IV. The role of surgery (PICO 4)

Author(s): Harris RC, Khan MS, Allen V, Moore DAJ, Fielding K, Grandjean L, and the LSHTM MDR-TB surgery systematic review group (11 November 2015)

Question: Surgery compared to no surgery for treatment of MDR or XDR TB

Setting: Georgia, Latvia, Russia, South Africa, South Korea and Turkey

Bibliography: (1) Harris RC, Khan MS, Martin LJ, Allen V, Moore DAJ, Fielding K, et al. and the LSHTM MDR-TB surgery systematic review group. The effect of surgery on the outcome of treatment for multidrug-resistant tuberculosis: a systematic review and meta-analysis. BMC Infect Dis. 2016;16(1). (2) Dravniece G, Cain KP, Holtz TH, Riekstina V, Leimane V, Zaleskis R. Adjunctive resectional lung surgery for extensively drug-resistant tuberculosis. Eur Respir J. 2009;34(1):180–183. (3) Gegia M, Kalandadze I, Kempker RR, Magee MJ, Blumberg HM. Adjunctive surgery improves treatment outcomes among patients with multidrug-resistant and extensively drug-resistant tuberculosis. Int J Infect Dis. 2012;16:e391–396. (4) Karagöz T, Yazicioğlu Moçin O, Pazarli P, Senol T, Yetiş Duman D, Duman G, et al. The treatment results of patients with multidrug resistant tuberculosis and factors affecting treatment outcome. Tuberk Toraks. 2009;57:383–392. (5) Keshavjee S, Gelmanova IY, Farmer PE, Mishustin SP, Strelis AK, Andreev YG, et al. Treatment of extensively drug-resistant tuberculosis in Tomsk, Russia: a retrospective cohort study. Lancet 2008;372:1403–1409. (6) Kim H-R, Hwang SS, Kim HJ, Lee SM, Yoo C-G, Kim YW, et al. Impact of extensive drug resistance on treatment outcomes in non-HIV-infected patients with multidrug-resistant tuberculosis. Clin Infect Dis. 2007;45(10):1290–1295. (7) Kim DH, Kim HJ, Park S-K, Kong S-J, Kim YS, Kim T-H, et al. Treatment outcomes and long-term survival in patients with extensively drug-resistant tuberculosis. Am J Respir Crit Care Med. 2008;178:1075–1082. (8) Kwak N, Kim HR, Yoo CG, Kim YW, Han SK, Yim JJ. Changes in treatment outcomes of multidrug-resistant tuberculosis. Int J Tuberc Lung Dis. 2015;19:525–530. (9) Kwon YS, Kim YH, Suh GY, Chung MP, Kim H, Kwon OJ, et al. Treatment outcomes for HIV-uninfected patients with multidrug-resistant and extensively drug-resistant tuberculosis. Clin Infect Dis. 2008;47:496–502. (10) Leimane V, Riekstina V, Holtz TH, Zarovska E, Skripconoka V, Thorpe LE, et al. Clinical outcome of individualised treatment of multidrug-resistant tuberculosis in Latvia: a retrospective cohort study. Lancet 2005;365:318–326. (11) Mitnick CD1, Shin SS, Seung KJ, Rich ML, Atwood SS, Furin JJ, et al. Comprehensive treatment of extensively drug-resistant tuberculosis. New Engl J Med 2008;359:563–574. (12) Shean KP, Willcox PA, Siwendu SN, Laserson KF, Gross L, Kammerer S, et al. Treatment outcome and follow-up of multidrug-resistant tuberculosis patients, West Coast/Winelands, South Africa, 1992–2002. Int J Tuberc Lung Dis. 2008;12(10):1182–1189. (13) Sklyuev S, Levin A, Tcheimach E, Krasnov D. PC-658–02 Complex treatment approach for patients with destructive pulmonary tuberculosis by application of endobronchial valve. Int J Tuberc Lung Dis. 2013;17(12, Supp.2):S329–330. (14) Tahaoğlu K, Törün T, Sevim T, Ataç G, Kir A, Karasulu L, et al. The treatment of multidrug-resistant tuberculosis in Turkey. N Engl J Med. 2001;345:170–174. (15) Törün T, Tahaoğlu K, Ozmen I, Sevim T, Ataç G, Kir A, et al. The role of surgery and fluoroquinolones in the treatment of multidrug-resistant tuberculosis. Int J Tuberc Lung Dis. 2007;11(9):979–985.

