2019	1 + L-CRTV 0.0033	S L-CRT + consol 0.0334	idation	1.00	[0.94; 1.07]	8.0%	8.9%	
$\begin{array}{l} \mbox{comparison} = \mbox{S-RT} + \mbox{cm}\\ \mbox{Bahadoer} \mbox{2021}\\ \mbox{Jin} \mbox{2022}\\ \mbox{Common effect model}\\ \mbox{Random effects model}\\ \mbox{Heterogeneity:} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	0.0032 0.1424 = 0.0017, p <	US L-CRT1 0.0232 0.0407	+	1.00 1.15 1.04 1.06	[0.96; 1.05] [1.06; 1.25] [1.00; 1.08] [0.99; 1.14]	16.5% 5.4% 21.8% 	12.9% 6.8% 19.8%	
comparison = L-CRT1 \ Conroy 2021	/S Inductior -0.0345	0.0322		0.97	[0.91; 1.03]	<mark>8</mark> .6%	9.3%	
comparison = S-RT + co Bujko (2) 2013	onsolidation 0.1191	0.0594	*	1.13	[1.00; 1.27]	2.5%	3.8%	
comparison = L-CRT2 \ Wang 2019	/S L-CRT + 0 -0.1636	consolidation 0.0881 ——+		0.85	[0.71; 1.01]	1.1%	1.9%	
Common effect model Random effects model				1.01 1.01	[0.99; 1.03] [0.99; 1.04]	100.0% 	 100.0%	
Heterogeneity: $l^2 = 46\%$, $\tau^2 = 0.0007$, $p = 0.03$ Residual heterogeneity: $l^2 = 60\%$, $\tau^2 = 0.0017$, $p = 0.04$ Test for subgroup differences (common effect): $\chi_8^2 = 12.32$, df = 8 ($p = 0.14$) Test for subgroup differences (random effects): $\chi_8^2 = 8.71$, df = 8 ($p = 0.37$)								

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



	V1	V2	V3	V4	V5	V6	V7
1	CHT			1.00 (0.90 to 1.11)	1.02 (0.92 to 1.14)		
2	0.95 (0.84 to 1.08)	Induction + L-CRT	1.00 (0.91 to 1.11)	1.04 (0.94 to 1.14)			
3	0.93 (0.81 to 1.08)	0.98 (0.89 to 1.07)	L-CRT + consolidation		1.18 (0.97 to 1.42)		
4	1.01 (0.92 to 1.11)	1.06 (0.97 to 1.16)	1.08 (0.96 to 1.21)	L-CRT1	0.99 (0.94 to 1.04)	0.94 (0.88 to 1.01)	1.07 (0.89 to 1.28)
5	1.01 (0.92 to 1.11)	1.06 (0.96 to 1.17)	1.08 (0.96 to 1.22)	1.00 (0.96 to 1.05)	L-CRT2	0.89 (0.77 to 1.02)	
6	0.94 (0.84 to 1.05)	0.99 (0.88 to 1.10)	1.01 (0.88 to 1.15)	0.93 (0.87 to 0.99)	0.93 (0.86 to 1.00)	S-RT + consolidation	
7	1.08 (0.88 to 1.32)	1.13 (0.93 to 1.38)	1.16 (0.93 to 1.43)	1.07 (0.89 to 1.28)	1.07 (0.89 to 1.28)	1.15 (0.95 to 1.39)	S-RTdelayed

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: other v (Random Effects	s 'L- Mod	CR1 el)	F1'	RR	95%-CI
L-CRT + consolidation S-RT + consolidation Induction + L-CRT CHT L-CRT2 S-RTdelayed	·				1.08 1.07 1.06 1.01 1.00 0.94	[0.96; 1.21] [1.01; 1.14] [0.97; 1.16] [0.92; 1.11] [0.95; 1.05] [0.78; 1.12]
0	.01 0.	51	2	4		

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

tau^2= 0.0016 ; tau= 0.0401

I^2= 56.51 % (0 % to 81.28 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

 Q df p-value

 Total
 13.80
 6
 0.0320

 Within designs
 9.24
 3
 0.0262

 Between designs
 4.55
 3
 0.2074

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:L-CRT2 0.41 2 0.8140 L-CRT1:S-RT + consolidation 8.83 1 0.0030

Between-designs Q statistic after detaching of single designs

Detached design Q df p-value Induction + L-CRT:L-CRT + consolidation 2.71 2 0.2584 L-CRT + consolidation:L-CRT2 2.71 2 0.2584 L-CRT1:Induction + L-CRT 2.71 2 0.2584 L-CRT1:L-CRT2 1.57 2 0.4564 L-CRT1:S-RT + consolidation 2.71 2 0.2583 L-CRT2:S-RT + consolidation 2.71 2 0.2583

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 2.06 3 0.5604 0.0446 0.0020

