NHS England

Evidence review: Stereotactic ablative radiotherapy (SABR) in patients with hepatocellular carcinoma

First published: July 2019

Updated: Not applicable

Prepared by: King's Technology Evaluation Centre on behalf of NHS England Specialised Commissioning

Contents

1 Introduction

This review focuses on patients who have hepatocellular carcinoma (HCC) which is the most common type of primary liver cancer. The incidence of HCC is increasing worldwide due to the increase in hepatitis C infection rates and a rise in non-alcohol fatty liver disease (Parikh et al. 2018).

Treatment of HCC depends on a number of factors associated with the patient's performance status, the size, and location of the lesion in the liver and prior liver function (Bruix and Sherman, 2011). Surgical resection, liver transplant, and local ablative treatments are available choices to treat early stage disease (Bruix et al. 2011). In some cases, however, severe comorbidities or advanced disease exclude treatment with surgery and liver transplant is not always available. More advanced disease can be treated with transarterial chemoembolization (TACE) but responses are often incomplete, and the treatment is associated with cumulative toxicity imposing a limit in the number of times a patient can undergo TACE[1]. Systemic chemotherapy such as the use of the oral tyrosine kinase inhibitor sorafenib can be used with palliative intent offering small improvements in local control (LC). For patients with HCC that are not candidates for any of the previously mentioned treatments SABR can be used to offer local ablation.

SABR for people with hepatocellular carcinoma has been investigated in the NHS England Commissioning through Evaluation (CtE) programme.

The objective of this review is to examine the clinical effectiveness, safety and cost effectiveness of SABR for the treatment of HCC compared with no treatment, local or systemic treatment.

2 Summary of results

Following a systematic search of medical databases (see section 3: Methodology), 7 studies were identified which met the inclusion criteria for this review. These included:

- 5 retrospective comparative cohort studies: one comparing SABR with sorafenib ((Bettinger et al., 2018)- 1023 patients) and 4 comparing SABR with radiofrequency ablation (RFA) ((Wahl et al., 2016, Rajyaguru et al., 2018, Parikh et al., 2018, Kim et al., 2019)- 224, 796, 64 and 773 patients respectively).
- 1 systematic review and meta-analysis of non-comparative studies ((Rim et al., 2019)- 1950 patients).
- 1 non-comparative prospective cohort study ((Klein et al., 2015)- 205 patients).
- There was no evidence that compared SABR with either standard fractionated radiotherapy or no treatment/best-supportive care.

[1] Note – TACE is excluded from this review in line with the inclusion criteria in the CtE project. See also section 10.

All five included comparative studies were selected because they used at least one statistical method to account for baseline imbalances between the two groups.

One of the included studies compared SABR with sorafenib reporting that SABR resulted in superior overall survival (OS) in comparison to sorafenib with a median OS of 17.0 (95% CI 10.8-23.2) months compared to 9.6 (95% CI 8.6-10.7) months, respectively (Bettinger et al. 2018). Progression-free survival (PFS) was a secondary outcome in patients with metastatic HCC. After propensity score matching, patients treated with SABR had an improved progression-free survival compared to patients treated with sorafenib (9.0 vs. 6.0 months).

The comparison of SABR and RFA showed that the two modalities result in equivalent 1- and 2-year OS of 70-80% and 40-50%, respectively. Higher OS rates for both SABR and RFA were reported by Kim et al. (2019). This finding suggests that differences in OS are mostly driven by patient characteristics and not due to the treatment effect. The results reported by the four studies examining this comparison (Wahl et al., 2016, Rajyaguru et al., 2018, Parikh et al., 2018, Kim et al., 2019) are in agreement with the pooled non-comparative results for SABR reported in the systematic review by Rim et al. (2019).

The strongest evidence from non-comparative studies on OS, LC and safety came from the Rim et al. (2019) systematic review which reported the analysis of 32 studies including 1950 patients with HCC and found that SABR resulted in 1-, 2-, and 3-year OS rates of 72.6% (95% CI 65.7-78.6), 57.8% (95% CI 50.9-64.4), and 48.3% (95% CI 40.3-56.5), respectively. The pooled analysis of LC rates, using random effects analyses, showed 1-, 2-, and 3-year LC rates of 85.7% (95% CI 80.1-90.0%), 83.6% (95% CI 77.4-88.3%), and 83.9% (95% CI 77.6-88.6%), respectively.

One non-comparative prospective cohort study reported quality of life (QoL) with SABR (Klein et al., 2015). The study included patients with HCC, intrahepatic cholangiocarcinoma, and liver metastases but presented separate results for the three cohorts. Treatment with SABR did not significantly affect QoL.

The meta-analysis reported toxicity rates from 23 of 33 included studies. The most commonly reported complications of grade ≥3 were gastrointestinal (GI) or hepatic toxicities with a pooled rate of 3.9% (95% CI 2.6-5.6%) for the former and 4.7% (95% CI 3.4-6.5%) for the latter. An association between Child-Pugh score and toxicity was found but not with either tumour size or radiation dose.

The meta-analysis also looked at separately the results of the three studies that reported high rates of grade ≥3 toxicity. The authors concluded that considering the pooled rates of complications and the fact that complications at high rates were mostly due to transient liver enzyme elevation and possibly caused by chronic liver disease, the use of SABR to treat patients with HCC was safe.

There are severe limitations to the evidence for SABR in people with HCC, with the evidence base mainly composed by retrospective non-comparative studies and high levels of heterogeneity in the included patient population and study designs. The inherent bias of retrospective comparisons cannot be completely addressed with statistical methods as evident by the Rajyaguru et al. (2018) study where failure to capture all important baseline characteristics during propensity-score matching resulted in wrong conclusions subsequently disproved by the literature.

3 Methodology

The methodology to undertake this review is specified by NHS England in their 'Guidance on conducting evidence reviews for Specialised Commissioning Products' (2016).

A description of the relevant Population, Intervention, Comparison, and Outcomes (PICO) to be included in this review was prepared by NHS England's Policy Working Group for the topic (see section 9 for PICO table).

The PICO criteria were used to search for relevant publications in EMBASE, MEDLINE and Cochrane CDSR and CENTRAL (see section 10 for search strategy).

The search dates for publications were between 01/01/2009 and 21/05/2019.

The searches retrieved 1275 records. Following de-duplication in EndNote X7, 859 records were assessed for relevance using the criteria from the PICO. Full text versions of papers which appeared potentially useful were obtained and reviewed to determine whether they were appropriate for inclusion. Papers which matched the PICO were selected for inclusion in this review.

Evidence from all 7 papers included was extracted and recorded in evidence summary tables, critically appraised and their quality assessed using National Service Framework for Long term Conditions (NSF-LTC) evidence assessment framework (see section 7 below).

The body of evidence for individual outcomes identified in the papers was graded and recorded in grade of evidence tables (see section 8).

4 Results

1. **In patients with HCC, what is the clinical effectiveness of stereotactic ablative body radiotherapy compared with no treatment, local treatment to the liver or systemic treatment?**

All 7 included studies reported on at least one clinical effectiveness outcome (Rim et al., 2019, Wahl et al., 2016, Rajyaguru et al., 2018, Parikh et al., 2018, Bettinger et al., 2018, Klein et al., 2015, Kim et al., 2019).

- **SABR vs. sorafenib (1023 patients)**

Overall survival

Bettinger et al. (2018) compared OS in 1023 patients treated with SABR with patients treated with sorafenib. After propensity matched scoring 190 patients (95 in each cohort) were analysed. Median OS in the SABR group was 18.1 (95% CI 10.3-25.9) months compared to 8.8 (95% CI 8.2-9.5) in the sorafenib group. After propensity-matched scoring adjusting for different baseline characteristics, the OS benefit for patients treated with SABR was still preserved with a median OS of 17.0 (95% CI 10.8-23.2) months compared to 9.6 (95% CI 8.6-10.7) months in sorafenib patients.

Progression free survival

Bettinger et al. (2018) also reported PFS with SABR as a secondary outcome in patients with metastatic HCC. PFS was defined from the day of starting sorafenib or SABR treatment until death or radiological progression. Data concerning PFS were available in 76.8% of patients treated with sorafenib and 75.5% of patients treated with SABR. After propensity score matching, patients treated with SABR had an improved progression-free survival compared to patients treated with sorafenib (9.0 vs. 6.0 months).

- **SABR vs. RFA (3033 patients)**

Overall survival

Wahl et al. (2016) performed propensity score matching in 224 patients to account for baseline characteristics imbalances between the two groups and reported OS at 1 and 2 years of 69.6% and 52.9% after RFA and 74.1% and 46.3% after SABR in patients with inoperable and not metastatic HCC, with no significant difference between treatment groups. Although the two groups were well balanced with respect to multiple clinical characteristics, patients undergoing SABR had received more prior treatments and were less likely to proceed to transplantation. There was also shorter follow-up in the SABR group, which could obscure late effects.

Kim et al. (2019) performed propensity score matching in 773 (668 in the RFA group and 105 in the SABR group) patients to account for baseline characteristics imbalances between the two groups and reported their analysis in patients with inoperable HCC. In the unmatched cohort, the 1- and 2-year OS rates were 88.5% and 74.8% for patients in the SABR group and 91.1% and 79.8% for patients in the RFA group, respectively (p= 0.504). In the matched cohort (190 patients, 95 in each cohort) the 1- and 2-year OS rates were 87.1% and 71.8% for the SABR group and 86.9% and 76.4% for the RFA group, respectively (p =0.667).