QUALITY ASSESSMENT							NO. OF PATIENTS		EFFECT			
NO. OF STUDIES	STUDY DESIGN	RISK OF BIAS	INCONSISTENCY	INDIRECTNESS	IMPRECISION	OTHER CONSIDERATIONS	SURGERY	NO SURGERY	RELATIVE (95% CL)	ABSOLUTE (95% CL)	QUALITY	IMPORTANCE
Cured (follow up: range 0.5 to 10 years; assessed with: WHO definition)												
5	observational studies	serious[a-g]	not serious[h]	not serious[i]	not serious	none[j]	118/157 (75.2%)	308/561 (54.9%)	**OR 3.03** (1.59 to 5.78)	**238 more per 1,000** (from 110 more to 327 more)	⨁◯◯◯ VERY LOW	CRITICAL
Successful outcome (follow up: range 0.25 to 7 years; assessed with: cure or treatment success, WHO definition)												
14	observational studies	serious[a-g,k,l]	not serious[m]	not serious[n]	not serious	none[i,o]	371/453 (81.9%)[p]	1197/2006 (59.7%)	**OR 2.62** (1.94 to 3.54)[p]	**198 more per 1,000** (from 145 more to 243 more)	⨁◯◯◯ VERY LOW	CRITICAL
Death (follow up: range 0.5 to 10 years; assessed with: all-cause mortality or TB mortality)												
5	observational studies	serious[a-f,k,q-s]	not serious[m]	serious[t]	serious[s]	none[j]	11/191 (5.8%)	52/720 (7.2%)	**OR 0.82** (0.41 to 1.64)	**12 fewer per 1,000** (from 41 fewer to 41 more)	⨁◯◯◯ VERY LOW	CRITICAL
Loss to follow up (previously default) (follow up: range 0.5 to 10 years; assessed with: WHO definition)												
4	observational studies	serious[a-f,u]	not serious[m]	not serious[v]	not serious	none[j,w]	6/156 (3.8%)	77/613 (12.6%)	**OR 0.35** (0.15 to 0.81)[x]	**78 fewer per 1,000** (from 21 fewer to 105 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Treatment failure (follow up: range 0.5 to 10 years; assessed with: WHO definition)												
5	observational studies	serious[a-g,k]	not serious[m]	not serious[v]	not serious	none[j,w]	8/191 (4.2%)	82/720 (11.4%)	**OR 0.38** (0.18 to 0.81)	**67 fewer per 1,000** (from 20 fewer to 91 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Transfer out (follow up: not reported)												
2	observational studies	serious[a-c,f,y,z]	not serious[aa]	not serious	not serious[aa]	none[z,bb]	0/139 (0.0%)	6/305 (2.0%)	not estimable		⨁◯◯◯ VERY LOW	CRITICAL

QUALITY ASSESSMENT							NO. OF PATIENTS		EFFECT			
NO. OF STUDIES	STUDY DESIGN	RISK OF BIAS	INCONSISTENCY	INDIRECTNESS	IMPRECISION	OTHER CONSIDERATIONS	SURGERY	NO SURGERY	RELATIVE (95% CL)	ABSOLUTE (95% CL)	QUALITY	IMPORTANCE
Relapse or relapse/failure – not reported												
-	-	-	-	-	-	-	-	-	-	see comment	-	
Adverse events from surgery (follow up: range 1.5 to 10 years)												
1	observational studies	serious[a,b,f]	not serious[cc]	not serious	not serious[cc]	publication bias strongly suspected[dd]	2/66 (3%) surgical patients died due to surgical complications.				⨁◯◯◯ VERY LOW	CRITICAL

CL: Confidence limits; OR: Odds ratio

[a] Do not address or adjust for confounders and some studies do not fully describe the population – Dravniece, et al. 2009; Karagoz, et al. 2009; Kim et al. 2007; Kwak et al. 2015; Kwon et al. 2008; Mitnick, et al. 2008; Shean, et al. 2008; Sklyuev, et al. 2013; Tahaoglu, et al. 2001; and Torun, et al. 2007.

