National Institute for Health and Care Excellence

Consultation

Renal and ureteric stones

Costing analysis of surgical treatments

NICE guideline

Economic analysis report

July 2018

Consultation

This guideline was developed by the National Guideline Centre

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and, where appropriate, their carer or guardian.

Local commissioners and providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Contents

1 Economic analyses: Surgical treatments

1.1 Introduction

Three subgroups were identified from the clinical evidence review comparing surgical interventions for people with renal stones, where the committee felt there is the most uncertainty in practice regarding choice of technique, and where the more expensive procedure was more effective. Full details of the published clinical evidence and the committee's discussion are in evidence report F: Surgical interventions.

The subgroups are:

- Ureteric stones in adults <10mm: ureteroscopy (URS) versus shockwave lithotripsy (SWL)
- Renal stones in adults <10mm: URS versus SWL
- Renal stones in adults 10-20mm: percutaneous nephrolithotomy (PCNL), versus URS, and SWL

The main consequence of the initial procedure having lower effectiveness is a higher rate of downstream procedures (either a repeat of the initial procedure or a different procedure). This will increase the intervention cost, and therefore to appropriately compare the cost difference between interventions it is important to take this into account. In addition, the committee concluded that other outcomes may also vary such as adverse events, and this could also impact overall costs. On this basis, the committee prioritised cost analyses based on the clinical evidence to explore the true cost differences between surgical options. The details of the analyses undertaken for each subgroup are described in this report.

It is also possible that health outcomes may vary due to the difference in initial effectiveness, because it will mean that in those people where the procedure is not effective initially, it will take longer for the stone to be removed. If having a stone reduces quality of life, this will result in lower QALYs for the less effective option. Therefore, in addition to the cost comparisons, exploratory analyses were undertaken to explore the QALY or quality of life difference that would be required to justify any remaining cost difference. Cost-utility analyses have not been undertaken due to multiple uncertainties in the data regarding the timing of downstream resource use for example – this is explained in more detail within the methods for each subgroup.

1.2 Ureteric stones <10mm, ureteroscopy (URS) versus shockwave lithotripsy (SWL)

1.2.1 Methods

1.2.1.1 Approach to analysis

The purpose of this analysis was to compare the total cost of a strategy that began with URS versus a strategy that began with SWL. The reason the interventions are described as 'strategies' is because the goal is to achieve stone free health status, and this could involve further downstream resources to achieve if the initial treatment was unsuccessful (either consisting of retreatment or ancillary procedures). Therefore 'strategies', that *begin* with the interventions of interest (URS and SWL), better reflect what the studies were trying to compare.

URS is a more expensive procedure than SWL. However, the clinical evidence review found that URS was associated with greater success in terms of people being stone-free and, presumably as a result, less retreatment and ancillary procedures. The reduced need for secondary procedures may offset the higher initial procedure costs and need to be taken into account when considering the overall cost difference between URS and SWL. In addition, the clinical evidence review found that readmission, and adverse events may be lower with SWL and this will also contribute to overall cost differences.

The clinical evidence review identified seven studies comparing URS with SWL in adults with ureteric stones <10mm. The studies are variable in many ways: the date of the studies; the length of follow up; whether the stone free outcome being reported is after the initial procedure or not, whether everyone is stone free at the end of the studies. Also in terms of the interventions themselves, it is not always clear how many sessions are being offered for SWL. A summary of the studies can be found in Appendix A:.

As touched on above, the studies were a combination of studies that measured the clinical outcome (number stone free) after the initial procedure, and studies that measured the clinical outcome after additional procedures, and sometimes the studies were unclear. A difficulty this creates is the impact it would have on the stone free rate outcome as one would assume that the more downstream treatments patients undertake, the more likely that a higher proportion of people in the trial will end up being stone free.

Pooling all the studies together in the clinical review therefore also creates a discord between the downstream resource use outcomes (of retreatment and ancillary procedures probabilities), and the stone free probabilities. As the downstream resource use are the resources that actually occurred and can be taken as fact, whereas the stone free outcome is influenced by where in the pathway it was measured, and the downstream resource use is either: a consequence of the stone free rate, or led to the stone free rate – depending on when it was measured which varies by study. Therefore, pooling studies for the stone free rate outcome means that that probability is not felt to be particularly reliable.

As a result, by costing up the retreatments and ancillary procedures reported in the studies, we could essentially assume that people would be stone free after taking the additional procedures into account. This assumption could also be supported by the fact that 4 out of the 7 studies included in the clinical review for this comparison were measuring outcomes before further treatment (as assumed from the language used in the studies such as 'initial' stone free rate).

However, because of concerns about heterogeneity in the data, as well as differences in how outcomes are being reported, and what it is possible to infer about the treatment pathway, multiple scenarios have been undertaken which are informed by different data and with differing assumptions. We undertook the following three analyses:

1. Cost comparison using only resource use reported in all trials. Assuming that this is the resource use required for everyone to be stone free.
2. Cost comparison using only studies where: everyone was stone free at the end of follow up *and* that also report initial stone free success.
3. Cost comparison using only studies that report more detail on the success of multiple lines of treatment.

A cost utility analysis was considered, but would be difficult because lots of assumptions would have to be made. For example, about what is happening at different points in the pathway regarding primary procedures and retreatments, because the trials are not clear. Therefore, we don't have clear data on effectiveness from these different time points in terms of how many people are stone free after the initial procedure, then how long the period is before they have other procedures etc, in order to apply utilities. We are also talking about a specific size stone group, some of which might be in pain and some not, but we might expect that there would be a difference in quality of life on average between people who still had a stone and people who didn't. As the interventions themselves are also different in their nature, then the effectiveness might be more immediate with URS because the stone is extracted at the time (in most cases). Whereas with SWL, the stone broken is up but still needs to be passed, which might not happen for a while afterwards (so more of a linear based increase in stone free people over time after the intervention). Additionally, the interventions also differ in their invasiveness - and as a result the recovery time - and may have different quality of life implications themselves, regardless of effectiveness. Therefore, there would be more assumptions to be made here if the effectiveness was modelled in detail to try and capture all QoL components. Utility data is also difficult to find as although we have picked up some SF-36 data, there are not many studies looking at the impact of having a stone versus not having a stone as most are from studies comparing treatments. The nature of the condition in terms of episodes of pain is also problematic from a QoL perspective because different tools may also come up with different QoL values depending on when the patients were asked. Different measures also used different recall periods (e.g. SF-36 asks you to think about the last 4 weeks, whereas EQ-5D asks you to think about your health state now – and it is unlikely people were asked their quality of life when they were having an acute pain episode).

A summary of the different scenarios can be found below in Table 1, with differences between them explained below the table to give a top level overview of how the scenarios differ. More specific details on each scenario can be found in later sections of this report.

Table 1: Scenario comparison

	Scenario 1	Scenario 2	Scenario 3
Approach	Using the clinical data on resource use reported in all 7 of the trials. Assuming that this is the resource use required for everyone to be stone free.	Using only studies where: • everyone was stone free at the end of follow up, and • that report initial stone free success rate. (3 studies)	Using only studies that report more detail on the success of multiple lines of treatment. (1 study)
Studies informing resource use (a)	Hendrikx 1999[3] Salem 2009 [8] Sarica 2017 [9] Zhang 2011[11] Kumar 2015A[4] Verze 2010 [10] Pearle 2001 [7] (All 7)	Salem 2009[8] Sarica 2017[9] Zhang 2011[11]	Hendrikx 1999[3]
Advantages of the scenario	It pools all the clinical data we have.	• Everyone being stone free at the end means that the assumption from scenario 1 is now true.	• Offers a different perspective on the treatment pathway as study explicitly includes possibility of 3 lines of treatment • The effectiveness of SWL is much lower than in scenario 2. So this

	Scenario 1	Scenario 2	Scenario 3
		• Initial effectiveness can be used in QALY work.	could be seen as more reflective of UK practice.
Analysis type	Cost comparison	• Cost comparison • Exploratory QALY work	• Cost comparison • Exploratory QALY work
Threshold/ exploratory/ Sensitivity analyses undertaken	• Varying SWL retreatment/ ancillary rates • Varying SWL costs • Pooling retreatment and ancillary procedures • Varying proportion of URS that get stents • Omitting readmission from the analysis • Pooled readmission and major adverse events • Pooled minor and major adverse events (readmissions also excluded)	• Exploratory QALY/QoL work. • Varying initial effectiveness of SWL • Varying SWL costs • Pooling retreatment and ancillary procedures • Varying proportion of URS that get stents	• Exploratory QALY/QoL work. • Including fourth line treatments for those not stone free at end • Varying initial effectiveness of SWL • Omitting readmission from the analysis • Pooled readmission and major adverse events • Pooled minor and major adverse events (readmissions also excluded) • Varying proportion of URS that get stents • Varying SWL costs

(a) Note that these are the studies being used to inform inputs for the scenarios as a whole, but different numbers of studies might be informing different inputs. Additionally, the adverse event outcomes of readmission, minor adverse events, and major adverse events are all being informed by the same studies regardless of scenarios, as these have little impact on the results and pooling more studies was also felt to be better.

Scenario 1 has the advantage of using all the available clinical data, with the assumption that costing up all the resource use will lead to everyone being stone free. This assumption means that this is the resource use difference needed for equivalent outcomes, and therefore this analysis is a cost comparison.

Limitations of this scenario include, that there may still be some difference in outcomes beyond the follow up of the trials. In 3 out of 7 studies not everyone was stone free at the end, in 1 study 93% were stone free with SWL, and 99% with URS. Therefore, if more resources are needed in the SWL arm for everyone to be stone free, then resource use may be being underestimated, in which case the incremental cost might be biasing against URS. However, in the majority of studies, either everyone was stone free at the end, or most people were stone free with small differences between the two groups.

Additionally, there are differences in the studies in terms of whether they reported number of people or number of procedures. Most studies reported the number of people who had retreatment or ancillary procedures rather than number of procedures, but it may be that some people had more than one procedure. In the absence of other information in the costing it was assumed that each person had only 1 procedure.

If URS is a higher cost strategy, this may still be justified by the higher effectiveness rate after the initial procedure (bearing in mind there may be some bias against URS), which would ultimately mean that more people were treated more quickly with URS. If having a stone impacts quality of life - then this could result in a difference in QALYs. There may also be differences in quality of life because of the interventions themselves and the overall number of procedures required. Quality of life is discussed more below when it is explained why a cost utility analysis was not considered appropriate.

This scenario does not have any exploratory QALY work because an average length of follow up would be needed for this, and the studies had different timeframes (ranging from 2 weeks to 3 months). Trying to pool the timeframes of the studies would require transformation of probabilities to rates, and rates have an inherent assumption that they are constant which would not be applicable to this situation, as resource use would not be constant once everyone in a trial was stone free for example.

Scenario 2 uses only 3 studies to inform resource use of retreatments and ancillary procedures, which are studies where everyone is stone free at the end, and also where initial stone free rates are reported. The advantage of using studies where everyone is stone free at the end is that the assumption made in scenario 1 can now be taken as fact, and therefore there is no concern about bias against URS, as these are the resources that would be needed for equivalent (100%) effectiveness of the two strategies.

Additionally, using studies where initial stone free rates are reported means that we have information about the initial part of the pathway. Although there are uncertainties as to where quality of life gains might come from, it can be taken to be true that there is a difference in quality of life between someone that is stone free and someone that is not stone free. Therefore, the difference in initial effectiveness between the two interventions leads to a difference in the number of people who are initially stone free, and therefore a difference in quality of life. Using this logic to infer that the incremental initial effectiveness would be the population contributing to the QALY gain between the interventions allows some exploratory QALY work, looking at the QALY or quality of life differences required for the most expensive intervention to be cost effective. This is explained further in section 1.2.1.5.2.

Disadvantages of this scenario are that it is using only 3 studies to inform inputs, which are fewer studies than in scenario 1. These studies also have the shortest follow up of the 7 studies, which might imply that treatments were given more aggressively in these trials and took place more quickly. This might not be representative of what would happen in clinical practice, and has an impact on the exploratory QALY work because there would have to be a very large quality of life gain during this short time period to reach the QALY needed to make the more expensive intervention cost effective.

Scenario 3 uses only 1 study to inform the resource use inputs on retreatments and ancillary procedures. This study has the benefit of breaking down the number of people that had different lines of treatment, in which case a person could have more than one procedure. This is more detailed than other studies. It also has the advantage of reporting effectiveness that is more reflective of UK practice, which the committee felt was a disadvantage in the clinical review as a whole for the surgery question with regards to SWL. Not everyone was stone free at the end of the trial, so the same assumption as scenario 1 is made, whereby this is the resource use needed to get everyone stone free.

Disadvantages however include that inputs for effectiveness of different lines of treatment are based only on a single study. This study is also from 1999. Although it is important to note that the clinical review did not define a date cut-off, it may be that experience has improved over time for certain techniques such as URS, or technology of SWL machines could have changed, so this is a limitation. Assumptions were made about what intervention should be assumed for interventions considered out of date, for example open surgery was a potential ancillary procedure, but this is not commonly used anymore and the committee decided that the modern day equivalent would be a URS.

Additionally, not everyone was stone free at the end of the trial, which means that we may be underestimating the resource use associated with SWL, as that is the less effective intervention, and therefore biasing against URS because the incremental cost difference estimated may be higher than it is in reality. To combat this, the success of third line treatment is known (there are three lines of treatment included in the trial), and so a sensitivity analysis is undertaken adding a fourth line of treatment and assuming that this would successfully lead to everyone being stone free.

Some exploratory QALY work is also undertaken in this scenario (through a hypothetical cost utility analysis) because the success of different lines of treatment is known. This alongside some assumptions being made about when, in the trial, these different lines of treatment would have taken place, means it can be calculated how long someone has a stone, and how long someone does not have a stone. The utility without a stone can be taken as the UK general population utility, so this allows us to estimate the unknown factor of what the utility of a non-stone free person would need to be to make URS cost effective.

1.2.1.2 Quantifying resource use

1.2.1.2.1 Scenario 1

The outcomes reported by each study can be seen in Table 49 in the appendix. There are 6 studies reporting retreatment and 5 reporting ancillary procedures probabilities, 1 study reporting readmission, 5 studies reporting minor adverse events, and 2 studies reporting major adverse events.

We assumed that all differences in resource use occurred within the time frame of the study. Studies varied in length of follow-up from 2 weeks to 3 months. It may be that shorter studies do not fully capture differences in resource use. However, this is not necessarily the case as it could be that follow up and subsequent retreatment and ancillary procedures were undertaken in a more aggressive nature and so took place more quickly. Additionally, we know that everyone was stone free at the end of the shorter trials.

An alternative would be to make an assumption about the rate of resource use after follow-up ended in shorter studies. However, given that procedure rates will reduce over time as people become stone free, this was not straightforward and the suggested approach was considered to be preferable.

Hendrikx 1999 reports the number of people that had different lines of treatment (e.g. one person could have had two ancillary procedures after a failed initial treatment) and so is overall reporting the number of procedures. In other studies: it is either clear that it is the number of people being reported (in which case one additional procedure per person is occurring in those that needed it), or it is unclear if people could have had more than one procedure because only information like the proportion of retreatments is being reported. Ideally, if an individual can have more than one procedure then rates is the more appropriate outcome to use rather than probabilities. However, as the studies have different timeframes, then as discussed above this wasn’t seen as appropriate (see more in the discussion section). For the purposes of the cost analysis, what we are interested in is the number of procedures to apply costs to. From that perspective, it doesn’t necessarily matter whether the number being reported from Hendrikx 1999 is the number of procedures, or whether in other studies it is the number of people.

Therefore the resource use from all the studies are being pooled to compare in the cost comparison. This may have methodological limitations, but was felt to be the best way to use all of the available data.

Table 2: Clinical inputs Error! Reference source not found.

Parameter	URS probability	SWL probability	SWL relative risk
Retreatment probability	0.02	0.11	5.01
Ancillary procedures probability	0.04	0.10	2.38
Readmission probability	0.13	0.06	0.5
Minor adverse events probability	0.02	0.02	0.67
Major adverse event probability	0.03	0.01	0.15 (peto odds ratio)

Note: for major adverse events an odds ratio was applied to the URS probability because there were no events in the SWL arm, and so a peto odds ratio was calculated from the clinical review meta-analysis for that outcome, rather than a relative risk.

URS was the control arm in the review and the absolute effect is used for the URS probability above, and the relative effect was then applied to the URS absolute effect to derive the SWL probability. Whereas in this analysis URS is the intervention as it is anticipated to be the more expensive intervention.

Failed access or failure of technology was another outcome that was reported by the clinical review. However because of uncertainty around whether this would be double counting resource use - as those whose URS failed, or the stone was not seen with SWL, are likely to have been counted under those having retreatments or ancillary procedures – this outcome has not been included in the costings, and applies to all scenarios.

1.2.1.2.2 Scenario 2

The 3 studies discussed above as being pooled for scenario 2 will be pooled for the stone free outcome, and the retreatment and ancillary procedures outcomes. The stone free probability is the initial stone free probability, and is only used in the exploratory QALY work, not the base case.

The adverse events and readmission probabilities will be informed by the same data as the previous scenario. This is because there will not be many studies to pool if only using the 3 studies for these other outcomes as not all studies report all outcomes (there would only be two studies for minor adverse events, only one study for major adverse events, and no studies for readmission). Additionally, these outcomes are only contributing a small difference to the overall cost difference between the comparators. As was seen from sensitivity analyses 5-7 in scenario 1, the readmission and adverse events outcomes were not making much difference to the overall results whether they were included/pooled. Therefore these are not subject to sensitivity analysis in this scenario as the impact is assumed to be the same.

The forest plots with the clinical data for this scenario can be found in Appendix A:.

Table 3: Clinical inputs Error! Reference source not found.

Parameter	URS probability	SWL probability	SWL relative risk	No. of studies informing outcome
Stone free probability (b)	*0.93*	*0.82*	*0.88*	*3*
Retreatment probability	0.01	0.06	4.1	3
Ancillary procedures probability	0.03	0.08	2.4	3
Readmission probability	0.13	0.06	0.5	1 (a)
Minor adverse events probability	0.02	0.01	0.67	5 (a)

Parameter	URS probability	SWL probability	SWL relative risk	No. of studies informing outcome
Major adverse events probability	0.03	0.01	0.15 (peto odds ratio) (c)	2 (a)

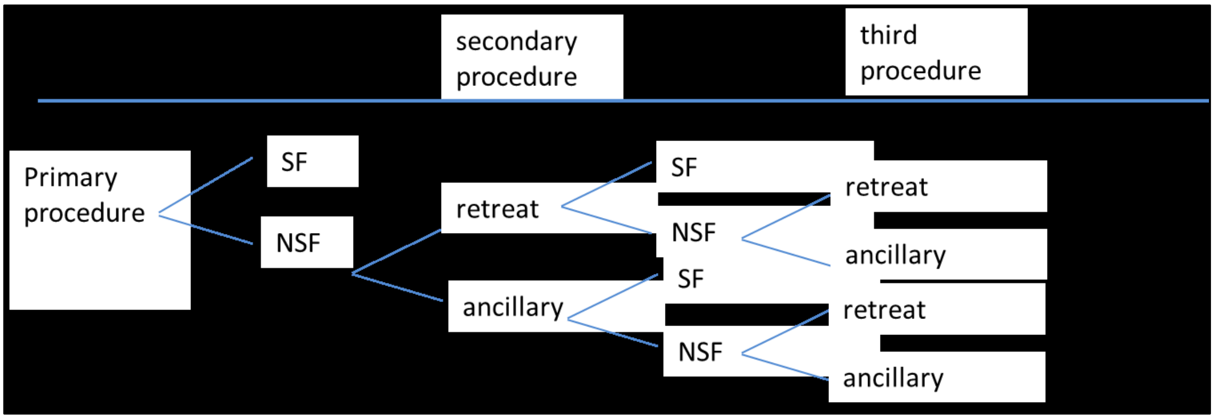
(a) The inputs for these outcomes are not from the 3 studies included in this analysis for the other outcomes.
(b) Only used in exploratory QALY work.
(c) For major adverse events an odds ratio was applied to the URS probability because there were no events in the SWL arm, and so a peto odds ratio was calculated from the clinical review meta-analysis for that outcome, rather than a relative risk.

Note that even though everyone is stone free at the end of the trials informing the stone free, retreatment, and ancillary procedures probabilities, the probabilities may not add up to 100%,and that is because some people not initially stone free may not have had further intervention in the trial. So everyone ends up stone free, but this may not be through a secondary procedure.

1.2.1.2.3 Scenario 3

A different structure to the analysis is used for scenario 3, as there are more lines of treatment being costed then a decision tree structure with branches reflecting pathways was more appropriate. See Figure 1.

Figure 1: Scenario 3 approach



SF = stone free, NSF = non stone free, retreat = retreatment, ancillary = ancillary procedure.

The raw data as reported in the study can be found below.

Table 4: Data from Hendrikx 1999 trial

	First line treatment	Second line treatment	Third line treatment	No. of people	total stone free
SWL intervention	SWL	-	-	38	
	SWL	SWL	-	5	
	SWL	SWL	URS	2	
	SWL	SWL	open surgery	1	
	SWL	URS	-	16	
	SWL	URS	URS	1	
	SWL	URS	open surgery	2	

	First line treatment	Second line treatment	Third line treatment	No. of people	total stone free
	SWL	open surgery	-	1	
	SWL	PCNL	-	3	
				69	64
URS intervention	URS	-	-	79	
	URS	SWL	-	5	
	URS	SWL	URS	1	
	URS	PCNL	-	2	
				87	86

The dashes are interpreted as the patient being stone free and not needing further intervention, therefore it is assumed that where no third line treatment is given the second line treatment was successful.

Open surgery is not something that is undertaken anymore in practice. This may reflect the age of the study. The opinion of the clinical experts was that it could be assumed this would be a URS in current practice for cost purposes, as that is likely to be the modern day equivalent.

From the above, it was possible to work out the probabilities associated with the different parts of the pathway. These can be seen in Table 5. The values in brackets show how the probability was derived using the raw data in Table 4.