Parikh et al. (2018) performed propensity score matching in 440 patients (408 treated with RFA and 32 with SABR) to account for baseline characteristics imbalances between the two groups and reported their analysis of patients with non-metastatic stage I or II HCC treated with SABR or RFA. In the unmatched cohort, patients undergoing SABR had worse survival than RFA-treated patients ($p < 0.001$). The 1-year OS for SABR-treated patients was 78.1% and 79.4% for RFA-treated patients. The 2-year OS was approximately 50% for both groups. However, 3-year survival was significantly longer in the RFA-treated cohort. After propensity-matched scoring (64 patients, 32 in each cohort), there was no significant difference in survival between SABR-treated and RFA-treated patients ($p = 0.30$) (Parikh et al., 2018).

Rajyaguru et al. (2018) performed propensity score matching in 796 patients to account for baseline characteristics imbalances between the two groups and analysed patients' data with inoperable non-metastatic HCC using the National Cancer Database, which includes about 70% of all newly diagnosed patients with cancer in the United States that undergone SABR or RFA as their primary treatment. In the propensity score matched and time to treatment matched analysis, RFA was associated with a significant OS benefit (HR 0.67; 95% CI, 0.55-0.81; p<0.001); the 5-year OS was 29.8% (95% CI, 24.5%-35.3%) in the RFA group versus 19.3% (95% CI, 13.5%-25.9%) in the SABR group (p<0.001).

Apart from one study (Rajyaguru et al., 2018), after adjusting for imbalances in the patients' characteristics with propensity matched scoring, SABR and RFA resulted in similar OS rates. In Rajyaguru et al. (2018), although propensity-score was used to match patients' baseline characteristics this did not include Child Pugh (CH-P) status, a variable associated with OS. In addition, 36% of the patients in the SABR cohort (n = 296) were treated with lower than standard radiotherapy dose (either 50-54 Gy in five fractions). A follow-up study that re-analysed the same data only including patients who received standard dose showed no difference in OS between the two cohorts (Shinde et al. 2018)[2].

Local control

Two of the included studies comparing SABR with RFA provided results on local control.

One retrospective comparative cohort study by Wahl et al. (2016) compared the LC rate between SABR and RFA. The 1- and 2-year LC was 83.6% and 80.2% for RFA-treated tumours and 97.4% and 83.8% for tumours treated with SABR. Twenty tumours (8%) treated with RFA showed residual disease after first ablation. Eight of these were re-ablated within 3 months of first treatment and were not counted as local failures. The authors used inverse probability of treatment weighting (IPTW) to adjust for potential imbalances in treatment assignment between the two groups. In IPTW univariate analysis, treatment modality was associated with local progression (HR, 2.63; p = 0.016). After adjusting for treatment type, tumour size was the only covariate predictive of local progression (HR, 1.36 per cm; p = 0.029).

The retrospective comparative cohort study by Kim et al. (2019) also compared the LC rate between SABR and RFA. In the matched cohort, the 1- and 2-year LC was 76.1% and 64.9% for RFA-treated tumours and 83.7% and 74.9% for tumours treated with SABR. In the multivariate analysis, treatment modality was significantly associated with LC, in favour of SABR (p = 0.004). Other independent predictors included alpha-fetoprotein (AFP) and Protein Induced by Vitamin K Absence or Antagonist-II (PIVKA-II) levels, tumour location (subphrenic region), and size (>2.0 cm).

- **Non-comparative SABR studies (2155 patients)**

Overall survival

[2] The study was published as a letter to the editor and therefore, not included in this review.

The meta-analysis by Rim et al. (2019) that included 32 observational single-arm studies involving 1950 patients with HCC who underwent SABR examined this outcome. Although the meta-analysis included studies with heterogeneous patient populations and study designs, the pooled analysis resulted in a patient cohort with similar characteristics to the scope of the review with a median proportion of patients with Child-Pugh class A of 82.3% (range: 47.9-100) and an overall median tumour size of 3.3 cm (range: 1.6-8.6). Pooled 1-, 2-, and 3-year OS rates were 72.6% (95% CI 65.7-78.6), 57.8% (95% CI 50.9-64.4), and 48.3% (95% CI 40.3-56.5), respectively.

Local control

The meta-analysis by Rim et al. (2019) also reported pooled 1-, 2-, and 3-year LC rates of 85.7% (95% CI 80.1-90.0%), 83.6% (95% CI 77.4-88.3%), and 83.9% (95% CI 77.6-88.6%), respectively. In subgroup analysis based on tumour size, lesions with less than 5cm diameter, had statistically significant better LC for 1-year, 2-year, and 3-year ($p < 0.001$, 0.001, and 0.001, respectively). In subgroup analysis based on radiation dose (median EQD2 estimates of 80 Gy10), the difference was marginally significant for 1-year LC rate ($p = 0.071$), and not significant for 2- and 3-year LC rates.

Quality of life

One prospective cohort study reported QoL with SABR (Klein et al., 2015). The study included 205 patients with HCC, intrahepatic cholangiocarcinoma, and liver metastases but presented separate results for the three cohorts. Although the main cohort consisted of patients with Child-Pugh A liver function, a small number of patients with HCC (n=10) with Child-Pugh B liver function were also treated. The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core-30 (EORTC QLQ-C30) and Functional Assessment of Cancer Therapy-Hepatobiliary (FACT-Hep) validated and cancer-specific questionnaires were used to assess QoL. No difference in baseline QoL ($p=0.17$) was seen between the HCC, liver metastases, and intrahepatic cholangiocarcinoma patients. The authors concluded that treatment with SABR in patients with liver cancer temporarily worsened appetite and fatigue at approximately 1 month after treatment but QoL returned to baseline levels at 1 year post treatment. Other QoL domains did not show significant change from baseline after SABR.

2. In patients with HCC, what is the safety of stereotactic ablative body radiotherapy compared with no treatment, local treatment or systematic treatment?

Three of the included studies provided results on toxicity as a secondary outcome.

One study was the meta-analysis of observational studies by Rim et al. (2019), one study was a comparative cohort comparing SABR with RFA (Wahl et al. 2016) and one comparative cohort study compared SABR with sorafenib (Bettinger et al., 2018). All studies used the Common Terminology Criteria for Adverse Events (CTCAE) criteria to record toxicity information. In most cases toxicity outcomes were reported as acute or late toxicity with the definition of the former varying from 1 to 3 months post treatment.

The most commonly reported toxicities were gastrointestinal (GI), haematologic, and hepatic. GI complications included gastric or duodenal ulcer, nausea and vomiting. Haematologic complications included abnormalities of white blood cells, platelets, and haemoglobin. Hepatic complications included abnormalities of liver function profile (alanine transaminase, aspartate aminotransferase, and bilirubin), albumin abnormalities, and liver decompensation (ascites, encephalopathy, and varices) (Rim et al., 2019).

- **SABR vs. sorafenib (1023 patients)**

Bettinger et al. (2018) compared the toxicity rates between SABR and sorafenib. Overall, 73.6% of sorafenib-treated patients experienced at least one adverse event at any grade. The most common adverse event was diarrhoea (39.3%), followed by hand-foot skin reaction (31.2%), fatigue (29.3%), weight loss (19.0%), and sorafenib-related hypertension (13.3%). A total of 19.4% of the patients had to stop sorafenib due to adverse events. For the group treated with SABR, 6.5% developed grade 2 toxicity mainly due to an increase in liver enzymes. Grade 3 toxicity was reported in 10.6% of the SABR-treated patients mainly due to an increase in liver enzymes, however, there were also 1 case of radiation-induced liver disease, 1 case of cholangitis and 2 cases of hepatic decompensation. Finally, grade 4 toxicity was reported in 2 cases (1.6%) as hepatic decompensation in 1 case and liver abscess in the other.

- **SABR vs. RFA (3033 patients)**

Wahl et al. (2016) compared the toxicity rates between SABR and RFA. Grade 3+ acute toxicity was 11% and 5% in the RFA and SABR groups respectively (p=0.31). The RFA complications were pneumothorax (n = 1), sepsis (n = 2), duodenal and colonic perforation (n = 2), and bleeding (n = 3) and resulted in two deaths. The SABR complications were radiation-induced liver disease (n = 1), GI bleeding (n = 1), and worsening ascites (n = 1) and there were no toxicity-related deaths. The rates of late grade 3+ biliary toxicity were similar in the RFA and SABR groups at 1 (2.3% v 3.3%; p=0.7) and 2 years (6% v 3.3%; p=0.38). The rates of late grade 3+ GI toxicity were also similar in the RFA and SABR groups at 1 (3.4% v 5.4%; p =0.49) and 2 years (6.4% v 8.3%; p =0.66). There were no late grade 5 adverse events in either group.

Kim et al. (2019) also compared the toxicity rates between SABR and RFA. Forty-three adverse events (6.5%) were observed in the RFA group. Complications included pleural effusion (n = 18), bile duct injury (n = 10), burn injury (n = 6), abscess formation (n = 4), hemoperitoneum (n = 3), haemothorax (n = 1), and pneumothorax (n = 1). There were 25 patients (3.7%) in the RFA group with grade 3 or 4 toxicities. All patients in the SABR group completed the scheduled treatment without severe toxicity. However, seven patients (6.7%) developed radiation-induced liver disease (RILD) in the SABR group. The overall toxicity rate did not differ significantly between the two treatment arms (p =0.850) and there were no cases of grade 5 toxicity.

