SHTAC response to comments from consultees

11th April 2013

General response

Patient populations
Comments were received from Heart Rhythm UK, BSH and ABHI regarding the three patient populations included in the assessment report.

The three populations were defined by the NICE scope, which the assessment report must adhere to. The report acknowledges that these are not distinct groupings.

As ABHI state, there are no accepted a priori criteria which could be used to categorise trials. Clinical experts were consulted in order to allocate trials to the population groups. NICE were informed of the difficulties encountered in allocating trials and of the decisions made to ensure that the assessment report adhered to the finalised scope. Pragmatic decisions were taken to allocate trials and to ensure all relevant RCT evidence comparing eligible interventions and comparators was included.

Where trials used different criteria to define their target population, this was noted and highlighted in the assessment report. The baseline characteristics of the trials allocated to each population were described in detail. Only those trials whose populations were considered similar enough were combined in a meta-analysis. Scenario analyses were presented in the economic evaluation to account for the different target populations.

Population 1:

- The clinical effectiveness section summarises all eligible trials comparing ICD vs OPT (or medical therapy). The target populations of the trials differed (e.g. previous cardiac arrest, myocardial infarction or cardiomyopathy). Results were therefore presented separately within the clinical effectiveness chapter for the different target populations and scenario analyses were conducted in the economic analyses.

Population 2:

- The clinical effectiveness section summarises all trials comparing CRT vs OPT. All trials included in TA120 are included here, no addition trials were identified. The COMPANION trial randomised participants to 3 groups (CRT-P, CRT-D and OPT).
- ABHI questioned whether the COMPANION trial belonged in population 2. This trial excluded people *meeting the general indications for an ICD*; therefore a pragmatic decision was taken to allocate the trial to population 2 (all trials in population 3 included people with an ICD indication). COMPANION was the only trial that randomised people to CRT-P and CRT-D (or to compare CRT-D vs OPT), therefore these comparisons were not combined in any meta-analyses with other trials; the results stand alone. COMPANION was used in the economic model to provide evidence on the comparison of CRT-P and CRT-D for both population 2 and population 3, as it was the best available evidence. Allocating this trial to population 3 in the clinical effectiveness chapter would not impact the results.

Population 3:

- The clinical effectiveness section summarises all trials comparing CRT-D vs ICD. All trials in this section included participants with an ICD indication and cardiac dyssynchrony. The NYHA class differed between the trials therefore data were presented by NYHA class in the forest plots. About half the trials included people with NYHA class II and therefore had less severe heart failure than those in the population 2 trials. Inevitably the severity of heart failure in these patients may have impacted on the outcomes. This was examined in the economic evaluation in the assessment report through sensitivity analyses.

The SHTAC assessment report relied on the use of aggregate trial data, which limits the approach that can be taken, especially with regards to subgroups. The focus of the assessment report was an evaluation of the effectiveness of the devices for specific patient groups identified by the NICE scope. In contrast, the ABHI submission accessed individual patient data (IPD) to produce a network meta-analysis (NMA) to inform their model. This allowed the identification of specific subgroups of patients that the different devices appeared to benefit.

Although the approach and focus of the SHTAC assessment report and ABHI submission differed, both approaches are reasonable.

Inclusion of device cross-over in treatment pathways

The ABHI and Heart Rhythm UK criticise the approach in the SHTAC analysis with regard to 'device crossover' in which patients start on a given therapy and may subsequently cross-over to a different device.

The model structure was developed to reflect the management of patients under current clinical practice, consisting of a simplistic approximation of the clinically plausible care pathways. Therefore, the model allows patients initially managed with OPT or CRT-P to have a device implanted or upgrade to a different device according to disease progression.

The model output obtained with this approach is intended to capture the impact of all treatments received by the patient over lifetime, instead of only those of the treatment initially allocated, providing a more realistic estimation of the consequences of the adoption of a particular technology as initial treatment.

This approach has been devised in consultation with our clinical experts as the most clinically plausible. This approach is consistent with the approach developed by TA95 and TA120.