[b] Retrospective observational studies do not have randomization and have inherent bias in who is offered surgery – Dravniece, et al. 2009; Karagoz, et al. 2009; Keshavjee, et al. 2008; Kim et al. 2007; Kim, et al. 2008; Kwak, et al. 2015; Kwon, et al. 2008; Leimane, et al. 2005; Mitnick, et al. 2008; Shean, et al. 2008; Tahaoglu, et al. 2001; and Torun, et al. 2007.

[c] Uncertainty in representativeness of study population – Dravniece, et al. 2009; Karagoz, et al. 2009; Kim et al. 2007; Kwak, et al. 2015; Kwon et al. 2008; Shean, et al. 2008; and Tahaoglu, et al. 2001

[d] No estimate of variability given – Dravniece, et al. (2009) and Tahaoglu, et al. (2001).

[e] Number of "lost to follow-up" reported, but characteristics not described – Tahaoglu, et al. (2001).

[f] Length of follow up not described or adjusted for in analysis – Dravniece, et al. 2009; Kim, et al. 2007; Kwak, et al. 2015; Kwon, et al. 2008; Leimane, et al. 2005; Mitnick, et al. 2008; Shean, et al. 2008; Tahaoglu, et al. 2001; and Torun, et al. 2007.

[g] In surgical studies, it is not possible to blind patients or the study team. Outcome assessors could be blinded, and is somewhat important for assessing cure using smear as an outcome indicator. However, personnel other than the diagnosing physician, generally conduct laboratory assessment. For treatment success/failure there is a risk of reporting bias due to lack of blinding where data are programmatic, as there may be over-reporting due to programmatic targets and could be biased by knowledge of surgical status.

[h] Moderate I-squared (54.2%) and overlapping CLs between studies, and are thus not downgraded.

[i] Some variation in duration of follow-up in outcome definition, however it is not downgraded as alone it is not classified as serious for this outcome.

[j] All studies are cohort based, and therefore there may be some confounding due to patient allocation to surgery or no surgery. Patients who are more unwell may be more likely to be recommended for surgery (therefore causing underestimate of effect size). However, the most sick are often not offered surgery as they may be too unwell or the disease may be too disseminated to allow surgery (therefore overestimating effect size). In addition, there may be variation in the population offered surgery by setting or surgeon. As there is a specific window for surgery, these biases may have an impact on estimation of effect size, though it is unclear whether they would bias the estimation in a particular direction, and are a reflection of the reality of the patient group offered surgery. Therefore, the reviewers decided not to upgrade or downgrade the rating.

[k] Reports number, but not summary statistics or precision for this specific outcome – Leimane, et al. (2005) and Mitnick, et al. (2008)

[l] Abstract only, outcome and patient characteristics not clearly described – Dravniece, et al. (2009)

[m] Low I-squared and overlapping CLs between studies, so not downgraded.

[n] Most studies followed WHO outcome definitions. Some variation in duration of follow up to assess outcome but not downgraded as alone is not classified as serious issue for this outcome.

[o] Empty lower right quadrant of funnel plot. However, it seems that smaller (less precise) studies are reporting lower effect estimate so if publication bias were to exist this would suggest the current estimate effect measure is conservative. Per protocol, studies with <10 surgical participants were excluded, therefore the very smallest of studies were not included. Plot is not sufficiently asymmetrical to raise serious concerns, and any bias would appear to cause an underestimate of effect, therefore quality is not downgraded.

[p] n=13 for OR estimates, but n=11 for numbers of patients summarized in the table, as only two studies report effect estimate rather than the number of patients with the outcome and the denominator.

[q] In surgical studies, it is not possible to blind patients or study team. Outcome assessors could be blinded, but unimportant in mortality outcome as no subjectivity in assessment.

[r] Time period of follow up very variable, and for patients with follow up for <2 years the follow up period is potentially insufficient for mortality outcome – Shean, et al. (2008) and Torun, et al. (2007).