- **Non-comparative SABR studies (2155 patients)**

The meta-analysis reported toxicity rates from 23 of 33 included cohort studies. The most commonly reported complications of grade ≥3 were GI or hepatic toxicities. For GI toxicities, the grade ≥3 events were less than 5% in 16 of 17 cohorts (94.1%) and were not reported in other 6 cohorts. The pooled rate using random effects analysis was 3.9% (95% CI 2.6-5.6%). For hepatic

toxicity, the rates of grade ≥3 complications were <10% in 23 of 24 cohorts (95.8%). The pooled rate was 4.7% (95% CI 3.4-6.5%). When tested in subgroup analysis neither tumour size nor radiation dose, were found to be statistically significant. Meta-regression analysis showed that CH-P class was significantly correlated with hepatic complications of grade ≥3 (p = 0.013).

The meta-analysis also looked at separately the results of the three studies that reported high rates of grade ≥3 toxicity. One study that reported high rates of hepatic toxicity (16.3%), all cases were transient elevation of liver enzymes. The authors assumed that possible risk factors were large tumour size and poor liver function (Scorsetti et al., 2015). Two other studies reported high rates of haematological (approximately 30% in both studies). The study by Kim et al. reported mostly thrombocytopenia (patients who experienced this complication had prior haematological problem). The authors concluded that considering the pooled rates of complications and the fact that complications at high rates were mostly transient and possibly caused by chronic liver disease, the use of SABR to treat patients with HCC was safe.

Treatment-related toxicity was a secondary outcome in all studies, therefore, it is unknown if any of them was adequately powered to detect a difference relative to a comparator (RFA or sorafenib). In addition, the retrospective design in most studies may lead to detection bias and inability to accurately capture toxicity events.

3. In patients with HCC, what is the cost effectiveness of stereotactic ablative body radiotherapy compared with no treatment, local or systematic treatment?

No eligible economic analyses comparing SABR with no treatment, local or systemic treatment were identified as part of this review.

4. From the evidence selected, are there any subgroups of patients who may benefit from stereotactic ablative body radiotherapy more than the wider population of interest?

- **SABR vs. sorafenib**

Bettinger et al. (2018) performed subgroup analyses based on the presence of portal vein thrombosis and extrahepatic metastases.

- Extrahepatic metastases

In the unmatched cohort, patients with extrahepatic metastases treated with SABR (only SABR of the hepatic tumour) showed a significantly improved OS compared to patients with sorafenib treatment (16.0 [6.7–25.4] vs. 7.6 [6.2–8.9] months, HR 0.43 [0.22– 0.84], p = 0.014). Also, in the matched cohort, the survival benefit of SABR treatment in metastatic patients was consistent (16.0 [6.6–25.4] vs. 10.0 [5.5-14.5] months, HR 0.38 [0.17-0.84], p = 0.018).

- Portal vein thrombosis

Patients with portal vein thrombosis treated with SABR had a median OS of 8.0 (4.3-11.7) compared to 6.1 (5.2-6.9) months in sorafenib-treated patients in the unmatched cohort (p = 0.330). After propensity score matching, there was no difference in OS between patients treated in either group (9.0 [2.9-15.1] vs. 6.0 [2.7-9.3] months, p = 0.568).

- **SABR vs. RFA**

Rajyaguru et al. (2018) performed exploratory subgroup analyses of the matched cohort, using as variables age, sex, clinical T stage, tumour size, tumour grade, Charlson-Deyo comorbidity score, and facility type. According to their analyses the overall advantages of RFA over SABR persisted across all subgroups, and no significant heterogeneity in HR was observed.

Kim et al. (2019) reported in their subgroup analyses based on tumour size, that SABR correlated with better local control in tumours larger than 2.0 cm (p =0.012) but not in tumours smaller than 2.0 cm (p = 0.061).

- **Non-comparative SABR studies**

The meta-analysis by Rim et al. (2019) performed subgroup analyses based on tumour size and radiotherapy dose.

  - Tumour size

The effect of tumour size (median value of 5 cm) was statistically significant for 1- and 2-year OS rates, and for 1-, 2- and 3-year LC rates.

  - Radiotherapy dose

The effect of radiotherapy dose (median EQD2[3] estimates of 80 Gy10), was not statistically significant for OS or LC. Neither tumour size or radiation dose influenced toxicity rates. The authors attributed the effect of tumour size on LC and OS to the fact that they categorised studies reporting high tumour invasion rates (>30%) into the subgroup of tumour size>5cm, and the higher tumour vascular invasive (TVI) rate might affect the difference in clinical outcomes. Larger tumours are also more likely to exhibit portal vein thrombosis as previously reported in other studies (Qi et al., 2014).

It should be noted that all subgroup analyses were retrospective and exploratory. Given the heterogeneity of study designs and included populations it is not possible from the current evidence to discern any subgroups of patients who may benefit from SABR more than the wider population.

5 Discussion

Seven studies provide evidence relevant to the scope of this review. There is grade C evidence from one retrospective, observational study that the OS and PFS following treatment with SABR is better than sorafenib. SABR resulted in superior OS in comparison to sorafenib with a median OS of 17.0 (95% CI 10.8-23.2) months compared to 9.6 (95% CI 8.6-10.7), respectively. Despite the

[3] Equivalent dose in Gray-2

use of a statistical method to account for differences in baseline patient characteristics, treatment allocation was not controlled and may be biased due to different factors such as the intrahepatic tumour burden, liver function, or the performance status of the included patients. In addition, it is difficult to overcome the institutional differences of the treatment decision policies in patients with advanced or recurrent HCCs after prior treatments. Due to the retrospective nature of data collection, it is difficult to assess the patients' performance status as well as the toxicity results. Finally, in both groups approximately a third of the patients had Child-Pugh score B and some patients presented with multifocal disease. Both these characteristics make the population less comparable to the scope of the review.

There is grade B and C evidence, respectively, that the OS and LC following treatment with SABR is similar to that achieved with RFA. All four comparative studies were selected because they used at least one statistical method to account for baseline imbalances between the two groups. The comparison of SABR and RFA showed that the two modalities result in equivalent 1- and 2-year OS of 70-80% and 40-50%, respectively. Higher OS rates for both SABR and RFA were reported by Kim et al. (2019). This finding suggests that differences in OS survival are mostly driven by patient characteristics and not due to the treatment effect. The SABR results reported by these studies are similar to the SABR pooled results reported by the meta-analysis.

There is grade B evidence provided by the meta-analysis of non-comparative studies that treatment with SABR achieves 1- and 2-year OS rates of 65.7-78.6% and 50.9-64.4%, respectively. Although the meta-analysis included studies with heterogeneous patient populations and study designs, the median proportion of patients with Child-Pugh class A was 82.3% and the overall median tumour size was 3.3 cm (range: 1.6-8.6 cm). There is some consistency between the results reported by Rim et al. (2019) and the evidence from the comparative studies: the 1-year 95% confidence intervals overlap with the same outcome reported in these studies. Differences in the included population and treatment could account for the different rates observed among studies. The results were less consistent for the 2-year OS rates.

Although OS was a primary outcome in most studies, none of them reported a sample size calculation it is therefore, unknown if they were adequately powered to detect a difference in the effect. In addition, all studies were retrospectively conducted with high risk of bias. The use of propensity matched scoring can improve the comparability of the two cohorts, however, it largely depends on the available information, and the clinical variables included in the matching.

Grade B evidence suggests that SABR for HCC is safe with low rates of severe toxicity mainly linked to transient increase of liver enzyme associated with liver toxicity, in some cases caused by chronic liver disease. Grade C evidence from a single study, suggests that SABR results in similar toxicity with RFA and lower toxicity when compared with sorafenib. For the latter, however, it should be noted that as the two modalities have different toxicity profiles a direct comparison is difficult. Treatment-related toxicity was a secondary outcome in all studies, therefore, it is unknown if any of them was adequately powered to detect a difference relative to a comparator (RFA or sorafenib). In addition, the retrospective design in most studies may lead to detection bias and inability to accurately capture toxicity events.

There is grade C evidence suggesting that SABR does not significantly influence QoL in patients with HCC. The evidence is provided by one prospective cohort study that included a heterogeneous population of patients with HCC, intrahepatic cholangiocarcinoma, and liver metastases but presented separate results for the three cohorts. The authors used two widely used and validated tools for concurrent comparison of QoL outcomes increasing the validity of their results. The maximum follow-up was, however, only 12 months and it is unknown what proportion of patients completed follow-up. The study did not report any sample size calculation, therefore, it is unknown if it was adequately powered to detect a difference between the different cohorts of patients. Multiple imputations were performed to account for missing data of eligible patients alive at follow-up. Patient compliance for questionnaire completion fell from 90% at baseline to 60% at 1-year post-treatment.

The main limitation of the current evidence is that most of the evidence comes from non-comparative observational studies. These studies include heterogeneous patient populations, and study designs that limit the generalisability of the results. The evidence from retrospective comparative studies even when using propensity score matching to account for baseline differences between SABR and its comparators suffer from the same limitations as the inherent biases of retrospective design, such as patient selection bias, lack of information on important baseline clinical characteristics and toxicity outcomes, cannot be fully addressed by statistical methods. This is evident by the Rajyaguru et al. (2018) study where failure to capture all important baseline characteristics during propensity-score matching resulted in wrong conclusions subsequently disproved by the literature. In addition, most studies had a relatively short follow-up schedule. Although some studies reported subgroup analysis, the low numbers of patients and the high risk of bias do not allow robust conclusions to be drawn.

The main implication from the available evidence is that the use of SABR in patients with HCC may result in similar OS rates compared to RFA and better OS rates compared to systemic treatment with sorafenib, however, there is considerable uncertainty about these findings.

In the future, prospective observational comparative studies or phase 3 randomised controlled trials are needed to confirm the benefits of SABR in comparison with other local or systemic treatments.