The assumptions for device upgrade are described comprehensively in the assessment report (section 5.4.3.4) Patients in population 1 who are assigned to OPT, can be upgraded to ICD implantation as a result of hospitalisation for major arrhythmia. Likewise patients initially assigned to CRT-P for population 2 and 3 can be upgraded to CRT-D if they experience serious arrhythmia. Furthermore, for population 2, patients receiving ICD initially can be upgraded to CRT-D after hospitalisation for heart failure.

We have modelled an 'OPT only' scenario, i.e. no upgrades (circulated to as a supplementary analysis). We considered modelling no upgrades in the other arms, however the number of device crossovers in all other arms is small (and so changes to these arms would have minimal effect on the model results) except for CRT-P in P3 model.

The CRT-P arm was included as comparator in P3 model to address the scope of the current appraisal. However, no clinical evidence was found for P3 patients with CRT-P and clinical advice

suggested CRT-P would not be appropriate in this group. Therefore, several assumptions on effect were required to include CRT-P in the P3 analysis (e.g. CRT-P assumed to have same risk of all-cause mortality and hospitalisation for severe arrhythmia as CRT-D), and results for the CRT-P arm should be regarded with caution. A scenario analysis assuming no upgrades in the CRT-P arm would also require several additional assumptions on what would happen to patients with CRT-P experiencing severe arrhythmia, and hence would not decrease uncertainty in this analysis.

Univariate sensitivity analyses were conducted for the effect of varying assumptions of treatment effect for CRT-P vs. CRT-D. For instance, when a relative risk of 1.2 for all-cause mortality was assumed for CRT-P vs. CRT-D, an ICER of £3,890/QALY was estimated for CRT-D vs. CRT-P. For a RR of 0.8, the ICER was £14,883/QALY for CRT-D vs. CRT-P (these results were not reported in the assessment report due to small impact in results, however can be found in the P3 model 'Sensitivity Analysis' spreadsheet).

Specific responses

Comments from ABHI: Technical critique of AG model

Derivation of treatment effects

ABHI criticise the use of trial data in population 2 for COMPANION and CARE-HF to inform CRT-P vs OPT but only COMPANION for the use of CRT-D vs OPT for the treatment effect for heart failure. ABHI state that patients in CARE-HF were much more likely to have wide QRS interval and LBBB than patients in COMPANION and identifies these as predictors of enhanced CRT efficacy. They state that this is likely to bias the treatment effect of CRT-P vs. CRT-D in favour of CRT-P.

However, we are unable to comment on the proportion of participants in CARE-HF and COMPANION with QRS >150 ms, or the proportion in CARE-HF with LBBB, as ABHI derived these from IPD. The reported median baseline QRS interval (160 msec) was the same in both trials, and distribution of NYHA class was similar. Therefore we took a pragmatic approach to include all the available evidence.

In view of the IPD data from ABHI, we acknowledge that the treatment effect from the COMPANION trial may be smaller than that expected if CARE-HF patients had received CRT-D.

This potential underestimation was explored in SA (Tables 127 to 129 of our report) by assuming a higher relative effect from CRT-D, i.e. a lower RR of HF death for CRT-D. For example, taking the case where the RR of HF death was at the 95% CI lower limit of RR HF death for CRT-D=0.47 (compared to RR = 0.73 baseline), the ICER for CRT-D vs OPT is £20,671/QALY compared to baseline of £27,899/QALY.

ABHI criticise the evidence used for population 3, in particular the assumptions regarding the relative treatment effect of CRT-P vs. CRT-D. For CRT-P vs CRT-D equivalence was assumed. This approach was taken due to the absence of evidence comparing CRT-D with OPT or CRT-P in population 3. As mentioned above, the CRT-P arm was included as comparator in P3 model to address the scope of the current appraisal. However, no clinical evidence was found for P3 patients with CRT-P and clinical advice suggested CRT-P would not be appropriate in this group. Therefore, several assumptions on effect were required to include CRT-P in the P3 analysis (e.g. CRT-P assumed to have same risk of all-cause mortality and hospitalisation for severe arrhythmia as CRT-D), and results for the CRT-P arm should be regarded with caution. As noted above, univariate sensitivity

analyses showing the effect of varying assumptions of treatment effect for CRT-P vs. CRT-D were conducted.

