Mirabegron for the treatment of symptoms associated with overactive bladder

STA REPORT

This report was commissioned by the NIHR HTA Programme as project number 11/138

BMJ Technology Assessment Group

Title: Mirabegron for the treatment of symptoms associated with overactive bladder

Produced by: BMJ Technology Assessment Group (BMJ-TAG)

Authors: Steve Edwards, Head of BMJ-TAG, London

Charlotta Karner, Health Technology Assessment Analysis Lead, BMJ-TAG, London

Nicola Trevor, Health Economics Lead, BMJ-TAG, London

Samantha Barton, Health Technology Assessment Analyst, BMJ-TAG, London

Leo Nherera, Health Economist, BMJ-TAG, London

Correspondence to: Steve Edwards, Head of BMJ-TAG, BMJ-TAG, BMJ Evidence Centre, BMJ Group, BMA House, Tavistock Square, London, WC1H 9JP.

Date completed: 09/01/2013

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 11/138

Declared competing interests of the authors:
None.

Acknowledgements:
The ERG would like to thank Dr Maneesh Ghei (Consultant Urologist) for providing clinical advice throughout the project. Thanks also to Dr Christopher Chapple (Consultant Urological Surgeon) and Dr Rizwan Hamid (Consultant Urological Surgeon) for providing feedback on the clinical sections of the report.

Rider on responsibility for report
The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows: Edwards SJ, Karner C, Trevor N, Barton S, Nherera L. Mirabegron for the treatment of symptoms associated with overactive bladder. BMJ-TAG, London, 2013.

Contributions of authors:

Steve Edwards	Project lead: supervised the production of the final report; carried out exploratory indirect comparisons; critical appraisal of the manufacturer's submission; critical appraisal of the clinical evidence; and critical appraisal of the economic evidence
Charlotta Karner	Critical appraisal of the manufacturer's submission; critical appraisal of the clinical evidence; cross checking of manufacturer's search strategies; and drafted the summary and clinical results sections

Nicola Trevor	Critical appraisal of the manufacturer's submission; critical appraisal of the economic model; critical appraisal of the economic evidence; and drafted the economic analysis sections
Samantha Barton	Critical appraisal of the manufacturer's submission; and critical appraisal of the clinical evidence
Leo Nherera	Critical appraisal of the manufacturer's submission; and critical appraisal of the economic evidence

All authors read and commented on draft versions of the ERG report

Table of contents

Abbreviations

BNF	British National Formulary
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CFB	Change from baseline
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence Interval
CSR	Clinical Study Report
DIC	Deviance information criterion
EQ-5D	EuroQol 5 dimensions questionnaire
EMA	European Medicines Agency
ER	Extended release
ERG	Evidence Review Group
FAS	Full analysis set
GP	General Practitioner
HR	Hazard ratio
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
IR	Immediate release
ITT	Intention-to-treat
IVRS	Interactive voice response system
kg	Kilogram
LYG	Life-years gained
MCMC	Markov Chain Monte Carlo
MD	Mean difference
mg	Milligram
mL	Millilitre
mm	Millimetre
MTC	Mixed treatment comparison
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
OAB	Overactive bladder
OR	Odds ratio
PPIUS	Patient Perception of Intensity of Urgency Scale

PSA	Probabilistic sensitivity analysis
QALY	Quality adjusted life year
QOF	Quality and Outcomes Framework
RCT	Randomised controlled trial
RR	Relative risk
SD	Standard deviation
SE	Standard error
SmPC	Summary of Product Characteristics
STA	Single Technology Appraisal
TEAE	Treatment-emergent adverse event
UK	United Kingdom
USA	United States of America
VAS	Visual analogue scale
vs	Versus

1 SUMMARY

1.1 Critique of the decision problem in the manufacturer's submission

The manufacturer of mirabegron (Myrbetriq®/Betanis®; Astellas) submitted to the National Institute for Health and Clinical Excellence (NICE) clinical and economic evidence in support of the effectiveness of mirabegron in the treatment of overactive bladder (OAB).

At the time of writing of the Evidence Review Group's (ERG) report, mirabegron does not have a European licence for use in OAB. However, in October 2012, the Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion on the use of mirabegron at doses of 50 mg (recommended dose) and 25 mg (for patients with renal or hepatic failure) for the symptomatic treatment of urgency, increased micturition frequency and/or urgency incontinence as may occur in adult patients with OAB.