6 Conclusion

Seven studies provide evidence relevant to the scope of this review. All evidence results described above are for an adult population. The most significant evidence is provided by the meta-analysis by Rim et al. (2019) that included 32 observational single-arm studies involving 1950 patients with HCC who underwent SABR. The analysis provides evidence for the clinical efficacy and safety of SABR. Both OS and LC were affected by tumour size, and radiation dose marginally affected LC. LC rates were better for smaller HCCs, and moderate efficacy was shown in treatment of tumours >5 cm. Reported rates of severe toxicity were low, and mainly due to hepatic or GI toxicity.

There is also grade B and C evidence, respectively, that the clinical efficacy of SABR is similar to that achieved with RFA and that is better than sorafenib. There is grade C evidence from a single study suggesting that SABR does not significantly affect QoL.

The main limitation of the current evidence is that most of the evidence comes from non-comparative (often retrospective) observational studies. These studies include heterogeneous patient populations, and study designs that limit the generalisability of the results. The evidence from retrospective comparative studies that used propensity score matching to account for baseline differences between SABR and RFA, and SABR and sorafenib, also suffer from the same limitations as the inherent biases of retrospective design, such as patient selection bias, lack of information on important baseline clinical characteristics and toxicity outcomes, cannot be fully addressed by statistical methods.

No published evidence exist that fit the scope on the cost-effectiveness of SABR compared with any of the comparators.

7 Evidence Summary Tables

Table 1: SABR vs. sorafenib

Study Design and Population Characteristics	Methodology	Results	Critical Appraisal Summary
Bettinger et al (2018) Retrospective comparative cohort study Multi-institution, Switzerland and Germany (6 centre Sorafenib, 13 centres SABR) Recruitment period: 2013-2017 1023 patients with primary unresectable HCC, 1-2 intrahepatic lesions, or multifocal HCC (3 or more lesions or diffuse growth pattern) SABR median Equivalent Dose in Gray-2 (EQD2) was 84.4Gy (36-124)	Median prescribed SABR dose was 44Gy (range: 21-66) Gy in 3-12 fractions. The median biologically equivalent dose (BED) (BED10) prescribed was 84.4Gy (range: 36-124). After propensity-score matching, 95 received Sorafenib and 95 SABR (overall 901 received Sorafenib, 122 received SABR). Following matching, the groups were similar in baseline characteristics. The following variables were included to match the patients: Child-Pugh score, prior surgery, radiofrequency ablation (RFA), transarterial chemoembolization (TACE), hepatic tumour burden, portal vein thrombosis (PVT), extrahepatic metastases, and Eastern Cooperative Oncology Group (ECOG) performance status. Median follow-up not reported.	Median overall survival: 16.0-months SABR vs. 9.6-months Sorafenib (p=0.005). Multivariable analysis showed SABR was a significant prognostic factor for OS (HR 0.53 [95%CI: 0.36-0.77], p=0.005). Higher EQD2 did not significantly influence OS rates. Sub-group extrahepatic lesions, median overall survival: 16.0-months SABR vs. 10.0-months Sorafenib; HR 0.38 [0.17–0.84], p=0.018. Progression free survival: 9.0-months SABR vs. 6.0-months Sorafenib (p=0.004)	Retrospective, observational study and therefore treatment allocation was not controlled and may be biased due to different factors such as the intrahepatic tumour burden, liver function, and especially the performance status (PS) of the patient. Recruitment period suggests that the cohort is more likely to be comparable with current practice. The propensity score matched analysis is very clear and the number of matched patients is relatively high. In both groups approximately 1/3 of the patients had CH-P score B. Also, some patients presented with multifocal disease. Both these characteristics make the population less comparable to the scope of the review. Some patients in the SABR group received less than standard radiation dose. 33 patients (27%) of the SABR cohort received additional treatment*, which could have

Study Design and Population Characteristics	Methodology	Results	Critical Appraisal Summary
		Toxicity: -Grade 1 = 1% SABR vs. 62% sorafenib -Grade 2 = 6.5% SABR vs. 50.4% sorafenib -Grade 3 = 9.8% SABR vs. 30.2% sorafenib -Grade 4 = 1.6% SABR vs. 2.4% sorafenib The most frequent side effects with sorafenib were diarrhoea, hand-foot skin reaction, fatigue, weight loss, and sorafenib-related hypertension. Sorafenib was stopped in 175 patients (19.4%) due to adverse events. Severe side effects associated with SABR were cholangitis, gastric ulcers with bleeding, and necrotic abscess.	confounded the OS results. However, the authors excluded those patients and the significant OS advantage for SABR remained. The reporting of the toxicity outcomes is very unclear, and no meaningful comparisons can be drawn. Adverse events were recorded using the CTCAE criteria.

CP; Child-Pugh score, HR; Hazard ratio PFS; progression free survival, OS; overall survival, LC; local control, PVT; portal vein thrombosis RFA; radiofrequency ablation, 95% CI = 95% confidence interval

Table 1: SABR vs. radiofrequency ablation

Study Design and Population Characteristics	Methodology	Results	Critical Appraisal Summary
Rajyaguru et al (2018) Retrospective comparative cohort study Date from the National Cancer Database, USA Recruitment period: 2004-2013 796 patients included in matched analyses (from 3,980 eligible patients) with non-advanced non-metastatic HCC; tumour size ≤5cm. Child Pugh status was unknown. Patients who received surgery or chemotherapy were excluded.	After propensity-score and time-to-treatment matching, 521 received RFA and 275 received SABR (overall 3684 received RFA, 296 received SABR). Following matching, the groups were similar in baseline characteristics. 42% of SABR patients received dose of 40-49Gy; 80% received their dose in 3-5 fractions (fx). Radiation dose was unknown in 14% of the patients. Of those with known dose, 26% received lower than standard dosing (< 40 Gy). Median 25.3-month follow-up.	Overall survival: RFA significantly better than SABR (hazard ratio 0.67 [0.55-0.81], p<0.001). 5-yr OS: 29.8% for RFA vs. 19.3% for SABR (p<0.001). Sub-group analysis of Tumour-Node-Metastases (TNM) status revealed a similar significant difference in survival between RFA and SABR.	Retrospective, observational study and therefore treatment allocation was not controlled and may be biased due to different clinical factors. The SABR group was heterogeneous in terms of radiation dose and fractionation schedule. As the Child Pugh status was unknown the analysis may also have included patients outside the scope of the review. Although propensity-score was used to match patients' baseline characteristics this did not include Child-Pugh status a variable associated with OS. 26% of the patients with recorded dose in the SABR cohort were treated with lower than standard radiotherapy dose (either 50-54 Gy in five fractions). A follow-up study that reanalysed the same data only including patients who received standard dose showed no difference in

Study Design and Population Characteristics	**Methodology**	**Results**	**Critical Appraisal Summary**
			OS between the two cohorts (Shinde et al., 2018)[4]. The long recruitment period means that practice may have changed over the course of the study, which could limit generalisability. The authors also carried out an inverse probability–weighted analysis, which confirmed the significantly different OS outcomes between the groups.

[4] The study was published as a letter to the editor and therefore, not included in this review.

Study Design and Population Characteristics	Methodology	Results	Critical Appraisal Summary
Parikh et al (2018) Retrospective comparative cohort study Data from Surveillance, Epidemiology, and End Results (SEER)-Medicare linked database, USA Recruitment period: 2004-2011 64 patients included in matched analyses (from 450 eligible patients) with non-metastatic stage I or II HCC; tumour size not reported. Child Pugh status was unknown. SABR dose not reported.	After propensity-score matching, 32 received RFA and 32 received SABR (overall 418 received RFA, 32 received SABR). Following matching, the groups were similar in baseline characteristics. Median follow-up (propensity score matched patients) was 594 days for RFA and 487 days for SABR.	Overall survival: 1-year: 78.1% SABR vs. 79.4% RFA 2-year: 40% SABR vs. 40% RFA (extracted from the Kaplan-Meier curve) Under propensity score matching, OS showed no significant differences between RFA and SABR (SABR HR 1.28 [95% CI 0.60-2.72], p=0.53). 90-day hospitalisation: Overall cohort analysis showed no significant differences between the groups.	Retrospective, observational study and therefore treatment allocation was not controlled and may be biased due to different clinical factors. In the overall cohort, 1-yr OS was similar (78.1% SABR, 79.4% RFA), but at 3-years there were significant differences (low certainty due to length of follow-up), with OS in the SABR group of approximately 16% and 50%* in the RFA group (SABR hazard ratio 1.80 [95% CI 1.15–2.82], p=0.01). However, the propensity score matched analysis showed no significant differences. It should be noted that although the propensity score matched cohort contained 64 patients compared to 450 patients in the overall cohort, the number treated with SABR was the same in both analyses (n=32). As the Child Pugh status was unknown the analysis may also have included patients outside the scope of the review. The long recruitment period means that practice may have changed over the course of the study, which could limit generalisability. *estimated from Kaplan-Meier graph