Table 5: Probabilities from Hendrikx 1999[3]

		URS	SWL
First line	Primary treatment success	0.91 (79/87)	0.55 (38/69)
	Primary treatment failure	0.09 (8/87)	0.45 (31/69)
Second line	Probability of having a retreatment	0.00 (0/8)	0.26 (8/31)
	Probability of retreatment success	0.00 (0/0)	0.63 (5/8)
	Probability of retreatment failure	0.00 (0/0)	0.38 (3/8)
	Probability of having an ancillary procedure	1.00 (8/8)	0.74 (23/31)
	Probability of ancillary procedure success	0.88 (7/8)	0.87 (20/23)
	Probability of ancillary procedure failure	0.13 (1/8)	0.13 (3/23)
Third line	Success of third line treatment	0 ((1-1)/1) [(b)]	0.17 ((6-5)/6) [(b)]
Ancillary procedure type ratios			
Second line	PCNL proportion	0.25 (2/8)	0.13 (3/23)
	URS proportion	0 [(a)]	0.87 (20/23)
	SWL proportion	0.75 (6/8)	0 [(a)]
Third line	URS proportion	1 (1/1)	1 (6/6)

(a) *These are by default retreatments not ancillary procedures and so can only have a probability of zero here.*
(b) *6 people have third line treatments, however only 5 are not stone free by the end of the trial, therefore the probability of success of the third line treatments is 1/6. For URS: 1 person has a third line treatment and there is also only 86 out of 87 not stone free at the end of the trial which means that individual still has a stone and the probability of success of third line treatment is zero.*

The adverse events and readmission probabilities will be informed by the same data as the other scenarios. These outcomes are only contributing a small difference to the overall cost difference between the comparators. Also, what the clinical review has defined as adverse events is stricter than what the studies report, as a selected number of adverse event outcomes are being reported that match those agreed by the committee on the protocol. Therefore it is likely that the adverse events are underestimates, and hence pooling from more studies is perhaps more appropriate than taking individual studies. On the other hand, readmission for example is only being informed by one study, and some of the patients that informed this outcome seem like the admissions were unrelated to the renal stone, so this may be an overestimate.

1.2.1.3 Resource use assumptions and unit costs

Assumptions regarding resource use are as follows:

- Retreatment was defined as the same treatment again (same as the primary treatment)
- Primary or retreatment SWL intervention is one session only (not always clear if this is the case from studies but agreed reasonable assumption with committee).
- Ancillary procedures were defined as a treatment different to the primary procedure.
- The studies capture all relevant differences in resource use and this is irrespective of study length
- Where the number of people having a procedure is reported (rather than the number of procedures) it is assumed that each person had 1 procedure.

The clinical studies informing the inputs were reviewed by the health economist to identify resource use detail. For example, how many sessions of SWL were undertaken as the primary procedure, what were considered ancillary procedures, or minor or major adverse events, in order to identify what information could be found from the studies and what would need assumptions where the studies were unclear.

Another assumption that has been inferred from the data, is about the sequencing of treatments. In the trials it is somewhat arbitrary whether someone was given a retreatment after the primary treatment failed, or were given an ancillary treatment, and this could be affecting costs. Practice in other countries may also be different to that of the UK as there were no UK trials. In some trials there may have been more than two lines of treatment, for example ancillary procedures may also apply to those who have failed a retreatment. This is likely to be the case in studies where it is difficult to distinguish if it is the number of people or number of procedures being reported. As the majority of the studies imply that ancillaries and retreatments are both second line treatments, an assumption has been made that the probability of both retreatment and ancillary procedures applies to the initial 1000 cohort for each strategy, which effectively means that both are mutually exclusive second line treatments. A sensitivity analysis explores pooling the retreatment and ancillary procedure rates and making assumptions about what the second line procedures are. This is to make sure that there is an element of being conservative towards SWL, as this is both a cheaper intervention and also has cheaper second line treatments compared to URS.

Resource use assumptions and unit costs used in the analysis are summarised in **Table 6** and **Table 7** below. These are discussed in more detail in the subsequent sections.

Table 6: Resource use assumptions

Parameter	URS	SWL
Number of initial sessions of SWL	-	1
Number of repeat sessions of SWL for retreatment	-	1
Proportion requiring stent insertion (b)	NA	1%
Proportion requiring stent removal (a)	70%	1%

Parameter	URS	SWL
Proportion requiring confirmation of stone passing	100%	100%
Proportion requiring ultrasound to confirm stone passing	10%	10%

(a) Applies to both URS as an initial procedure and URS as a repeat or ancillary procedure.
(b) This is not applicable for URS as the stent is inserted at the time of the URS procedure.

Table 7: Unit costs

Parameter	NHS reference cost description	Cost
SWL cost (per session)	LB36Z Extracorporeal Lithotripsy Day case schedule	£452
URS cost	Elective schedule: Weighted average of LB65C, LB65D and LB65E, Major Endoscopic, Kidney or Ureter Procedures, 19 years and over. (a) = £2,605 Day case schedule: Weighted average of LB65C, LB65D and LB65E, Major Endoscopic, Kidney or Ureter Procedures, 19 years and over. (a) = £1,739	£2,172 (b)
PCNL	Weighted average of LB75A, LB75B, Percutaneous nephrolithotomy Elective schedule (a)	£5,195
Stent insertion, stent removal (d)	LB09D Intermediate Endoscopic Ureter Procedures, 19 years and over Day case schedule	£1,018
Follow up consultation	WF01A Consultant led - Non-Admitted Face to Face Attendance, Follow-Up, Urology (c)	£103
Ultrasound cost	RD40Z Ultrasound Scan with duration of less than 20 minutes, without contrast Outpatient schedule	£52
Re-admission	VB03Z Emergency Medicine, Category 3 Investigation with Category 1-3 Treatment. Type 01 admitted	£307
Minor adverse event	VB09Z Emergency Medicine, Category 1 Investigation with Category 1-2 Treatment Type 01 non admitted	£119
Major adverse event	VB03Z Emergency Medicine, Category 3 Investigation with Category 1-3 Treatment. Type 01 admitted	£307

Source: NHS references costs 2016-17[6]
SWL: shockwave lithotripsy; URS: Ureteroscopy; PCNL: percutaneous nephrolithotomy.
(a) Includes all complication categories, and is weighted by activity with excess bed days incorporated.
(b) 50% elective and 50% day case cost as was decided by committee to reflect UK practice.

(c) Note that the cost of an x-ray is included in the consultation cost as x-ray imaging is bundled because "plain film x-rays that are part of an admission or outpatient attendance are not reported separately due to their high volume and low cost" – NHS reference costs 2016-17 guidance.
(d) Stent insertion and stent removal OPCS codes were provided by the guideline committee. Using the 'HRG 2016-17 engagement grouper code to group' spreadsheet - these OPCS codes were found to map to the same HRG code.

1.2.1.3.1 Primary procedures

SWL

The studies are not very clear on how many sessions of SWL were included as the primary procedure: Kumar 2015A[4] stated there could have been up to 4 sessions, but this again was unclear in terms of how many counted as initial and how many as retreatments. Verze 2010[10] reported the proportion of people that needed a certain number of sessions to clear the stone. This led to an average of 1.6 sessions. However the study combined people with a stone of <10mm and >10mm, so it is unclear the number of sessions that apply to the former group, which is the group we are interested in.

Economic subgroup opinion was that for a ureteric stone, up to 2 sessions of SWL would be attempted in UK practice. Therefore, it was agreed that 1 session could be assumed for the initial procedure (and another 1 session to be assumed for those that have a retreatment).

This was explored in sensitivity analyses, whereby 2 sessions were assumed for the primary treatment instead of 1. This was in order to increase the cost of SWL to be more conservative towards SWL, and see what impact this would have on the results.

The cost of SWL was taken from NHS reference costs where this was interpreted as the cost per session. There were concerns highlighted from the committee about the NHS reference cost, as although it is an average, it is likely to be skewed towards a lower cost by the high volume centres. Similarly, if SWL was more available then machines only being used for small volumes could drive up the NHS reference cost. However, as this was the best source currently available, it was used with these caveats in mind, and threshold analyses undertaken around the cost per session.

URS

For URS, it was decided with the subgroup that the cost of URS would be a cost based on 50% day case cost and 50% elective cost, as this reflects UK practice and was arrived at based on national audit data (from GC member access to audit). Using the relative activity from NHS reference costs may be less accurate because the HRG covers various kidney procedures and may include other procedures as well as URS.

Stent usage

It is common for a stent to be inserted during the URS procedure, which protects from kidney obstruction due to ureteric swelling in the post-operative phase. This would accrue a cost in terms of needing to remove the stent at a later date (the insertion would happen at the time of the URS and so should be included in NHS reference cost for URS). The proportion reported as requiring a stent ranged from all URS patients in some studies to a much lower proportion in other studies where it was only used as a way of treating particular adverse events, or unclear in other studies. Consensus within the economic subgroup was that it would be most appropriate to use national audit data of UK practice where 70% of those who had URS had a stent put in (from GC member access to audit). This proportion is assumed to apply to all instances where a URS is a treatment option, whether that is as the primary procedure or retreatment in the URS strategy, or as an ancillary procedure in the SWL strategy. The proportion used is tested in sensitivity analyses.

It is rare for a stent to be needed after an SWL procedure, however it was not felt appropriate to have it as zero. It was agreed with the committee that 1% of people that had an SWL would have a stent, and this would need a procedure to both insert and remove the stent, and applies to both to the primary procedure and retreatment SWL.

The cost of a stent was identified by mapping the OPCS code provided by a GC member that they would use for reimbursement of a stent procedure. It was discussed with the committee how not everyone uses the same code and some clinicians might use the cost of an outpatient appointment for example, however given the variation in coding practice – the most conservative estimate was used.

Confirmation of stone passing

Other resource use that should be considered includes the resources involved in confirming whether the stone has passed/been removed. In the model, it is assumed there will be a follow-up involving an outpatient appointment and imaging after every SWL and URS procedure and any other ancillary procedures e.g. percutaneous nephrolithotomy (PCNL)).

For SWL, the opinion of the economic subgroup was that everyone that has SWL would be followed up with a consultation and also imaging to confirm the stone has passed. With a URS, it may be that because URS is invasive, the clinician can see whether they have removed the stone. However, it was agreed with the economic subgroup that in the base case analysis everyone would be followed up.

The type of imaging is usually a plain x-ray in most cases (90%) but might also be an ultrasound (10%). These proportions apply to both strategies. As x-ray is a bundled cost (because of its high volume and low cost it is not reported separately to an admission or outpatient attendance in NHS reference costs), then this does not have to be costed separately, and so only the ultrasound cost will be applied to 10% of people followed up.

Confirmation of stone passing is also assumed to happen after every procedure (e.g. after every session of SWL, every URS, or every other procedure that might take place that is ancillary/additional e.g. PCNL.

The cost of follow up was assumed to be a consultant led - Non-Admitted Face to Face Attendance follow up for urology.

1.2.1.3.2 Retreatments and ancillary procedures

For the SWL strategy, as previously mentioned, a retreatment consists of another session of SWL. In terms of ancillary procedures, the view of the economic subgroup was that a failed SWL would always be followed with a URS, therefore URS is assumed to be the ancillary procedure in the SWL group.

For the URS strategy, the view of the subgroup was that if a patient failed a URS then a repeat URS would be performed. This is the definition of retreatment however, so for ancillary procedures a different procedure applies. The subgroup agreed that in the small proportion of cases where URS failed, a PCNL would be undertaken, and so this was assumed to be the ancillary procedure in the URS group.

PCNL unit cost was taken from NHS reference costs, and is the most expensive procedure type.

A sensitivity analysis will be undertaken where the retreatment and ancillary probabilities are pooled and assumptions made about what the secondary procedures would be, as a way of varying costs (only applies to scenarios 1 and 2). More detailed can be found on this in sensitivity analyses 3 and10.

1.2.1.3.3 Treating adverse events

A minor adverse event could result in either a GP appointment or hospital attendance, as minor adverse events are likely to be a fever, UTI, or bleeding for example. The cost of an A&E attendance has been assumed because the URS arm has more adverse events, and so using a higher cost means we are being more conservative to URS by increasing its cost compared to SWL, in order to be more certain about its potential cost effectiveness. As if it is cost effective with higher costs being used, then it will be even more cost effective if costs were lower.

Based on subgroup opinion, a re-admission has been classified as an emergency admission with a category 3 investigation and category 1-3 treatment. This is used to represent that a readmission is likely to involve a radiological or surgical investigation, for example an ultrasound is classified as a category 3 investigation, and minor surgery or insertion of a catheter is a category 3 treatment.

A major adverse event was assumed to have the same cost as a re-admission, because assuming a major adverse event is a Clavien Dindo grade 3-4 involves hospitalisation and an intervention surgical or radiological.

It is possible that there may be some double counting between major adverse events and readmission, so these are subject to sensitivity analysis by pooling the outcomes and also omitting readmission. It is important to note that the cost of an admission at A&E does not include any surgical interventions that may follow after a patient has been admitted such as if a URS is needed. This is a limitation, however would involve making further assumptions about what would happen upon readmission. As this would only apply to a small proportion of people - and it is also possible that we may have overestimated minor adverse event costs, as well as potentially underestimated major adverse event costs, but on the other hand as there may be some double counting – then this is unlikely to have a big impact on the results as these impacts would balance each other out.

It is important to note that for some outcomes the clinical data only comes from a small number of studies. For re-admission, only one study reported this. It is also possible that the re-admissions reported in the study are not related to the actual intervention, so the clinical review data needs to be interpreted with caution (in this particular case – 1 of the 2 re-admissions in the SWL arm were for a herniated lumbar disk, and 1 of the 4 in the URS arm was for exacerbation of diabetes). When dealing with such small numbers, any small changes (for example, 2 events instead of 1) will have a large impact on the absolute effect. A sensitivity analysis omits re-admission to see the impact this has.

1.2.1.4 Utilities

For exploratory QALY work, the following utilities were used as estimates: for people without a renal stones, the UK general population EQ-5D (3L) quality of life from the Health Survey for England 2014 was used (0.874). This was the latest version of the survey available that had EQ-5D data.

For those with a renal stone, an estimate of quality of life was taken from the Pickard study, which was a large UK RCT included in the medical expulsive therapy review. This had a population of ureteric stones <10mm. The baseline EQ-5D value was reported for each intervention being compared in the trial so a weighted average was calculated (0.684)

1.2.1.5 Sensitivity analyses

1.2.1.5.1 Scenario 1

1. **Varying the retreatment and ancillary procedure probabilities of SWL**

Committee opinion was that the effectiveness rates of SWL seen in the studies included in the clinical review are higher than seen in UK practice. One UK audit found that stone free rates for ureteric stones were around 60%[2]. However this is based on a mixed number of sessions, rather than just after 1 session for example (whereas in our analysis we assume 1 session initially). Therefore effectiveness could be lower than this.

One potential explanation could be that that the level of effectiveness would be achievable if SWL was used more as operators would have greater experience, however to reflect current UK practice, the retreatment and ancillary procedure probabilities are being varied in a sensitivity analysis, as this is a reverse method of reducing the effectiveness of SWL. URS rates were not varied.

Table 8 shows the values that were used in the sensitivity analysis. The retreatment and ancillary procedure values are a range between the values from the base case, which were generally seen as low by the committee, and values closer to UK practice (down to effectiveness of SWL of 40%, which was considered anecdotally to be a very low effectiveness from the committees experience in practice). Ten increments were chosen between the range of values, with equal increases in the increments.

The ratio of ancillaries to retreatments from the base case are kept the same in each row.

Table 8: Sensitivity analysis 1 - SWL retreatment and ancillary procedure probabilities

	Probability of needing retreatment	Probability of needing ancillary procedure	Effectiveness of initial SWL this suggests (a)
Base case values	10.9%	9.9%	79.3%
	12.9%	11.8%	75.3%
	15.0%	13.6%	71.4%
	17.0%	15.5%	67.5%
	19.1%	17.4%	63.6%
	21.1%	19.2%	59.6%
	23.2%	21.1%	55.7%
	25.2%	23.0%	51.8%
	27.3%	24.9%	47.9%
	29.3%	26.7%	43.9%
Suggested UK practice values	31.41%	29%	40.0%

(a) Based on the following equation assuming everyone stone free at the end: 100% - retreatment probability – ancillary probability

2. **Varying cost of SWL**

The base case assumption of 1 session for the initial treatment is altered to 2 sessions. The number of repeat sessions for retreatment still remains at 1 session as in the base case. See section 1.2.1.3.1 for more explanation about this sensitivity analysis.

Additionally, a threshold analysis is undertaken to see what the cost of an SWL session (both initial and repeat) would have to be to make the comparators cost neutral.

3. Pooling outcomes of retreatment and ancillary

Studies differ as to how they treat people who have failed first line treatment, in terms of whether they retreat people or whether an ancillary procedure is undertaken instead. This could also be explaining the heterogeneity in the forest plots for the retreatments and ancillary procedures. The retreatment and ancillary procedures undertaken in the SWL group are also lower cost procedures than the ones in the URS group. Therefore it's possible that even if there are higher retreatment and ancillary probabilities in the SWL arm, these may add up to lower costs than the secondary procedures in the URS arm.

To avoid any potential bias that SWL may have cheaper primary interventions as well as secondary interventions, the retreatment and ancillary procedure rates were pooled, and the type of secondary procedure varied. The meta analysed pooled probabilities are summarised in Table 9; the forest plot is available in the appendix.

Table 9: Pooled retreatment and ancillary procedure probabilities

	URS	SWL	SWL relative risk	Source
Probability	0.06	0.23	3.64	Pooling of 6 studies: Hendrikx 1999 [3] Salem 2009[8] Sarica 2017 [9] Zhang 2011[11] Kumar 2015A[4] Verze 2010[10]

The proportional divide between retreatment and ancillary procedures for both URS and SWL was found by taking the retreatment and ancillary procedure probabilities and dividing by the sum of both probabilities (see Table 10). These proportions were then multiplied by the pooled procedure probabilities to derive the probability of each type of secondary procedure (see Table 11).

The secondary procedures were varied in various ways, as shown in Table 10. The type of ancillary procedure is being tested, as well as increasing the proportion that have a URS as a secondary procedure until everyone gets a URS as the secondary procedure in both groups. This will mean that the total secondary procedures should have a higher cost for SWL if there are more of them for SWL, and they are also of the same type of procedure for both interventions. All other inputs are kept the same as the base case (such as the proportion of URS procedures that have a stent).

Table 10: SA3 values - proportions allocated to types of secondary procedures

	Interventions					
	SWL			URS		
Procedure	SWL (retreat)	URS (anc)	PCNL (anc)	URS (retreat)	SWL (anc)	PCNL (anc)
Base case proportions	52%	48%	-	34%	-	66%
Assuming different ancillary	52%	-	48%	34%	66%	-
Assuming 50% split between different ancillaries	52%	24%	24%	34%	33%	33%
Increasing proportion that have URS as secondary treatment in increments, until 100% of the secondary procedures are URS in both groups.	42%	58%	-	47%	-	53%
	31%	69%	-	61%	-	39%
	21%	79%	-	74%	-	26%
	10%	90%	-	87%	-	13%

	Interventions					
	0%	100%	-	100%	-	0%

Table 11: SA3 values – probability of each type of secondary procedure

	Interventions					
	SWL			URS		
Procedure	SWL (retreat)	URS (anc)	PCNL (anc)	URS (retreat)	SWL (anc)	PCNL (anc)
Base case proportions (a)	12.0%	11.0%	0.0%	2.2%	0.0%	4.2%
Assuming different ancillary	12.0%	0.0%	11.0%	2.2%	4.2%	0.0%
Assuming 50% split between different ancillaries	12.0%	5.5%	5.5%	2.2%	2.1%	2.1%
Increasing proportion that have URS as secondary treatment in increments, until 100% of the secondary procedures are URS in both groups.	9.6%	13.4%	0.0%	3.0%	0.0%	3.3%
	7.2%	15.8%	0.0%	3.8%	0.0%	2.5%
	4.8%	18.2%	0.0%	4.7%	0.0%	1.7%
	2.4%	20.6%	0.0%	5.5%	0.0%	0.8%
	0.0%	23.0%	0.0%	6.3%	0.0%	0.0%

(a) Note that the SWL figures are slightly different to the main base case inputs because of differences in how the meta-analysis weights the studies using the pooled retreatment and ancillary data.

4. Varying proportion of URS patients that have stents post-surgery

In the base case, it is assumed that 70% of people that have a URS have a stent inserted, and therefore have to have it removed later. This is an estimate from UK audit data because studies were very variable from everyone having a stent to only those that had complications, which would be a much smaller proportion.

The assumption on stent use for URS creates higher costs for the URS strategy. As this guideline has also recommended that stents not be used, then it is expected the use of stents will reduce in practice, and so testing this assumption is also important from that perspective.

This value has been varied in a sensitivity analysis from between 0% to 100%, going up in increments of 20%.

5. Omitting re-admission outcome from the analysis

As there is a risk that readmission and major adverse events are double counted, as a major adverse event could also result in a readmission, then the outcome of readmission has been omitted from the analysis to see the impact of this on the results. The outcome itself is only based on one study showing there are around twice as many readmissions for URS than for SWL.

6. Pool major adverse events and re-admission

As discussed above, to counteract any potential double counting, these two outcomes have been pooled, which would reduce the impact of the readmission outcome as the study will be pooled with other studies creating a larger pool of people in the denominator. As the same cost was applied to both these outcomes then this will still be applied to the pooled analysis.

The table below reports the inputs for the pooled outcomes.

Table 12: Pooled major adverse events and re-admission

	URS	SWL	SWL relative risk	Source
Probability	0.04	0.01	0.19	Pooling of 3 studies: Hendrikx 1999 [3] Zhang 2011[11] Pearle 2001[7]

7. **Pool major and minor adverse events**

There is some heterogeneity present in the minor adverse events, with different studies being on different sides of the forest plot. This is probably due to the small number of events taking place and the studies being small themselves and underpowered to detect differences in the outcomes. Minor and major adverse events have been pooled to provide an outcome of overall adverse events. This pooled outcome is reported below.

