NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Appraisal consultation document

Selective internal radiation therapies (SIRTs) for treating hepatocellular carcinoma

The Department of Health and Social Care has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using QuiremSpheres, SIR-Spheres and TheraSphere in the NHS in England. The appraisal committee has considered the evidence submitted and the views of non-company consultees and commentators, clinical experts and patient experts.

This document has been prepared for consultation with the consultees. It summarises the evidence and views that have been considered, and sets out the recommendations made by the committee. NICE invites comments from the consultees and commentators for this appraisal and the public. This document should be read along with the evidence (see the committee papers).

The appraisal committee is interested in receiving comments on the following:

- Has all of the relevant evidence been taken into account?
- Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
- Are the recommendations sound and a suitable basis for guidance to the NHS?
- Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of race, gender, disability, religion or belief, sexual orientation, age, gender reassignment, pregnancy and maternity?

Issue date: November 2019

Note that this document is not NICE's final guidance on these technologies. The recommendations in section 1 may change after consultation.

After consultation:

- The appraisal committee will meet again to consider the evidence, this appraisal consultation document and comments from the consultees.
- At that meeting, the committee will also consider comments made by people who are not consultees.
- After considering these comments, the committee will prepare the final appraisal document.
- Subject to any appeal by consultees, the final appraisal document may be used as the basis for NICE's guidance on using QuiremSpheres, SIR-Spheres and TheraSphere in the NHS in England.

For further details, see NICE's guide to the processes of technology appraisal.

The key dates for this appraisal are:

Closing date for comments: 8 January 2020

Second appraisal committee meeting: 22 January 2020

Details of membership of the appraisal committee are given in section 5.

1 Recommendations

1.1 The selective internal radiation therapies (SIRTs) QuiremSpheres, SIR-Spheres and TheraSphere are not recommended, within their CE marking, for treating hepatocellular carcinoma in adults.

Why the committee made these recommendations

Treatment for hepatocellular carcinoma (HCC) depends on the stage of cancer and the liver function. It includes surgery, ablation, transarterial therapies, chemotherapy (such as lenvatinib and sorafenib) and best supportive care. Treatment does not cure the disease for most people.

QuiremSpheres, SIR-Spheres and TheraSphere are SIRTs. These are small radioactive beads that are injected into the liver's blood supply to treat liver cancer. In clinical trials, SIR-Spheres has not been shown to improve survival compared with available treatment options. There is very limited clinical evidence to compare the effectiveness of QuiremSpheres and TheraSphere with other treatments. Also, there are not enough data to compare the effectiveness of the 3 SIRTs with each other.

There is not enough evidence to consider SIRTs a cost-effective use of NHS resources for early and intermediate stage HCC. For people with advanced stage HCC, the economic analysis shows that SIRTs are less clinically effective and cost more than lenvatinib or sorafenib. Because of this, SIRTs are not recommended.

2 Information about QuiremSphere, SIR-Spheres and TheraSphere

QuiremSpheres (Terumo Europe)	
CE marking	QuiremSpheres received its CE mark on 1 April 2015. It is classified as an Active Implantable Medical Device (AIMD) by Council Directive 90/385/EEC. It is indicated for treating unresectable liver tumours.

Dosage in the CE mark	The company has stated that the typical number of particles that are given by QuiremSpheres is approximately 20 to 30 million.
Price	The company has stated that the cost of QuiremSpheres is £9,896 for a single treatment. The company has a commercial arrangement (simple discount patient access scheme), which would have applied if the technology had been recommended.

SIR-Spheres (SIRTEX)	
CE marking	SIR-Spheres received its CE mark as a class III active medical device in October 2002. It is indicated for treating advanced inoperable liver tumours.
Dosage in the CE mark	SIR-Spheres is given through a catheter to the hepatic artery. It is supplied at 3 GBq yttrium-90 per vial in 5 ml water for injection in a shielded shipping vial. Each vial contains 40 to 80 million microspheres, ranging from 20 to 60 micrometres in diameter (median diameter 32.5 micrometres). The maximum range of beta emission in tissue is 11 mm with a mean of 2.5 mm. The average number of particles implanted is 30×10^6 to 60×10^6.
Price	The company has stated that the cost of SIR-Spheres is £8,000 for a single treatment. Costs may vary in different settings because of negotiated procurement discounts.

TheraSphere (BTG)	
CE marking	TheraSphere received its CE mark as a class III active medical device in September 2014. It is indicated for treating hepatic neoplasia.
Dosage in the CE mark	TheraSphere is given through a catheter to the hepatic artery. It is supplied in 6 dose sizes: 3 GBq, 5 GBq, 7 GBq, 10 GBq, 15 GBq or 20 GBq in 0.6 ml pyrogen-free water supplied in a 1 ml vial, inside an acrylic shield. Custom dose sizes are also available in increments of 0.5 GBq between 3 GBq and 20 GBq. A single treatment with TheraSphere contains 1.2 to 8 million microspheres. The recommended dose to the liver is 80 Gy to 150 Gy.
Price	The company has stated that the cost of Thera-Spheres is £8,000 for a single treatment. Costs may vary in different settings because of negotiated procurement discounts.