Study Design and Population Characteristics	Methodology	Results	Critical Appraisal Summary
Wahl et al. 2016 Retrospective comparative cohort study Single centre, USA Recruitment period: 2004-2012 224 patients with inoperable, non-metastatic HCC (332 discrete liver tumours) Mean Child-Pugh score: 6.2 SABR, 6.9 RFA SABR median age = 62 years, 85.7% men RFA median age = 60 years, 72.7% men	Patients were identified from a prospective departmental database. Typically, RFA was the first choice for tumours smaller than 3 to 4 cm. SABR was first choice for tumours not visualised by ultrasound, abutting a vessel or the luminal GI tract, or after RFA failure. Freedom from local progression (FFLP) and toxicity were retrospectively analysed. SABR median dose: Patients were treated with either three (46%) or five (53%) fractions delivered two to three times per week with median doses of 30 or 50 Gy with a range of 27 to 60 Gy, Median follow up: SABR 13 months, RFA 20 months.	FFLP 1 year = 97.4% SABR vs. 83.6% RFA 2 years = 83.8% SABR vs. 80.2% RFA Increasing tumour size predicted for FFLP in patients treated with RFA (HR 1.54 per cm; P= 0.006), but not with SABR (HR, 1.21 per cm; P= .617). For tumours ≥2 cm, there was decreased FFLP for RFA compared with SABR (HR, 3.35; P = 0.025). After adjusting for treatment type, tumour size was the only covariate predictive of LC (HR, 1.36 per cm; p= 0.029. OS 1 year = 74.0% SABR vs. 70.0% RFA 2 years = 46.0% SABR vs. 53.0% RFAAcute grade 3+ complications occurred after 11% and 5% of RFA and SABR treatments, respectively (P= 0.31). Late Grade ≥3 biliary:	Non-randomised. Due to the nature of the intervention, blinding was not possible. However, inverse probability of treatment weighting was used to control for differences in baseline characteristics. Although the two treatment populations were well balanced with respect to multiple factors, patients undergoing SABR had, on the average, received more prior treatments, and were less likely to proceed to transplantation. This may have biased the OS results. The two groups were well matched in terms of tumour size (median 1.8 vs. 2.2 cm, RFA and SABR respectively). LC was defined as the absence of progressive disease by the Response Evaluation Criteria in Solid Tumors (RECIST) criteria within or at the PTV margin for patients receiving SABR and the absence of recurrence within or adjacent to the ablation zone for patients receiving RFA. Adverse events were defined as grade 3+ events according to the CTCAE criteria during the 30 days after treatment (acute) or at all later time points (late biliary and luminal GI toxicity). Follow-up was shorter in the SABR group.

Study Design and Population Characteristics	**Methodology**	**Results**	**Critical Appraisal Summary**
		1-year=3.3% SABR vs. 2.3% RFA 2-years=3.3% vs. 6.0% RFA -Late Grade ≥3 GI: 1-year=5.4% SABR vs. 3.4% RFA 2-years=8.3% SABR vs. 6.4% RFA -Late Grade 5=0 for SABR and RFA. For SABR the toxicities were radiation-induced liver disease (n = 1), GI bleeding (n = 1), and worsening ascites (n = 1). For RFA complications included pneumothorax (n = 1), sepsis (n = 2), duodenal and colonic perforation (n = 2), and bleeding (n = 3) and resulted in two deaths.	

Study Design and Population Characteristics	Methodology	Results	Critical Appraisal Summary
Kim et al. 2019 Retrospective comparative cohort study Single centre, Korea Recruitment period: 2012-2016 773 patients with inoperable, non-metastatic HCC (850 discrete liver tumours) Child-Pugh score A = 94.7% SABR median age = 63 years, 84.2% men RFA median age = 67 years, 87.4% men Median tumour size: 2.4 in SABR and 2.1 in the RFA group	The SABR cohort of this study is a re-analysis of the Kim et al. study included in the systematic review by Rim et al. 2019. FFLP and toxicity were retrospectively analysed. SABR dose: 61.9% of the patients received a total of 60 Gy in 4 fractions and 21.0% received 52 Gy/4 fractions. Median follow up: SABR 21.9 months, RFA 21.6 months.	FFLP 1 year = 83.7% SABR vs. 74.9% RFA 2 years = 76.1% SABR vs. 64.9% RFA After subgroup analysis based on tumour size in the matched cohort, SABR was attributed to better local control only in tumours larger than 2.0 cm but not in tumours smaller than 2.0 cm (p = 0.635) OS 1 year = 87.1% SABR vs. 86.9% RFA 2 years = 71.8% SABR vs. 76.4% RFA Toxicity The overall toxicity rate did not differ significantly between the two treatment arms (p = 0.850). Acute -Grade 1= 28.6% SABR vs. 13.4% RFA*	Non-randomised. Due to the nature of the intervention, blinding was not possible. However, inverse probability of treatment weighting was used to control for differences in baseline characteristics. Recruitment period suggests that the cohort is more likely to be comparable with current practice. Although the two treatment populations were well balanced with respect to multiple factors, patients undergoing SABR had, on the average, received more prior treatments. This may have biased the OS results. The two groups were well matched in terms of tumour size (median 2.1 vs. 2.4 cm, RFA and SABR respectively). LC was defined as the absence of progressive disease by the RECIST criteria within or at the PTV margin for patients receiving SABR and the absence of recurrence within or adjacent to the ablation zone for patients receiving RFA. Adverse events were defined according to the CTCAE criteria. Follow-up was similar in the two groups.

Study Design **and Population Characteristics**	**Methodology**	**Results**	**Critical Appraisal Summary**
		-Grade 2 = 2.9% SABR vs. 13.4% RFA* -Grade 3 or 4= 0% SABR vs. 3.6% RFA - Grade 5 adverse events = 0 in both groups Late -Grade 1= 11.5% SABR vs. 7.2% RFA -Grade 2 = 3.9% SABR vs. 3.2% RFA -Grade 3 or 4= 0% in both groups - Grade 5 adverse events = 0 in both groups *results for RFA are presented together for grade 1 and 2 in the publication.	
CP; Child-Pugh score, FFLP; Freedom from local progression, HR; Hazard ratio PFS; progression free survival, OS; overall survival, LC; local control, PVT; portal vein thrombosis RFA; radiofrequency ablation, 95% CI = 95% confidence interval			

Table 2: Non-comparative studies

Study Design and Population Characteristics	Methodology	Results	Critical Appraisal Summary
Rim et al (2019) Systematic review and meta-analysis 32 studies (33 cohorts) including 1950 Patients with HCC with median tumour size of 3.3cm (ranging from 1.6-8.6cm) Median EQD2 was 83.3Gy (ranging from 48-114.8Gy) Median CP-A score: 82.3% (ranging from 47.9–100%)	A systematic search of PubMed, Medline, Embase, and Cochrane was undertaken. Inclusion criteria: SABR in <10fx, at least 10 patients treated with SABR, reporting survival or local control. No date limits were used (search date was 23-Apr-2018)	**Pooled analyses using random effects model** Overall survival: 1-yr: 72.6% (95% CI 65.7–78.6) 2-yr: 57.8% (95% CI 50.9–64.4) 3-yr: 48.3% (95% CI 40.3–56.5) Tumour size >5cm significantly associated with 1-yr OS ($p<0.001$) Under meta-regression tumour size was significantly correlated with 1-, 2-, and 3- year OS ranges ($p < 0.0001$, $p = 0.0022$, and $p = 0.0002$). Local control: 1-yr: 85.7% (95% CI 80.1–90.0) 2-yr: 83.6% (95% CI 77.4–88.3) 3-yr: 83.9% (95% CI 77.6–88.6)	The systematic review methods were well reported and reproducible, although the search strategy used was overly simplistic. The majority (85%) of the included studies were retrospective. The authors report potential publication biases (using Egger's test quantitatively and visual inspections of funnel plots), which could have influenced several outcomes. However, the authors presented trimmed results using the Duval and Tweedie method. Recruitment period suggests that the cohort is more likely to be comparable with current practice. Included studies were mostly retrospective single centre studies with high risk of bias for patient selection and outcomes detection. Complications reported within 3 months after the end of radiotherapy were classified as acute complications, and those reported later than 3 months or described as 'late complication' were classified as late complications.

Study Design and Population Characteristics	Methodology	Results	Critical Appraisal Summary
		Tumour size >5cm significantly associated with 1-, 2- and 3-yr OS (p<0.001 for all). 1-year LC was also influenced by radiation dose (median EQD2 estimates of 80 Gy10). Complications: Grade ≥3 complications GI: 3.9% (95% CI 2.6–5.6) Grade ≥3 complications hepatic: 4.7% (95% CI 3.4–6.5) Under meta-regression, CP-A status was a significant factor for hepatic toxicity (p=0.013). Neither tumour size nor dose were significant factors.	
Klein et al (2015) Prospective non-comparative cohort study Single centre, Canada Recruitment period: 2003-2011	Median radiation dose = 37Gy (5.1-60) Radiation dose was unknown in 6 patients of the patients. Maximum follow-up 12 months (median unknown).	QoL: FACT-Hep: Baseline= 137.4 1-month=129.7	Prospective, observational study and therefore treatment allocation was not controlled and may be biased due to different clinical factors. The SABR group was heterogeneous including patients with HCC, cholangiocarcinoma, and liver metastases.