ABHI criticises the analyses conducted for population 1, on the basis that the evidence chosen does not represent the full evidence base for population 1 patients. The aim for population 1 was not for either trial to represent the full evidence base, but to present representative scenarios for the different target populations.

Derivation of survival curves

ABHI criticise the SHTAC analysis as it applies relative risk as if they are hazard ratio in a number of instances in the model. We acknowledge that we have assumed that the hazard ratio is representative of the relative risk and have applied this in some instances in the model. We consider that the hazard ratio provides a good approximation of the relative risk and using this approach does not have a material effect on the model results.

The ABHI state that the proportions alive at any given point in the SHTAC models do not reflect the data on which the models were largely parameterised. They raise concerns on the fit to OPT arm in population 2 and consider that this lead to an inflation in survival. In the validation appendices we acknowledge that the fit of the model to the trial data underestimates the benefit of CRT compared to OPT and therefore may be a conservative fit (this would reduce the ICER to less than the current base case of £27,900 vs. OPT).

ABHI criticise the age specific scaling factor applied to all-cause mortality data to reflect increasing risk of death in population 1 and population 3. They consider that including this scaling factor has accelerated the rate of death, potentially double counting some mortality and biasing against treatments with a better survival profile.

We note that we have included this scaling factor to be consistent with approach taken in TA120. Sensitivity analysis to the age-specific relative risks of all-cause mortality showed the results robustness to these parameters (removing the scaling factor gives slightly more favourable ICERs £1-2000 less for population 2 and 3 and almost no difference for population 1).

Use of RAFT mortality data in population 3

ABHI comment that they do not consider the RAFT trial to be representative of population 3 patients.

The proportions of participants in each NYHA class varied between the trials in population 3, however as can be seen in Table 53 of the assessment report, only a minority were in class I or IV. RAFT was the only trial demonstrating a statistically significant benefit of CRT-D vs. ICD and strongly influenced the results of the meta-analyses.

The mortality data from RAFT were used in the base case; RAFT included participants with NYHA class II (80%) and NYHA class III (20%). We considered that the RAFT trial was the best trial evidence available for this population. We conducted scenario analysis using mortality data from MADIT-CRT, which included participants with NYHA class I (15%) and class II (85%).

We also undertook sensitivity analyses using the upper and lower confidence limits of baseline mortality data from the RAFT trial. There was little impact on the results with cost effectiveness of CRT-D vs. OPT varying between £35,257 - £35,655 / QALY (these results were not reported in the assessment report due to small impact in results, however can be found in the P3 model 'Sensitivity Analysis' spreadsheet).

Lack of face validity of results for population 3

ABHI and Heart Rhythm UK criticise the face validity of population 3 results and contend that population 3 would be expected to benefit more from CRT-D than population 2, and therefore CRT-D should be more cost effective in population 3.

The reason for differences between the ICERs in these groups is that we have used different populations with different heart failure severity and life expectancies in population 2 and population 3, which makes it difficult to compare results between populations in this way. Due to the limited evidence base, the RAFT trial provided the best evidence for population 3, and as described above the majority of participants had NYHA class II. The P2 and P3 cost-effectiveness results for CRT-D vs OPT are £27,899/QALY for P2 and £35,193/QALY for P3.

To explore the effect of using a higher baseline risk for population 3, we undertook sensitivity analyses using the confidence limits of baseline mortality data from the RAFT trial. There was little impact on the results with cost effectiveness of CRT-D vs. OPT of £35,655 / QALY for a higher baseline mortality risk (these results were not reported in the assessment report due to small impact in results, however can be found in the P3 model 'Sensitivity analysis' spreadsheet).

ABHI criticised the face validity for population 3 because the use of ICD therapy results in the generation of fewer lifetime QALYs than with patients treated with OPT alone.