Table 13: Pooled major and minor adverse events

	URS	SWL	SWL relative risk	Source
Probability	0.04	0.02	0.36	Pooling of 5 studies: Hendrikx 1999[3] Kumar 2015A[4] Salem 2009[8] Zhang 2011[11] Pearle 2001[7]

Major and minor adverse events had different costs in the base case analysis as major adverse events were expected to lead to an admission. The cost of major adverse events is applied to be conservative. Conversely, it may be possible that the cost of a major adverse event has been underestimated because the cost used is based on the emergency medicine schedule in NHS reference costs, which does not include the cost of any procedures after the patient is admitted. For example, if it results in additional surgery being needed, then the cost attached to the NHS reference cost used would be low. Therefore, although there are more minor adverse events than major, the cost used is likely represent an average of what might be involved to treat some minor adverse events and a very small amount of major adverse events.

Additionally, readmissions will also be zero for both strategies in this sensitivity analysis as it has already been discussed how this might be double counting with major adverse events, and omitting it has also had little impact on the results.

1.2.1.5.2 Scenario 2

Threshold analysis on QALYs

Taking the incremental cost per person, and dividing this by the NICE threshold gives the QALY gain that would be needed to make the more expensive intervention (in this case URS) cost effective. This involves a simple rearrangement of the ICER equation:

Figure 2: equation rearranged to find incremental QALY gain needed

$$QALYs(URS) - QALYs(SWL) = \frac{Costs(URS) - Costs(SWL)}{ICER}$$

Where: Costs(URS) = total costs for URS; QALYs(URS) = total QALYs for URS

Where the ICER = £20,000.

The 3 trials being pooled have a short time frame: two studies are 2 weeks long and one is 4 weeks long. For simplicity we shall assume they are all 2 weeks in order to have a timeframe to use in QALY calculations. Given the short time frame, this means the quality of life difference that would need to be gained from the URS group over the SWL group would have to be quite large, because once that is divided by a small timeframe this will shrink the total QALY gain.

As a QALY is made up of a combination of length of life and quality of life, then dividing the QALY needed by the time period that this would apply to leaves the quality of life part of the QALY. This would tell us on average, the quality of life gain that someone who had a URS, instead of an SWL, will need to have over a 2 week period, to make URS cost effective. There are however lots of caveats to this – see the sensitivity analysis results for more detail, and also the next section on exploratory QALY calculations where the source of this quality of life gain is explored further.

Exploratory QALY calculations

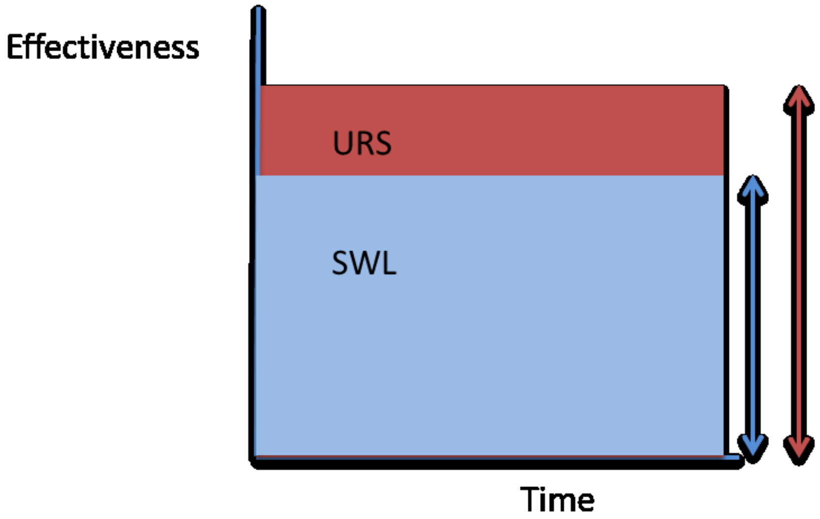
It is uncertain where the quality of life gain might come from when comparing URS to SWL. One source of the gain is likely to be because of a difference in quality of life between people who still have a stone and people who no longer have a stone, and therefore because there are more stone free people in the URS group. Some arguments we can take as fact in order to help us work through some exploratory calculations around quality of life are:

- Initial treatments have different probabilities of success
- A stone free person will have a higher QoL than a non-stone free person
- Over time further treatments will result in those not initially stone free becoming stone free
- The QALY gain is therefore likely to come from the difference in numbers of people who are initially stone free between the two interventions.

There are however a lot of unknowns: there was no data identified on the QoL decrement of someone with a stone; we don’t know the timeframe that QoL would apply to in terms of how long people are in pain, and the frequency of their episodes; and we don’t know how long people will be waiting for further treatments if their treatment is initially unsuccessful. However, the principles in the bullet points outlined above can be taken to be true.

The larger the difference in (initial) effectiveness between the two interventions, the more additional people that would be successfully treated with URS who will then accrue this QoL gain from not having a stone anymore. This is demonstrated in the diagram below.

Figure 3: Difference in effectiveness of treatments



We know the initial effectiveness of the interventions, as we are using studies that specifically tell us this.

So even though we assumed that the effectiveness of the two strategies would end up being the same after we accounted for the additional procedures of retreatments and ancillaries, Figure 3 is actually assuming that this difference in stone free people between the initial interventions remains for the whole time period of the trials. Effectively this means that people have their other procedures at the end of the trials. In reality, people won't have their retreatments right at the end of the studies, so if anything, we are being generous towards URS because we are saying this incremental effectiveness will remain throughout the time period. But that won't be the case because as people have their other treatments, which happens over time, then the red block of effectiveness difference will get narrower and narrower. However as the timeframe of the studies is very short, this assumption may also not be too far from the truth.

In summary, this exploratory QALY work uses the assumption that those who are initially stone free are the ones that are accruing all the QoL. Therefore by apportioning the average QoL difference between the two groups (calculated in the previous threshold analysis) only to the people we think this difference will come from, we can work out the QoL (or utility) difference needed between a stone free person and a non-stone free person to make URS cost effective.

Other sensitivity analyses

8. **Varying the initial effectiveness of SWL**

Effectiveness rates of SWL are varied in order to allow the extrapolation of the exploratory QALY work to a range of SWL effectiveness levels, to further assess the feasibility of the cost effectiveness of URS.

As we know, from the data being used in this scenario analysis, that everyone is stone free at the end of the trials, it would make sense that the sum of the probabilities of initially stone free, retreatment, and ancillary procedures would add to 100%. However, this is not the case, because in one of the studies not everyone that failed the initial treatment needed other treatments:

URS = 0.93 + 0.01 + 0.03 = 0.977

SWL = 0.82 + 0.06 + 0.08 = 0.954

This is possibly because they already passed their stone without needing further treatment. The study doesn't say.

Because of this, in this sensitivity analysis, the probability of not needing any further treatment (around 5% for SWL) is kept the same, and only the retreatment and ancillary probabilities of SWL are varied (as a consequence of the initial effectiveness), so the probabilities still all add to 100%.

Table 14 shows the values that were used in the sensitivity analysis. The effectiveness values range between the values from the base case (which for the stone free rate of SWL was generally seen as too high by the committee) and values closer to UK practice chosen by the committee. Ten equal increments within this range were used in sensitivity analyses. The ratio of ancillaries to retreatments is kept the same as in the base case.

Table 14: Sensitivity analysis 8 - figures used

	Initial effectiveness of SWL	Probability of needing retreatment for SWL	Probability of needing ancillary procedure for SWL
Base case value	82.0%	5.8%	7.6%
	77.8%	7.6%	10.0%
	73.6%	9.5%	12.4%
	69.4%	11.3%	14.8%
	65.2%	13.1%	17.1%
	61.0%	14.9%	19.5%
	56.8%	16.7%	21.9%
	52.6%	18.6%	24.3%
	48.4%	20.4%	26.7%
	44.2%	22.2%	29.1%
Suggested UK practice values	40.0%	24.0%	31%

The QALY exploratory analysis was undertaken for each of the sensitivity analysis values, to explore whether URS is likely to become cost effective using the same assumption that the difference in initial effectiveness between the treatments is leading to the quality of life differences.

9. Varying cost of SWL

The base case assumption of 1 session for the initial treatment of SWL is altered to 2 sessions.

Additionally a threshold analysis is undertaken to see what the cost of an SWL session would have to be to make the comparators cost neutral.

10. Pooling retreatment and ancillary procedures

As in scenario 1, there is a lot of heterogeneity from the forest plots for the retreatments and ancillary procedures outcomes, which could be explained by the fact that different studies have different criteria for which procedure to undertake after a failed procedure.

To avoid any potential bias that SWL may have cheaper primary interventions as well as secondary interventions, the retreatment and ancillary procedure rates are pooled, and the type of secondary procedure is varied.

Table 15: Pooled retreatment and ancillary procedure probabilities

	URS	SWL	SWL relative risk	Source
Probability	0.05	0.12	2.72	Pooling of same 3 studies as the base case of scenario 2.

The proportional divide between retreatment and ancillary procedures for both URS and SWL was found by taking the retreatment and ancillary procedure probabilities and dividing by the sum of both probabilities (see Table 16). These proportions were then multiplied by the pooled procedure probabilities to derive the probability of each type of secondary procedure (see Table 17).

These probabilities were varied in various ways, as shown in Table 16. The type of ancillary procedure is being tested, as well as increasing the proportion that have a URS as a secondary procedure until everyone gets a URS as the secondary procedure in both groups. This will mean that the total secondary procedures should have a higher cost for SWL if there are more of them for SWL, and they are also of the same type of procedure for both interventions. All other inputs are kept the same as the base case (such as the proportion of URS procedures that have a stent).

Table 16: SA10 values - proportions allocated to types of secondary procedures

	Interventions					
	SWL			URS		
Procedure	SWL (retreat)	URS (anc)	PCNL (anc)	URS (retreat)	SWL (anc)	PCNL (anc)
Base case proportions	43%	57%	-	31%	-	69%
Assuming different ancillary	43%	-	57%	31%	69%	-
Assuming 50% split between different ancillaries	43%	28%	28%	31%	34%	34%
Increasing proportion that have URS as secondary treatment in increments, until 100% of the secondary procedures are URS in both groups.	42%	58%	-	47%	-	53%
	31%	69%	-	61%	-	39%
	21%	79%	-	74%	-	26%
	10%	90%	-	87%	-	13%
	0%	100%	-	100%	0%	0%

This leads to the following probabilities of having each type of secondary procedure:

Table 17: SA10 values – probability of each type of secondary procedure

	Interventions					
	SWL			URS		
Procedure	SWL (retreat)	URS (anc)	PCNL (anc)	URS (retreat)	SWL (anc)	PCNL (anc)
Base case proportions (a)	5.4%	7.0%	0.0%	1.4%	0.0%	3.1%
Assuming different ancillary	5.4%	0.0%	7.0%	1.4%	3.1%	0.0%
Assuming 50% split between different ancillaries	5.4%	3.5%	3.5%	1.4%	1.6%	1.6%
Increasing proportion that have URS as secondary treatment in increments, until 100% of the secondary procedures are URS in both groups.	5.2%	7.2%	0.0%	2.2%	0.0%	2.4%
	3.9%	8.5%	0.0%	2.8%	0.0%	1.8%
	2.6%	9.8%	0.0%	3.4%	0.0%	1.2%
	1.3%	11.1%	0.0%	3.9%	0.0%	0.6%

	Interventions					
	0.0%	12.4%	0.0%	4.5%	0.0%	0.0%

(b) Note that the SWL figures are slightly different to the main base case inputs because of differences in how the meta-analysis weights the studies using the pooled retreatment and ancillary data.

11. Varying proportion of URS patients that have stents post-surgery

As in scenario 1, the probability of needing a stent after a URS is varied between 0% and 100% in 20% increments (base-case analysis value 70%).

1.2.1.5.3 Scenario 3

Exploratory QALY calculations

Making estimations about the time points at which further treatment was happening, as well as about the utility of those with and without stones, allows a cost utility analysis to be undertaken, applying these utilities to the time points at which treatments are occurring.

The study (Hendrikx 1999)[3] had a timeframe of 3 months, and reports up to 3 lines of treatment. In terms of when further procedures of either retreatments or ancillaries might have occurred, the study states “the outcome [of stone free or not] was determined on a regularly planned outpatient visit 2 weeks after treatment in 77.5% and 12 weeks after treatment in 16.6%, and during an intercurrent visit or after calling in 5.7%”. It could be inferred from this, that for most people the outcome was determined at 2 weeks, it may then take some time before retreatments/further treatments can be organised. We could take a time point of say 4 weeks when second line treatments may occur, if there is a delay for further treatment to happen. Extrapolating this even further, then third line treatments may be occurring at around 8 weeks. This is a simplified assumption based on what can be inferred from the study.

The quality of life of someone who is stone free is based on the EQ-5D 3L from the Health survey for England 2014 which was the latest available.

The quality of life of someone without a stone could then be subject to a threshold analysis to find what the utility of a non stone free person would need to be to make URS cost effective, at the NICE threshold of £20,000 per QALY.

Other sensitivity analyses

12. Including the fourth line treatments

As we know not everyone is stone free at the end of the trial, then in this sensitivity analysis it is assumed those who are not stone free go on to have fourth line treatments, which are assumed to be URS in both arms and are also assumed to be successful.

13. Varying the effectiveness of the primary SWL

The effectiveness of SWL is lower from the clinical data informing this scenario analysis than in the other scenario analyses. However, because the results are likely to be sensitive to this, as it is the initial difference in effectiveness that creates the biggest divide in costs because it impacts the need for downstream treatment. This is tested in a sensitivity analysis.

The base case value of 55% effectiveness is taken to be the midpoint, and this is varied to as low as 40%, and to as high as the effectiveness of URS (91%).

Table 18: Values used in SA13 – varying effectiveness of primary SWL

	Effectiveness of primary SWL
Same as URS	0.91
	0.84
	0.77
	0.69
	0.62
Base case	0.55
	0.52
	0.49
	0.46
	0.43
Low effectiveness	0.4

The effectiveness of URS is not varied as this is similar to the effectiveness in practice in the committee's opinion.

The total QALYs and ICER derived from the exploratory cost utility analysis were also reported for each of the effectiveness values above.

It is anticipated that the incremental costs will be smaller for this scenario than in scenario 1, because the effectiveness of SWL is lower and therefore more further procedures are needed, making that strategy more expensive than in the other scenarios. Because of this, the following sensitivity analyses around readmission and adverse events that were undertaken for scenario 1 are also undertaken here to see what impact these will have on the incremental cost:

14. **Omitting readmission from the analysis**
15. **Pooled readmission and major adverse events**
16. **Pooled minor and major adverse events (readmissions also excluded)**

(see sections 1.2.1.3.3 and 1.2.1.5.1 for more detail on the rationale for these sensitivity analyses).

17. **Varying the proportion of URS that get stents**

The proportion of those who have URS that get stents is assumed to be 70% in the base case. This is varied in a sensitivity analysis in the same way as it was for the other scenarios.

18. **Varying cost of SWL**

The base case assumption of 1 session for the initial treatment is altered to 2 sessions.

Additionally a threshold analysis is undertaken to see what the cost of an SWL session would have to be to make the comparators cost neutral.

1.2.1.6 Computations

The costing was constructed in Microsoft Excel 2010. Resource use was multiplied by costs and summed to create the total cost of the strategies. A cohort of 1000 people per strategy was assumed so that the probabilities could be applied to generate numbers of people that each event would apply to e.g. retreatment, minor adverse events.

Probabilities were derived from the clinical review where possible and supplemented with assumptions where clinical review data was not available.

The total costs for a cohort of 1000 people per strategy were divided by 1000 to find the cost per person for each strategy. An incremental was then taken. In some scenarios, in order to back-calculate to find the QALYs that would be needed to make URS a cost effective intervention: the incremental cost was divided by the NICE threshold of £20,000 per QALY. This value was then back-calculated even further by dividing it by the most common time frame of the trials pooled, in order to derive the QoL difference that would need to be achieved to make URS cost effective over SWL. In scenario 3, cost effectiveness was explored through an assumption based exploratory cost utility analysis.

1.2.1.7 Analysis validation

The analysis was developed in consultation with the committee: analysis structure, inputs and results were presented to and discussed with the committee for clinical validation and interpretation.

The analysis was systematically checked by the health economist undertaking the analysis; this included inputting null and extreme values and checking that results were plausible given inputs. The analysis was peer reviewed by a second experienced health economist from the NGC, this included systematic checking of the calculations.

1.2.1.8 Estimation of cost effectiveness

The widely used cost-effectiveness metric is the incremental cost-effectiveness ratio (ICER). This is calculated by dividing the difference in costs associated with 2 alternatives by the difference in QALYs. The decision rule then applied is that if the ICER falls below a given cost per QALY threshold the result is considered to be cost effective. If both costs are lower and QALYs are higher the option is said to dominate and an ICER is not calculated.

$$ICER = \frac{Costs(B) - Costs(A)}{QALYs(B) - QALYs(A)}$$ Where: Costs(A) = total costs for option A; QALYs(A) = total QALYs for option A	Cost effective if: • ICER < Threshold

In this analysis, QALYs were obtained in an exploratory way by back-calculating from the incremental cost (in the above formula) to help with decision making, as the analysis is a costing comparison. This was done using the same cost per QALY threshold that would be used for the above decision rule.

1.2.1.9 Interpreting Results

NICE's report 'Social value judgements: principles for the development of NICE guidance'[5] sets out the principles that committees should consider when judging whether an intervention offers good value for money. In general, an intervention was considered to be cost effective if either of the following criteria applied (given that the estimate was considered plausible):

- The intervention dominated other relevant strategies (that is, it was both less costly in terms of resource use and more clinically effective compared with all the other relevant alternative strategies), or
- The intervention costs less than £20,000 per quality-adjusted life-year (QALY) gained compared with the next best strategy.

Given that this is a cost analysis, no ICERs were derived in the base case, but exploratory QALY work (or exploratory CUA) was undertaken to explore whether the QALY gains that would be needed to make the more expensive intervention cost effective at the £20,000 level would be feasible. The feasibility of the gains were discussed with the (see discussion section).

1.2.2 Results

1.2.2.1 Overview

A summary of the results of all scenarios is provided below for comparison across the scenarios, as they are all alternative ways of looking at the same problem.

Table 19: Comparison of results across scenarios

	Scenario 1	Scenario 2	Scenario 3
URS total cost	£3,329	£3,252	£3,240
SWL total cost	£961	£865	£2,028
Base case incremental cost	£2,368	£2,387	£1,212
Exploratory QALY work	NA	QALY gains needed are high (implausible in base case). These do become more plausible if both time and effectiveness difference are varied.	Exploratory CUA showed high ICER. This was still above threshold when primary effectiveness varied.

Scenario 3 has the lowest incremental cost because it has the largest difference in effectiveness between URS and SWL of all the scenarios, and therefore the largest difference in downstream treatments, increasing the cost of SWL. Another reason for this smaller difference in total incremental cost in scenario 3, is because the procedures defined as ancillary procedures in the single study informing scenario 3, tended to be cheaper procedures for URS than what has been assumed in the base cases of scenarios 1 and 2.

1.2.2.2 Scenario 1

1.2.2.2.1 Base case

Table 20 shows the total costs per person of a strategy starting with URS versus a strategy starting with SWL. This is taking into account all the resource use that has been costed of retreatments, ancillary procedures, major and minor adverse events, and readmissions, as well as assumptions made such as about stent use.

Table 20: Scenario 1 - Total costs of strategies

Strategy	Total cost per person
URS	£3,329
SWL	£961
Incremental	**£2,368**

The cost of an SWL strategy was much cheaper overall, even though it involved more retreatments and ancillary procedures. Overall, it would cost an additional £2,368 to get a person stone free with the URS strategy than with the SWL strategy. This is because the types of procedures being used in the SWL strategy are lower cost than those in the URS strategy: retreatment is of a cheaper intervention, and ancillary procedures are also of a cheaper procedure (URS vs PCNL). Therefore despite the proportion receiving these being higher in the SWL arm, the total costs of these parameters end up being similar (see Table 21).

A breakdown of the different components can be seen in Table 21.

Follow-up costs are higher for the SWL arm because more people have further procedures and need to be followed up. Adverse events and readmissions are all higher for URS.

The main driver for the cost difference is really the initial intervention cost, this includes the stents as part of the intervention which applies proportionally more to the URS arm which is also contributing to the cost of the procedure because of the invasive nature of the intervention causing swelling, meaning more people need stents.

Although there is a higher retreatment probability in the SWL arm, the cost of retreatment in the URS arm is larger because a URS is around 5 times the cost of an SWL. So a much larger cost to a smaller group of people is outweighing a smaller cost applied to a larger group of people. Additionally, the cost of removing the stent applies to 70% of those who have a URS, which adds to the cost. In fact, if the stent costs in either strategy were not included, then the retreatment for URS would be slightly cheaper.

The ancillary procedure costs are higher for the SWL strategy because the ancillary probability is higher for SWL, and the ancillary for an SWL is around half the cost of an ancillary for a URS. Therefore there is a smaller cost difference between the types of ancillary procedures for each intervention compared to retreatments, meaning that the number of people affected is outweighing the difference in cost between the types of ancillaries. There is also the cost of stents to consider when a ureteroscopy is undertaken. Without any stent costs for URS, SWL ancillary procedures would be slightly cheaper.