Study Design **and Population Characteristics**	**Methodology**	**Results**	**Critical Appraisal Summary**
205 patients with hepatocellular carcinoma (HCC= 99 patients), liver metastases, or intrahepatic cholangiocarcinoma; tumour size median=133cm. 93% Child-Pugh score A Median age = 67 years, 66% men		3-months=133.4 6-months=133.6 12-months=135.1 QLQ-C30: Baseline=65.8 1-month=61.7 3-months=62.9 6-months=58.6 12-months=64.5 Higher baseline QoL scores were associated with improved survival.	Some of the patients received low radiation dose (minimum dose 5.1Gy). Multiple imputations were performed on missing data of eligible patients alive at follow-up. Two widely used and validated tools in this population were chosen for concurrent comparison of QoL outcomes: FACT-Hep and the EORTC QLQ-C30. QoL was evaluated at each visit. Maximum follow-up was only 12 months and it is unknown what proportion of patients completed follow-up.
CP; Child-Pugh score, QLQ-C30; European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire Core-30, FACT-Hep; Functional Assessment of Cancer Therapy-Hepatobiliary , HR; Hazard ratio PFS; progression free survival, QoL; Quality of life, OS; overall survival, LC; local control, PVT; portal vein thrombosis RFA; radiofrequency ablation, 95% CI = 95% confidence interval			

8 Grade of evidence table

Use of SABR to treat HCC (vs. sorafenib)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
Overall survival	Bettinger, 2018	6	Direct	C	Median overall survival (OS) is the length of time from either the date of diagnosis or the start of treatment, that half of the patients in a group of patients diagnosed with the disease are still alive. The best evidence on median OS is provided by the retrospective observational study by Bettinger et al. (2018) that included 190 patients in the matched cohort and compared SABR to sorafenib. Median OS in the SABR group was 18.1 (95% CI 10.3-25.9) months compared to 8.8 (95% CI 8.2-9.5) in the sorafenib group. Given the alternative treatment options for patients with HCC overall survival is a clinically meaningful outcome. The study performed a retrospective comparison between the two groups. Despite the use of a statistical method (propensity score matching) to account for baseline differences among the participants, patient selection bias cannot be excluded between the two cohorts. In both groups the inclusion of patients with Child Pugh score B will make the population less comparable to the scope of the review. Overall, there is considerable uncertainty about this outcome and

Use of SABR to treat HCC (vs. sorafenib)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
					additional randomised controlled studies will need to verify this finding.
Progression-free survival	Bettinger, 2018	6	Direct	C	Progression free survival (PFS) is the length of time during which the disease does not worsen, or the proportion of patients without worsening disease at a defined follow-up point after beginning treatment. PFS was defined from the day of starting sorafenib or SABR treatment until death or radiological progression. The best evidence on PFS is provided by the retrospective observational study by Bettinger et al. (2018) that included 190 patients in the matched cohort and compared SABR to sorafenib. Median PFS in the SABR group was 9.0-months (95% CI 5.8-12.2) months compared to 6.0-months (95% CI 4.8-7.2) in the sorafenib group (p=0.004).In patients with metastatic disease treatment is not given with curative intent and secondary outcomes such as PFS are clinically meaningful. The study performed a retrospective comparison between the two groups. Despite the use of a statistical method (propensity score matching) to account for baseline differences among the participants, patient selection bias cannot be excluded between the two cohorts. In both groups the inclusion of patients with Child Pugh score B will make the population less comparable to the scope of the review. Overall, there is

Use of SABR to treat HCC (vs. sorafenib)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
					considerable uncertainty about this outcome and additional randomised controlled studies will need to verify this finding.
Toxicity	Bettinger, 2018	6	Direct	C	Toxicity is defined based on the number and severity of adverse events a patient can experience after undergoing treatment. Treatment-related toxicity in patients with cancer is usually recorded and graded according to the Common Toxicity Criteria Adverse Events (CTCAE). The best evidence on toxicity is provided by the retrospective observational study by Bettinger et al. (2018) that included 190 patients in the matched cohort and compared SABR to sorafenib. 73.6% of sorafenib-treated patients experienced at least one adverse event at any grade. For the group treated with SABR, 6.5% developed grade 2 toxicity. Grade 3 toxicity was reported in 10.6% of the SABR-treated patients. However, it should be noted that as the two modalities have different toxicity profiles a direct comparison is difficult. Given that alternative treatment options with different toxicity profiles exist for patients with HCC, toxicity is clinically meaningful outcome. The study performed a retrospective comparison between the two groups. Despite the use of a statistical method (propensity score matching) to account for

Use of SABR to treat HCC (vs. sorafenib)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
					baseline differences among the participants, patient selection and outcome detection bias cannot be excluded between the two cohorts. In both groups the inclusion of patients with Child Pugh score B will make the population less comparable to the scope of the review. Overall, there is considerable uncertainty about this outcome and additional randomised controlled studies will need to verify this finding.
LC = local control PFS = progression free survival QoL = quality of life OS = overall survival					

Use of SABR to treat HCC (vs. radiofrequency ablation)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
Actuarial overall survival	Rajyaguru, 2018 Parikh, 2018 Wahl, 2016 Kim, 2019	5 5 6 6	Direct Direct Direct Direct	B	Actuarial overall survival (OS) is reported as the proportion of patients surviving at a defined follow-up point, such as 1- or 2-years after beginning treatment.

Use of SABR to treat HCC (vs. radiofrequency ablation)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
					Studies reported mainly OS at 1- and 2-years post treatment. The best evidence on OS is provided by the retrospective observational study by Wahl et al (2016), which included 224 patients and compared SABR with radiofrequency ablation (RFA). It reported OS at 1 and 2 years of 69.6% and 52.9% after RFA and 74.1% and 46.3% after SABR. Given the alternative treatment options with the possibility of curative intent for patients with non-metastatic HCC overall survival is a clinically meaningful outcome for patients. The study performed a retrospective comparison between the two groups. Despite the use of a statistical method (propensity score matching) to account for baseline differences among the participants, patient selection and outcome detection bias cannot be excluded between the two cohorts. The 1- and 2-years LC rates reported by Wahl et al. (2016) are comparable to the SABR results reported by the Rim et al. (2019) meta-analysis of non-comparative studies. Overall, there is some uncertainty about this outcome.
Local control	Wahl, 2016 Kim, 2019	6 6	Direct Direct	C	Local control (LC) is the proportion of patients for which the treated lesion does not increase in size at a defined follow-up point after beginning treatment.

Use of SABR to treat HCC (vs. radiofrequency ablation)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
					The best evidence on LC is provided by the retrospective observational study by Wahl et al (2016) that included 224 patients and compared SABR with radiofrequency ablation (RFA). The study reported LC at 1 and 2 years of 97.4% and 83.6% with SABR and 83.6% and 80.2% with RFA. After adjusting for tumour size LC was no statistically different between the two groups. The clinical benefit to the patient group is that a less invasive treatment such as SABR can provide equivalent results. The study performed a retrospective comparison between the two groups. Despite the use of a statistical method (propensity score matching) to account for baseline differences among the participants, patient selection and outcome detection bias cannot be excluded between the two cohorts. The 1- and 2-years LC rates reported by Wahl et al. (2016) are comparable to the SABR results reported by the Rim et al. (2019) meta-analysis of non-comparative studies. Overall there is some uncertainty about this outcome.
Toxicity	Wahl, 2016 Kim, 2019	6 6	Direct Direct	B	Toxicity is defined based on the number and severity of adverse events a patient can experience after undergoing treatment. Treatment-related toxicity in patients with cancer is usually recorded and graded

Use of SABR to treat HCC (vs. radiofrequency ablation)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
					according to the Common Toxicity Criteria Adverse Events (CTCAE). The best evidence on toxicity is provided by the retrospective observational study by Wahl et al (2016) that included 224 patients and compared SABR with RFA. The rates of late grade 3+ GI toxicity in the study were similar in the RFA and SABR groups at 1 (3.4% v 5.4%; p =0.49) and 2 years (6.4% v 8.3%; p =0.66). There were no late grade 5 adverse events in either group. Given the alternative treatment options with different toxicity profiles for patients with HCC, toxicity is clinically meaningful outcome. This outcome is even more important for patients with advanced disease that treatment-related toxicity may result in significant impairment of their quality of life. The study performed a retrospective comparison between the two groups. Despite the use of a statistical method (propensity score matching) to account for baseline differences among the participants, patient selection and outcome detection bias cannot be excluded between the two cohorts. Overall, there is considerable uncertainty about this outcome and additional randomised control studies will need to verify this finding.

Use of SABR to treat HCC (vs. radiofrequency ablation)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
LC = local control PFS = progression free survival QoL = quality of life RFA = radiofrequency ablation OS = overall survival					

Use of SABR to treat HCC (non-comparative studies)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
Actuarial overall survival	Rim, 2019	7	Direct	B	Actuarial overall survival is reported as the proportion of patients surviving at a defined follow-up point, such as 1- or 2-years after beginning treatment. The best non-comparative evidence on actuarial survival is provided by the Rim et al. (2019) systematic review and meta-analysis that included 32 observational single-arm studies (n=1950 patients) and reported 1-year OS of 72.6% (95% CI 65.7–78.6) and 2-year at 57.8% (95% CI 50.9–64.4). Given the alternative treatment options with the potential of curative intent for patients without

Use of SABR to treat HCC (non-comparative studies)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
					metastatic HCC, overall survival is a clinically meaningful outcome. Actuarial overall survival was a primary outcome in a number of the studies included in the systematic review. However, almost none of them reported sample size calculations. There is some consistency between the results reported by Rim et al. (2019) and the OS evidence for SABR provided by Wahl et al. (2016) and Parikh et al. (2018) as the 1-year 95%CI show overlap with the same outcome reported in these studies. Differences in the included population and treatment could account for the different rates observed among studies. The results were less consistent for the 2-year OS rates.
Local control	Rim, 2019	7	Direct	B	Local control (LC) is the proportion of patients for which the treated lesion does not increase in size at a defined follow-up point after beginning treatment. The best non-comparative evidence on LC is provided by the Rim et al. (2019) systematic review and meta-analysis that included 32 observational single-arm studies (n=1950 patients) and reported 1-year LC of 85.7% (95% CI 80.1-90.0) and 2-years LC of 83.6% (95% CI 77.4-88.3).