A significant proportion of P3 patients initially managed with OPT alone are estimated to be referred to CRT-D over their lifetime. These patients will therefore benefit from lower risk of death and of hospitalisation for HF than patients with ICD + OPT. We have also conducted a 'OPT only' scenario with no allowable device upgrades or cross overs for OPT patients (see SHTAC Additional scenario analysis).

No quality of life benefit of CRT-P in population 3

ABHI state that the failure to model a QoL benefit with CRT-devices in population 3 appears to contradict SHTAC's own systematic review findings.

We acknowledge that there was an improvement in QoL score for CRT-D, although this was for MLWHF, and it is uncertain whether there would also be an improvement in EQ-5D QoL, which is used in the economic model.

Data were not available from the RAFT trial (used to model the base case) that showed the progression of heart failure. However, we have investigated the effect of a beneficial effect of CRT devices on patients' HF progression and consequently on HRQoL (Table 145). The ICERs obtained with this scenario are similar to those of the base case analysis.

Transplant mortality

ABHI state that 20% of transplant patients are still alive after 34 years.

We are unclear how the ABHI has calculated that 20% of transplant patients are still alive at age 100 years. However, the numbers of patients who have transplant is very small < 0.1%, and so changes to the mortality rate for this group of patients have no effect on the model results.

Utility data

ABHI questioned the relevance of the study used for NYHA class specific utility data as they were not from an average device population.

The NYHA class specific utility data were taken from a study identified by our systematic review as being of the best quality for those with heart failure.

We have explored the use of different utility estimates with scenario analysis and these results are similar to those of the base case analysis.

Systematic inflation of costs in the device arm

ABHI criticised the results generated for population 3, and that the higher cost of CRT-P when compared to CRT-D appear implausible.

Very high rates of crossover to CRT-D

As noted by ABHI, there are high rates of crossover to CRT-D for patients originally allocated to CRT-P. Patients originally allocated to CRT-P crossover to CRT-D if they experience hospitalisation for arrhythmia. They criticise the hospitalisation rate for non-fatal arrhythmias which is much higher for population 3 than population 1 and 2.

The probability of hospitalisation for arrhythmia, as per our report, for population 3 of 0.029 (0.015-0.042) was derived from population 3 trials. For population 2, the estimate from the MIRACLE RCT participants of 0.0075 was used. The lack of estimates for population 1 led us to taking the conservative assumption of the same estimate as for P2 (the lowest risk estimate). A scenario was conducted for population 1 using population 3 estimates (Assessment report p. 344) that lead to a minimal change to the results (ICER ICD vs OPT of £18,185/QALY).

As noted above, CRT-P has been included within the P3 model to be consistent with the NICE scope and clinical advice suggested CRT-P wouldn't be appropriate in this group. Therefore in this population there is no clinical evidence available for CRT-P in P3 and this required several assumptions on effect. We consider that the results for CRT-P arm should be regarded with caution.

ABHI question the modelling approach with regard to the modelling of arrhythmic and heart failure hospitalisations.

We acknowledge that the hospitalisation health state combines patients hospitalised due to heart failure and non-fatal severe arrhythmia and this may be leading to an overestimation of the upgrades to CRT-D. As the probabilities of upgrade for hospitalised patients receiving ICD, CRT-P or OPT to CRT-D are very small, this is unlikely to have a relevant impact on results.

When the baseline probability of non-fatal arrhythmia is set to zero, CRT-D upgrades in the ICD arm still occur for patients with ICD or being managed with OPT (due to unsuccessful ICD implants) who experience hospitalisation due to heart failure and are subsequently referred to CRT-D. The number of CRT-D upgrades increases when the probability of arrhythmia-related hospitalisation is input because patients managed with OPT can then be referred directly to CRT-D and because those with an ICD who are hospitalised (irrespective of the cause) can be upgraded according to the probability derived from Essebag and colleagues (0.0068). This estimate has been subject to sensitivity analysis (95% CI 0.0006, 0.0129) and the cost-effectiveness results were very similar to the base case (and therefore not reported in the AG report).