Table 21: Scenario 1 - Total costs of strategies

Parameter	URS strategy cost (per 1000)	SWL strategy cost (per 1000)
Primary intervention cost	£2,884,276 (£712,270 = stent costs)	£471,973 (£20,351 = stent costs) (c)
Retreatment	£62,475 (£15,428 = stent costs)	£51,218 (£2,208 = stent costs) (c)
Ancillary procedures	£215,682 (b)	£284,991 (£70,378= stent costs)
Follow up (consultation + imaging) (a)	£115,063	£130,664
Readmission	£38,387	£19,194
Minor adverse events	£2,661	£1,783
Major adverse events	£10,352	£1,553
TOTAL	**£3,328,895**	**£961,376**

(a) Includes follow up after primary intervention, retreatment, and ancillary procedures. With 10% having an ultrasound.
(b) No stent costs assumed for PCNL
(c) For SWL stent cost is both for insertion and removal

1.2.2.2.2 SA1

Table 22 shows the results of varying the retreatment and ancillary procedure probabilities of SWL. A pattern can be seen that as the probabilities increase, the cost of the SWL strategy increases, which leads to a lower incremental cost. This was as expected, although the incremental cost is still fairly high, meaning there is still a large additional cost with URS of getting everyone stone free.

Table 22: Scenario 1 - SA1 results - varying retreatment/ancillary procedure rates

Probability of needing retreatment	Probability of needing ancillary procedure	Total cost of SWL strategy per pt	Incremental cost (URS - SWL)
10.9%	9.9%	£961	£2,368
12.9%	11.8%	£1,029	£2,300

Probability of needing retreatment	Probability of needing ancillary procedure	Total cost of SWL strategy per pt	Incremental cost (URS - SWL)
15.0%	13.6%	£1,097	£2,232
17.0%	15.5%	£1,165	£2,164
19.1%	17.4%	£1,233	£2,096
21.1%	19.2%	£1,301	£2,028
23.2%	21.1%	£1,369	£1,960
25.2%	23.0%	£1,437	£1,892
27.3%	24.9%	£1,505	£1,824
29.3%	26.7%	£1,573	£1,756
31.41%	29%	£1,641	£1,688

1.2.2.2.3 SA2

When two initial sessions of SWL were assumed as the intervention, then this led to a smaller incremental cost of £1,916.

Additionally, a threshold analysis was undertaken to find the cost of an SWL session that would make the interventions cost neutral. This was found to be £2,587, which would be nearly 6 times the base case cost of a session.

1.2.2.2.4 SA3

The three final columns of Table 23 below show the results when the types of secondary procedures are pooled and varied.

Note that the base case results here are not the same as those in the base case analysis, because the pooling of studies in the meta-analysis has led to slightly different probabilities of secondary events for SWL, than if the probabilities of the separate outcomes from the base case were to be summed.

Table 23: Scenario 1 - SA3 results – pooling retreatment and ancillary probabilities

	Interventions						Total cost		Incremental cost
	SWL			URS			URS	SWL	
Procedure	SWL (retreat)	URS (anc)	PCNL (anc)	URS (retreat)	SWL (anc)	PCNL (anc)			
1. Base case proportions	12.0%	11.0%	0.0%	2.2%	0.0%	4.2%	£3,329	£1,001	£2,328
2. Assuming different ancillary	12.0%	0.0%	11.0%	2.2%	4.2%	0.0%	£3,133	£1,254	£1,879
3. Assuming 50% split between different ancillaries	12.0%	5.5%	5.5%	2.2%	2.1%	2.1%	£3,231	£1,127	£2,104
4. Increasing proportion that have URS as secondary treatment in increments, until 100% of the secondary	9.6%	13.4%	0.0%	3.0%	0.0%	3.3%	£3,310	£1,059	£2,251
	7.2%	15.8%	0.0%	3.8%	0.0%	2.5%	£3,291	£1,117	£2,174
	4.8%	18.2%	0.0%	4.7%	0.0%	1.7%	£3,271	£1,175	£2,097
	2.4%	20.6%	0.0%	5.5%	0.0%	0.8%	£3,252	£1,233	£2,019
	0.0%	23.0%	0.0%	6.3%	0.0%	0.0%	£3,233	£1,291	£1,942

	Interventions						Total cost		Incremental cost
procedures are URS in both groups.									

Assuming a different ancillary procedure than the base case (number 2 in the table) – i.e. PCNL as the ancillary for SWL, rather than URS, leads to a lower incremental cost, because a PCNL is the most expensive type of procedure, and SWL assumed as the ancillary procedure for URS is the cheapest procedure.

As the proportion of secondary procedures being URS for either intervention increases, this also leads to a reduction in the incremental cost because the cost of SWL is increasing.

1.2.2.2.5 SA4

Varying the proportion of URS procedures that have stents, leads to the results in Table 24. As expected, the higher the proportion of stents inserted following a URS, the higher the incremental cost, this is because there are more URS procedures in the URS strategy, as everyone has it as a primary procedure and some also have retreatments. Therefore a higher stent proportion will drive up the URS strategy costs at a much faster rate than the SWL cost.

Table 24: Scenario 1 - SA4 results - varying stent use

Proportion of URS procedures that have a stent	Total cost URS	Total cost SWL	Incremental cost
70%	£3,329	£961	£2,368
0%	£2,601	£891	£1,710
20%	£2,809	£911	£1,898
40%	£3,017	£931	£2,086
60%	£3,225	£951	£2,274
80%	£3,433	£971	£2,461
100%	£3,641	£992	£2,649

1.2.2.2.6 SA5

Omitting the readmission outcome, because of the uncertainty around that as it was only from one study, and also because of potential double counting with major adverse events, led to total costs of £3,291 for URS, £942 for SWL, and an incremental cost of £2,348. This is not very different to that of the base case. Relatively speaking there was a much higher readmission cost for URS than for SWL, however readmission costs were only a small part of the total cost of the strategies, hence the change not being very impactful from omitting this outcome.

1.2.2.2.7 SA6

Pooling readmission and major adverse events again only led to small changes in the results: total costs of £3,293 for URS, £943 for SWL, and an incremental cost of £2,350.

1.2.2.2.8 SA7

Pooling minor and major adverse events, and applying the cost of major adverse events, as well as not including the re-admission outcome, led to total costs of £3,291 for URS, £944 for SWL, and an incremental cost of £2,347. This was also similar to the base case cost.

1.2.2.3 Scenario 2

1.2.2.3.1 Base case

Pooling the studies: Salem 2009, Sarica 2017 and Zhang 2011 in a scenario together, as these all report initial stone free rates and everyone is stone free at the end, provide the results in the table below. The incremental cost of getting someone stone free with a URS strategy versus an SWL strategy is similar to that of scenario 1.

Table 25: Scenario 2 - Total costs of strategies

Strategy	Total cost per person
URS	£3,252
SWL	£865
Incremental	**£2,387**

A breakdown of the different components can be seen in Table 26. The main differences with scenario 1 is that lower retreatment and ancillary probabilities from the studies pooled in this analysis have led to lower costs, which makes sense. The retreatment costs for the URS strategy would be higher than for URS even without the stent costs. The ancillary procedures cost for SWL would also be slightly higher than that of its URS counterpart even without stent costs.

Overall, using only a subset of the analyses of scenario 1 does not lead to major changes in the results because the main driver of the cost difference is still the initial procedure cost.

Table 26: Total costs of strategies

Parameter	URS strategy cost (per 1000)	SWL strategy cost (per 1000)
Primary intervention cost	£2,884,276 (£712,270 = stent costs)	£471,973 (£20,351 = stent costs) (c)
Retreatment	£40,970 (£10,117= stent costs)	£27,487 (£1,185= stent costs) (c)
Ancillary procedures	£162,347 (b)	£219,926 (£54,311= stent costs)
Follow up (consultation + imaging) (a)	£113,145	£122,781
Readmission	£38,387	£19,194
Minor adverse events	£2,661	£1,783
Major adverse events	£10,352	£1,553
TOTAL	**£3,252,138**	**£864,696**

(a) Includes follow up after primary intervention, retreatment, and ancillary procedures. With 10% having an ultrasound.
(b) No stent costs assumed for PCNL
(c) For SWL stent cost is both for insertion and removal

1.2.2.3.2 Sensitivity analyses

Threshold analysis on QALYs

Taking the incremental cost per person, and dividing this by the NICE threshold gives the QALY gain that would be needed to make URS cost effective:

Equation 1: QALY gain needed to make URS cost effective = £2,387/£20,000 threshold = **0.12**

Given the short time frame of the trials, a QALY gain of 0.12 seems high. As a QALY combines both length of life with quality of life, then for a short time frame this means the quality of life difference that would need to be gained from the URS over the SWL group would have to be quite large, because once that is divided by a small timeframe this will shrink the total QALY gain.

If we take the 2 week time frame we are assuming all studies have, see section 1.2.1.5.2, (as a fraction of a year this is = 2/52 = 0.038) and divide the QALY gain from equation 1 by this timeframe; this will give us the additional quality of life gain that URS needs to provide (over and above that of SWL) to make it cost effective:

Equation 2: QoL gain needed for URS = 0.12/0.038 = **3.1**

This tells us that on average, someone who has a URS instead of an SWL, will need to have a quality of life gain of 3.1 over a 2 week period, to make URS a cost effective choice. Not everyone will experience this gain as this is the average gain everyone in the cohort will have to get. Also, this isn't necessarily going to apply consistently throughout the 2 week period because people will tend to have episodes of pain and some of the time be pain free. So there are lots of reasons as to why we cannot be confident about where the quality of life gain will come from, or whether this will be achievable.

This quality of life gain is explored further in the next section.

Exploratory QALY calculations

As discussed in section 1.2.1.5.2, one source of the quality of life gain is likely to be because of a difference in quality of life between people who still have a stone and people who no longer have a stone.

The larger the difference in (initial) effectiveness between the two interventions, the more additional people that would be successfully treated with URS who will then accrue this QoL gain from not having a stone anymore.

We know from the studies we have pooled for this scenario, what the initial effectiveness of the interventions is: URS initial success probability = 93.2%, SWL initial success probability = 82%. This gives an incremental success probability of URS over SWL of 11.2%.

It is this 11.2% who are stone free with URS but would not be stone free with SWL, that we think this quality of life difference is going to come from. Moreover, as mentioned in section 1.2.1.5.2, we are assuming that this difference in stone free people between the initial interventions remains for the whole time period of the trials – effectively this means that people have their other procedures at the end of the trials.

From our previous calculations we now know two things:

- The QoL gain needed over 2 weeks for URS to be cost effective (from equation 2),
- And that this average will most likely come from the difference in initial effectiveness (of 11.2%).

Therefore by apportioning this average QoL difference only to the people we think this difference will come from, we can work out the QoL difference needed between a stone free person and a non-stone free person to make URS cost effective.

We can also demonstrate this through the following equations, where the **quality of life difference between a stone free health state and non-stone free health state needed to make URS cost effective** can be represented by X:

11.2% * X = 3.1

88.8% * 0 = 0

3.1

Working out X = 3.1/11.2% = **27.8**

Quality of life is usually on a zero to one scale, but there are also states considered worse than death on the EQ-5D, as a score of 3 on each domain will lead to a utility value of -0.594. The best possible health state of full health has a utility value of 1. So even if we take the quality of life with renal stones to be the worst possible state on the EQ-5D, and the quality of life without a stone to be the best possible state: this difference is much smaller than the difference in QoL we worked out that is needed of 27.8. Therefore 27.8 is physically impossible to achieve.

As another example to demonstrate if there are any feasible QALY gains from using URS:

- If we take QoL with no stone to be 0.874 (see section 1.2.1.4)
- If we take QoL with a stone to be 0.684 (see section 1.2.1.4)

This creates a utility gain of 0.19, which is still much smaller than the QoL difference calculated from the worst level of SWL effectiveness we have assumed in sensitivity analysis 8 (see section 0). And also bearing in mind again that this difference won't apply for the whole time period we are assuming of 2 weeks, people might not be waiting for further treatments if unsuccessful for 2 weeks and they would not all be in pain that whole time.

If we take this example further and choose some extreme incremental effectiveness values between URS and SWL of say 40%, and multiply this by the utility gain of 0.19, and the timeframe of 2 weeks, this provides a QALY gain of 0.0029. Which is very small compared to those we have calculated is needed of 0.12.

Another way to use this information is to calculate the incremental cost that would be needed to make URS cost effective at the £20,000 threshold, if we used the above QALY gain. This requires another simple rearrangement of the ICER equation:

Figure 4: equation rearranged to find incremental QALY gain needed

$$\frac{ICER}{QALYs\ (URS) - QALYS\ (SWL)} = Costs(URS) - costs\ (SWL)$$

Where: Costs(URS) = total costs for URS; QALYs(URS) = total QALYs for URS

£20,000/0.0029 = £58.46.

The incremental cost needed to make the QALY gain calculated achievable is substantially smaller than the incremental cost found in the base case analysis. This also demonstrates that the difference in costs between the two strategies is large enough that it is highly unlikely URS would be cost effective.

Other sensitivity analyses

SA8

The effectiveness of the primary SWL treatment in the pathway was varied in order to undertake the exploratory QALY work for each effectiveness value being tested, as this difference in the initial effectiveness of the primary treatments was assumed to be where the QALY gains would come from.

Table 27 shows the results. The effectiveness of the initial SWL was varied down to 40%. The lower the effectiveness of SWL then the higher the cost of the SWL strategy because more downstream resources are being consumed, and therefore the incremental cost decreased, but not enough to make the quality of life difference between a stone free and non stone free person ever reach a level that would be considered feasible. In other words, even at the lowest tested SWL effectiveness level of 40%, the quality of life gain needed was still higher than the difference between the minimum and maximum values on the EQ-5D (1.594). Hence why all the values are shaded in pink.

Any achievable differences in QALYs would be very small, particularly because of the timeframe of the studies being so short. If in reality people were waiting for a few months for another treatment if their first treatment had failed, then those 11.2% of people from the base case that failed with an SWL but would not have failed with a URS, are waiting for a longer period of time, and still have a stone.

To explore further this relationship between the effectiveness of initial SWL and the time spent waiting for further treatment, if failed - and how those factors impact the QoL difference needed between the two health states: a two-way sensitivity analysis was undertaken. The results of this are shown in Table 28.

Along the rows the difference in initial effectiveness between the interventions (and that has a corresponding cost impact) is being varied, and in the columns the timeframe is being varied. In the body of the matrix is the difference in QoL needed between a stone free health state and non-stone free health state to make URS cost effective. The red areas show where the values are not physically possible because they are larger than the maximum range possible on the EQ-5D (the maximum range being 1.594).

This analysis tells us that as SWL reduces in effectiveness, and as people have to wait longer for further treatments, then the QoL difference needed between someone with and without a stone is shrinking. And that is because more people are failing SWL, and they are also having to wait longer, so there are more people with reduced quality of life, and for longer periods, therefore the gains needed with URS are starting to look more possible.

However, it still remains that even with the worst effectiveness tested, and waiting for the longest time tested here, a QoL difference is needed of 0.38 of someone living with a stone compared to someone without. It’s possible this may be feasible, but is still higher than the difference calculated earlier using QoL sources (0.19), and again there are caveats as the values calculated here are averages spread across the whole population in question.

We are also talking about the same period of time of waiting for further treatments regardless of the treatments, when in practice waiting times would be variable for the different procedures. For example you may wait less time for an SWL than you would for surgery, which would reduce the QoL gain possible from surgery as more people would become stone free over time, therefore reducing the QoL gain possible for URS.

Hence although the analysis is reflecting what might be considered more realistic effectiveness and waiting time, which has led to more possible values, that does not mean that those are feasible values, because there are still a lot of unknowns.

Table 27: Scenario 2 – SA8 results – varying initial effectiveness of SWL

				RESULTS		EXPLORATORY QALY CALCULATIONS			
	Initial effectiveness	probability of needing retreatment	probability of needing ancillary procedure	Total cost of SWL strategy per pt	Incremental cost (URS - SWL)	QALY gain needed	QoL gain needed (assuming 2 wk time horizon)	Effectiveness difference with URS (I.e. proportion that QoL gain applies to)	Difference in QoL needed between a stone free and non stone free person (a)
Base case value	82.0%	5.8%	7.6%	£865	£2,387	0.12	3.10	11.2%	27.76
	77.8%	7.6%	10.0%	£947	£2,306	0.12	3.00	15.4%	19.49
	73.6%	9.5%	12.4%	£1,028	£2,224	0.11	2.89	19.6%	14.76
	69.4%	11.3%	14.8%	£1,110	£2,142	0.11	2.78	23.8%	11.71
	65.2%	13.1%	17.1%	£1,192	£2,060	0.10	2.68	28.0%	9.57
	61.0%	14.9%	19.5%	£1,274	£1,978	0.10	2.57	32.2%	7.99
	56.8%	16.7%	21.9%	£1,356	£1,897	0.09	2.47	36.4%	6.78
	52.6%	18.6%	24.3%	£1,437	£1,815	0.09	2.36	40.6%	5.81
	48.4%	20.4%	26.7%	£1,519	£1,733	0.09	2.25	44.8%	5.03
	44.2%	22.2%	29.1%	£1,601	£1,651	0.08	2.15	49.0%	4.38
Suggested UK practice values	40.0%	24.0%	31%	£1,683	£1,569	0.08	2.04	53.2%	3.84

a) Note the maximum range of the EQ-5D is 1.594. Cells that are red that the QoL difference needed is more than the maximumEQ-5D difference.

Table 28: 2 way sensitivity analysis varying time to further treatment and initial SWL effectiveness

		Time to retreatments					
Cost difference	**Difference in effectiveness between primary URS and SWL that corresponds to the cost difference**	***2 weeks***	***4 weeks***	***8 weeks***	***12 weeks***	***16 weeks***	***20 weeks***
£2,387	*11.2%*	27.76	13.88	6.94	4.63	3.47	2.78
£2,306	*15.4%*	19.49	9.74	4.87	3.25	2.44	1.95
£2,224	*19.6%*	14.76	7.38	3.69	2.46	1.85	1.48

		Time to retreatments					
Cost difference	**Difference in effectiveness between primary URS and SWL that corresponds to the cost difference**	***2 weeks***	***4 weeks***	***8 weeks***	***12 weeks***	***16 weeks***	***20 weeks***
£2,142	*23.8%*	11.71	5.85	2.93	1.95	1.46	1.17
£2,060	*28.0%*	9.57	4.79	2.39	1.60	1.20	0.96
£1,978	*32.2%*	7.99	4.00	2.00	1.33	1.00	0.80
£1,897	*36.4%*	6.78	3.39	1.69	1.13	0.85	0.68
£1,815	*40.6%*	5.81	2.91	1.45	0.97	0.73	0.58
£1,733	*44.8%*	5.03	2.52	1.26	0.84	0.63	0.50
£1,651	*49.0%*	4.38	2.19	1.10	0.73	0.55	0.44
£1,569	*53.2%*	3.84	1.92	0.96	0.64	0.48	0.38

(a) In the body of the matrix rows (columns3 to 8) is the difference in QoL needed between a stone free health state and non-stone free health state to make URS cost effective. The red areas show where the values are not physically possible because they are larger than the maximum range possible on the EQ-5D (the maximum range being 1.594).

SA9

The number of SWL sessions as part of the primary treatment was varied from 1 to 2, and this led to an incremental cost of £1,936.

Additionally a threshold analysis is undertaken to see what the cost of an SWL session would have to be to make the comparators cost neutral. This was found to be £2,708 per session. This is more than the cost of URS we included in the costing. One might also assume that even if more investment in SWL machinery meant that the NHS reference cost for SWL was driven up – because maybe some areas would not see as many patients as others so the average cost per person might increase – it is uncertain whether this would go up by over 5 times to approach the cost of URS. Once incorporating the cost of removing a stent, this also drives up the cost of an average URS.

SA10

The three final columns of the table below show the results when the types of secondary procedures are pooled and varied.

Note that the base case results here are not the same as those in the base case analysis, because the pooling of studies in the meta-analysis has led to slightly different probabilities of secondary events for SWL than if the probabilities of the separate outcomes from the base case were to be summed.

Table 29: Scenario 2 – SA10 results - pooling retreatment and ancillary probabilities

	Interventions						Total cost		Incremental cost
	SWL			URS			URS	SWL	
Procedure	SWL (retreat)	URS (anc)	PCNL (anc)	URS (retreat)	SWL (anc)	PCNL (anc)			
1. Base case proportions	5.4%	7.0%	-	1.4%	-	3.1%	£3,252	£844	£2,409
2. Assuming different ancillary	5.4%	-	7.0%	1.4%	3.1%	-	£3,105	£1,006	£2,099
3. Assuming 50% split between different ancillaries	5.4%	3.5%	3.5%	1.4%	1.6%	1.6%	£3,178	£925	£2,254
4. Increasing proportion that have URS as secondary treatment in increments, until 100% of the secondary procedures are URS in both groups.	5.8%	6.6%	-	2.3%	-	2.2%	£3,235	£848	£2,387
	4.4%	8.0%	-	2.9%	-	1.7%	£3,221	£879	£2,342
	2.9%	9.5%	-	3.4%	-	1.1%	£3,208	£910	£2,297
	1.5%	10.9%	-	4.0%	-	0.6%	£3,194	£941	£2,252
	0.0%	12.4%	-	4.5%	-	0.0%	£3,180	£973	£2,207

Assuming a different ancillary procedure than the base case (number 2 in the table) – i.e. PCNL as the ancillary for SWL rather than URS, leads to a lower incremental cost, because a PCNL is the most expensive type of procedure, and SWL assumed as the ancillary procedure for URS is the cheapest procedure. This leads to a higher total cost for the SWL strategy because there are more ancillaries in that strategy, which reduces the incremental cost.

As the proportion of secondary procedures being URS for either intervention increases, this also leads to a reduction in the incremental cost, as the SWL total cost has increased by more than the URS total cost has decreased, creating a smaller gap in costs between the two strategies.

SA11

Varying the proportion of URS procedures that have stents (whether this is a URS that is a retreatment or a URS that is an ancillary procedure in the SWL strategy), leads to the results in Table 30. As expected, the higher the proportion of stents inserted following a URS, the higher the incremental cost, this is because there are more URS procedures in the URS strategy as everyone has it as a primary procedure, and so the total cost is being driven up by more in the URS strategy, creating a larger incremental cost difference.