Use of SABR to treat HCC (non-comparative studies)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
					The clinical benefit to the patient group is that a less invasive treatment such as SABR can provide good LC results. LC was a secondary outcome in most of the studies included in the meta-analysis. There is some consistency between the results reported by Rim et al. (2019) and the LC evidence for SABR provided by Wahl et al. (2016) and Parikh et al. (2018) as the 1-year 95%CI show overlap with the same outcome reported in these studies.
Toxicity	Rim, 2019	7	Direct	B	Toxicity is defined based on the number and severity of adverse events a patient can experience after undergoing treatment. Treatment-related toxicity in patients with cancer is usually recorded and graded according to the Common Toxicity Criteria Adverse Events (CTCAE). Treatment-related toxicity was a secondary outcome in all studies. The best non-comparative evidence on actuarial survival is provided by the Rim et al. (2019) systematic review and meta-analysis that included 32 observational single-arm studies (n=1950 patients) and reported grade ≥3 GI and hepatic complications of 3.9% and 4.7%, respectively.

Use of SABR to treat HCC (non-comparative studies)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
					Given the alternative treatment options with different toxicity profiles for patients with HCC, toxicity is clinically meaningful outcome. This outcome is even more important for patients with advanced disease that treatment-related toxicity may result in significant impairment of their quality of life. There is consistency between the results reported by Rim et al. (2019) and the evidence from comparative studies with grade ≥3 rates <10%.

Quality of life	Klein, 2015	6	Direct	C	Quality of life (QoL) is a composite patient-reported outcome that captures the impact of an intervention on a patient's psychology and everyday life activities. The best evidence on QoL is provided by the prospective, non-comparative observational study by Klein et al (2015) that included 99 patients with hepatocellular carcinoma and captured QoL outcomes up to 12 months post SABR treatment. The study did not report a difference in QoL between baseline (137.4) and after SABR treatment (3 months = 133.4, 12 months = 135.1) using the Functional Assessment of Cancer Therapy-Hepatobiliary (FACT-Hep) checklist. No difference was also reported using the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core-30 (EORTC QLQ-C30) checklist with baseline=65.8, 3-months=62.9 and 12-months=64.5.One of the factors weighting in treatment decisions for HCC is the possible impact that treatment may have on their quality of life. Given that SABR is less invasive than other forms of treatment for HCC this is a clinically important outcome. The SABR group was heterogeneous including patients with HCC, cholangiocarcinoma, and liver metastases. Maximum follow-up was only 12 months and it is unknown what proportion of patients completed follow-up. Overall there is considerable uncertainty about this outcome.
EORTC QLQ-C30 = European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core-30					

Use of SABR to treat HCC (non-comparative studies)					
Outcome Measure	**Reference**	**Quality of Evidence Score**	**Applicability**	**Grade of Evidence**	**Interpretation of Evidence**
FACT-Hep = Functional Assessment of Cancer Therapy-Hepatobiliary LC = local control PFS = progression free survival QoL = quality of life OS = overall survival					

9 Literature search terms – PICO table

P –Population and Indication Describe the relevant population and indication provided previously including if necessary, disease severity or duration, previous treatment, new or recurrent symptoms, any specific co-morbidities and other population factors (for example, age range). Add details of any subgroups or stratifications for which separate evidence may be required.	Patients of all ages with localised hepatocellular carcinoma (HCC) with or without low burden metastatic disease who are/ have: • unsuitable for surgery (resection or transplant) • unsuitable for or refractory to radiofrequency ablation • unsuitable for or refractory to transhepatic arterial chemo-embolisation (TACE) • have five or fewer discrete intrahepatic parenchymal foci of HCC • HCC tumour 6 cm or less • Low burden of disease defined as: extrahepatic metastases or malignant lymph nodes (that enhance with typical features of HCC <3.0 cm in sum of maximal diameters (e.g. 2 lung lesions <3cm in total diameter). • Childs-Pugh Class A only (Childs Pugh scoring system classification).
I – Intervention Describe the intervention details provided previously including if necessary, details of treatment, mode of delivery, size/frequency/duration of dose, position of intervention in treatment pathway (e.g. first/second line/salvage) and any background/ concomitant medication	Stereotactic Ablative Body Radiotherapy (SBRT or SABR) alone [include studies that deliver SABR in 10 or less fractions; the NHS uses 5 fractions and studies using 5 fractions are therefore of particular relevance]

C – Comparators
What is/are the main alternative/s to compare with the intervention being considered?
Describe the comparator details provided previously including if necessary, details of treatment, mode of delivery, size/frequency/duration of dose, position of intervention in treatment pathway (e.g. first/second line/salvage) and any background/ concomitant medication
	• Targeted/ biological agents
	  ○ Sorafenib
	  ○ Lenvatinib
	• Thermal ablation (radiofrequency ablation or microwave ablation)
	• Standard fractionated radiotherapy
	• No treatment (best supportive care)

O – Outcomes Outcomes should be patient focussed and relate to those detailed in the PPP and the Research Questions covering clinical effectiveness, safety and cost effectiveness as required.	*Critical to decision-making:* • 1 year overall survival • 2 year overall survival • Median overall survival • Local control (i.e. tumour regression/ resolution OR no tumour progression within the lesion treated/ treatment field)

Examples will be topic specific but might include intermediate or short-term outcomes; mortality; morbidity; quality of life; treatment complications; adverse effects; rates of relapse; late morbidity and re-admission; return to work, physical and social functioning, resource use.	• Progression free survival • Side effects and adverse events (including but not limited to change (reduction) in liver function (Child-Pugh score) • Acute and late radiotherapy toxicity (including but not limited to fatigue, nausea, diarrhoea and gastro-intestinal ulcers or perforations). • Quality of Life. *Important to decision-making:* • Cost effectiveness.
Inclusion criteria	
Study design	Systematic reviews, randomised controlled trials, controlled clinical trials, cohort studies. If no higher level quality evidence is found, case series can be considered.
Language	English only
Patients	Human studies only
Age	All ages
Date limits	2009-2019
Exclusion criteria	
Publication type	Conference abstracts, non-systematic reviews, narrative reviews, commentaries, letters and editorials
Study design	Case reports, resource utilisation studies
In addition to the above criteria, any study with a patient population of <30 patients was also excluded.	

10 Search Strategy

Total number of references: 1275

Total following de-duplication: 859

- Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily 1946 to May 16, 2019
- 21st May 2019

1	(primary and ((hepatocellular or liver) adj3 (cancer* or carcinoma* or tumour* or mass* or growth* or lesion*))).tw.	17844
2	(Fibrolamellar adj3 (HCC or hepatocell* or carcinoma*)).tw.	558
3	Carcinoma, Hepatocellular/ or HCC.tw. or "hepatocellular carcinoma".kw.	93734
4	or/1-3	102389
5	(SABR or SBRT or SABRT or SRS or stereotactic ablati* or stereotactic body radio* or stereotactic radio*).tw.	17806
6	(arc therap* or vmat).tw.	2923
7	(hypofraction* or hypo-fraction* or hypo fraction*).tw.	3145
8	(cyber knife* or cyberknife* or gamma knife* or gammaknife*).tw.	5518
9	exp Radiosurgery/ or radiosurg*.tw.	18195
10	or/5-9	31828
11	4 and 10	537
12	limit 11 to yr="2009 -Current"	486
13	(editorial or letter or case report or comment or news).pt.	1907311
14	12 not 13	465

- Embase 1974 to 2019 Week 20
- 21st May 2019

1	(primary and ((hepatocellular or liver) adj3 (cancer* or carcinoma* or tumour* or mass* or growth* or lesion*))).tw.	26865
2	(Fibrolamellar adj3 (HCC or hepatocell* or carcinoma*)).tw.	744
3	liver cell carcinoma/ or HCC.tw. or "hepatocellular carcinoma".kw.	153610
4	or/1-3	164347
5	(SABR or SBRT or SABRT or SRS or stereotactic ablati* or stereotactic body radio* or stereotactic radio*).tw.	30875
6	(arc therap* or vmat).tw.	7757
7	(hypofraction* or hypo-fraction* or hypo fraction*).tw.	6716
8	(cyber knife* or cyberknife* or gamma knife* or gammaknife*).tw.	8591
9	gamma knife radiosurgery/ or stereotactic body radiation therapy/ or stereotactic radiosurgery/	23796
10	or/5-9	53120
11	4 and 10	1451
12	limit 11 to yr="2009 -Current"	1386
13	(editorial or letter or case report or comment or news or conference abstract or Conference Paper or Conference Review).pt.	5838065
14	12 not 13	712

- Cochrane (CDSR and CENTRAL)
- 21st May 2019

ID	Search	Hits
#1	(hepatocellular carcinoma) OR (liver NEAR/3 (cancer* OR carcinoma* OR tumour* OR mass* OR growth* OR lesion*))	7334
#2	Fibrolamellar NEAR3 (HCC or hepatocell* or carcinoma*)	7
#3	[mh "Carcinoma, Hepatocellular"] OR HCC OR "hepatocellular carcinoma":kw	3742
#4	(Chenelle et al.-#3)	7661
#5	(SABR or SBRT or SABRT or SRS or "stereotactic ablati*" or "stereotactic body radio*" or "stereotactic radio*"):ti,ab,kw	1400

#6	(arc therap* or vmat):ti,ab,kw	816
#7	(hypofraction* or hypo-fraction* or hypo fraction*):ti,ab,kw	833
#8	(cyber knife* or cyberknife* or gamma knife* or gammaknife*):ti,ab,kw	208
#9	[mh "Radiosurgery"] or radiosurg*	789
#10	(Siva et al.-#9)	3412
#11	#4 and #10 with Cochrane Library publication date from Jan 2009 to present	98

11 Evidence selection

- Total number of publications reviewed: 859
- Total number of publications considered relevant: 127
- Total number of publications selected for inclusion in this briefing: 7

References from the PWG supplied in the PPP		Paper selection decision and rationale if excluded
1	Bujold A, Massey CA, Kim JJ, et al. Sequential phase I and II trials of stereotactic body radiotherapy for locally advanced hepatocellular carcinoma. J Clin Oncol 2013;31(13):1631–1639	This study was excluded because it was included in the systematic review and meta-analysis by Rim et al. (2019).
2	Yoon, S., Ryoo, B., Lee, S., Kim, J., Shin, J., An, J., Lee, H. and Lim, Y. (2018). Efficacy and Safety of Transarterial Chemoembolization Plus External Beam Radiotherapy vs Sorafenib in Hepatocellular Carcinoma With Macroscopic Vascular Invasion. *JAMA Oncology*, 4(5), p.661.	This study was excluded because it did not meet the scope as outlined in section 9 (intervention).
3	Wahl, D., Stenmark, M., Tao, Y., Pollom, E., Caoili, E., Lawrence, T., Schipper, M. and Feng, M. (2016). Outcomes After Stereotactic Body Radiotherapy or Radiofrequency Ablation for Hepatocellular Carcinoma. *Journal of Clinical Oncology*, 34(5), pp.452-459.	This study was included in the review.