The AG conducted an additional scenario setting the probability of upgrade of hospitalised patients receiving any treatment to CRT-D to zero and the results are also very similar to the base case (ICER for CRT-D vs. OPT £35,157 / QALY compared to the base case of £35,193 / QALY) as these probabilities are very small.

Incorrect modelling of device related infections and lead displacement

ABHI state that the proportions of displacement are much higher than expected (43% of CRT-P and CRT-D patients in population 3, compared to 10.3% in TA120) according to the detailed results sheet. In addition, ABHI criticise the assumption used in the model whereby the complication event probability is constant each month. They consider this unrealistic as patients are at an event risk of such a displacement for a short period of time after implant, after which, the risk is much lower.

SHTAC used the same approach as in TA120 but whilst estimates in TA120 were based upon expert opinion, those used in the SHTAC model were based upon trial estimates. SHTAC acknowledge that the event rates are higher for the numbers of lead displacements than in TA120.

SHTAC acknowledge the limitation of the approach which assumes constant probabilities for events and has conducted a new sensitivity analysis to show the effect of a lower event rate (similar to that used in TA120) on the model results (not previously reported in the AG report).

Scenario analysis using same probability per cycle for lead displacement with CRT (0.0015) as previous model for TA120 are shown in Table 1, and show marginal changes to the results (slight improvement in ICER of about £1000).

Table 1. Lead displacement scenario analysis results for Population 2

Strategy	Cost (£)	Life-years	QALYs	Incremental cost (£)	Incremental QALYs	ICER (£/QALY gained)
vs next best option [a]						
OPT	7,605	4.86	3.48	-	-	-
CRT-P + OPT	25,645	5.52	4.18	18,040	0.69	25,994
CRT-D + OPT	37,050	7.23	4.59	11,405	0.41	27,504
vs OPT						
CRT-D + OPT	37,050	7.23	4.59	29,445	1.11	26,559

Discounted costs and benefits; QALY, Quality-adjusted life year. ICER, Incremental cost-effectiveness ratio
[a] Treatments compared with the preceding best option, i.e. the preceding treatment, which is neither dominated or extendedly dominated

We also note that the rates of infection (0.0006) and surgical failure (0.084 and 0.087 for CRT-P and CRT-D, respectively) used in the SHTAC model are lower than in the TA120 AR (0.0022 for infection and 0.0938 for surgical failure, for both devices).

ABHI states surgical failure is assumed to occur every cycle in the AG model and therefore the AG approach exaggerates device failure. We restate that surgical failure is modelled only in patients that undergo surgery (due to initial implantation, device replacement or infection) in the cycle that follows a surgical intervention (assumption listed in AG report p.304) so limiting the effect upon the ICER.

ABHI also criticise the cost for surgical failures. They consider that the vast majority of CRT-D system failures relate to pacing or sensing problems requiring the reposition or replacement of a lead, not a new defibrillation device (total cost of ICD implant is £15,248).

Following clinical experts advice (as reported on AG report pages 301 and 302), the model assumes that patients experiencing CRT-P surgical failure (i.e. unsuccessful CRT-P implantation) or ICD surgical failure are assumed to return to being managed with OPT alone. Those with unsuccessful CRT-P who

go back to OPT management and experience severe arrhythmia are assumed to have an ICD implant and patients who survive an unsuccessful CRT-D implantation are assumed to undergo ICD implantation. We acknowledge that an overestimation of the cost associated to CRT-D surgical failure in population 3 model due to the ICD implantation cost is possible and have therefore conducted a new scenario where this cost is null (not previously reported in the AG report, see results in Table 2 below). Results for this scenario show that the costs associated with this assumption are relatively small and that it has a minimal impact on the relative cost-effectiveness of the different strategies.