3 Committee discussion

The appraisal committee (section 5) considered evidence from a number of sources. See the committee papers for full details of the evidence.

Potential new treatment option

People with hepatocellular carcinoma would welcome a new treatment option

3.1 Hepatocellular carcinoma (HCC) is the most common form of liver cancer in England. Treatment depends on the location and stage of the cancer, and how well the liver is functioning. Treatment options include surgery or ablation in early disease, transarterial therapies in intermediate stage disease, and chemotherapy in advanced stage disease, as well as best supportive care. Treatment does not cure the disease for many people. Patient experts explained that HCC can have a substantial impact on quality of life. People with HCC and their carers live with uncertainty and hopelessness. Often people with HCC also live with stigma and isolation because of underlying causes of disease such as alcohol. Clinical experts highlighted that people with advanced HCC have a poor prognosis with median life expectancy of less than 12 months. The committee concluded that people with HCC would welcome a new treatment option.

People with HCC and portal vein thrombosis are a relevant subgroup

3.2 The clinical experts explained that portal vein involvement, such as portal vein thrombosis (PVT), is a common comorbidity that might negatively affect prognosis. PVT happens when a blood clot narrows the vein that takes blood to the liver from the intestines. The committee understood that people with PVT were included in the NICE scope for this appraisal. It concluded that evidence for people with HCC and PVT should be considered.

This appraisal assesses 3 selective internal radiation therapies (SIRTs) for treating hepatocellular carcinoma

3.3 QuiremSpheres, SIR-Spheres and TheraSphere are SIRTs. These are small radioactive beads that are injected into the liver's blood supply to treat liver cancer. The 3 SIRTs are medical devices with CE marks for their licenced indications. QuiremSpheres is indicated for treating unresectable liver tumours, SIR-Spheres for treating advanced inoperable liver tumours and TheraSphere for treating hepatic neoplasia. The committee was aware that the scope for the appraisal was narrower than the CE marks, because it only included unresectable hepatocellular carcinoma. The committee agreed that the 3 SIRTs should be compared with each other and with available treatments to assess their cost effectiveness for treating hepatocellular carcinoma.

SIRTs might have fewer and less severe side effects than other treatment options

3.4 Clinical and patient experts stated that there were fewer and less severe side effects with SIRTs than with other treatments. Also, side effects from SIRTs are temporary, whereas side effects from chemotherapies such as sorafenib and lenvatinib can continue for the whole treatment course. The clinical experts also stated that SIRTs might extend life expectancy in advanced stage disease. The committee agreed that SIRTs might have fewer and less severe side effects than current treatments.

SIRTs are already used in the NHS, but not for HCC

3.5 The clinical experts and NHS England explained that SIRTs are available in some specialist centres across England for other cancers (such as metastatic colorectal cancer). The clinical experts explained that SIRTs for HCC have been used in England through compassionate schemes, but are not currently available through routine commissioning. The committee understood that SIRTs are currently not commissioned for HCC in the NHS but that the infrastructure exists in some specialist centres.

Clinical management

Stage of cancer and liver function characterises the disease and therefore people with HCC are a heterogenous population

3.6 There are different causes of HCC, including cirrhosis, alcohol, fatty liver disease and hepatitis. Therefore, people with HCC are a heterogenous population and their disease is characterised by both cancer and liver function. Treatment choice is multifaceted because both the cancer and liver function affect treatment outcomes. Clinical experts advised that clinicians use the Barcelona Clinic Liver Cancer (BCLC) staging system and the Child-Pugh score to help treatment decisions.

- BCLC staging looks at the number and size of tumours in the liver. There are 5 stages: very early stage (BCLC 0), early stage (BCLC A), intermediate stage (BCLC B), intermediate stage (BCLC C) and terminal stage (BCLC D). The committee agreed that stages A, B and C align with the scope for this appraisal.
- The Child-Pugh score looks at the liver function. It has 5 components: serum albumin levels, bilirubin levels, time for blood to clot, presence of ascites (fluid in the peritoneal cavity) and presence of hepatic encephalopathy. There are 3 classes: class A (the liver is working normally), class B (mild to moderate liver damage), class C (severe liver damage). Clinical experts advised that the BCLC stage and the Child-Pugh score together inform treatment choice. People with BCLC A to C can have either good liver function (Child-Pugh A) or mild to moderate liver damage (Child-Pugh B).
- More recently an alternative measure, the albumin-bilirubin (ALBI) grade, was developed to look at liver function. The committee was aware that in previous NICE guidance for HCC, the Child-Pugh score was used as a criterion for treatment, but that ALBI was not. The committee noted that both might help to inform treatment decisions. The clinical experts advised that ALBI is less frequently used for this

purpose, and that Child-Pugh is expected to be the measure of choice for the foreseeable future.