[s] Pooled CLs cross the null. Event rate is low and post hoc optimal information size calculation indicated number included in assessment of this outcome is too low to give sufficient power.

[t] Variation between studies in outcome definition used (all-cause versus TB-only). Unclear/variable period over which death was assessed (e.g. died during treatment, within six months of completion, or after two years).

[u] In surgical studies, it is not possible to blind patients or study team. Outcome assessors could be blinded, but where data are programmatic they are unlikely to be. This could introduce underestimate in reporting of default, but this bias is unlikely to vary between study groups.

[v] Mostly use WHO definition, minor variation in definition in some studies, but sufficiently direct not to downgrade.

[w] OR (similar to relative risk given the infrequency of the event) is <0.5 and the upper confidence limit would still provide a clinically significant benefit, therefore this would be considered a large effect size. However, the quality is not upgraded as according to GRADE methodology this should not be done if the risk of bias is serious.

[x] n=2 studies had no patients lost to follow-up in the surgery group, so 0.5 has been added to all cells in order that a CL can be calculated. The summary OR restricted to the 2 studies that had at least one patient lost to follow-up in each group is 0.47 (95% CL: 0.18, 1.24).

[y] Although reported separately, unlikely that clear differentiation has been made between “loss to follow-up” and “transfer out”.

[z] Suspected underreporting of outcome, but uncertain as to how this would impact the conclusions.

[aa] No pooled estimate, so insufficient evidence to assess.

[bb] Only two publications, so not possible to assess publication bias, but given how few report this outcome publication bias may be plausible.

[cc] One study and no comparator group so not possible to estimate.

[dd] Likely that complications occurred in other studies, but have either not been reported or have been included in all-cause deaths.

Author(s): Fox GJ, Mitnick CD, Benedetti A, Chan ED, Becerra M, Chiang C-Y, Keshavjee S, Koh W-J, Shiraishi Y, Viiklepp P, Yim J-J, Pasvol G, Robert J, Shim TS, Shin SS, Menzies R (11 November 2015)

Question: Elective partial lung resection compared to no surgery for patients on treatment for MDR-TB

Setting: Which types of surgery encompassed (lobectomy, segmentectomy, wedge resection)? Definition of non-response and adverse outcome of surgery; definition of extensive disease; how specialized were the centres/practitioners which provided surgery (external validity)? Under which conditions to indicate resection surgery and when to contraindicate; before or after culture conversion.

Bibliography: Fox GJ, Mitnick CD, Benedetti A, Chan ED, Becerra M, Chiang C-Y, et al. Surgery as an adjunctive treatment for multidrug-resistant tuberculosis: an individual patient data meta-analysis. Clin Infect Dis. 2016;62(7):887–95.

QUALITY ASSESSMENT							NO. OF PATIENTS		EFFECT		CERTAINTY OF EVIDENCE	IMPORTANCE
NO. OF STUDIES	STUDY DESIGN	RISK OF BIAS	INCONSISTENCY	INDIRECTNESS	IMPRECISION	OTHER CONSIDERATIONS	ELECTIVE PARTIAL LUNG RESECTION	NO SURGERY	RELATIVE (95% CL)	ABSOLUTE (95% CL)		
Success versus treatment failure or relapse (assessed with: Individual patient data meta-analysis)												
26[a]	observational studies[b]	not serious[c]	not serious[d]	not serious[e]	not serious[f]	none	185/204 (90.7%)[g]	1134/1398 (81.1%)[h]	**OR 2.4** (0.4 to 15.6)[i]	**100 more per 1000** (from 174 more to 179 fewer)	⨁⨁◯◯ LOW	CRITICAL
Success versus treatment failure or relapse or death (assessed with: Individual patient data meta-analysis)												
26[a]	observational studies[b]	not serious[c]	not serious[d]	not serious[e]	not serious[f]	none	185/214 (86.4%)[j]	1134/1702 (66.6%)[k]	**OR 2.0** (0.4 to 9.5)[i]	**133 more per 1000** (from 222 fewer to 284 more)	⨁⨁◯◯ LOW	CRITICAL
Success versus treatment failure or relapse or death or loss to follow-up (assessed with: Individual patient data meta-analysis)												
26[a]	observational studies[b]	not serious[c]	not serious[d]	not serious[e]	not serious[f]	none	185/229 (80.8%)[l]	1134/2193 (51.7%)[m]	**OR 3.5** (1.5 to 8.1)[i]	**272 more per 1000** (from 99 more to 380 more)	⨁⨁◯◯ LOW	CRITICAL