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k prop nma direct indir. RoR z p-value
L-CRT1:CHT 1 0.81 0.99 1.00 0.96 1.05 0.36 0.7159
CHT:L-CRT2 1 0.78 1.01 1.02 0.98 1.04 0.36 0.7159Induction + L-CRT:L-CRT + consolidation 1 0.82 0.98 1.00 0.88 1.14 1.09 0.2743
L-CRT1:Induction + L-CRT 1 0.83 0.94 0.97 0.84 1.14 1.09 0.2743
L-CRT + consolidation:L-CRT2 1 0.39 1.08 1.18 1.03 1.14 1.09 0.2743
L-CRT1:L-CRT2 4 0.87 1.00 0.99 1.09 0.91 -1.28 0.1998
L-CRT1:S-RT + consolidation 2 0.82 0.93 0.94 0.88 1.07 0.76 0.4468
L-CRT2:S-RT + consolidation 1 0.27 0.93 0.89 0.95 0.94 -0.76 0.4468

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence prop - Direct evidence proportion nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect)

z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
S-RT + consolidation	0.7965
L-CRT + consolidation	0.7953
Induction + L-CRT	0.6967
CHT	0.4112
L-CRT1	0.3256
L-CRT2	0.3090
S-RTdelayed	0.1657

eAppendix 26. Overall survival at 5 years

Characteristics of the network

Number of treatments:

7

Number of studies:

8

Number of individuals included:

5625

Number of individuals randomized to each treatment:

	Treatment name	N. individuals randomized
1	CHT	585
2	Induction + L-CRT	54
3	L-CRT1	2747
4	L-CRT2	1249
5	L-RT	367
6	S-RT + consolidation	462
7	S-RTearly	161

Pairwise meta-analysis

Risk ratios above 1 favor the first treatment of the comparison.

Study	logRR SE	(logRR)	Risk Ratio	RR	95%-CI	Weight (common) (Weight random)	
comparison = L-CRT2 Aschele 2011 Gerard (2) 2010 Schmoll 2021 Common effect model Random effects model Heterogeneity: / ² = 80%, τ	VS L-CRT1 0.0338 0.1236 -0.0359 ² = 0.0050, <i>p</i>	0.0379 0.0427 0.0288 < 0.01		1.03 1.13 0.96 1.02 1.04	[0.96; 1.11] [1.04; 1.23] [0.91; 1.02] [0.98; 1.06] [0.95; 1.14]	10.5% 8.3% 18.2% 37.0%	13.4% 11.6% 17.8% 42.8%	
comparison = L-RTVS Gerard 2006	0.0056	0.0508		1.01	[0.91; 1.11]	5.8%	9. 1%	
comparison = L-CRT1 Ngan 2012	VS S-RTear -0.0663	ly 0.0693 —		0.94	[0.82; 1.07]	3.1%	5.7%	
comparison = Inductio Fernandez-Martos 2015	n + L-CRT \ -0.0731	/S L-CRT2 0.0963		0.93	[0.77; 1.12]	1.6%	3.2%	
comparison = S-RT + c Bahadoer 2021	onsolidatio 0.0170	n VS L-CRT1 0.0322		1.02	[0.96; 1.08]	14.6%	16.0%	
comparison = L-CRT1 Schrag 2023	VS CHT 0.0074	0.0200		1.01	[0.97; 1.05]	37.9%	23.2%	
Common effect model Random effects model	I			1.01 1.01	[0.99; 1.03] [0.98; 1.05]	100.0% 	 100.0%	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$								

Network map The thickness of lines is proportional to the number of studies comparing the two treatments and the size of circles is proportional to the number of individuals for each treatment.



	V1	V2	V3	V4	V5	V6	V7
1	CHT		0.99 (0.86 to 1.15)				
2	1.03 (0.77 to 1.38)	Induction + L-CRT		0.93 (0.74 to 1.18)			
3	0.99 (0.86 to 1.15)	0.96 (0.75 to 1.24)	L-CRT1	0.96 (0.88 to 1.06)	0.99 (0.84 to 1.18)	0.94 (0.77 to 1.14)	0.98 (0.84 to 1.15)
4	0.96 (0.81 to 1.14)	0.93 (0.74 to 1.18)	0.96 (0.88 to 1.06)	L-CRT2			
5	0.99 (0.79 to 1.24)	0.96 (0.71 to 1.30)	0.99 (0.84 to 1.18)	1.03 (0.85 to 1.25)	L-RT		
6	0.93 (0.73 to 1.18)	0.90 (0.66 to 1.24)	0.94 (0.77 to 1.14)	0.97 (0.78 to 1.20)	0.94 (0.73 to 1.22)	S-RTearly	
7	0.98 (0.79 to 1.20)	0.95 (0.71 to 1.27)	0.98 (0.84 to 1.15)	1.02 (0.85 to 1.22)	0.99 (0.79 to 1.24)	1.05 (0.82 to 1.35)	S-RT + consolidation

Forest plot L-CRT-1 was used as a common comparator. RRs above 1 favor the strategy over the common comparator (L-CRT1).