Treatment of HCC differs between the 3 BCLC stages and is influenced by Child-Pugh score

3.7 Treatment options include ablation and transplant in early disease, and conventional transarterial therapies (CTT) such as transarterial chemoembolisation (TACE) or transarterial embolisation (TAE) in intermediate stage disease. In advanced stage disease, treatment options are chemotherapy or systemic therapy with sorafenib or lenvatinib or regorafenib. In some people the aim of treatment might be to reduce the tumour size ('downstaging') to allow subsequent transplantation that could cure the disease. The committee understood that people with HCC have different treatment options depending on the stage of their disease as assessed by BCLC and Child-Pugh.

There are 3 distinct subgroups relevant to this appraisal

3.8 The committee concluded that there are 3 subgroups relevant for this appraisal:

- People for whom liver transplant is appropriate, including people with BCLC A and Child-Pugh A or B.
- People for whom CTT is appropriate, including people with BCLC B and Child-Pugh A or B.
- People for whom CTT is inappropriate, including people with BCLC C and Child-Pugh A or B.

In people with early stage disease, ablation and transplant are the standard of care in current NHS practice in England

3.9 Treatment options for early stage disease (BCLC A) are ablation and transplant. However, 1 clinical expert explained that transplants might not be available for people with good liver function (Child-Pugh A). The

committee concluded that both ablation and transplant are the standard of care for people with early stage disease in clinical practice in England.

In people with intermediate stage disease, CTTs are the standard of care in current NHS practice in England

3.10 Treatment for intermediate stage disease (BCLC B) are CTTs including transarterial chemoembolisation (TACE), drug-eluting bead transarterial chemoembolisation (DEB-TACE) and transarterial embolisation (TAE). The committee accepted that all CTTs available in the NHS in England are appropriate comparators for people with intermediate stage disease.

In people with advanced stage disease, sorafenib is the standard of care in current NHS practice in England

3.11 Systemic therapies, sorafenib and lenvatinib are both recommended for advanced HCC (BCLC C) in people with Child-Pugh grade A liver impairment (NICE technology appraisal guidance on sorafenib for treating advanced hepatocellular carcinoma and lenvatinib for untreated advanced hepatocellular carcinoma). Regorafenib is only recommended after treatment with sorafenib (NICE technology appraisal guidance on regorafenib for previously treated advanced hepatocellular carcinoma). The committee understood that sorafenib is the standard of care in clinical practice in England because there are subsequent treatments available after progression with sorafenib. Lenvatinib is now rarely used. The committee concluded that sorafenib is the appropriate comparator for SIRTs in people with advanced stage disease and with Child-Pugh grade A.

Clinical evidence

The systematic review included non-randomised controlled trials (RCTs) when not enough RCT evidence was identified

3.12 The assessment group (AG) did a systematic review of the clinical evidence on SIRTs and comparators. The research protocol is registered

on PROSPERO, the international prospective register of systematic reviews in health and social care; registration number CRD42019128383. RCTs were eligible for inclusion in the review. The AG had identified all the RCTs that were also identified by the companies in their submissions. The committee was aware of non-RCT evidence and agreed with the AG's approach to only include non-RCT evidence in the review when there was not enough RCT evidence. The committee understood that some studies might include a mixed population. It agreed to exclude these studies from the network meta-analyses if they did not provide separate results for the 3 subgroups of interest (see section 3.8). The committee used the AG's report for its decision making. This was because it included evidence for all 3 SIRTs and so was more comprehensive than the companies' submissions.

There is not enough clinical evidence for QuiremSpheres in the 3 subgroups relevant to this appraisal

3.13 The clinical evidence for QuiremSpheres came from 1 retrospective case series including 9 people that showed a 56% response rate. The committee heard that a mixed population was included, and results were only presented for the whole study population. The committee concluded that the single, small retrospective study did not provide enough data to assess clinical effectiveness of QuiremSpheres in any of the 3 subgroups relevant to this appraisal (see section 3.8).

There is limited randomised clinical evidence with a high risk of bias for TheraSphere compared with TACE for people when transplant is appropriate

3.14 The committee heard that 2 small RCTs (PREMIERE and Kulik 2014) for TheraSphere were identified that included people for whom transplant is appropriate (see section 3.8). The committee was also aware of 10 non-RCT studies, including 7 prospective comparative studies that included people from the 3 subgroups relevant to this appraisal. The PREMIERE study was done in the US and included 45 people for whom transplant would be appropriate. It compared TheraSphere with TACE as an

alternative to prepare for transplant. The AG advised that PREMIERE had a high risk of bias because of concerns with randomisation and potential deviations from the intended interventions. Also, the baseline characteristics were different in the 2 arms so that people in the TACE arm had better prognosis than people in the TheraSphere arm. Overall survival of people who had a transplant was numerically, but not statistically significantly, longer in the TheraSphere arm. The median overall survival was 18.6 months (95% confidence interval [CI] 7.4 to 32.5) compared with 17.7 months (95% CI 7.4 to 32.5). The committee concluded that there was limited evidence, with a high risk of bias, to establish whether TheraSphere was better than TACE in people for whom transplant is appropriate.