Table 30: Scenario 2 – SA11 results - varying stent use

Proportion of URS procedures that have a stent	Total cost URS	Total cost SWL	Incremental cost
70%	£3,252	£865	£2,387
0%	£2,530	£810	£1,719
20%	£2,736	£826	£1,910
40%	£2,943	£841	£2,101
60%	£3,149	£857	£2,292
80%	£3,355	£872	£2,483
100%	£3,562	£888	£2,674

1.2.2.4 Scenario 3

1.2.2.4.1 Base case

In this scenario analysis, only the Hendrikx study was used. This was because this study was slightly different to the others in a number of ways, firstly, it reported both second and third line treatments, and therefore where these further lines of treatment were pooled together for scenario 1, this is technically not appropriate as if a person can have multiple procedures then these values should be reported as rates, but was undertaken for simplicity. Secondly, this breakdown of lines of treatment meant more information was available about the success of treatments at individual parts of the pathway, which required a different structure in the analysis. Additionally this also represented an effectiveness of SWL closer to UK practice anecdotally reported by the committee.

Table 31: Scenario 3 - Total costs of strategies

Strategy	Total cost per person
URS	£3,240
SWL	£2,028
Incremental	**£1,212**

The table with the breakdown of the results can be found below.

Table 32: Scenario 3 base case results (per 1000)

	First line	Second line	Third line	First line	Second line	Third line
	URS			SWL		
Primary intervention	£2,884,276			£471,973		
Retreatment		£0	£33,153		£54,722	
Ancillary procedures		£151,978			£1,061,897	£250,807
Follow up (consult + imaging)	£108,226	£9,952	£1,244	£108,226	£48,623	£9,411
Readmission to hospital	£38,387			£19,194		
Minor adverse events	£2,661			£1,783		
Major adverse events	£10,352			£1,553		
TOTAL	**£3,240,227**			**£2,028,187**		

The results for the two comparisons are closer together than the other scenarios.

This is being driven by a big difference in the ancillary procedure probabilities: there are many more ancillary procedures for SWL, which reflects that the success probability of the initial procedures is further apart than in the other scenarios.

The result is also being driven by the types of ancillary procedures in each strategy: for SWL, some of these (13%) are PCNL which is the most expensive procedure. On the contrary, 75% of the ancillary procedures for URS are SWL, which is the cheapest procedure, and around 10 times cheaper than a PCNL.

What this also means is that URS is possibly closer to being cost effective based on the results of this analysis because a smaller incremental cost means fewer incremental QALYs are needed to derive an ICER below the NICE threshold.

The possible cost effectiveness is explored further in the next section.

1.2.2.4.2 Exploratory QALY work

As mentioned in section 1.2.1.5.3, if we make assumptions about when second and third line treatments are occurring, and if we also knew what the utilities of people with and without a stone, then this would allow a cost utility analysis to be undertaken.

The quality of life of someone with and without a stone were those reported in section 1.2.1.4.

This led to the results in Table 33 below.

Table 33: Scenario 3 - exploratory cost utility analysis results

	Total cost per pt	Total QALY per pt	ICER
URS	£3,240	0.217	
SWL	£2,028	0.209	
Incremental	£1,212	0.008	
			£155,049

The ICER calculated based on the exploratory cost utility analysis is significantly higher than the NICE threshold of £20,000.

It is also important to note that the quality of life value used to represent someone with a stone (from the Pickard study), was felt to be higher than the quality of life of living with a stone in reality, as GC opinion was that the Pickard study (which is a well known UK study) only recruited patients who were not in pain. However as discussed below QoL is explored in more detail through threshold analysis.

Using the goal seek function in excel, it was possible to calculate what the decrement in quality of life from having a stone would need to be to make URS cost effective. This was found to be 1.47. Subtracting this from the quality of life used for someone without a stone means that the quality of life of living with a stone would need to be -0.596.

The lowest possible quality of life score on the EQ-5D is -0.594 (it is negative because there are states on the EQ-5D considered worse than death), therefore -0.596 is just outside of the feasible range. It is possible that someone may consider their renal stones pain to be worse than death, as the opinion of the committee is that it is one of the worst imaginable pains. However, it is unlikely that someone would be in this kind of pain the entire time that they have a stone and are waiting for treatment, as pain tends to be episodic. Hence, even if the utility calculated was within the feasible range on the EQ-5D, this utility would not apply for the whole time period that it was applied in the analysis, and therefore it is likely that the QALY gain would be lower than demonstrated. Additionally, if someone was in unimaginable pain then it is likely they would be treated urgently, and therefore perhaps are not the population of this model.

1.2.2.4.3 Sensitivity analysis

SA12

If the fourth line interventions were costed for those that were not stone free at the end of the trial, the incremental cost is £1,030 per person.

This is slightly lower than that of the base case because there are more non stone free people at the end of the trial in the SWL arm, so including the treatment of those people would increase the total SWL cost by more than the increase in the total URS cost, thereby making the incremental cost smaller.

SA13

The effectiveness of the primary SWL was varied as this was felt to be very variable in practice and dependent on a number of factors such as the experience and training of the operator, and also anecdotally if it is a mobile or fixed site lithotripter for example.

The effectiveness was varied in equal increments from the same effectiveness as a URS treatment, down to 40% effectiveness. With the midpoint being the base case SWL effectiveness.

As expected, the lower the probability of success, the higher the total cost of the SWL strategy, because further procedures are being needed.

The results from the exploratory cost utility analysis are also recorded below for each level of effectiveness. As the effectiveness of SWL decreases, the total QALYs from the SWL strategy also decrease because there are fewer people clearing their stone from the first go, which means there are fewer people accruing the quality of life associated with being stone free for the whole period of the trial. A larger incremental QALY and smaller incremental cost from lower SWL effectiveness therefore reduce the ICER, but still not to a level that would be considered cost effective.

However, as previously discussed from the base case results above, the quality of life calculated is likely to be an overestimate because it may not apply for the whole time period calculated in the exploratory cost utility analysis. Other factors however that have not been considered include what else might be affecting the QALYs, for example – whether there is any quality of life impact from the nature of the interventions themselves, and the use of stents in the URS arm which has a quality of life reduction from being uncomfortable in nature. These factors are discussed in detail in section 1.2.3.2

Table 34: Scenario 3 - SA13 results – varying effectiveness of primary SWL

		RESULTS		EXPLORATORY QALY CALCULATIONS		
	Effectiveness of primary SWL	**Total cost of SWL strategy per pt**	**Incremental cost (URS - SWL)**	**Total QALY SWL strategy per pt**	**Incremental QALY (URS - SWL)**	**ICER**
Same as URS	0.91	£894	£2,346	0.217	0.000	£15,369,168
	0.84	£1,121	£2,119	0.215	0.002	£1,257,174
	0.77	£1,348	£1,892	0.213	0.003	£587,945
	0.69	£1,575	£1,666	0.212	0.005	£350,538
	0.62	£1,801	£1,439	0.210	0.006	£228,951
Base case	0.55	£2,028	£1,212	0.209	0.008	£155,049
	0.52	£2,124	£1,116	0.208	0.008	£131,903
	0.49	£2,219	£1,021	0.208	0.009	£112,043
	0.46	£2,315	£925	0.207	0.010	£94,815
	0.43	£2,411	£829	0.206	0.010	£79,729
Low effectiveness	0.4	£2,506	£734	0.206	0.011	£66,408

SA14

Omitting readmission from the analysis led to a slight reduction in the incremental cost, to £1,183.

SA15

Pooling readmission and major adverse events in case there was any double counting led to a similar incremental cost to the bae case of £1,194.

SA16

Pooling minor and major adverse events (and also excluding readmissions) again led to a similar incremental cost to the base case of £1,192.

The costs of the adverse events and readmissions are a small component of the total overall cost which is why they are not having very much of an impact.

SA17

When the stent use is zero, this reduces the URS strategy cost by a larger amount than the change in the SWL strategy cost, as more stents are being used in the URS strategy. This leads to a lower incremental cost. When the stent proportion is high, this increases the incremental cost because it makes the URS strategy more costly.

Table 35: Scenario 3 – SA17 results - varying stent use

Proportion of URS procedures that have a stent	Total cost URS	Total cost SWL	Incremental cost
70%	£3,240	£2,028	£1,212
0%	£2,520	£1,760	£760
20%	£2,726	£1,836	£889
40%	£2,931	£1,913	£1,018
60%	£3,137	£1,990	£1,147
80%	£3,343	£2,067	£1,277
100%	£3,549	£2,143	£1,406

SA18

Assuming two initial sessions of SWL led to an incremental cost of £760. This is lower than the base case because this is essentially saying that the primary SWL procedure is twice as expensive, which leads to a smaller gap between the two strategies. Dividing this by the incremental QALY derived from the exploratory cost utility analysis creates an ICER beyond £20,000 (around £97,000).

The threshold analysis on what the cost of an SWL session would need to be to make the strategies cost neutral is £1,609. This is more than 3 times the current cost of a session.

1.2.3 Discussion

1.2.3.1 Summary of results

Scenario 1, which was a costing analysis using resource use data from all the studies included in the clinical review, showed a cost difference of over £2,000. This is considered a reasonable estimate of the overall average cost difference to get someone stone free with URS versus getting someone stone free SWL as this was the case in the majority of studies (4 out of 7). In the other 3 studies, one stone free rate at the end of the study was high and very similar (95% URS, 96% SWL), in the other two URS was favoured (99% URS, 93% SWL, and 88% URS, 85% SWL). Given this, it might be the case that the estimate is slightly biased against SWL but it is considered unlikely that this would have a substantial impact on the cost difference estimate. Various sensitivity analysis were undertaken showing that the magnitude of cost difference between the strategies was sensitive to the probabilities associated with further treatments, and the types of procedures these are, because of the range in costs of different procedures. The incremental cost was also affected by the resource use assumptions such as the proportion of patients that have a stent following a URS procedure. In no sensitivity analysis did URS ever become cheaper than SWL. The analysis however was not very sensitive to the adverse event or readmission outcomes, as these are only a small proportion of the overall costs.

Scenario 2 used only studies that reported the initial stone free probability, and where everyone was stone free at the end (3 studies). This was to remove the potential bias in the resource use estimates in scenario 1 where in some studies not everyone was stone-free at the end of follow up, and so that exploratory QALY calculations could be undertaken using the initial stone free rate. The studies were also all of a similar timeframe which was helpful for QALY threshold analysis and exploratory QALY work. Although this scenario has the disadvantage of not using all the clinical studies available. This scenario found a similar overall incremental cost to that of scenario 1. The QALY threshold analysis and exploratory QALY work found that the QALY difference would have to be very high for URS to be cost effective. Using the initial effectiveness difference as the population that would derive the quality of life benefit from URS over SWL showed that this QoL difference was beyond feasible levels (based on the difference between the maximum and minimum on the EQ-5D). This is however dependent on the time period in question. A 2-way sensitivity analysis varying the effectiveness of initial SWL and time to further treatments after primary treatment failed, showed that there were some plausible values arising. Although it still remains to be seen whether these would be feasible values. Sensitivity analysis also showed that the magnitude of the cost difference was affected by the probabilities of further treatments, and what these treatments were, as well as to resource use assumptions.

Scenario 3 used only one study to inform the clinical parameters of the analysis. This offers a different perspective on the treatment pathway as the study explicitly includes the possibility of 3 lines of treatment, and the effectiveness of SWL is much lower than in scenario 2 so this could be seen as more reflective of UK practice. This showed a cost difference of around £1,200, which was closer together than the other scenarios because there is a bigger difference in the ancillary procedure probabilities between the interventions in this scenario, and the result is also being driven by the types of ancillary procedures in each strategy. An exploratory cost utility analysis was undertaken based on assumptions about when the different lines of treatment were occurring, and a general UK population utility to represent the utility of people who were stone free. This was used as a way to derive what the utility of someone who still had a stone would need to be to make URS cost effective. This showed that the utility of a non-stone free person would need to be slightly beyond the worst possible state of health on the EQ-5D, making it unfeasible that URS would be cost effective (see the discussion section for more information). Varying the primary treatment effectiveness also showed that the incremental cost and therefore cost effectiveness was sensitive to this because the ICERs decreased, but still not to a level close to the £20,000 threshold. Other parameters that varied the magnitude of the cost difference were the proportion of people that have stents after URS.

1.2.3.2 Limitations and interpretation

The clinical review for this population stratum included seven clinical studies. On closer inspection of the studies for the purposes of the economic analysis, differences were identified that led to multiple scenarios being undertaken, to ensure that any potential biases were tested as much as possible. Differences between the studies included: their length of follow up; the point in the pathway that outcomes were being reported (i.e. after the first procedure or after further procedures), the clarity with which they reported whether the outcomes were in terms of numbers of procedures or people.

The main difference that led to different scenarios being undertaken, was the fact that the studies included in the stone free outcome in the clinical review (all seven studies) were a mix of outcomes time points: some being after the initial procedure, and some being at the end after further procedures, and others being unclear. This would affect the stone free outcome as the more procedures people have, then the more likely people will be stone free. So pooling these studies for the stone free outcome leads to a lack of clarity because the outcome is not telling us the stone free probability after the initial intervention, which is what we are interested in if we want to know how effective the different interventions being compared are.

Scenario 1 approached a costing analysis with the assumption that if all the retreatments and ancillary procedures reported in the studies were costed up, this would lead to 100% effectiveness in both strategies. This is implying that the intervention that had more downstream resource use must be initially less effective. In the studies informing this scenario, not everyone is stone free at the end, which means a limitation from this assumption is that there may be some costs that haven't been captured. This is more likely to be a bias favouring the SWL strategy, which is initially less effective, and therefore likely to have more people that are not stone free at the end. It may be plausible that people pass their stone sometime after a procedure, without needing further intervention. However, for a ureteric stone, it is very unlikely that a stone which didn't pass would be left without further intervention, as a ureteric stone can cause an obstruction. This potential underestimation of some costs is part of the reason why scenario 2 is undertaken as a way to identify the incremental cost of the strategies using only studies where everyone is stone free at the end. Additionally, sensitivity analyses were undertaken varying inputs in ways to make the analysis more conservative to SWL, such as varying what the secondary procedures were, and lowering the proportion that get stents following a URS.

As touched on earlier, the studies in scenario 1 are also different in terms of whether they are reporting the number of procedures that occur or the number of people. One study in particular (Hendrikx 1999) reports how many people had two or three lines of treatment, which means that some people had more than one further procedure if they had both a second line treatment and third line treatment. This study is therefore reporting the number of procedures that had taken place. In the other studies, some are reporting the number of people (which can be identified because the number initially stone free, in addition to those who have a retreatment or ancillary, sum to the number of participants in the study), and some are unclear. Pooling all these studies together to derive the probability of a retreatment or an ancillary procedure therefore creates some methodological difficulties. If an individual can have more than one procedure, then rates should be used, as a probability cannot be more than one. In the studies where the number of people having procedures is being reported, or if it is unclear, then people could be assumed to be having one procedure each, and rates would still be applicable. However, it is assumed that all the resource use that is going to happen, is happening within the time frame of the trials, and trials have different follow up periods. Additionally, everyone is stone free at the end of the trials with shorter follow up. Therefore, extrapolating resource use beyond the timeframe of the shorter trials conflicts with the assumption behind rates - that they are constant. As the rates of resource use would vary as people become stone free, and therefore the properties of rates do not suit our particular problem. What is of interest from a cost perspective is the number of procedures taking place, as this is what we want to apply costs to. Hence, it does not necessarily matter if it is procedures or people being reported, as for the purposes of the analysis we are trying to find out the level of resource use. The approach taken is to use the data as probabilities and pool all the available data. Applying this method in order to include the Hendrikx 1999 data within the analysis means that there is some potential overestimation of the probabilities, because the Hendrikx study is driving them up as people can have multiple procedures. This is more so in the SWL arm, however if anything, we want to be more conservative to SWL, so overestimating resource use is better than underestimating it, as the effectiveness of SWL is believed to be lower in practice which would lead to a higher probability of further treatments in reality. Scenario 3 uses only the Hendrikx study which allows the probability associated with each part of the pathway to be used and gets around the rate problem if the study is analysed on its own.

Scenario 2 has the limitation that it is informed by fewer studies than scenario 1. The three studies included for the retreatment, ancillary, and stone free outcomes (used only for QALY work) were all of short time periods. As everyone was stone free at the end, this may be because treatment was more intensive and happened more quickly in these studies. It is perhaps odd that the shorter duration studies from the seven included in the clinical review are the ones where people are stone free, rather than in the longer studies. Pooling from fewer studies means there is potentially more uncertainty around the estimates.

Part of the rationale behind scenario 2 was including studies where the stone free probability was the initial stone free probability in order to allow some exploratory QALY work to be undertaken. An assumption was made that the QoL gain between the two strategies would come from the difference in initial effectiveness of the interventions i.e. those people who were stone free with a primary URS, who would not be stone free with a primary SWL. This assumption was based on the logic that different effectiveness would lead to different numbers of people being stone free, and someone with a stone would have a different quality of life to someone without a stone. In addition to this, further treatments would lead to more people being stone free over time. If the further treatments were assumed to happen at the end of the trials, then this difference in initial effectiveness is likely to apply for the whole time horizon of the trials, and therefore would be the main source of QALY gain between the two strategies. Given that a cost utility analysis would require yet more assumptions, it was felt that using the method described to try and infer cost effectiveness would be a useful exercise to demonstrate the potential likelihood of URS being cost effective.

There are however may caveats to the exploratory QALY approach that need to be taken into consideration: it is uncertain how the QALYs over time are going to look in terms of what the quality of life differences actually are, how long they apply for and the frequency of peoples pain episodes, and when further treatments are happening. So we can only estimate whether URS is likely to be cost effective. It is also important to remember that we are referring to the general ureteric <10mm stone population here, which will be a mix of people with different levels and frequency of symptoms. Which is why even if QoL differences between a stone free and non-stone free person are physically possible, this does not mean these are feasible values, when considering the average population in question.

Part of the reason that such a large quality of life difference was derived through the exploratory work is because the timeframe being used for the studies is so short. Multiplying the QoL gains by a very short time frame creates a very small QALY, as the units are quality adjusted life *years*. And thus the Qol gain has to be very large in order to derive the QALY gain needed to make URS cost effective. If say it would not be true that all the further treatments for those initially unsuccessful would happen at the end of the trial, then the two week timeframe used could be varied. However bringing that forward, say to 1 week instead of 2, would only increase the QoL gain needed even more. However, given that the time period is so small then it is likely that undertaking these retreatments in the time period of the trials is actually not feasible in the NHS, and therefore the timeframe was varied in a sensitivity analysis to see if this would lead to any plausible QoL values. Sensitivity analysis 8 demonstrated that when varying effectiveness levels of SWL and time to further treatments following primary treatment, then although QoL differences between a stone free and non-stone free person could be within plausible ranges, these may still not be feasible.

The likely waiting times in the NHS for retreatment of either procedure is important here and has a bearing on what quality of life differences are likely to be feasible. The time frame that has been used in the exploratory QALY work for scenarios 2 and 3 is the time between having failed a retreatment and having further treatment, and this is the same regardless of strategy. Note that this is not the time to the primary treatment (which would also be a factor in practice that would be considered when a clinician is considering treatment options). Waiting times are variable within the NHS for both SWL and URS. This is dependent on many local factors such as availability of equipment and staff. For SWL specifically, whether a fixed site lithotripter or mobile one is available can lead to differences in waiting times. URS waiting times are also variable because of staffing and theatre list arrangements. Anecdotally, having a fixed site lithotripter means SWL could be undertaken in a shorter space of time than waiting for a mobile machine which tends to come to each hospital once a month. If SWL has a shorter waiting time than URS for example, then multiple retreatments might be undertaken within the same timeframe of waiting for surgery, which would close the gap in effectiveness between the two interventions. Additionally, further treatment after a failed treatment would be seen as less of a priority in the NHS than primary treatment, in which case waiting time could be many weeks. The longer the waiting time, the more time that people are living with a stone having failed the less effective treatment, and the more QALYs the initially effective treatment would accrue.

The feasibility of quality of life gains demonstrated to the committee from the analyses were discussed at length, however it is difficult to come up with a single estimate of what feasible time to retreatment might be given the variability across the country, or the quality of life of living with a stone.

The logic of who the QALY gains would come from was based on those people who are stone free with a URS but who would not have been with an SWL. However there may also be differences in QoL between the two interventions that haven't been considered. For example, because of the nature of the interventions themselves. Perhaps URS has a higher initial decrement in QoL because it is invasive and involves general anaesthetic, but outweighing this might be the fact that there could be a shorter recovery time as it gets rid of the stone in one go. Alternatively, SWL may have a higher decrement in QoL because people remember the SWL treatment as being worse as they were not asleep for it, and remember the uncomfortable nature of the shockwave treatment. However it is also more convenient for patients as they can arrange a time around their daily routine for the sessions. Although these effects are only likely to last for a short amount of time for the recovery period and so may be negligible, they would apply to the whole populations in both strategies because everyone has the primary treatment (which is different to the QoL decrement from still having a stone which obviously only applies to those that still have a stone).

Another issue is that people are more likely to have stents inserted after a URS, and stents are uncomfortable and therefore have a quality of life impact (with side effects like frequent need to urinate for example). This can interfere with people's daily activities e.g. people that have stents who work often take time off, and therefore a URS may have a QoL impact for a longer period of time than purely the recovery from the procedure itself. This means that to have an achievable QALY gain for URS, the effectiveness difference between SWL and URS needs to be larger, in order for the QoL gain from the additional stone free individuals to counteract the QoL loss from stents. A recommendation has been made to discourage the use of stents after surgery as there was no evidence of benefit, therefore as the recommendation is implemented then there would be fewer people experiencing the QoL impact of stents.

Other factors influencing quality of life that haven't been considered include the impact of an untreated ureteric stone, who this population actually are and whether they would be in pain as mentioned above, and what kind of pain it is i.e. severe and episodic or constant. These are all factors that could influence the quality of life and how it might be captured. It could be argued that if a patient is not in pain i.e. an asymptomatic stone, then there is no benefit to treatment. The population in question however is likely to be people who are having planned treatment, and therefore those that are considered emergency cases because they are obstructed/infected would be outside the general ureteric population being discussed here. The risk of obstruction is difficult to quantify as generally these are people that are excluded from trials. The goal from a clinical perspective is to treat a ureteric stone a soon as possible because if obstruction was to develop then animal studies have shown that a kidney can die within 6 weeks of obstruction if not treated. The risk of obstruction is not something that could be included in the analysis as it could not be quantified, but this was a concern the committee raised with regards to the less effective intervention of SWL. Someone with a ureteric stone could be a pain, but again as pain is episodic and variable then it is still uncertain what the quality of life on average of someone with a stone would be.