QUALITY ASSESSMENT							NO. OF PATIENTS		EFFECT		CERTAINTY OF EVIDENCE	IMPORTANCE
NO. OF STUDIES	STUDY DESIGN	RISK OF BIAS	INCONSISTENCY	INDIRECTNESS	IMPRECISION	OTHER CONSIDERATIONS	ELECTIVE PARTIAL LUNG RESECTION	NO SURGERY	RELATIVE (95% CL)	ABSOLUTE (95% CL)		
Death versus treatment failure or relapse or success (assessed with: Individual patient data meta-analysis)												
26[a]	observational studies[b]	not serious[c]	not serious[d]	not serious[e]	not serious[f]	none	10/214 (4.7%)	304/1702 (17.9%)	**OR 0.6** (0.2 to 2.2)[i]	**63 fewer per 1000** (from 137 fewer to 145 more)	⨁⨁◯◯ LOW	CRITICAL

CL: confidence limits; OR: odds ratio

[a] 26 studies include 18 studies where surgery was performed, and eight studies where surgery was not performed.

[b] *Limitations*. All data are from observational studies. The background medication regimen and the quality of surgery and other care are expected to differ between the studies. Bias expected because the decision to operate and the type of surgery are usually closely linked to prognostic factors such as severity/seriousness of the condition, the extent of resistance pattern, effectiveness of the medical options available and the patient response to treatment.

[c] *Risk of bias.* All included studies are observational, and selection bias is a substantial risk. Patient selection for surgery may be biased towards patients with more favourable prognostic factors or the opposite. Length of treatment differed substantially between surgical and non-surgical patients, suggesting that differences in the background medical regimens may also affect outcomes; although this and other measured potential confounders were included in the adjusted analysis of effect.

[d] *Inconsistency.* Based on estimated I^2_R. Estimates for the first two outcomes (success versus treatment failure or relapse +/- death) were very similar but OR for success increases when individuals who were lost to follow-up were included in the analysis.

[e] *Indirectness.* No indirectness expected given that all patients were on treatment for MDR-/XDR-TB. The outcomes (success, treatment failure, relapse and death) were among those scored as critical by the Guideline Development Group; loss to follow up was not one of the specified outcomes but is relevant to the question.

[f] *Imprecision.* 95% confidence limits for effect estimate applied with adjustment.

[g] Pooled proportion 93% (89%–97%).

[h] Pooled proportion 77% (69%–85%).

[i] *Adjusted effect estimates.* The method of adjustment was one to one propensity score matching between surgical patients and non-surgical patients, from non-surgical studies.

[j] Pooled proportion 90% (86%–94%).

[k] Pooled proportion 64% (54%–73%).

[l] Pooled proportion 66% (62%–70%).

[m] Pooled proportion 51% (40%–62%).

Author(s): Fox GJ, Mitnick CD, Benedetti A, Chan ED, Becerra M, Chiang C-Y., Keshavjee S, Koh W-J, Shiraishi Y, Viiklepp P, Yim J-J, Pasvol G, Robert J, Shim TS, Shin SS, Menzies R

Question: Elective pneumonectomy compared to no surgery for patients on treatment of MDR-TB.

Setting: Before or after culture conversion; which comparison group would have (a) failure / relapse, (b) failure / relapse / death, and (c) failure / relapse / death / loss to follow up.

Bibliography: Fox GJ, Mitnick CD, Benedetti A, Chan ED, Becerra M, Chiang C-Y, et al. Surgery as an adjunctive treatment for multidrug-resistant tuberculosis: an individual patient data metaanalysis. Clin Infect Dis. 2016;62(7):887–95.