There is limited evidence with high risk of bias for TheraSphere compared with TheraSphere with sorafenib in people when transplant is appropriate

3.15 The study by Kulik 2014 was done in the US and included 20 people for whom transplant would be appropriate. It compared TheraSphere with TheraSphere and sorafenib in combination. The AG had some concerns with the randomisation process, potential deviations from the intended interventions and measurement of outcomes. The baseline characteristics were different in the 2 arms so that people in the TheraSphere plus sorafenib arm had a better prognosis. There was no difference in overall survival between the 2 arms (3 deaths in the TheraSphere arm, 2 deaths in the combination arm). The committee was aware that TheraSphere with sorafenib in combination was not included in the licence of sorafenib or the CE mark of TheraSphere. The committee concluded that there was limited evidence with high risk of bias to establish whether TheraSphere is better than TheraSphere with sorafenib in people when transplant is appropriate.

Non-randomised evidence comparing TheraSphere with non-SIRT treatments is not robust and should not be used for decision making

3.16 Of the 7 prospective comparative non-RCTs, only 4 reported overall survival or progression-free survival. Of these, 2 compared TheraSphere with TACE or DEB-TACE across the 3 subgroups. The AG's assessment suggested that both studies had high risk of bias and differences in baseline characteristics. The committee concluded that results from these studies might be unreliable for decision making. Another study compared TheraSphere with TheraSphere and sorafenib in combination, in people for whom CTT is inappropriate. This study also had a high risk of bias and was only published as an abstract. The remaining prospective study was done in people for whom CTT is inappropriate. This compared TheraSphere in people with PVT with TheraSphere in people without PVT and best supportive care. The AG advised that this study had a high risk of bias, and that the people in the treatment arms had very different baseline characteristics. Because of this, the committee concluded that these studies should not be used for decision making. It also concluded that there was not enough evidence to establish whether TheraSphere is better than other treatments in people for whom CTT is appropriate and in people for whom CTT is inappropriate.

There were no data identified to establish the clinical effectiveness of SIR-Spheres compared with non-SIRT treatments in people for whom transplant is appropriate

3.17 The AG identified 1 RCT comparing SIR-Spheres with TACE (SIR-TACE) that included people for whom transplant was appropriate. SIR-TACE was done in Germany and Spain, and included 28 people with early, intermediate and late stage disease. Only overall results for the mixed population were available. The AG assessed that the study had a high risk of bias because of the randomisation process, missing outcome data and measurement of the outcome. Only overall results were published, and the company could not provide subgroup-specific data. The committee

concluded that there were insufficient data to establish whether SIR-Spheres are better than TACE in people when transplant is appropriate.

It is unclear whether SIR-Spheres is better than DEB-TACE or TACE in people for whom CTT is appropriate

3.18 The AG identified 2 RCTs that compared SIR-Spheres with TACE (SIR-TACE) or DEB-TACE (Pitton 2015) that included people for whom CTT is appropriate in their trial populations. SIR-TACE is described in section 3.17. Pitton 2015 was done in Germany and included 24 people with intermediate stage disease (BCLC B). Overall survival and progression-free survival were longer in the DEB-TACE arm compared with SIR-Spheres arm, but this was not statistically significant (788 days compared with 592 days and 216 days compared with 180 days, respectively). Based on the identified evidence, the committee concluded that it could not establish whether SIR-Spheres was better than TACE or DEB-TACE in people for whom CTT is appropriate.

SARAH and SIRveNIB may not be generalisable to the NHS in England, but they are preferable to non-randomised evidence in people when CTT is inappropriate

3.19 The AG identified 2 RCTs comparing SIR-Spheres with sorafenib (SARAH and SIRveNIB) in people for whom CTT is inappropriate:

- SARAH was done in France between 2011 and 2015 and included a heterogenous population of people with HCC. This included, for example, people with advanced HCC, people with HCC that were previously treated with 2 treatments of TACE and people with Child-Pugh A or B. There was no difference in overall survival or progression-free survival between the treatment arms. The median overall survival was 8.0 months (95% CI 6.7 to 9.9) for SIR-Spheres and 9.9 months (95% CI 8.7 to 11.4) for sorafenib, with hazard ratios (HRs) of 1.15 (95% CI 0.94 to 1.41) for the intention-to-treat (ITT) population and 0.99

(95% CI 0.79 to 1.24) for the per-protocol (PP) population. The median progression-free survival was 4.1 months (95% CI 3.8 to 4.6) for SIR-Spheres and 3.7 months (95% CI 3.3 to 5.4) for sorafenib with a HR of 1.03 (95% CI 0.85 to 1.25) for the ITT population. More adverse events were reported with sorafenib than SIR-Spheres. A post-hoc analysis of SARAH focused on people with ALBI grade 1 and low tumour burden (equal or less than 25% tumour burden). Again, there was no difference in overall or progression-free survival between the treatment arms. The median overall survival was 21.9 months (95% CI 15.2 to 32.5) for SIR-Spheres and 17.0 months (95% CI 11.6 to 20.8) for sorafenib, with a HR of 0.73 (95% CI 0.44 to 1.21). The median progression-free survival HR was 0.65 (95% CI 0.41 to 1.02). The clinical experts advised that the SARAH trial had more people with a high tumour burden, PVT and impaired liver function than people seen in clinical practice in England. The committee understood that because of this, people in the SARAH trial had poorer prognosis than people seen in clinical practice in England. It concluded that results from the SARAH trial may not be generalisable to people seen in the NHS in England.