In essence, the above are just examples, but there may be factors on the health outcomes side that have not been captured, and therefore the exploratory QALY work needs to be interpreted with caution. The results however show that the gains being calculated as needed are beyond feasible levels which provide some reassurance that URS is unlikely to be cost effective.

Scenario 3 used only the Hendrikx 1999 study to inform the effectiveness and resource use outcomes. This study provided more information than the other studies in terms of effectiveness of the first, second and third lines of treatment. Therefore a different analysis structure was needed whereby the effectiveness of the different lines of treatment were directly included in a decision tree structure. This allowed an exploratory cost utility analysis to be undertaken as a sensitivity analysis, on the basis that it was assumed when the different lines of treatment would occur, and the utility of a stone free individual would be that of the UK general population. It could then be calculated what the utility of someone that still had a stone would have to be to make the URS cost effective. When the different lines of treatment would occur was based on a combination of information from the paper and assumptions. If people are waiting longer to have their additional treatments, then this would favour the URS arm more because those who are not initially stone free with SWL (of which there are more) have to wait longer until they are stone free which reduces the total QALY from the SWL arm, making the incremental QALY larger, and reducing the ICER. However if the opposite were true and additional treatments happened sooner, then this increases the total QALY from the SWL arm by more than in the URS arm because all the people having second and third line treatments are having them sooner, and so those that become stone free from those procedures derive a QoL gain sooner. This reduces the incremental QALY gain and increases the ICER of URS vs SWL.

Only one study was used for this analysis which is a limitation. This study is also from 1999, and the study states a mobile lithotripter was used. The date of the study influenced the procedures being used as some of the procedures were open surgery, which is rarely used in current practice, so this was assumed to be URS in the analysis. These may have different effectiveness levels in reality. Effectiveness of SWL can vary depending on whether a fixed site or mobile lithotripter is being used, with mobile lithotripters anecdotally being less effective. However as this study reported more detailed information, it was thought useful to have a scenario with a different approach to scenarios 1 and 2. It also reports lower effectiveness than the studies used in scenarios 1 and 2 and therefore is more reflective of UK practice.

Linking to the above point, a concern of the committee in general was that the clinical data was not felt to be reflective of the accuracy of SWL in practice (particularly for scenarios 1 and 2). SWL is not current practice and there is a lack of availability in many regions of the country. Therefore effectiveness at the current time in England tends to be variable and can be much lower than identified from the review, e.g. one source from UK audit data showed an effectiveness for ureteric stones of around 60%[2]. This can be for a number of reasons such as the effectiveness being operator dependent and what training that operator might have had. One argument however may be that the effectiveness reflected in the review may well reflect effectiveness in the UK if the technology was widely available. The effectiveness of SWL was tested in sensitivity analyses in all the scenarios.

There may also be limitations with regards to some of the costs used, particularly the cost of SWL. As this is an intervention not currently widely available, then providing this in all areas to ensure timely access to it may be a difficult investment decision if there is not felt to be enough people that would use it to make that a worthwhile investment. More people using the equipment would reduce the cost per person (as purchase and running costs would have to be annuitized and spread over the uses that would be gotten from the machine). NHS reference costs include costs on a full absorption basis, which means that the purchase and running costs are included in the cost per procedure that is reported. So although this is an average cost that would include the cost per person from those that submitted data to NHS reference costs, if SWL was more widely available then without adequate people using them that may well drive up the average in NHS reference costs. One limitation therefore of the reference cost is that it might be being skewed by the larger volume centres. This was tested in sensitivity analyses such as assuming 2 initial SWL sessions rather than 1, which is essentially the same as assuming it is twice as expensive, and found that the cost of a session of SWL would have to range from £1,600 to around £2,600 across the scenarios to make the two strategies cost neutral.

1.2.3.3 Generalisability to other populations or settings

This analysis was only using the adult data, as with children there are other considerations such as general anaesthetic being needed for SWL as well as URS, therefore the risks associated with that and the fact that SWL may require more repeat sessions means that this may be less preferable than surgery for those reasons. Children's services also tend to be more centrally located with children being treated in specialist centres because renal stones are rarer in children, therefore all the alternatives are probably more accessible than in adults. Therefore there are other considerations in children that make the results of this analysis difficult to generalise in children.

In terms of the generalisability to other stone sizes or other types of stones, the effectiveness of the interventions may or may not be different for other population subgroups. However as the conclusion is generally that URS is not feasibly cost effective, even in sensitivity analyses that tested alternative levels of effectiveness, then it is possible that this conclusion could also be extrapolated to other size stones. Particularly for renal stones for example, further intervention after a failed intervention is not always required because the stone does not have such as much of a risk of obstruction. Therefore undertaking an initial SWL would still stop some people from having to go on to further treatments if it was successful, and it is much cheaper than a URS. Renal stones are also less painful than ureteric stones, so the quality of life impact from still having a stone is not as significant.

For larger stone sizes, such as larger ureteric stones, there may also be more benefits to patients from being treated with an intervention more likely to be successful on the first attempt because patients are likely to have more pain and be at higher risk of complications such as obstruction. Therefore, benefits with URS may be higher. Again given the large differences in cost between strategies, even if there was a larger effectiveness difference between interventions, it would be difficult to say whether the benefits of URS would be large enough to make that a cost effective intervention.

1.2.3.4 Comparisons with published studies

No published economic evaluations were included that compare different surgical interventions in this population based on the RCTs included in the clinical review for the guideline.

The committee are aware of an ongoing study with accompanying cost effectiveness analysis (TISU trial) comparing SWL with URS in adults with ureteric stones.

1.2.3.5 Conclusions

The committee agreed that URS is unlikely to be cost effective for ureteric stones <10mm. URS was more effective than SWL in terms of stone-free rate after the initial procedure, but was also associated with higher costs even when the difference in downstream costs between interventions were taken into account. It was considered unlikely that the benefits of URS in terms of higher initial effectiveness will result in QALY gains that would justify this cost.

The committee discussed that there may be short term implementation costs of recommending SWL due to it not currently being widely used in the NHS, but that there are likely to be many options for implementation that may minimise these costs. For example, having good referral systems to refer patients to a hospital with a lithotripter may mean additional machines are not needed, and currently there is believed to be less waiting time for SWL than being on a surgery waiting list, so existing capacity may be available. Alternatively, perhaps more investment in mobile lithotripters could be an option rather than needing fixed site lithotripters in all hospitals (or regions) (however the effectiveness between mobile and fixed can differ which has not been addressed here). Other resources may be affected however, because if more machines will be available then there may not currently be enough trained staff to undertake the procedures such as ultrasonographers/radiographers.

1.2.3.6 Implications for future research

This analysis has added to research informing on the cost and potential cost effectiveness of strategies starting with different interventions for treating small ureteric stones. This is an important topic because it highlights the important trade-off between effectiveness and cost. It will also hopefully encourage further cost effectiveness research from a UK setting particularly alongside a UK trial that would be able to capture effectiveness and resource use in UK practice as well as quality of life data, which is lacking.

1.3 Renal stones <10mm, and Renal stones 10-20mm

1.3.1 Introduction and approach

1.3.1.1 Approach to analysis

There were three populations mentioned in the introduction (1.1) as groups where there is uncertainty in practice regarding choice of technique, and where the more expensive procedure was more effective.

Given how:

- the ureteric <10mm analysis showed that URS is unlikely to be cost effective, even when larger effectiveness differences were assumed between the strategies,
- and also comparing across the clinical review data for the three groups, which showed the effectiveness not to be too dissimilar

It was inferred that simpler cost offset calculations would be adequate in helping to infer the likelihood of the cost effectiveness of the more expensive treatments.

The cost offset calculations only incorporate the cost of the initial interventions, and retreatment and ancillary procedures. What is being tested as to whether costs offset each other is the difference in initial intervention costs traded off against the difference in downstream resource use of retreatments and ancillary procedures. As the more expensive intervention is also more effective, which in turn leads to lower downstream resource use. Therefore the purpose is to see whether the downstream resource use will offset the difference in upfront intervention costs. Note that it is not clear if this is the cost that would make everyone stone free, as this depends on the endpoint of the studies that the clinical data is a summary of. So there are limitations to the approach in terms of potential underestimation of cost, however these calculations are meant to be interpreted as informal cost calculations using the available clinical data. As discussed further below in the specific stone subgroups, it may not be the case that the aim is to get someone stone free, as this depends on their symptoms and size of stone for example.

Additionally, as the ureteric analysis was a costing analysis primarily, with the QALY work being exploratory, then the conclusion can only be an estimate of whether the intervention is feasibly cost effective, and therefore simpler costing calculations would still allow exploratory work around the feasibility of cost effectiveness.

Furthermore, as another potential source of data to assist in illustrating the costs of an SWL strategy, UK audit data from the BAUS Endourology national ESWL practice and outcomes audit[1] was analysed and costs applied to identify the cost of treating people with SWL using real data.

1.3.1.2 Quantifying resource use

1.3.1.2.1 Clinical review data

A comparison of the clinical data across the three subgroups, summarised from the clinical review, is shown below in Table 36.

Table 36: Comparison of clinical data across population subgroups

	Subgroup		
Intervention	**Ureteric <10mm**	**Renal <10mm**	**Renal 10-20mm (b)**
Stone free probabilities (a)			
PCNL	NA	- (one small study only; SWL vs PCNL)	93% - 96%
URS	93%	88%	90% - 91%
SWL	84%	84%	61% - 75%
Retreatment probabilities			
PCNL	NA	-	1.2% - 2.7%
URS	2.9%	5.7%	1.6% - 9.5%
SWL	14.5%	34%	22.4% - 56.6%
Ancillary procedure probabilities			
PCNL	NA	-	1.7% - 5.1%
URS	4.1%	3.9%	6.1% - 9.3%
SWL	9.4%	9.3%	10.1% - 18.8%

Note that these figures are based on the clinical evidence summaries in the review (chapter F), whereas in the ureteric costing analysis the raw data has been used, and so the probabilities may not appear exactly the same as those used in the ureteric model. Note that another reason these renal cost analyses are simpler than the ureteric analysis is because the clinical studies for the renal subgroups were not explored in detail to find out about resource use or analyse the raw data, as was done in the ureteric analysis.
(a) Note that uncertainties around when the stone free outcomes are being measured (i.e. after initial procedures or after further procedures) in the studies can vary the above effectiveness levels.
(b) There were 3 pairwise comparisons in this strata hence a range of probabilities.

1.3.1.2.2 BAUS Endourology national ESWL practice and outcomes audit

A snapshot audit of current ESWL practice across the UK in 2017 was undertaken[1]. This involved all units undertaking ESWL across the country being asked to recruit 10 consecutive new patients with renal stones attending for ESWL and submit data over a 6 month period (note that ESWL and SWL are the same thing and the terms are used interchangeably here).

The raw data was obtained through the committee, and analysed to crudely obtain the cost of SWL treatment by costing up the resource use involved in providing SWL including the primary treatments and downstream resource use. Note that as this audit only includes renal stones, a similar analysis could not be undertaken for the ureteric analysis.

In total there were 141 patients suitable for evaluation in the dataset, with 101 patients having renal stones <=10mm, and 40 having renal stones 10-20mm.

The dataset reports information such as: The size of the stone, the number of SWL treatments administered in total (what is referred to as sessions in this document), the interval between the first and last SWL treatment (in days), the status at review 3 months and 6 months following the first SWL treatment, and the subsequent management decision following the 3 and 6 month reviews.

The status of the patient at review is broken down into 4 categories: 'stone free', 'stone fragments <2mm in maximal diameter', 'stone fragments 2-4mm in maximal diameter', and 'stone fragments >4mm in maximal diameter'. In practice, the definition of stone free can vary, but generally clinicians consider the stone being broken down into fragments less than 3mm as the patient being stone free. Therefore stone free using this dataset has been defined as patients in the 'stone free' and 'stone fragments <2mm in maximal diameter' category. This may be an underestimate of those considered stone free as there will be people with fragments between 2-3mm in the 'stone fragments 2-4mm in maximal diameter' group, however, it is not possible to identify those people.

The 3 month status of the patient and subsequent management decided at 3 months are the source of information on resource use, which costs were attached to. It is acknowledged that omitting the 6 month data may lead to an underestimate of the resource use of an SWL strategy if further resource use is consumed after 3 months. However, at 6 months more people were lost to follow up or the status was blank which would have led to fewer patients having outcomes that could be costed. Additionally, as the subsequent management at 3 months was included in the costings, which included those who had interventions planned, then if the 6 month outcome was that the intervention had been undertaken, then this would have already been accounted for. Therefore this was unlikely to make a large difference.

The dataset also reported the imaging modality used to confirm stone passage, but this was not included in the costings because assumptions would have to be made about whether follow up is after every SWL session or just at the end of the course. Therefore it is acknowledged that there are costs omitted as part of the treatment pathway, however these costings should be interpreted as illustrative costs of an SWL pathway using 'real' data, and is therefore more exploratory given that this data was available.

1.3.1.3 Resource use assumptions and unit costs

These costings are more simplistic than the ureteric analysis, as other resource use such as stents, follow up, and adverse events are not included.

The costs used are only the cost of the procedures (SWL, URS, PCNL) and are the same as those reported in section 1.2.1.3.

As with the ureteric analysis, the retreatments and ancillary procedures are assumed to be mutually exclusive rather than sequential. In other words, those people who fail primary treatment have either a retreatment or an ancillary, rather than only those who fail the retreatment going on to have an ancillary procedure. If anything, this approach is more favourable to the more expensive intervention, because the probablity of an ancillary treatment is higher if applied as a secondary treatment, than if applied only to the proportion that have retreatments, and as there are more secondary treatments for the less effective intervention - this would lead to a larger difference in downstream costs and therefore upfront costs are more likely to be offset.

Assumptions have also been made about the number of sessions that constitute a primary treatment of SWL, and how many constitute a retreatment. Primary session might imply a single session, however generally a course of treatment is offered which can be up to 3 or 4 sessions for a renal stone. Studies generally state ‘a maximum of X number of sessions was offered’, however it is not always clear how many sessions each person had, unless they report the average number, as not everyone would need up to 4 as some people would be stone free with 1 session, some with 2 etc. Therefore based on this, the retreatment outcome has been interpreted as the number who had more than 1 session, and conversely the primary treatment is assumed to be a single session. As the studies are not always clear, then if retreatment was meant as additional sessions after a course of treatment has failed, costs of SWL may be being understimated. On the other hand to assume everyone had a certain number might be overestimating. The number of sessions included as retreatment is discussed in the relevant sections for each subgroup below as this differs between the subgroups. However the number of primary and retreatment is varied to see how this affects the costs.

Although an ancillary was technically defined in the ureteric analysis as being a different procedure to the primary procedure, different scenarios are presented in the results where the ancillary is assumed to be different types of procedures to see the impact this has on the costs.

1.3.1.4 Sensitivity analysis

Exploratory QALY work will be undertaken for the cost offset clauclations, similar to the methods described in section 1.2.1.5.2.

The scenario with the ancillary treatment assumed to be URS for both interventions will be taken as the ‘base case’ value to be used for the exploratory QALY work, because this is the most likely scenario in practice.

The exploratory QALY work involves some back-calculations based on the cost offset values and the NICE threshold to determine the incremetnal QALY needed to make URS cost effective. Taking this a step further and dividing out the time horizon leads to the QoL difference needed per patient to make URS cost effective. The time horizon for the amjority of studies across both renal stone groups was 3 months, but results for a time horizon of 1,2 and 4 months will also be reported to see how this impacts the QoL.

If we knew the difference in initial effectivenes between the interventions then we could also calculate the QoL difference needed between a stone free and non stone free health state. However as the data is an average from the clinical review then we are not certain what the initial effectiveness of a single SWL session is, so an effectiveness difference of 20% against URS will be assumed for renal stones <10mm, and 30% for renal stones 10-20mm, for illustration purposes.

Other sensitivity analyses will also be explored such as the number of sessions being costed for SWL.

1.3.2 Renal stones <10mm, URS versus SWL

1.3.2.1 Methods

1.3.2.1.1 Clinical review data

For this subgroup only URS and SWL are being compared as there was only clinical evidence from one small study that had a PCNL comparator. Additionally this would not commonly be a first line treatment for this size of stone.

Table 36 shows that the effectiveness of URS and SWL is closer together for small renal stones than it was for small ureteric stones because the effectiveness of URS has decreased. This implies that there will be less benefit of URS above that of SWL compared to the ureteric group, as fewer people will be initially stone free with URS, and so there will be less people achieving an increase in QoL early on in the pathway. Also as more resource use would be required downstream in the URS arm to get everyone stone free, then this would lead to higher costs also. The result of this is likely to be an even bigger cost divide between the interventions and a smaller difference in QALYs, compared to the ureteric group.

On the other hand, small renal stones may not necessarily require further intervention if treatment has failed and the patient is asymptomatic, as the stone does not cause as much concern with regards to possible risk of obstruction, because there is more room for the stone to move in the kidney. In which case, if only primary intervention was undertaken (i.e. a course of SWL and primary URS) then there would not be downstream resource use of SWL (because it is initially less effective) to help offset the cost of the strategies. Furthermore, if only primary intervention was undertaken, there are still factors we are unsure of in order to estimate the cost difference, such as the success of SWL after different numbers of sessions. As in theory, a course of SWL for renal stones can involve up to 3 or 4 sessions in practice, whereas only 2 sessions would be offered for a ureteric stone. Therefore the more sessions consumed, the closer the costs between SWL and URS become. However this logic only applies if everyone gets multiple sessions, but this may not be the case as for some people they may become stone free after a single session, some after 2 sessions, and so on.

The assumptions made about the number of sessions being assumed for the primary treatment was discussed in section 1.3.1.3.

The number of sessions that count as a retreatment is assumed to be 2 in the <10mm renal stones group, therefore those that have a retreatment have another 2 sessions.

1.3.2.1.2 BAUS Endourology national ESWL practice and outcomes audit

Table 37 below provides a breakdown of the data for the 101 patients, broken down by the number of SWL sessions.

Table 37: Renal stones <10mm, ESWL BAUS data, outcome by number of sessions

No. of sessions	No. of people	3 month status (a)		
		No. stone free (<2mm) after each number of sessions (%)	No. with 2-4mm stone (%)	No. with >4mm stone (%)
1	34	26 (76%)	4 (12%)	4 (12%)
2	50	16 (32%)	15 (30%)	19 (38%)
3	13	4 (31%)	5 (38%)	4 (31%)
4	4	2 (50%)	0 (0%)	2 (50%)

(a) Total of 48 stone free, and 53 not stone free (stone of >2mm).

The average number of SWL sessions across the 101 individuals at 3 months was 1.87. The average stone free probability at 3 months was 48%, which is the average across individuals having different numbers of sessions. So 1.87 sessions led to 48% of people being stone free. Therefore, to get everyone stone free would require more sessions.

Subsequent management at 3 months for those who are and are not stone free is summarised in the table below. The possible courses of action for subsequent management have been defined by the dataset.

Table 38: Subsequent management at 3 months – renal stones <10mm

	Stone free (<2mm)		Not stone free (>2mm)	
Management	**No. of people**	**Percentage**	**No. of people**	**Percentage**
Flexible URS undertaken	0	0.0%	4	7.5%
Flexible URS planned	0	0.0%	5	9.4%
Rigid URS undertaken	0	0.0%	4	7.5%
Rigid URS planned	0	0.0%	1	1.9%
PCNL planned	0	0.0%	1	1.9%
Further SWL	1	1.9%	9	17.0%
Clinical Review/Observation	26	49.1%	26	49.1%
Discharged	21	39.6%	2	3.8%
Missing	0	0.0%	1	1.9%

Costs were applied to the types of management based on those used in the ureteric analysis (1.2.1.3): both rigid and flexible URS were applied the cost of URS as this is an average unit cost regardless of the method used for the procedure. Planned PCNL was applied the cost of a PCNL, further SWL was applied the cost of a single session of SWL, clinic review/observation was applied a follow up cost. Those discharged were not applied a cost. Missing data was not included in the analysis. As the missing value was actually in the >4mm group, it is possible that further treatment may have happened, however 31% of those in the >4mm group had clinic review/observation. As only one patient had missing data, this is also not likely to have a large impact on the overall cost.

1.3.2.2 Results

1.3.2.2.1 Using clinical review data

The results of the cost offset calculations are shown in Table 39.

The probability of retreatment and ancillary procedures from Table 36 were multiplied by unit costs to derive the values in the second and third column of Table 39.

Table 39: Renal stones <10mm cost, offset calculations

	URS	SWL	Incremental cost (b)	Cost offset
Cost of primary procedures				
	£2,172	£452	£1,720	
Cost of retreatment				
	£124	£307 (a)	-£184	
Cost of ancillary procedure				
1. If ancillary = SWL	£18	£42	-£24	**£1,512**
2. If ancillary = URS	£85	£202	-£117	**£1,419**
3. If ancillary = PCNL	£203	£483	-£281	**£1,256**
4. Ancillary = URS for SWL, and PCNL for URS	£203	£202	£0.61	**£1,537**
Pooling retreatments and ancillary probabilities				
URS as secondary procedure	£209	£940	-£732	**£988**

(a) This includes 2 sessions.
(b) Calculated as URS minus SWL.

Assuming a single session for the primary intervention and 2 for retreatment means that each person would have on average 1.7 sessions, as 100% receive 1 session, and 34% receive 2 sessions.