Table 2. Scenario without CRT-D failure-related costs results for Population 3

Strategy	Cost (£)	Life-years	QALYs	Incremental			ICER (£/QALY gained)
				Cost (£)	Life-years	QALYs	
Base case [a]							
vs next best option [b]							
ICD + OPT	39,719	7.45	5.57	-	-	-	-
OPT	40,006	7.59	5.67	287	0.14	0.10	2,824
CRT-P + OPT	51,202	7.96	5.94	11,196	0.37	0.27	Extendedly dominated
CRT-D + OPT	50,911	8.01	5.98	10,906	0.42	0.31	35,193
No CRT-D surgical failure-related costs							
vs next best option [b]							
ICD + OPT	39,620	7.45	5.57	-	-	-	-
OPT	38,034	7.59	5.67	-1,586	0.14	0.10	Dominant
CRT-P + OPT	49,134	7.96	5.94	11,099	0.37	0.27	Extendedly dominated
CRT-D + OPT	48,372	8.01	5.98	10,338	0.42	0.31	33,361

[a] replicated from report Table 141, [b] Treatments compared with the preceding best option, i.e. the preceding treatment, which is neither dominated or extendedly dominated; Discounted costs and benefits; QALY – quality-adjusted life year; ICER – incremental cost-effectiveness ratio

Device longevity data

i) ABHI criticised the SHTAC analysis for using a fixed period of time for device longevity. They claim the rate of failure is not consistent.

SHTAC adopted a pragmatic approach to deal with device longevity, consistent with the approach taken in the TA120 analysis. We have tested the effect of different device lifetimes in sensitivity analyses.

ii) ABHI criticised the use of mean device survival time from the previous CRT appraisal.

For the base case analysis, SHTAC used mean device longevity estimates derived from data reported in the ABHI submission. However, SHTAC's clinical advice questioned the device longevity estimates, presented in the ABHI submission. Therefore we considered it appropriate to show the effect of the estimates used in the previous appraisal in sensitivity analyses (Table 131, Table 144).

Specific responses to Heart Rhythm UK

Heart device longevity
Heart Rhythm UK commented that they were surprised that SHTAC had disregarded the UK device longevity estimates from the ABHI submission.

In fact we have used UK device longevity estimates from the ABHI (Table 107), and tested these in sensitivity analyses (Table 131).

page 15, paragraph 2
Heart Rhythm UK quote from the Executive Summary "One trial reporting hospitalisations found higher rates per 1000 months follow-up among people with ICDs (11.3 vs 9.4, p=0.09), with higher heart failure hospitalisations (19.9% vs 14.9%)" and suggests it refers to the DAVID trial.

However, this sentence is in the section 'People with a remote MI' in the Executive Summary and refers to the MADIT II trial.

page 14, paragraph 4
Hearth Rhythm UK refer to the CABG Patch trial and state that it is not relevant to current ICD treatment in the UK and that its results should not be used to infer the clinical effectiveness or risk associated with current practice.

This comment refers to the Executive Summary. The CABG Patch trial met the inclusion criteria for the systematic review. Details of the trial are described in section 4.2.1.1 'characteristics of included studies'.

page 62, paragraph 3
Heart Rhythm UK suggest checking baseline characteristics relating to NYHA class.

On checking an error was discovered in Table 11 and related text for the CIDS trial, which should read that around 50% had no congestive heart failure, 40% had NYHA class I or II and 11% had NYHA class III or IV . This does not affect the analysis or results.

ABHI clarification of points arising from the SHTAC review of the ABHI model: SHTAC's response