- SIRveNIB was done in the Asia-Pacific region between 2010 and 2018. The clinical experts explained that results from SIRveNIB might not be generalisable to the NHS in England. This was because in the Asia-Pacific region HCC is often caused by hepatitis B and C, whereas in the UK fatty liver disease and alcohol are the most common causes. There was no difference in overall survival or progression-free survival between the treatment arms. The median overall survival was 8.8 months for SIR-Spheres and 10.0 months for sorafenib, with HRs of 1.12 (95% CI 0.9 to 1.4) for the ITT population and 0.86 (95% CI 0.7 to 1.1) for the PP population. The median progression-free survival was 5.8 months for SIR-Spheres and 5.1 months for sorafenib, with HRs of 0.89 (95% CI 0.7 to 1.1) for the ITT population and 0.73 (95% CI 0.6 to 0.9) for the PP population. More adverse events were reported with

sorafenib than SIR-Spheres. The committee concluded that results from the SIRveNIB may not be generalisable to people seen in the NHS.

- The committee considered including non-RCT evidence identified by the AG. The AG assessed the 3 non-RCT studies as having a high risk of bias. So the committee concluded that the RCT evidence from SARAH and SIRveNIB was preferable for decision making in people for whom CTT is inappropriate.

There is no evidence in people for whom transplant is appropriate and in people for whom CTT is appropriate to compare the 3 SIRTs' effectiveness

3.20 The clinical evidence for comparative effectiveness of the 3 SIRTs came from 5 retrospective studies that reported overall survival or progression-free survival. Of these, 4 compared SIR-Spheres with TheraSphere and 1 small study of 30 people compared all 3 SIRTs. The AG advised that most of these studies had a high risk of bias. None of the studies included people for whom transplant was appropriate. The study comparing all 3 SIRTs potentially included people for whom CTTs were appropriate but there were no results presented for this subgroup. The committee concluded that there was no evidence identified for people when transplant or CTT was appropriate.

There is not enough direct evidence for people when CTT is inappropriate to compare the 3 SIRTs' effectiveness, so mixed treatment comparison should be considered

3.21 The AG identified 5 retrospective studies that included people for whom CTT is inappropriate (see section 3.20). The study comparing all 3 SIRTs also included people for whom CTTs were appropriate, but no results for subgroups were presented. The committee was aware that the populations were different across these studies and acknowledged that this meant results were difficult to compare. The committee was also aware that the baseline characteristics were different in most studies, and that this might affect prognosis and outcomes between the arms. In 2

studies that compared TheraSphere with SIR-Spheres, there was no difference in overall survival. In van der Gucht et al. (2017, n=77), the median overall survival was 7.0 months for TheraSphere (95% CI 1.6 to 12.4) compared with 7.7 months for SIR-Spheres (95% CI 7.2 to 8.2). In Bhangoo et al. (2015, n=17) the median overall survival for TheraSphere was 8.4 months (95% CI 1.3 to 21.1) compared with 7.8 months for SIR-Spheres (95% CI 2.3 to 12.5). In 2 studies (Biederman et al. 2015 and Biederman et al. 2016) that compared TheraSphere with SIR-Spheres in people with PVT, overall survival was better in the TheraSphere arm than the SIR-Spheres arm. The committee concluded that there was not enough direct evidence to establish the relative effectiveness of the 3 SIRTs in people with HCC, and so decided to consider mixed treatment comparisons for decision making.

Mixed treatment comparisons

Data are not robust enough to provide a meaningful comparison between treatment options when transplant is appropriate

3.22 The AG assessed the feasibility of a mixed treatment comparison to estimate comparative effectiveness between available treatment options in people when transplant is appropriate. There are 2 RCTs that could be included in this analysis. Both were done in the US and compared TheraSphere with TACE (n=45) or with a combination of TheraSphere and sorafenib (n=20). The committee agreed that the evidence base was small, and not generalisable to people seen in the NHS (see section 3.9). Because of limited data, results from the mixed treatment comparison would be very uncertain. The committee concluded that a mixed treatment comparison in this population would not help decision making for the subgroup in whom transplant is appropriate.

The comparative effectiveness of treatment options for people for whom CTT is appropriate is very uncertain, and so is not suitable for decision making

3.23 After consultation on the assessment report, the AG did a mixed treatment comparison in people for whom CTT was appropriate. There were 6 RCTs that could be included in this analysis: 5 compared different CTTs with each other and 1 compared SIR-Spheres with DEB-TACE (n=24). The AG also included 1 retrospective study that compared SIR-Spheres with TheraSphere (n=77). From this study, only a subgroup of 35 people with early or intermediate HCC could be included in the analysis. The study had a high risk of bias because its 2 treatment groups were not similar at baseline (people with small tumour volumes were preferentially treated with TheraSphere). The committee agreed that there was little evidence to link SIR-Spheres and TheraSphere to the network of treatments. Results from the mixed treatment comparison for overall survival and progression-free survival were uncertain (see Table 1 and Table 2). The committee concluded that the results from the mixed treatment comparison in this population were uncertain, and that there was not enough evidence to compare SIR-Spheres with TheraSphere, and the SIRTs with TACE, DEB-TACE and TAE, in this population.