Treatment	Comparison: othe (Random Effec	r vs 'L- ts Mod	CRT el)	'1' RR	95%-CI
S-RTearly L-CRT2 S-RT + consolidatior L-RT CHT Induction + L-CRT	ı			1.07 1.04 1.02 1.01 0.99 0.96	[0.88; 1.30] [0.95; 1.14] [0.87; 1.19] [0.85; 1.19] [0.86; 1.15] [0.75; 1.24]
C	.01	0.5 1	2	4	

Assessment of heterogeneity and consistency

Global heterogeneity

We interpreted tau² as follows: heterogeneity low with tau² \leq 0.010, moderate with 0.010<tau² \leq 0.242, high with tau²>0.242, and I² statistics as follows: not important (0%-40%), moderate (30%-60%), substantial (50%-90%), considerable (75%-100%).

tau^2= 0.0051 ; tau= 0.0712

I^2= 79.58 % (35.15 % to 93.57 %)

Consistency: global approach

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q statistics to assess homogeneity / consistency

Q df p-valueTotal9.8020.0075Within designs9.8020.0075Between designs0.000--

Design-specific decomposition of within-designs Q statistic

Design Q df p-value L-CRT1:L-CRT2 9.80 2 0.0075

Q statistic to assess consistency under the assumption of a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within Between designs 0.00 0 -- 0.0712 0.0051

Consistency: local approach

Separate indirect from direct evidence (SIDE) using back-calculation method

Random effects model:

comparison k nma direct indir. RoR z p-value

Legend:

comparison - Treatment comparison k - Number of studies providing direct evidence nma - Estimated treatment effect (RR) in network meta-analysis direct - Estimated treatment effect (RR) derived from direct evidence indir. - Estimated treatment effect (RR) derived from indirect evidence RoR - Ratio of Ratios (direct versus indirect) z - z-value of test for disagreement (direct versus indirect) p-value - p-value of test for disagreement (direct versus indirect)

Netrank Higher p-score values indicate higher ranking of treatments

	P-score
S-RTearly	0.6908
L-CRT2	0.6342
S-RT + consolidation	0.5198
L-RT	0.4734
L-CRT1	0.4180
CHT	0.4144
Induction + L-CRT	0.3494

eReferences

- 1. Higgins JPT, Thomas J, Chandler J. *Cochrane Handbook for Systematic Reviews of Interventions, 2nd Edition.* (Wiley, ed.). Chichester; 2019. https://training.cochrane.org/handbook/current. Accessed March 7, 2023.
- 2. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ*. 1995;310(6977):452. doi:10.1136/BMJ.310.6977.452
- 3. Veroniki AA, Bender R, Glasziou P, Straus SE, Tricco AC. The number needed to treat in pairwise and network meta-analysis and its graphical representation. *J Clin Epidemiol*. 2019;111:11-22. doi:10.1016/J.JCLINEPI.2019.03.007
- 4. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med.* 2004;23(20):3105-3124. doi:10.1002/sim.1875
- 5. Shim SR, Kim SJ, Lee J, Rücker G. Network meta-analysis: application and practice using R software. *Epidemiol Health*. 2019;41:e2019013. doi:10.4178/EPIH.E2019013
- 6. Higgins J, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.3* (*Updated February 2022*).; 2022. www.training.cochrane.org/handbook.
- 7. Cipriani A, Higgins JPT, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med.* 2013;159(2):130-137. doi:10.7326/0003-4819-159-2-201307160-00008
- 8. Ostuzzi G, Schneider-Thoma J, Tedeschi F, Leucht S, Barbui C. Crossroads of methodological choices in research synthesis: insights from two network meta-analyses on preventing relapse in schizophrenia. *BMJ Ment Heal*. 2023;26(1):e300677. doi:10.1136/BMJMENT-2023-300677
- 9. Ostuzzi G, Bertolini F, Tedeschi F, et al. Oral and long-acting antipsychotics for relapse prevention in schizophrenia-spectrum disorders: a network meta-analysis of 92 randomized trials including 22,645 participants. *World Psychiatry*. 2022;21(2):295-307. doi:10.1002/wps.20972
- Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol*. 1997;50(6):683-691. doi:10.1016/S0895-4356(97)00049-8
- Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods*. 2012;3(2):98-110. doi:10.1002/JRSM.1044
- 12. White IR, Barrett JK, Jackson D, Higgins JPT. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Res Synth Methods*. 2012;3(2):111-125. doi:10.1002/JRSM.1045
- 13. Chaimani A, Higgins JPT, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. *PLoS One*. 2013;8(10). doi:10.1371/JOURNAL.PONE.0076654
- 14. White IR. Network meta-analysis. *Stata J.* 2015;15(4):951-985. doi:10.1177/1536867x1501500403
- 15. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-634. doi:10.1136/BMJ.315.7109.629
- 16. Salanti G, Giovane C Del, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the quality of evidence from a network meta-analysis. *PLoS One*. 2014;9(7). doi:10.1371/journal.pone.0099682
- 17. Nikolakopoulou A, Higgins JPT, Papakonstantinou T, et al. Cinema: An approach for assessing confidence in the results of a network meta-analysis. *PLoS Med*. 2020;17(4). doi:10.1371/JOURNAL.PMED.1003082