12 References

BETTINGER, D., PINATO, D. J., SCHULTHEISS, M., SHARMA, R., RIMASSA, L., PRESSIANI, T., BURLONE, M. E., PIRISI, M., KUDO, M., PARK, J. W., BUETTNER, N., NEUMANN-HAEFELIN, C., BOETTLER, T., ABBASI-SENGER, N., ALHEIT, H., BAUS, W., BLANCK, O., GERUM, S., GUCKENBERGER, M., HABERMEHL, D., OSTHEIMER, C., RIESTERER, O., TAMIHARDJA, J., GROSU, A. L., THIMME, R., BRUNNER, T. B. & GKIKA, E. 2018. Stereotactic Body Radiation Therapy as an Alternative Treatment for Patients with Hepatocellular Carcinoma Compared to Sorafenib: A Propensity Score Analysis. *Liver Cancer*, 1-14.

BRUIX, J. & SHERMAN, M. 2011. Management of hepatocellular carcinoma: an update. *Hepatology,* 53, 1020-2.

CHENELLE, C. T., ITAGAKI, T., FISHER, D. F., BERRA, L. & KACMAREK, R. M. 2017. Performance of the PneuX System: A Bench Study Comparison With 4 Other Endotracheal Tube Cuffs. *Respiratory Care,* 62, 102.

KIM, N., KIM, H. J., WON, J. Y., KIM, D. Y., HAN, K. H., JUNG, I. & SEONG, J. 2019. Retrospective analysis of stereotactic body radiation therapy efficacy over radiofrequency ablation for hepatocellular carcinoma. *Radiotherapy & Oncology,* 131, 81-87.

KLEIN, J., DAWSON, L. A., JIANG, H., KIM, J., DINNIWELL, R., BRIERLEY, J., WONG, R., LOCKWOOD, G. & RINGASH, J. 2015. Prospective Longitudinal Assessment of Quality of Life for Liver Cancer Patients Treated With Stereotactic Body Radiation Therapy. *International Journal of Radiation Oncology, Biology, Physics,* 93, 16-25.

PARIKH, N. D., MARSHALL, V. D., GREEN, M., LAWRENCE, T. S., RAZUMILAVA, N., OWEN, D., SINGAL, A. G. & FENG, M. 2018. Effectiveness and cost of radiofrequency ablation and stereotactic body radiotherapy for treatment of early-stage hepatocellular carcinoma: An analysis of SEER-medicare. *Journal of Medical Imaging & Radiation Oncology,* 62, 673-681.

QI, X., TANG, Y., AN, D., BAI, M., SHI, X., WANG, J., HAN, G. & FAN, D. 2014. Radiofrequency ablation versus hepatic resection for small hepatocellular carcinoma: a meta-analysis of randomized controlled trials. *J Clin Gastroenterol,* 48, 450-7.

RAJYAGURU, D. J., BORGERT, A. J., SMITH, A. L., THOMES, R. M., CONWAY, P. D., HALFDANARSON, T. R., TRUTY, M. J., KURUP, A. N. & GO, R. S. 2018. Radiofrequency Ablation Versus Stereotactic Body Radiotherapy for Localized Hepatocellular Carcinoma in Nonsurgically Managed Patients: Analysis of the National Cancer Database. *Journal of Clinical Oncology,* 36, 600-608.

RIM, C. H., KIM, H. J. & SEONG, J. 2019. Clinical feasibility and efficacy of stereotactic body radiotherapy for hepatocellular carcinoma: A systematic review and meta-analysis of observational studies. *Radiotherapy & Oncology,* 131, 135-144.

SCORSETTI, M., COMITO, T., COZZI, L., CLERICI, E., TOZZI, A., FRANZESE, C., NAVARRIA, P., FOGLIATA, A., TOMATIS, S., D'AGOSTINO, G., IFTODE, C., MANCOSU, P., CERIANI, R. & TORZILLI, G. 2015. The challenge of inoperable hepatocellular carcinoma (HCC): results of a single-institutional experience on stereotactic body radiation therapy (SBRT). *Journal of Cancer Research & Clinical Oncology,* 141, 1301-9.

SHINDE, A., JONES, B. L., CHEN, Y.-J. & AMINI, A. 2018. Radiofrequency Ablation Versus Stereotactic Body Radiotherapy for Localized Hepatocellular Carcinoma: Does Radiation Dose Make a Difference? *Journal of clinical oncology: official journal of the American Society of Clinical Oncology,* 36, 2566.

SIVA, S., BRESSEL, M., MURPHY, D. G., SHAW, M., CHANDER, S., VIOLET, J., TAI, K. H., UDOVICICH, C., LIM, A., SELBIE, L., HOFMAN, M. S., KRON, T., MOON, D., GOAD, J., LAWRENTSCHUK, N. & FOROUDI, F. 2018. Stereotactic Abative Body Radiotherapy (SABR) for Oligometastatic Prostate Cancer: A Prospective Clinical Trial. *European Urology,* 74, 455-462.

WAHL, D. R., STENMARK, M. H., TAO, Y., POLLOM, E. L., CAOILI, E. M., LAWRENCE, T. S., SCHIPPER, M. J. & FENG, M. 2016. Outcomes After Stereotactic Body Radiotherapy or Radiofrequency Ablation for Hepatocellular Carcinoma. *Journal of Clinical Oncology,* 34**,** 452-9.

13 Appendices

13.1 Quality of evidence scores

13.1.1 Comparative studies

Rajyaguru et al. 2018

Each quality item is scored as follows: · **Yes= 2** · **In part = 1** · **No= 0**	**Score**
1. Are the research questions/aims and design clearly stated?	**2**
2. Is the research design appropriate for the aims and objectives of the research?	**1**
3. Are the methods clearly described?	**2**
4. Is the data adequate to support the authors' interpretation/conclusions?	**0**
5. Are the results generalizable?	**0**
Total	**5**

Parikh et al. 2018

Each quality item is scored as follows: · **Yes= 2** · **In part = 1** · **No= 0**	**Score**
1. Are the research questions/aims and design clearly stated?	**2**
2. Is the research design appropriate for the aims and objectives of the research?	**1**
3. Are the methods clearly described?	**1**
4. Is the data adequate to support the authors' interpretation/conclusions?	**1**
5. Are the results generalizable?	**0**
Total	**5**

Bettinger et al (2018

Each quality item is scored as follows: · **Yes= 2** · **In part = 1** · **No= 0**	**Score**
1. Are the research questions/aims and design clearly stated?	**2**

2. Is the research design appropriate for the aims and objectives of the research?	**1**
3. Are the methods clearly described?	**1**
4. Is the data adequate to support the authors' interpretation/conclusions?	**1**
5. Are the results generalizable?	**1**
Total	**6**

Wahl et al. 2016

Each quality item is scored as follows: · **Yes= 2** · **In part = 1** · **No= 0**	**Score**
1. Are the research questions/aims and design clearly stated?	**2**
2. Is the research design appropriate for the aims and objectives of the research?	**1**
3. Are the methods clearly described?	**1**
4. Is the data adequate to support the authors' interpretation/conclusions?	**1**
5. Are the results generalizable?	**1**
Total	**6**

Kim et al. 2019

Each quality item is scored as follows: · **Yes= 2** · **In part = 1** · **No= 0**	**Score**
1. Are the research questions/aims and design clearly stated?	**2**
2. Is the research design appropriate for the aims and objectives of the research?	**1**
3. Are the methods clearly described?	**1**
4. Is the data adequate to support the authors' interpretation/conclusions?	**1**
5. Are the results generalizable?	**1**
Total	**6**

13.1.2 Non-comparative studies

Rim et al (2019)

Each quality item is scored as follows: · **Yes= 2** · **In part = 1** · **No= 0**	**Score**

1. Are the research questions/aims and design clearly stated?	**2**
2. Is the research design appropriate for the aims and objectives of the research?	**2**
3. Are the methods clearly described?	**1**
4. Is the data adequate to support the authors' interpretation/conclusions?	**1**
5. Are the results generalizable?	**1**
Total	**7**

Klein et al (2015)

Each quality item is scored as follows: · **Yes= 2** · **In part = 1** · **No= 0**	**Score**
1. Are the research questions/aims and design clearly stated?	**2**
2. Is the research design appropriate for the aims and objectives of the research?	**1**
3. Are the methods clearly described?	**2**
4. Is the data adequate to support the authors' interpretation/conclusions?	**1**
5. Are the results generalizable?	**0**
Total	**6**