Table 1 Mixed treatment comparison for overall survival, (HRs of less than 1 indicate better overall survival)

	TACE comparator	SIR-Spheres comparator	TheraSphere comparator	DEB-TACE comparator	TAE comparator

TACE treatment mean HR (95% CI)	–	–	–	–	–
SIR-Spheres treatment mean HR (95% CI)	1.06 (0.21 to 3.31)	–	–	–	–
Thera-Sphere treatment mean HR (95% CI)	1.02 (0.13 to 3.77)	0.96 (0.34 to 2.18)	–	–	–
DEB-TACE treatment mean HR (95% CI)	0.88 (0.29 to 2.09)	0.95 (0.35 to 2.56)	1.41 (0.28 to 4.34)	–	–
TAE treatment mean HR (95% CI)	0.98 (0.61 to 1.57)	1.60 (0.27 to 5.25)	2.08 (0.24 to 8.01)	1.48 (0.42 to 3.77)	–

Table 2 Mixed treatment comparison for progression-free survival, (HRs of less than 1 indicate better progression-free survival)

	TACE comparator	SIR-Spheres comparator	TheraSphere comparator	DEB-TACE comparator	TAE comparator
TACE treatment mean HR (95% CI)	–	–	–	–	–
SIR-Spheres treatment mean HR (95% CI)	1.20 (0.22 to 3.82)	–	–	–	–
Thera-Sphere treatment mean HR (95% CI)	1.14 (0.15 to 4.20)	0.95 (0.36 to 2.05)	–	–	–
DEB-TACE treatment mean HR (95% CI)	0.86 (0.26 to 2.15)	0.92 (0.31 to 2.12)	0.94 (0.26 to 3.44)	–	–
TAE treatment mean HR (95% CI)	0.87 (0.61 to 1.20)	0.93 (0.21 to 4.05)	1.58 (0.20 to 5.97)	1.35 (0.38 to 3.50)	–

The comparative effectiveness of treatment options in people for whom CTT is inappropriate is uncertain, but is useful for decision making

3.24 The AG did a mixed treatment comparison to estimate comparative effectiveness between available treatment options in people when CTT is inappropriate. There were 3 RCTs included in this analysis. Of these, 1 RCT compared lenvatinib with sorafenib and 2 compared sorafenib with SIR-Spheres. To include TheraSphere in the network, 2 retrospective studies comparing TheraSphere with SIR-Spheres were included in sensitivity analyses. In the main analysis, in people for whom CTT is inappropriate and with Child-Pugh grade A, there was no evidence of a difference between SIR-Spheres, sorafenib and lenvatinib. The mean HR in the PP population for SIR-Spheres compared with sorafenib was 0.94

(95% credible interval [CrI] 0.77 to 1.14), for lenvatinib compared with sorafenib it was 1.06 (95% CrI 0.79 to 1.4), and for lenvatinib compared with SIR-Spheres the HR was 1.14 (95% CrI 0.79 to 1.58). In the ITT population for SIR-Spheres compared with sorafenib the HR was 1.13 (95% CI 0.96 to 1.32), for lenvatinib compared with sorafenib or SIR-Spheres the HRs were 1.06 (95% CI 0.79 to 1.4) or 0.92 (95% CI 0.67 to 1.29) respectively. A value of less than 1 indicates better overall survival. When the retrospective evidence was included, TheraSphere was shown to be more effective than SIR-Spheres, sorafenib and lenvatinib. The mean HR in the PP population for SIR-Spheres compared with sorafenib was 0.94 (95% CrI 0.77 to 1.13) For lenvatinib compared with sorafenib or SIR-Spheres it was 1.06 (95% CrI 0.79 to 1.4) or 1.13 (95% CrI 0.79 to 1.57) respectively. The mean HRs for TheraSpheres compared with sorafenib or SIR-Spheres or lenvatinib were 0.41 (95% CrI 0.20 to 0.77) or 0.44 (95% CrI 0.20 to 0.84) or 0.4 (95% CrI 0.18 to 0.78) respectively. In the ITT population the mean HR for SIR-Spheres compared with sorafenib was 1.13 (95% CrI 0.96 to 1.32), for lenvatinib compared with sorafenib or SIR-Spheres the HRs were 1.06 (95% CrI 0.79 to 1.4) or 0.95 (95% CrI 0.67 to 1.29) respectively. For TheraSpheres compared with sorafenib or SIR-Spheres or lenvatinib the HRs were 0.47 (95% CrI 0.21 to 0.88) or 0.41 (95% CrI 0.20 to 0.77) or 0.45 (95% CrI 0.20 to 0.89) respectively. In an alternative analysis with a wider population, SIR-Spheres was less effective than sorafenib. In the corresponding sensitivity analysis including the retrospective evidence, TheraSphere was again more effective than SIR-Spheres, sorafenib and lenvatinib. The AG assessment suggested that the retrospective studies had a high risk of bias and uncertain results (see section 3.16). The committee agreed that the retrospective studies should not be included in the analysis because of the risk of bias. It agreed that the comparative effectiveness results based on RCT evidence could be used in a cost-effectiveness analysis. The committee concluded that the estimates of comparative effectiveness were uncertain, but were suitable to inform decision making.