As an example to illustrate how the cost offsets were calculated, if we take the 'if ancillary =SWL' scenario, then the incremental primary procedure cost plus the retreatment cost difference (which is a saving of -£184), plus the ancillary cost difference (which is a saving of -£24) equals £1,512. So the primary cost difference has been offset by £208. The magnitude of the cost offset figures tells us that the primary cost difference is never entirely being offset by the downstream savings of the more effective alternative. The primary procedure incremental cost has been offset the most when retreatment and ancillary probabilities were pooled, with URS was assumed to be the secondary procedure. This is because pooling the probabilities means that the 34% retreatment rate (which had a cost of 2 SWL sessions attached to it) is now being applied the cost of a URS (which costs more than twice as much), which helps drive up the cost of the SWL downstream resource use, and has more of an impact on the cost offset.

Comparing the incremental costs with those from the ureteric analysis (note the cost offset is the same as incremental cost) show that these are slightly lower. There are many factors omitted here such as stent costs, which added a substantial amount to the URS strategy in the ureteric analysis. Overall, the conclusion is similar to that in the ureteric analysis that there is a large cost difference between the strategies even when considering downstream resource use.

Sensitivity analysis

If we assume that for primary treatment of SWL everyone gets 2 sessions, and a retreatment is 1 session, then the average number of sessions per person is now 2.34, and the cost offset for the ancillary scenario labelled 2 in Table 39, reduces to £1,121.

Rearranging the ICER equation to work out the QALY gain needed to make URS cost effective leads to a value of 0.07. Table 40 details the remainder of the exploratory QALY work.

Table 40: Renal stones <10mm - exploratory QALY work

	Timeframe			
	1 month	**2 months**	**3 months**	**4 months**
Time as a proportion of a year	0.083	0.167	0.250	0.333
QoL difference needed per patient to make URS cost effective (a)	0.851	0.426	0.284	0.213
QoL difference between a stone free health state and non stone free health state (b)	4.26	2.13	1.42	1.06

(a) Derived from dividing the QALY gain of 0.07 by the time.
(b) Derived from dividing the QoL difference needed per patient by the population this would apply to (assumed to be 0.2)

The QoL difference needed between a stone free and non-stone free health is within the possible range on the EQ-5D if the time horizon is longer than 3 months. However as discussed for the ureteric analysis, there are many caveats as to why this is likely to be an overestimate.

1.3.2.2.2 BAUS Endourology national ESWL practice and outcomes audit

The breakdown of the costs is shown in Table 41. The average number of sessions was multiplied by the cost of SWL to find the cost of the primary treatment. And the subsequent 3 month management of those who were and were not stone free was costed up by multiplying the probability by its respective cost.

Table 41: Renal stones <10mm, ESWL BAUS data, total cost

	Number applied to	Proportion applied to	Cost
Primary treatment			
	101	100%	£85,357
Subsequent management at 3 months			
Stone free (<2mm)	*48*		
Flexible URS undertaken	0	0.0%	£0
Flexible URS planned	0	0.0%	£0
Rigid URS undertaken	0	0.0%	£0
Rigid URS planned	0	0.0%	£0
PCNL planned	0	0.0%	£0
Further SWL	1	1.9%	£409
Clinical Review/Observation	26	49.1%	£2,426
Discharged	21	39.6%	£0
Not stone free (>2mm) [(a)]	*53*		
Flexible URS undertaken	4	7.5%	£8,688
Flexible URS planned	5	9.4%	£10,860
Rigid URS undertaken	4	7.5%	£8,688
Rigid URS planned	1	1.9%	£2,172
PCNL planned	1	1.9%	£5,195
Further SWL	9	17.0%	£4,065
Clinical Review/Observation	26	49.1%	£2,679
Discharged	2	3.8%	£0
Total cost			£130,539
Total cost per person			**£1,292**

(a) Note that the percentages may not add to 100% because there was 1 missing record in this group.

The total cost of an SWL strategy per person based on the audit data was found to be £1,292.

This cost is roughly in line with the total cost of SWL from the crude cost analysis in the previous section (if sum up the primary treatment cost, retreatment and ancillary cost for each scenario then this ranges from £800 to around £1,400). It is important to note that there are various inputs omitted such as adverse events and use of stents for example, as this is a crude and simple analysis of audit data.

If we compare this cost to the ureteric analysis, it is slightly higher than the cost of the SWL strategy in scenarios 1 and 2. This is likely to be because the effectiveness of SWL is lower here, therefore downstream resource use is also higher. Additionally, the cost per person of the primary treatment is higher as there are more sessions assumed here. The difference with scenario 3 in terms of the cost of the SWL strategy being higher there is also likely to be explained by the fact that different subsequent procedures are being used, such as more PCNL use in scenario 3 of the ureteric analysis.

Overall, we know that the cost of a URS strategy is going to be more than £2,200 per person as that is the rough cost of a URS procedure. Once taking into account the downstream resource use also consumed there, then using real data has also shown that there is likely to be a substantial cost difference between the two strategies.

1.3.3 Renal stones 10-20mm, PCNL versus URS versus SWL

1.3.3.1 Methods

1.3.3.1.1 Clinical review data

For the larger renal stones subgroup, there was data in the clinical review for all three types of surgery because there were three pairwise comparisons: SWL vs URS, URS vs PCNL, and SWL vs PCNL, hence why the summary values in Table 36 are presented as a range.

If we rank interventions in order of their cost, then focus on one pairwise comparison at a time, starting with the most expensive (PCNL) compared to the next most expensive (URS). Table 36 shows that the difference in effectiveness in terms of stone free rate is not too dissimilar between PCNL and URS. The retreatment and ancillary procedure probabilities show that URS has slightly higher probabilities but this can vary depending on the pairwise comparison that the data was taken from. PCNL is also more than twice as expensive as a URS.

Taking the next pairwise comparison of URS versus SWL for this subgroup, and looking at Table 36, tells us that the effectiveness difference is larger than that of the other subgroups in the table. This may be because the effectiveness of SWL reduces with the size of the stone. There is also a large variation in SWL retreatment and ancillary rates, depending on which pairwise comparison these are from, but as expected, SWL leads to more downstream resource use which we assume is a consequence of lower effectiveness.

Cost offset calculations are undertaken for these two pairwise comparisons.

Comparing PCNL to SWL was not deemed necessary because there is such a large difference in the primary costs of treatments alone that it can be inferred PCNL is highly unlikely to be cost effective against SWL, even though it is considered more effective.

The clinical data for each of the two comparisons was based on taking the midpoint of the ranges in Table 36 for the retreatment and ancillary procedure outcomes. These are presented below in Table 42.

Table 42: Clinical data for renal stones 10-20mm group

Intervention	Renal 10-20mm
Retreatment probabilities	
PCNL	2.0%
URS	5.6%
SWL	39.5%
Ancillary procedure probabilities	
PCNL	3.4%
URS	7.7%
SWL	14.5%

With a larger renal stone it is considered more likely that there will be further intervention following a failed procedure, because of the size of the stone, unlike a small renal stone. There may therefore be a higher chance of the primary cost difference being offset.

Renal stones are likely to have up to 3 or 4 sessions of SWL offered as a course of treatment. The number of sessions assumed for the primary intervention was a single session, as discussed in section 1.3.1.3. The number of sessions that count as a retreatment is assumed to be 3 in this size renal stone group because the stone is larger, therefore those that have a retreatment have another 3 sessions. As with the small renal stones, this is an estimate as there is uncertainty around the success of different numbers of sessions. Assuming a single primary session and 3 further sessions for those who need retreatment leads to an average of 2.2 sessions.

1.3.3.1.2 BAUS Endourology national ESWL practice and outcomes audit

Table 43 below provides a breakdown of the data for the 40 patients with renal stones 10-20mm in the dataset, broken down by the number of SWL sessions.

Table 43: Renal stones 10-20mm, ESWL BAUS data, outcome by number of sessions

No. of sessions	No. of people	3 month status (a)		
		No. stone free (<2mm) after each number of sessions (%)	No. with 2-4mm stone (%)	No. with >4mm stone (%)
1	13	7 (54%)	2 (15%)	4 (31%)
2	12	6 (50%)	0 (0%)	5 (42%)
3	10	1 (10%)	2 (20%)	7 (70%)
4	4	0 (0%)	0 (0%)	4 (100%
6	1	0 (0%)	0 (0%)	1 (100%)

(a) *Total of 14 stone free, and 25 not stone free (stone of >2mm). Note these do not add to 40 because there is 1 record with missing data in the 2 sessions group.*

The average number of sessions across the 40 individuals was 2.2. This coincidentally is the same average number of sessions per person that were assumed in the cost offset calculations for this subgroup. However the actual average number of sessions to clear the stone may be higher in reality for either the cost offset calculations or using the BAUS data, because not everyone is stone free at the end of the trials, or at 3 months in the dataset.

The average stone free rate at 3 months was 35%. This is lower than the small renal stones group using the same dataset.

Subsequent management at 3 months for those who are and are not stone free is summarised in the table below. The possible courses of action for subsequent management have been defined by the dataset.

Table 44: Subsequent management at 3 months – renal stones 10-20mm

	Stone free (<2mm)		Not stone free (>2mm)	
Management	No. of people	Percentage	No. of people	Percentage
Flexible URS undertaken	0	0.0%	4	16.0%
Flexible URS planned	0	0.0%	2	8.0%
PCNL planned	0	0.0%	1	4.0%
Further SWL	2	14.3%	9	36.0%
Clinical Review/Observation	4	28.6%	8	32.0%
Discharged	7	50.0%	1	4.0%
Missing	1	7.1%	0	0.0%

(a) *Note that the number of people in the table still add to 29 because there is 1 record where the status at 3 months and subsequent management fields are missing. This is on top of the 1 record where subsequent* *management data is missing in the <2mm group in the above table; therefore there are two individuals with missing data out of the 40 people in total*

As in the renal stones <10mm group, costs were applied to the types of management using the costs used in the ureteric analysis. Those who were discharged or where data was missing were not applied any costs. This may underestimate costs, however for the records missing management data, this was in the <2mm group where the majority of people were either discharged or reviewed. For the other missing record it is not clear what the outcome of that individual was in order to attribute a particular management or a judgement about whether they were stone free or not.

1.3.3.2 Results

1.3.3.2.1 Using clinical review data

The results for PCNL versus URS, and URS versus SWL, are presented separately.

PCNL vs URS

For PCNL versus URS, three scenarios around the ancillary procedure probability have been assumed, as well as retreatment and ancillary probability pooled and URS assumed to be the secondary procedure overall.

Table 45: Renal stones 10-20mm, cost offset calculations – PCNL vs URS

	PCNL	URS	Incremental cost (a)	Cost offset
Cost of primary procedures				
	£5,195	£2,172	£3,023	
Cost of retreatment				
	£104	£122	-£18	
Cost of ancillary procedure				
1. If ancillary = SWL	£15	£35	-£19	**£2,986**
2. If ancillary = URS	£74	£167	-£93	**£2,912**
3. If ancillary = PCNL	£177	£400	-£223	**£2,782**
Pooling retreatments and ancillary probabilities				
URS as secondary procedure	£117	£289	-£172	**£2,851**

(a) Calculated as PCNL minus URS.

The results show that regardless of what procedure might be assumed as an ancillary procedure, the procedural cost difference is very unlikely to be offset by any downstream savings from PCNL needing less downstream resource use. This is because the differences in downstream resource use between the two procedures are very small that this is having a negligible impact on the incremental initial intervention cost because PCNL is much more expensive.

This tells us that PCNL is highly unlikely to be cost effective, because the small effectiveness difference between the interventions is unlikely to create a large enough QALY gain to justify the large additional cost of PCNL.

URS vs SWL

Table 46: Renal stones 10-20mm, cost offset calculations – URS vs SWL

	URS	SWL	Incremental cost (b)	Cost offset
Cost of primary procedures				
	£2,172	£452	£1,720	
Cost of retreatment				
	£122	£536 (a)	-£414	
Cost of ancillary procedure				
1. If ancillary = SWL	£35	£66	-£31	**£1,275**
2. If ancillary = URS	£167	£315	-£148	**£1,158**
3. If ancillary = PCNL	£400	£753	-£353	**£953**
4. Ancillary = URS for SWL, and PCNL for URS	£400	£315	£85	**£1,391**
Pooling retreatments and ancillary probabilities				
URS as secondary procedure	£289	£1,173	-£884	**£836**

(b) This includes 3 sessions.
(c) Calculated as URS minus SWL.

As in the small renal stone analysis, the incremental cost of the interventions themselves are not being fully offset by the difference in downstream resource use. This is because the difference in intervention cost is still substantial. As the number of sessions of SWL increase this closes the gap between the cost of SWL and URS (so the range between of cost offsets is lower than the renal <10mm cost offsets), however as mentioned, multiple sessions may not apply to all individuals. This is explored further when the BAUS audit data is costed for this stone group in the next section.

Sensitivity analysis

As a sensitivity analysis, if everyone got 2 sessions of SWL as the primary treatment, and retreatment was also assumed to be 2 sessions, then this would lead to an average of 2.8 sessions, and the cost offset for ancillary scenario 2 to fall to £885.

Rearranging the ICER equation to work out the QALY gain needed to make URS cost effective leads to a value of 0.06. Table 47 details the remainder of the exploratory QALY work.

Table 47: Renal stones 10-20mm - exploratory QALY work

	Timeframe			
	1 month	2 months	3 months	4 months
Time as a proportion of a year	0.083	0.167	0.250	0.333
QoL difference needed per patient to make URS cost effective (a)	0.695	0.347	0.232	0.174
QoL difference between a stone free health state and non stone free health state (b)	2.32	1.16	0.77	0.58

(a) Derived from dividing the QALY gain of 0.06 by the time.
(b) Derived from dividing the QoL difference needed per patient by the population this would apply to (assumed to be 0.3)

The QoL difference needed between a stone free and non-stone free health state is within the possible range on the EQ-5D if the time horizon is longer than 2 months. Note that the assumption is that this timeframe is the time between the primary treatment and further treatment, assuming this happens at the end of the time horizon. It is difficult to infer if the difference is feasible however. A large renal stone is likely to have more of an impact on quality of life than a small renal stone might for example, but perhaps not as much as a ureteric stone. Also, all the caveats discussed for the ureteric analysis also still apply, for example, someone may not be in pain all the time but have very poor quality of life during a pain episode. Overall quality of life depends on the frequency of pain episodes, and how many people have symptoms, as some people may not, but we are talking about average quality of life across the whole population with this size renal stone. Based on all the above it is difficult to make a judgement about whether this calculated gain is likely, as it may be within the realms of possibility, but still be unfeasible.

1.3.3.2.2 BAUS Endourology national ESWL practice and outcomes audit

The breakdown of the costs are shown in Table 48. The average number of sessions was multiplied by the cost of SWL to find the cost of the primary treatment. And the subsequent 3 month management of those who were and were not stone free was costed up.

Table 48: Renal stones 10-20mm, ESWL BAUS data, total cost

	Number applied to	Proportion applied to	Cost
Primary treatment	*40 (a)*	100%	£40,194
Stone free (<2mm) (b)	*14*		
Flexible URS undertaken	0	0.0%	£0
Flexible URS planned	0	0.0%	£0
PCNL planned	0	0.0%	£0
Further SWL	2	14.3%	£903
Clinical Review/Observation	4	28.6%	£412
Discharged	7	50.0%	£0
Not stone free (>2mm)	*25*		
Flexible URS undertaken	4	16.0%	£8,688
Flexible URS planned	2	8.0%	£4,344
PCNL planned	1	4.0%	£5,195
Further SWL	9	36.0%	£4,065
Clinical Review/Observation	8	32.0%	£824
Discharged	1	4.0%	£0
Total cost			£64,626
Total cost per person			**£1,616 (c)**

(a) The sum of the number of people with <2mm and >2mm stone fragments sum to 39 as there is 1 missing record that cannot be categorised by stone fragment size.
(b) Note that summing the numbers for all the subsequent management in those with stone fragments <2mm is equal to 13 not 14, as one record from this group does not report subsequent management and was omitted.
(c) This is the total cost divided by all 40 individuals.

The total cost of an SWL strategy per person was £1,616.

This is similar to the top end of the total costs in the SWL strategy in the cost-offset calculations (as there were various scenarios based on ancillary procedure types).

This is higher than the renal stone <10mm group, and this might be explained by the lower overall effectiveness, so there are more people having further procedures as a proportion of the overall cohort. Additionally, of those who are not stone free, there are higher probabilities of further SWL for example. Although it may appear as the same management strategy leads to the same cost (e.g. further SWL in >2mm group of Table 41 and Table 48, in Table 48 it is being spread over fewer people). Furthermore, the primary treatment cost per person is higher in the larger renal stones group because the average number of sessions per person is higher. It is important to note that there are various inputs omitted such as adverse events and use of stents for example, as this is a crude and simple analysis of audit data.

We know that the cost of a URS strategy will be higher than the cost of the intervention itself, which means it is still highly likely that a URS strategy will have a higher cost. The audit data only extended to SWL data, and therefore we do not have audit data on the real success of URS to compare. Again whether a cost difference can be justified by adequate benefit of URS depends on many factors, and it is uncertain whether URS is cost effectiveness against SWL in this stone group. However as demonstrated earlier, PCNL is unlikely to be cost effective.

1.3.4 Discussion

The conclusions of the ureteric analysis showed that higher downstream resource use (even when tested in sensitivity analysis at higher levels) did not offset the incremental intervention cost. This conclusion meant that it could be inferred, also given the similarities in clinical data across the three population groups, that the conclusion was also likely to be the same for the other two groups. Therefore simpler cost calculations were undertaken for the renal stone populations, to demonstrate how upfront costs were unlikely to be offset by differences in downstream costs. This still involved some exploratory QALY work using the same methods as the ureteric analysis to determine the QoL difference needed between a stone free and non-stone free health state.

The results for both groups showed that there remained a significant cost difference between the interventions even when considering downstream resource use of retreatments and ancillary procedures assuming various scenarios associated with what ancillary/secondary procedures might be. For large renal stones, in particular PCNL was unlikely to be cost effective compared to URS (or SWL) because of how much more costly this is than the other interventions and the high effectiveness means there is little room for cost offsets. The exploratory QALY work also showed that over longer timeframes there can be quality of life differences between health states that are possible, but the feasibility of these are uncertain.

There are many caveats to the costs estimated as this only included three parameters of the primary treatment cost, the retreatment cost, and the ancillary procedure costs. This is not as detailed as the ureteric analysis which was a full analysis in terms of extensive sensitivity analysis. Also, parameters omitted in these simple calculations include: the cost of stents, follow up (and imaging) costs, and adverse events. Assumptions are also made about the number of SWL sessions which may be an over or under estimate depending on the average number of sessions in reality.

There are factors that we can extrapolate from the ureteric analysis to allow us to consider the cost effectiveness of the alternatives for the renal stone subgroups. For example, renal stones are likely to have a smaller quality of life impact than ureteric stones, because of the location of the stone. Therefore even if the incremental costs of interventions being compared were smaller than that suggested in the ureteric analysis, it would still be unlikely that the more expensive intervention would be cost effective because there are smaller gains to be had from clearing the stone earlier with the more effective intervention. As with the caveats around the ureteric analysis exploratory QALY work, the same would apply here in terms of there being many unknowns as to what the quality of life of someone with a renal stone (of different sizes) would be on average, and how long people would be waiting for further treatments after they have failed.

Alongside the cost offset calculations, access to UK audit data on resource use of SWL patients meant that another source of cost illustrations for the less effective intervention was available, using effectiveness levels more reflective of UK practice. BAUS ESWL audit data contained data on 141 patients with renal stones and their practice and outcomes, was analysed to cost up resource use involved in providing SWL using real UK data that would reflect practice. These costings of the resource use in the dataset showed that the cost of a strategy beginning with SWL is similar to that predicted in the cost offset calculations. They were however higher than the SWL strategy cost demonstrated in the ureteric analysis because of lower effectiveness (35-48%), and consequently more downstream resource use from the audit data than the clinical review, but this is still likely to be less than a strategy with URS. The limitations of these crude costings must be acknowledged in terms of: only the resource use up until that undertaken by (or subsequently planned at) 3 months was costed, which means there may be an underestimate of costs. Costs such as imaging or follow up appointments to confirm stone passage, or other resource use that was not included as a 'management decision' in the dataset, was not costed. i.e. no assumptions were made beyond the resource use reported in the dataset. The imaging modality was reported that was used to confirm stone free status and in the majority of cases this was x-ray, but assumptions would have to be made about whether follow up is after every SWL session or just at the end of the course and this would be going beyond the dataset. Costing up the resource use from the BAUS dataset serves to illustrate the real costs involved in an SWL pathway, however without a similar dataset for URS, we can only estimate what the incremental cost might be between the two strategies, and as the cost of a URS strategy would be at minimum the cost of the surgery, then there is still likely to be a cost difference between the two strategies. Whether the additional cost would be justified by additional benefit of the more effective intervention remained uncertain.

In conclusion, for renal stones <10mm and 10-20mm, it has been estimated through more informal calculations that there is still likely to be a significant cost difference between different interventions, and taking into account subsequent downstream resource use, which would be higher for a less effective intervention, still does not make the comparators cost neutral. Whether there is adequate benefit to justify the higher cost of a more expensive strategy is uncertain, but given that a stone in the ureter is of more clinical concern and likely to cause more pain, the benefits from the renal stone subgroups would be expected to be smaller, therefore the conclusion is likely to be the same as the ureteric analysis whereby URS is unlikely to be cost effective. There are many uncertainties around this depending on factors in clinical practice such as waiting time, variability of the effectiveness of SWL, and patient factors such as uncertainty of quality of life impact. But these are factors that have been explored in more detail in the ureteric analysis, and were not found to alter the main conclusion.

References

1. British Association of Urological Surgeons. Section of Endourology - National Audit of ESWL Practice. 2018. Available from: https://www.baus.org.uk/professionals/sections/endourology/research/default.aspx Last accessed: 16th May 2018.