QUALITY ASSESSMENT							NO. OF PATIENTS		EFFECT		CERTAINTY OF EVIDENCE	IMPORTANCE
NO. OF STUDIES	STUDY DESIGN	RISK OF BIAS	INCONSISTENCY	INDIRECTNESS	IMPRECISION	OTHER CONSIDERATIONS	ELECTIVE PNEUMONEC-TOMY	NO SURGERY	RELATIVE (95% CL)	ABSOLUTE (95% CL)		
Success versus treatment failure or relapse (assessed with: individual patient data meta-analysis)												
26[a]	observational studies	not serious[b]	not serious[c]	not serious[d]	not serious[e]	none	72/91 (79.1%)[f]	1134/1398 (81.1%)[g]	**OR 0.8** (0.1 to 6.0)[h]	**4 fewer per 100** (from 15 more to 51 fewer)	⨁⨁◯◯ LOW	CRITICAL
Success versus treatment failure or relapse or death (assessed with: individual patient data meta-analysis)												
26[a]	observational studies	not serious[b]	not serious[c]	not serious[d]	not serious[e]	none	72/105 (68.6%)[i]	1134/1702 (66.6%)[j]	**OR 0.7** (0.1 to 3.0)[h]	**8 fewer per 100** (from 19 more to 50 fewer)	⨁⨁◯◯ LOW	CRITICAL
Success versus treatment failure or relapse or death or loss to follow-up (assessed with: individual patient data meta-analysis)												
26[a]	observational studies	not serious[c]	not serious[c]	not serious[d]	not serious[e]	none	72/117 (61.5%)[k]	1134/2193 (51.7%)[l]	**OR 1.4** (0.7 to 3.2)[h]	**83 more per 1000** (from 89 fewer to 257 more)	⨁⨁◯◯ LOW	CRITICAL

QUALITY ASSESSMENT							NO. OF PATIENTS		EFFECT		CERTAINTY OF EVIDENCE	IMPORTANCE
NO. OF STUDIES	STUDY DESIGN	RISK OF BIAS	INCONSISTENCY	INDIRECTNESS	IMPRECISION	OTHER CONSIDERATIONS	ELECTIVE PNEUMONEC-TOMY	NO SURGERY	RELATIVE (95% CL)	ABSOLUTE (95% CL)		
Death versus success or treatment failure or relapse (assessed with: individual patient data meta-analysis)												
26[a]	observational studies	not serious[b]	not serious[c]	not serious[d]	not serious[e]	none	14/105 (13.3%)	304/1702 (17.9%)	**OR 1.8** (0.6 to 5.1)[h]	**103 more per 1000** (from 63 fewer to 347 more)	⨁⨁◯◯ LOW	CRITICAL

CL: confidence limits; OR: odds ratio

[a] 26 studies include 18 studies where surgery was performed, and eight studies where surgery was not performed.

[b] *Risk of bias.* All included studies are observational, and selection bias is a substantial risk. Patient selection for surgery may be biased towards patients with more favourable prognostic factors or the opposite. Length of treatment differed substantially between surgical and non-surgical patients, suggesting that differences in the background medical regimens may also affect outcomes; although this and other measured potential confounders were included in the adjusted analysis of effect.

[c] *Inconsistency.* Based on estimated I-squared R. Estimates for the first two outcomes (success versus treatment failure or relapse +/- death) were very similar but OR for success increases when individuals who were lost to follow-up were included in the analysis.

[d] *Indirectness.* No indirectness expected given that all patients were on treatment for MDR-/XDR-TB. The outcomes (success, treatment failure, relapse and death) were among those scored as critical by the Guideline Development Group; loss to follow up was not one of the specified outcomes but is relevant to the question.

[e] *Imprecision.* 95% confidence limits for effect estimate applied with adjustment.

[f] Pooled proportion 79% (71%–88%).

[g] Pooled proportion 77% (6%–85%).

[h] *Effect estimates.* Adjusted effect estimates applying one to one propensity score matching between surgical patients and non-surgical patients from non-surgical studies.

[i] Pooled proportion 69% (60%–78%).

[j] Pooled proportion 64% (54%–73%).

[k] Pooled proportion 62% (54%–71%).

[l] Pooled proportion 51% (40%–62%).