Cost-effectiveness evidence

The AG's model was used for decision making

3.25 There were 2 companies that included economic analyses in their submissions. For SIR-Spheres, the company submitted a cost-minimisation analysis for people when CTT was appropriate, and a cost-utility analysis for people when CTT was inappropriate. The cost-utility analysis was restricted to people with ALBI grade 1 and low tumour burden, a subpopulation from the SARAH trial (see section 3.19). For TheraSphere the company submitted 2 cost-utility analyses, 1 for people when CTT was appropriate and 1 for people when CTT was inappropriate. The committee acknowledged the submission of the companies' models. It noted that the AG model used a similar structure (see section 3.26) as the companies' cost-utility analyses, and that the AG used inputs from the company models, such as costs and treatment frequency. The committee concluded that there was not enough evidence to support an economic analysis in people for whom CTT was appropriate (see section 3.23). It also concluded that when CTT was inappropriate the AG model was the most comprehensive analysis, because it included all 3 SIRTs as specified in the NICE scope (see section 3.3).

The structure of the AG model for people when CTT is appropriate is acceptable for decision making

3.26 The AG did a cost-utility analysis for people with unresectable intermediate (BCLC stage B) or advanced (BCLC stage C) HCC, when CTT is inappropriate, with or without macroscopic vascular invasion but without extrahepatic disease. The model consisted of a decision tree and partitioned survival model with 3 health states. The decision tree represented the outcome of the work-up procedure that happens before SIRT. The partitioned survival model was like that used by the companies. The interventions were SIR-Spheres, TheraSphere and QuiremSpheres, which were assumed to have equal effectiveness in the base case. The comparators were sorafenib and lenvatinib. Because sorafenib and

lenvatinib are recommended only for people with Child-Pugh grade A, the base case was restricted to this population. The committee concluded that the model structure was acceptable for decision making.

Sorafenib is the only relevant comparator for cost effectiveness in people for whom CTT is inappropriate

3.27 In line with the NICE scope, the AG included sorafenib and lenvatinib as comparators in the model. The AG used the hazard ratio from the mixed treatment comparison to include lenvatinib in the model and assumed proportional hazards over time. Therefore, they chose the Weibull function to model overall survival and progression-free survival, even though it was not the best-fitting function. Following consultation on the AG report, sorafenib was considered to be the only relevant comparator (see section 3.11). The generalised gamma was used to fit overall survival and progression-free survival in the revised base case, because the proportional hazards assumption was no longer needed. The committee concluded that sorafenib is the only appropriate comparator, and that the best-fitting function should be used to estimate overall survival and progression-free survival.

There is not enough robust data for the ALBI grade 1 and low tumour burden subgroup for decision making

3.28 The AG presented scenario analyses that restricted the population to people with ALBI grade 1 and low tumour burden. The clinical experts explained that ALBI grade could be a more objective measure than Child-Pugh score for liver dysfunction and that people with ALBI grade 1 have good liver function. However, this measure is not routinely used in the NHS, and the Child-Pugh score is expected to be the standard assessment method for liver dysfunction for the foreseeable future (see section 3.6). The committee was aware that clinical outcomes for the ALBI grade 1, low tumour burden subgroup came from a post-hoc analysis of the SARAH trial (n=85) (see section 3.19). It agreed that this analysis was not robust because the subgroup was not prespecified and the numbers

were small. It concluded that it had not seen sufficiently robust data in this subgroup, but agreed that more evidence may be useful for decision making.

Usually, only 1 lobe is treated at a time in people with bilobar disease

3.29 HCC can be unilobar (tumour is in 1 lobe of the liver) or bilobar (tumours in both lobes of the liver). The clinical experts explained that people with bilobar disease have a higher risk of liver impairment, and therefore usually only 1 lobe is treated at a time. The same lobe might be treated twice to reduce the size of the tumour. The committee concluded that it is not appropriate for a model to assume that both lobes are treated simultaneously in bilobar disease.

Downstaging of HCC might benefit some people with advanced HCC, but the proportion of people and subsequent outcomes are uncertain

3.30 The clinical experts explained that downstaging might be a treatment aim for some people who have SIRT, to potentially allow for subsequent liver transplantation. However, these people are rarely included in clinical trials because trials mainly include people with advanced stage disease. This means there is limited evidence on downstaging and overall survival in advanced HCC. The committee concluded that downstaging may be an appropriate consideration for a small proportion of people with advanced HCC, so the base-case model should include downstaging. However, the proportion of people who have tumours that downstage, and subsequent outcomes, are uncertain.