2. Doherty R, Manley K, Gordon S, Irving S, Kumar S, Masood J et al. Current ESWL practice and outcomes in the UK: a multicentre snapshot. Journal of Clinical Urology. 2017; 10(4):340-6

3. Hendrikx AJ, Strijbos WE, Knijff DW, Kums JJ, Doesburg WH, Lemmens WA. Treatment for extended-mid and distal ureteral stones: SWL or ureteroscopy? Results of a multicenter study. Journal of Endourology. 1999; 13(10):727-33

4. Kumar A, Nanda B, Kumar N, Kumar R, Vasudeva P, Mohanty NK. A prospective randomized comparison between shockwave lithotripsy and semirigid ureteroscopy for upper ureteral stones <2?cm: a single center experience. Journal of Endourology. 2015; 29(1):47-51

5. National Institute for Health and Clinical Excellence. Social value judgements: principles for the development of NICE guidance. London. National Institute for Health and Clinical Excellence, 2008. Available from: https://www.nice.org.uk/media/default/about/what-we-do/research-and-development/social-value-judgements-principles-for-the-development-of-nice-guidance.pdf

6. NHS Improvement. Reference costs 2016/17: highlights, analysis and introduction to the data. London. 2017. Available from: https://improvement.nhs.uk/resources/reference-costs/

7. Pearle MS, Nadler R, Bercowsky E, Chen C, Dunn M, Figenshau RS et al. Prospective randomized trial comparing shock wave lithotripsy and ureteroscopy for management of distal ureteral calculi. Journal of Urology. 2001; 166(4):1255-60

8. Salem HK. A prospective randomized study comparing shock wave lithotripsy and semirigid ureteroscopy for the management of proximal ureteral calculi. Urology. 2009; 74(6):1216-21

9. Sarica K, Eryildirim B, Sahin C, Türko?lu Ö K, Tuncer M, Coskun A et al. Emergency management of ureteral stones: evaluation of two different approaches with an emphasis on patients' life quality. Archivio Italiano di Urologia, Andrologia. 2017; 88(3):201-5

10. Verze P, Imbimbo C, Cancelmo G, Creta M, Palmieri A, Mangiapia F et al. Extracorporeal shockwave lithotripsy vs ureteroscopy as first-line therapy for patients with single, distal ureteric stones: a prospective randomized study. BJU International. 2010; 106(11):1748-52

11. Zhang J, Shi Q, Wang GZ, Wang F, Jiang N. Cost-effectiveness analysis of ureteroscopic laser lithotripsy and shock wave lithotripsy in the management of ureteral calculi in eastern China. Urologia Internationalis. 2011; 86(4):470-5

Appendix A: Clinical data

A.1 Summary: ureteric stones clinical studies

Table 49: Summary of clinical data reported in trials

	Hendrikx 1999	Salem 2009	Sarica 2017	Zhang 2011	Kumar 2015A	Verze 2010	Pearle 2001
Population	Extended-mid or distal ureteral stone can be either more or less than 5mm, not successfully treated conservatively (2 weeks in same position). Most stones were 6-10mm but there were some more than 11mm.	Solitary unilateral radiopaque calculi 5-20mm (split into a <10mm and >10mm groups). Proximal ureteric stones. Mean stone size 6.2-6.8mm.	People with acute colic pain due to a single obstructing opaque upper ureteral stone 5-10mm. Emergency treatment within 24 hours after the onset of pain was performed.	Ureteral calculi that failed to pass spontaneously for 4 weeks with or without MET. Mean stone size 8.7mm.	People with a single upper ureteral radioopaque calculus <2cm who were planned for either SWL or URS. Split into <1cm and 102cm.	Patients with solitary unilateral radioopaque distal ureteric stone with a size of 0.5-1.5cm. split into <1cm and >1cm.	Patients with distal ureteral calculus below the bony pelvis 15mm or less.
Intervention details	**SWL**: Mobile lithotriptor. HM4 Dornier machine Patients who had to wait > 2 weeks for SWL had URS. **URS:** Ureterorenoscopy was performed in combination with a pulsed dye laser or electrohydraulic lithotripsy. Semirigid ureterenoscopes of 7.0F to 9.5F were used.	**SWL:** in situ without stenting. Dornier HM3. Max shock waves 3000. **URS:** 8.5-11F semirigid URS. Intracorporeal lithotripsy (Swiss Lithoclast EMS) was used to fragment the stones which were then extracted by forceps.	**SWL:** electromagnetic lithotripter (compact Sigma). **URS:** semi-rigid ureteroscopy was performed with 8Fr ureteroscope.	**SWL:** done in situ using Dornier compact S lithotripter. Average of 2,900 shock waves. **URS:** 8.5-9.5F semirigid ureteroscope in combination with holmium YAG laser intracorporeal lithotripsy.	**SWL:** using Dornier compact delta. 300 shock waves max. **URS:** performed with 6/7.5F semirigid ureteroscope. A holmium laser was used for intracorporeal lithotripsy.	**SWL:** using Modulith SLX-X electromagnetic lithotripter. **URS:** using a Storz semirigid ureteroscope of 7.5-9.5 F diameter. Stones were fragmented with the Swiss Lithoclast Master lithotripter and/or extracted via forceps.	**SWL:** HM3 lithotriptor with patient prone on a modified stryker frame. Up to a total of 2,400 shock waves were used. **URS**: 6.9F semirigid ureteroscope and in some an 11.5F rigid uretersocope. Some stones extracted whole

	Hendrikx 1999	Salem 2009	Sarica 2017	Zhang 2011	Kumar 2015A	Verze 2010	Pearle 2001
							and some fragmented.
Definition of stone free	Stone free on a plain film. Stone disintegration defined as good if stone particles were less than 5mm.	When the patient is stone free without any residual fragments.	Not defined, but talks of completely stone free and people with residual fragments having further treatments. Stone free therefore implies no residual fragments.	Stone free defined as when patients had no residual fragments. Repeat or ancillary procedures performed when fragments were more than 5mm.	Stone free status defined as radiologic absence of stone, asymptomatic, with fragment <3mm.	Defined as absence of residual lithiasis at plain radiography.	Not defined
Reporting number stone free after initial procedure?	Yes	Yes	Yes	Yes	Unclear [(a)]	No	Unclear
Is everyone stone free at the end of the trial?	No (SWL: SF at 3 months = 93%, URS: SF at 3 months = 99%)	Yes	Yes	Yes	No (SWL: SF at 3 months = 85%, URS: SF at 3 months = 88%)	No (SWL: SF at 3 months = 96%, URS: SF at 3 months = 95%)	Yes
Length of follow up	3 months	2 weeks	4 weeks	2 weeks	3 months	3 months	3 months
Information on number of sessions for primary SWL	Seems to imply stone free outcome reported after 1 SWL session from the use of the word 'initial'.	Seems to imply stone free outcome reported after 1 SWL session from the use of the word 'initial'.	Says that SWL was a 'single session' in the discussion.	Seems to imply stone free outcome reported after 1 SWL session from the use of the word 'initial'.	Unclear how many sessions on average. States a max of 4 sessions of were given. 'retreatment SWL given for incomplete clearance' implies this is the criteria for retreatment.	1.58 average number of sessions to clear the stone. Reported for the whole study not just the small stone group; "in 70 cases (55.11%) one SWL session was enough to clear stone, in 31.49% it was 2, and in 13.38% it was 3"	No mention of number of sessions.
Reporting retreatments?	Yes	Yes	Yes	Yes	Yes	Yes	No

	Hendrikx 1999	Salem 2009	Sarica 2017	Zhang 2011	Kumar 2015A	Verze 2010	Pearle 2001
Reporting ancillary procedures?	Yes	Yes	Yes	Yes	Yes	No	No
Reporting outcome as number of people or procedures?	Procedures	People	People	People	Unclear	Unclear	People
Protocol for further treatment after primary procedures (i.e. retreatments and ancillary procedures)	No special criteria. SWL: most people failing SWL had ancillaries. URS: most people failing primary URS got ancillary procedures except for 1.	No criteria stated. SWL: mostly retreatments	No criteria stated. SWL: those in SWL arm that failed were treated with a URS (once). URS: those in URS arm that failed were treated with URS again (once).	SWL: repeat SWL or ancillary of URS were performed when residual fragments remained. URS: URS patients had auxiliary SWL when there were residual fragments.	SWL: Retreatment of SWL given for incomplete clearance. URS and PCNL were auxiliary procedures after failed SWL. URS: not specified	SWL: If significant fragments remained after first SWL treatment, patients were instructed to return for a further session. The need for URS or ureterolithotomy in SWL group were recorded as treatment failures. URS: The need for SWL or ureterolithotomy in URS group were recorded as treatment failures.	NA

(a) This is labelled as unclear because the study talks about how stone free after initial SWL was defined, so it's not clear if the stone free rate at 3 months that is reported is the initial one, or because 3 months is at the end of the trial then it could be the stone free rate after additional procedures. The retreatment and ancillary rates are reported as percentages, and for SWL are quite high, and seeing as the stone free probability is quite high also then it is probable that the stone free rate is following those further treatments.

A.2 Forest plots: ureteric analysis

Scenario 1:

Figure 5: Retreatment

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Weight	Risk Ratio M-H, Random, 95% CI
Hendrikx 1999	8	69	1	87	16.8%	10.09 [1.29, 78.72]
Kumar 2015A	25	53	3	49	24.0%	7.70 [2.48, 23.92]
Salem 2009	10	58	0	52	12.2%	18.86 [1.13, 314.18]
Sarica 2017	0	34	5	31	12.0%	0.08 [0.00, 1.44]
Verze 2010	8	69	3	66	22.8%	2.55 [0.71, 9.20]
Zhang 2011	20	257	0	269	12.2%	42.91 [2.61, 705.74]
Total (95% CI)		**540**		**554**	**100.0%**	**5.01 [1.39, 18.04]**
Total events	71		12			

Heterogeneity: Tau² = 1.44; Chi² = 13.10, df = 5 (P = 0.02); I² = 62%
Test for overall effect: Z = 2.47 (P = 0.01)

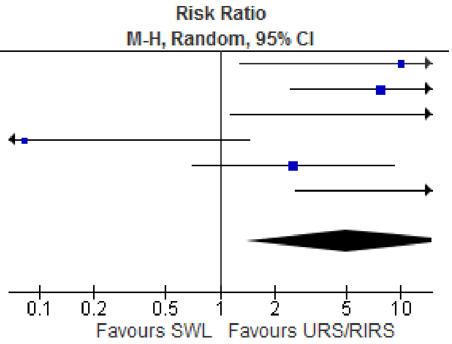


Figure 6: Ancillary procedures

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Weight	Risk Ratio M-H, Random, 95% CI
Hendrikx 1999	29	69	8	87	28.2%	4.57 [2.23, 9.35]
Kumar 2015A	9	53	4	49	24.8%	2.08 [0.68, 6.32]
Salem 2009	2	58	0	52	10.7%	4.49 [0.22, 91.45]
Sarica 2017	9	34	0	31	11.7%	17.37 [1.05, 286.56]
Verze 2010	0	69	0	66		Not estimable
Zhang 2011	4	257	11	269	24.6%	0.38 [0.12, 1.18]
Total (95% CI)		**540**		**554**	**100.0%**	**2.38 [0.71, 8.02]**
Total events	53		23			

Heterogeneity: Tau² = 1.23; Chi² = 15.43, df = 4 (P = 0.004); I² = 74%
Test for overall effect: Z = 1.40 (P = 0.16)

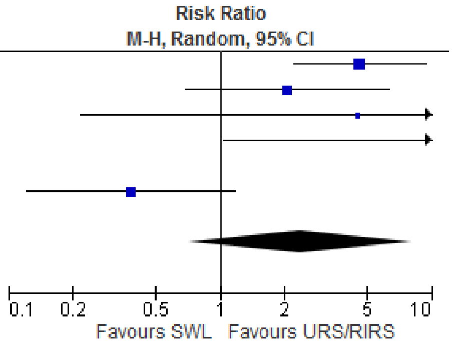


Figure 7: Readmission to hospital

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Risk Ratio M-H, Fixed, 95% CI
Pearle 2001	2	32	4	32	0.50 [0.10, 2.54]

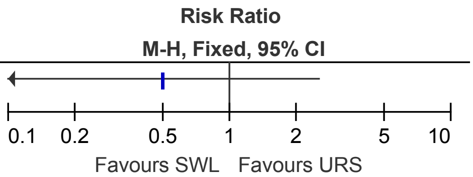
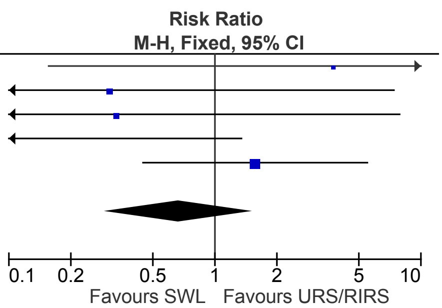


Figure 8: Minor adverse events

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Hendrikx 1999	1	69	0	87	3.2%	3.77 [0.16, 91.16]
Kumar 2015A	0	53	1	49	11.2%	0.31 [0.01, 7.40]
Pearle 2001	0	32	1	32	10.8%	0.33 [0.01, 7.89]
Salem 2009	0	100	6	100	46.7%	0.08 [0.00, 1.35]
Zhang 2011	6	257	4	269	28.1%	1.57 [0.45, 5.50]
Total (95% CI)		**511**		**537**	**100.0%**	**0.67 [0.29, 1.52]**
Total events	7		12			

Heterogeneity: Chi² = 5.52, df = 4 (P = 0.24); I² = 28%
Test for overall effect: Z = 0.96 (P = 0.33)

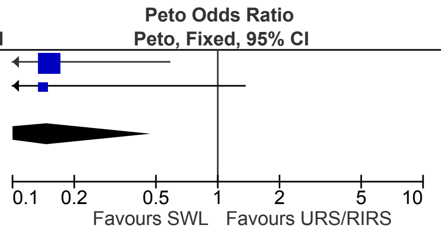


Note that for the Salem 2009 study; the total number of participants for adverse event outcomes is higher than for the stone free, retreatment, and ancillary outcome because the adverse events were not separated by stone size in the study, and as the majority of stones were <10mm, the outcomes has been included in the <10mm strata.

Figure 9: Major adverse events

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Hendrikx 1999	0	69	9	87	73.8%	0.15 [0.04, 0.58]
Zhang 2011	0	257	3	269	26.2%	0.14 [0.01, 1.36]
Total (95% CI)		**326**		**356**	**100.0%**	**0.15 [0.05, 0.47]**
Total events	0		12			

Heterogeneity: Chi² = 0.00, df = 1 (P = 0.96); I² = 0%
Test for overall effect: Z = 3.23 (P = 0.001)



Sensitivity analyses:

Figure 10: Retreatment and ancillary procedures pooled

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Hendrikx 1999	37	69	9	87	22.9%	5.18 [2.69, 9.99]
Kumar 2015A	34	53	7	49	20.9%	4.49 [2.20, 9.18]
Salem 2009	12	58	0	52	1.5%	22.46 [1.36, 370.14]
Sarica 2017	9	34	5	31	15.0%	1.64 [0.62, 4.37]
Verze 2010	8	69	3	66	8.8%	2.55 [0.71, 9.20]
Zhang 2011	24	257	11	269	30.9%	2.28 [1.14, 4.57]
Total (95% CI)		**540**		**554**	**100.0%**	**3.64 [2.58, 5.13]**
Total events	124		35			

Heterogeneity: Chi² = 7.64, df = 5 (P = 0.18); I² = 35%
Test for overall effect: Z = 7.37 (P < 0.00001)

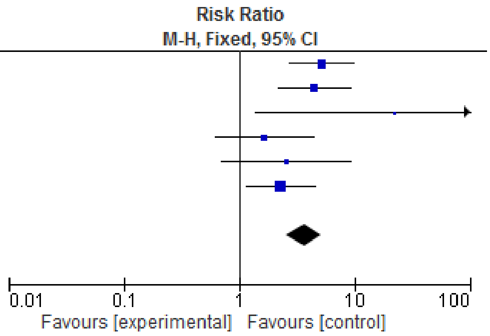


Figure 11: Readmission and major adverse events pooled

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Zhang 2011	0	257	3	269	21.6%	0.15 [0.01, 2.88]
Pearle 2001	2	32	4	32	25.3%	0.50 [0.10, 2.54]
Hendrikx 1999	0	69	9	87	53.1%	0.07 [0.00, 1.12]
Total (95% CI)		**358**		**388**	**100.0%**	**0.19 [0.06, 0.66]**
Total events	2		16			

Heterogeneity: Chi² = 1.89, df = 2 (P = 0.39); I² = 0%
Test for overall effect: Z = 2.61 (P = 0.009)

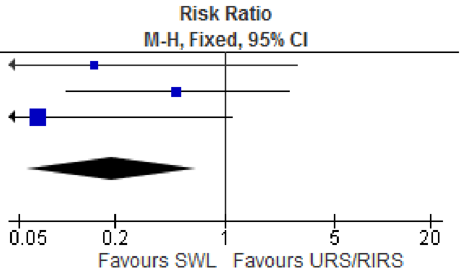
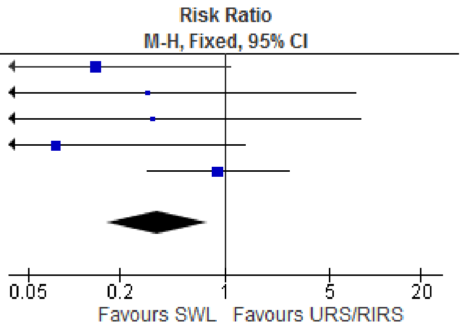


Figure 12: Pooled major and minor adverse events

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Hendrikx 1999	1	69	9	87	32.7%	0.14 [0.02, 1.08]
Kumar 2015A	0	53	1	49	6.4%	0.31 [0.01, 7.40]
Pearle 2001	0	32	1	32	6.2%	0.33 [0.01, 7.89]
Salem 2009	0	100	6	100	26.7%	0.08 [0.00, 1.35]
Zhang 2011	6	257	7	269	28.1%	0.90 [0.31, 2.63]
Total (95% CI)		**511**		**537**	**100.0%**	**0.36 [0.17, 0.78]**
Total events	7		24			

Heterogeneity: Chi² = 4.72, df = 4 (P = 0.32); I² = 15%
Test for overall effect: Z = 2.60 (P = 0.009)



Scenario 2:

Figure 13: Stone free probability

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Weight	Risk Ratio M-H, Random, 95% CI
Salem 2009	46	58	52	52	34.0%	0.80 [0.70, 0.91]
Sarica 2017	25	34	26	31	17.8%	0.88 [0.68, 1.13]
Zhang 2011	227	257	250	269	48.2%	0.95 [0.90, 1.00]
Total (95% CI)		**349**		**352**	**100.0%**	**0.88 [0.77, 1.01]**
Total events	298		328			

Heterogeneity: Tau² = 0.01; Chi² = 5.96, df = 2 (P = 0.05); I² = 66%
Test for overall effect: Z = 1.87 (P = 0.06)

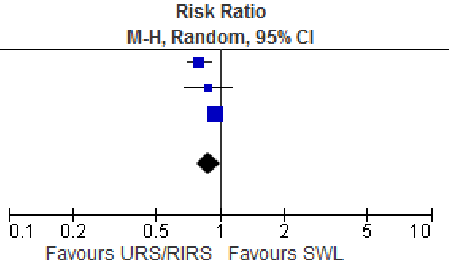


Figure 14: Retreatment probability

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Weight	Risk Ratio M-H, Random, 95% CI
Salem 2009	10	58	0	52	33.4%	18.86 [1.13, 314.18]
Sarica 2017	0	34	5	31	33.2%	0.08 [0.00, 1.44]
Zhang 2011	20	257	0	269	33.4%	42.91 [2.61, 705.74]
Total (95% CI)		**349**		**352**	**100.0%**	**4.10 [0.09, 187.47]**
Total events	30		5			

Heterogeneity: Tau² = 9.34; Chi² = 11.00, df = 2 (P = 0.004); I² = 82%
Test for overall effect: Z = 0.72 (P = 0.47)

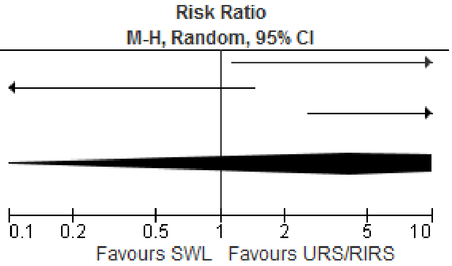
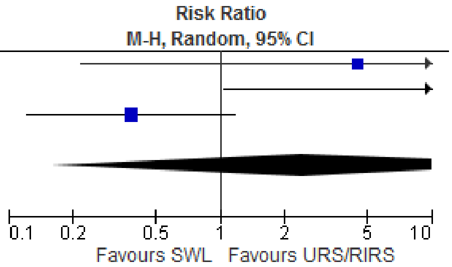


Figure 15: Ancillary procedures probability

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Weight	Risk Ratio M-H, Random, 95% CI
Salem 2009	2	58	0	52	28.7%	4.49 [0.22, 91.45]
Sarica 2017	9	34	0	31	30.1%	17.37 [1.05, 286.56]
Zhang 2011	4	257	11	269	41.2%	0.38 [0.12, 1.18]
Total (95% CI)		**349**		**352**	**100.0%**	**2.44 [0.16, 36.92]**
Total events	15		11			

Heterogeneity: Tau² = 4.33; Chi² = 8.51, df = 2 (P = 0.01); I² = 77%
Test for overall effect: Z = 0.64 (P = 0.52)



Sensitivity analyses:

Figure 16: retreatment and ancillaries pooled

Study or Subgroup	SWL Events	SWL Total	URS/RIRS Events	URS/RIRS Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Salem 2009	12	58	0	52	3.2%	22.46 [1.36, 370.14]
Sarica 2017	9	34	5	31	31.7%	1.64 [0.62, 4.37]
Zhang 2011	24	257	11	269	65.1%	2.28 [1.14, 4.57]
Total (95% CI)		**349**		**352**	**100.0%**	**2.72 [1.58, 4.69]**
Total events	45		16			

Heterogeneity: Chi² = 3.45, df = 2 (P = 0.18); I² = 42%
Test for overall effect: Z = 3.62 (P = 0.0003)