IPD and NMA

In reviewing the comments that SHTAC made concerning the IPD and NMA presented in the ABHI's submission, it does not appear that SHTAC has been unfairly critical. First, section 4.1 of the MS (page 100), which outlines the basis for the IPD analysis of the three outcomes of all-cause mortality, all-cause hospitalisation and health related quality of life conducted in section 4, states that this is the first network meta-analysis of individual patient data in the field of CRT/ICD devices and does not specify that it only relates to all-cause mortality. Having mapped the network of evidence for the different outcomes, it is evident that those for all-cause hospitalisations and health related quality of life do not provide sufficient evidence for all the comparisons presented in figure 12 that are encompassed in the network meta-analysis for all-cause mortality. Although the evidence base and the comparisons possible are more limited for all-cause hospitalisations and health related quality of life, the basic methodology used in the analyses is similar (i.e. meta-regression). This is identified in the ABHI's submission on page 141. Importantly, it does not affect the outcomes of the analyses presented or the interpretation of the results. Second, SHTAC has not 'used concerns about one aspect (the HRQoL analysis) to cast doubt on all aspects – in particular the estimates of treatment effect modifiers'. SHTAC has looked at the analysis of each outcome separately, critically appraising the methodology used and trying to provide appropriate context for assessing the results presented. In specific instances SHTAC has provided an indication of where uncertainties are in the analysis due to a lack of presentation of information about the analyses undertaken in the ABHI submission (e.g. limited information on fitting of parametric distributions used for survival analysis) and in others SHTAC has highlighted limitations presented by the ABHI itself concerning interpretation of the findings (particularly those concerning the treatment effect modifiers). Importantly, it should be noted that SHTAC states that it considers that 'the steps taken seem appropriate and the results presented appear reasonable given the note of caution provided in the MS throughout all three analyses' (see page 249 Assessment Report). However, use of the results should recognise the uncertainties in the analyses and the potential difficulties in relating its findings to the original scope.

SHTAC would agree that the IPD NMA presented in the ABHI submission, like the meta-analyses in SHTAC's assessment report, provides evidence that the different devices appear clinically effective. However, given the different approaches taken, it is less clear which should form the basis for developing subsequent guidance.

ABHI corrected treatment effect presentation

The re-analysis of the predicted treatment effects for the different sub-groups, having corrected the error in the previous ABHI submission concerning the failure to take proper account of the covariance between the model parameters (identified by ABHI post-submission), resulted in a narrowing of the confidence intervals around the hazard ratios for the comparisons with OPT (see Table 3). As SHTAC does not have access to the IPD analyses, the error can't be verified. Although this increased the number of comparisons where there was a statistically significant benefit, the groups identified differed little from those that were shown to benefit significantly or that were on the margins of statistical significance in the previous SHTAC assessment. In the re-analysis ICDs were shown to provide a statistically significant benefit for all males irrespective of age, QRS or LBBB status and for women aged <60years with a QRS ≥120 to <150 and no LBBB. CRT-D benefitted a wider group of patients. Benefits that were statistically significant or on the margins of statistical significance were reported for males and females of all ages with a QRS ≥120 with or without LBBB. In contrast, CRT-P only had a statistically significant effect for males and females aged ≥60 years with a QRS of ≥150ms with LBBB.

Table 3 Updated hazard ratios (95% confidence intervals) for all-cause mortality from NMA with covariables for the comparisons between the different devices and OPT following revised analysis by ABHI

Non-LBBB					
QRS	**Device**	**Sex and Age Groups**			
		Male <60yrs	**Male ≥60yrs**	**Female <60yrs**	**Female ≥60yrs**
<120	ICD	[redacted]	[redacted]	[redacted]	[redacted]
≥120 to <150	ICD	[redacted]	[redacted]	[redacted]	[redacted]
	CRT-D	[redacted]	[redacted]	[redacted]	[redacted]
	CRT-P	[redacted]	[redacted]	[redacted]	[redacted]
≥150	ICD	[redacted]	[redacted]	[redacted]	[redacted]
	CRT-D	[redacted]	[redacted]	[redacted]	[redacted]
	CRT-P	[redacted]	[redacted]	[redacted]	[redacted]
LBBB					
QRS	**Device**	**Sex and Age Groups**			
		Male <60yrs	**Male ≥60yrs**	**Female <60yrs**	**Female ≥60yrs**
≥120 to <150	ICD	[redacted]	[redacted]	[redacted]	[redacted]
	CRT-D	[redacted]	[redacted]	[redacted]	[redacted]
	CRT-P	[redacted]	[redacted]	[redacted]	[redacted]
≥150	ICD	[redacted]	[redacted]	[redacted]	[redacted]
	CRT-D	[redacted]	[redacted]	[redacted]	[redacted]
	CRT-P	[redacted]	[redacted]	[redacted]	[redacted]