Some aspects of health-related quality of life might not be captured in the utility values

3.31 Both the SARAH and SIRveNIB trials collected data on health-related quality of life. SARAH used the European Organisation for Research and Treatment of Cancer Quality-of-life Questionnaire Core 30 (EORTC-QLQ-C30) questionnaire, which the company mapped onto the EQ-5D scale using the Longworth et al. algorithm. The AG used these

estimates in its model. The committee noted that utility values were similar between SIRTs and systemic therapies (sorafenib or lenvatinib) for the following disease states: progression-free survival, progressive disease and post-transplant. There were only small differences in utilities between progression-free survival and progressive disease. The clinical experts explained that people who have sorafenib for a long time may have a long-lasting negative effect on their quality of life. SIRTs are given in 1 procedure, meaning there is a shorter duration of impaired health-related quality of life. The committee considered that the potential difference in long-term quality of life might not be captured in clinical trial results because quality-of-life data are collected at fixed time points. It acknowledged that the cancer, liver function and other comorbidities affect health-related quality of life in people with HCC. The committee concluded that some aspects of health-related quality of life might not be captured in the utility values, but it was not presented with evidence comparing this benefit with the relevant non-SIRT comparator, sorafenib.

Cost-effectiveness results

In the AG's model sorafenib dominated SIRTs in all plausible scenarios using confidential patient access schemes for QuiremSpheres and sorafenib

3.32 The probabilistic base case of the AG model, including confidential patient access schemes for QuiremSpheres and sorafenib, showed that all SIRTs were less effective and more expensive than sorafenib (exact incremental cost-effectiveness ratios [ICERs] are confidential and cannot be reported here). Because of uncertainties in the evidence, the AG presented 17 scenario analyses, for example using alternative functions to model overall survival and progression-free survival (see section 3.27). The committee noted that ICERs did not change much if alternative functions were used. The committee also accepted that alternative costs and utility values did not have a big effect on ICERs. It acknowledged that in scenarios that restricted the population to people with ALBI grade 1 and low tumour burden (see section 3.28), TheraSphere was more cost

effective than sorafenib. However, the committee agreed that such scenarios are not plausible because this population is not relevant to NHS practice in England (see section 3.28). TheraSphere was also more effective in the scenario that included retrospective studies with high risk of bias. The committee agreed that this scenario should not be considered because of the high risk of bias and uncertainty of the data (see section 3.24). The committee agreed that while the modelling may not capture all health-related quality-of-life outcomes for people with HCC and SIRT, this was unlikely to change the cost-effectiveness estimates for SIRTs enough to change its conclusions. The committee concluded that sorafenib dominated SIRTs in all plausible scenarios. Therefore, it did not consider SIRTs to be a cost-effective use of NHS resources for treating unresectable HCC.

End of life

The end-of-life criteria are not met

3.33 The committee considered the advice about life-extending treatments for people with a short life expectancy in NICE's guide to the methods of technology appraisal.

- When transplant or CTT is appropriate, people have a life expectancy of more than 24 months. This means that the life-expectancy criterion (that is, the treatment is indicated for patients with a short life expectancy, normally less than 24 months) was not met for these subgroups.
- When CTT is inappropriate, in advanced stage disease, people have a poor prognosis with a life expectancy of less than 24 months. Therefore, the short life-expectancy criterion was met for this subgroup.
- In all plausible scenarios, there was no increase in the modelled undiscounted life expectancy with SIRTs compared with sorafenib. The committee concluded that the life-extending criterion (that is, there is sufficient evidence that the treatment could extend life, normally by a

mean value of at least an additional 3 months, compared with current NHS treatment) was not met.

Because both parts of the criteria were not met, the committee concluded that the end-of-life criteria were not met.

Innovation

No evidence was identified showing additional benefits of SIRT, above those captured in the cost-effectiveness analysis

3.34 The companies considered SIRTs to be innovative because they offer a more personalised treatment option. The patient experts stated that SIRTs would be a substantial change in treating HCC because they could offer a chance for subsequent curative treatment for people who would not otherwise have this option. The committee concluded it was not shown evidence of any additional benefits that were not captured in the measurement of quality-adjusted life years in the model.

Equality

3.35 No equality or social judgement issues were identified.

4 Proposed date for review of guidance

4.1 NICE proposes that the guidance on this technology is considered for review by the guidance executive 3 years after publication of the guidance. NICE welcomes comment on this proposed date. The guidance executive will decide whether the technology should be reviewed based on information gathered by NICE, and in consultation with consultees and commentators.

Stephen O'Brien
Chair, appraisal committee
November 2019

5 Appraisal committee members and NICE project team

Appraisal committee members

The 4 technology appraisal committees are standing advisory committees of NICE. This topic was considered by committee C.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

The minutes of each appraisal committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

NICE project team

Each technology appraisal is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the appraisal), a technical adviser and a project manager.

Verena Wolfram
Technical lead

Jamie Elvidge
Technical adviser

Louise Jafferally
Project manager

ISBN: [to be added at publication]