The direct clinical evidence described in the MS is derived from three randomised controlled trials (RCTs): ARIES; CAPRICORN; and SCORPIO. All three RCTs evaluated the clinical effectiveness of mirabegron at doses of 25, 50 or 100 mg versus placebo. Based on the doses indicated in the positive opinion issued by the CHMP, the ERG considers that data on mirabegron 100 mg are unlikely to be relevant to the decision problem that is the focus of this Single Technology Appraisal (STA), and chose not to present these data. The manufacturer presented data for all outcomes listed in the final scope, and for additional clinical outcomes such as number of incontinence episodes (any involuntary leakage of urine, with or without accompanying or preceding urgency), and volume voided during micturition. The ERG chose to report data for only the outcomes listed in the scope and the additional clinical outcome of frequency of incontinence episodes as this outcome was a key driver in the manufacturer's economic model.

The direct evidence submitted predominantly compares the effects of mirabegron versus placebo, which, in the context of the comparisons of interest, the ERG considers does not fully address the decision problem. The scope issued by NICE lists comparators of interest as:

- oxybutynin (including modified-release preparations);
- tolterodine;
- fesoterodine;
- solifenacin;
- trospium.

The SCORPIO trial included an active control of tolterodine, but the manufacturer did not present relative clinical effectiveness data for the comparison of mirabegron versus tolterodine, indicating that

SCORPIO was not powered to evaluate the superiority or non-inferiority of mirabegron versus tolterodine.

The manufacturer carried out a systematic review of the literature to identify studies that could potentially inform a mixed treatment comparison (MTC). In addition to the three trials submitted as direct clinical evidence, the manufacturer identified publications on 40 trials in OAB evaluating interventions listed as comparators of interest in the scope and that were used to construct networks to evaluate the comparative clinical effectiveness of mirabegron. The ERG notes that the manufacturer excluded studies evaluating non-oral preparations of interventions, which the ERG considers to be a deviation from the final scope, which specified modified-release formulations of oxybutynin (available as a transdermal patch) as a comparator of interest.

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer

ARIES, CAPRICORN, and SCORPIO were multiple arm trials in patients with symptoms of OAB. SCORPIO and ARIES evaluated the clinical effectiveness of mirabegron at doses of 50 and 100 mg versus placebo, and CAPRICORN evaluated mirabegron at doses of 50 and 25 mg, again against placebo. The primary outcomes in the three trials were change from baseline (CFB) in frequency of micturition and in frequency of incontinence episodes. SCORPIO included an additional treatment group of an active control (tolterodine), the comparative results of which the manufacturer chose not to report as SCORPIO was not powered to evaluate the superiority or non-inferiority of mirabegron versus tolterodine. The ERG acknowledges the manufacturer's point but considers that exclusion of data from the tolterodine treatment group in SCORPIO resulted in a lack of direct evidence relevant to the decision problem.

For SCORPIO, ARIES, and CAPRICORN the manufacturer presents data on outcomes assessed based on CFB for the individual treatment groups within each trial, and the difference between mirabegron (50 mg and 25 mg) and placebo at the end point. Considering the three trials submitted as direct evidence (ARIES, CAPRICORN, and SCORPIO), mirabegron 50 mg was found to be more effective than placebo at reducing all clinical outcomes evaluated, with most differences reaching statistical significance: urinary frequency per 24 hours; frequency of incontinence per 24 hours; frequency of urgency urinary incontinence per 24 hours; level of urgency; number of urgency episodes per 24 hours; and nocturia. However, results from the tolterodine active control group from SCORPIO suggest that mirabegron 50 mg is of similar clinical effectiveness to tolterodine, with no statistically significant differences noted between the two active treatments for the outcomes reported.

The manufacturer did not perform a meta-analysis of the included trials, but presented the results of a pre-specified pooled analysis of ARIES, CAPRICORN, and SCORPIO for the comparison of

mirabegron versus placebo. The results of the manufacturer's analysis support the results from the individual RCTs, with mirabegron 50 mg being statistically significantly more effective than placebo at improving all clinical outcomes evaluated. However, as noted earlier, the ERG considers that comparison versus placebo is not as relevant to the decision problem as comparison against interventions specified in the final scope issued by NICE. The ERG carried out an independent meta-analysis of the available data (additional data provided by the manufacturer during clarification) for the comparison of mirabegron versus tolterodine, focusing on the primary clinical outcomes of frequency of urination and of incontinence episodes.

The manufacturer also synthesised data from ARIES, CAPRICORN, and SCORPIO for the subgroups of male versus female, and previously treated for OAB versus treatment naïve, which were specified as subgroups of interest in the final scope. The manufacturer reported data for primary outcomes evaluated in the three key trials submitted as direct clinical evidence, that is, frequency of micturition per 24 hours and of incontinence episodes per 24 hours, for the comparison of mirabegron versus placebo. Analyses of comparative clinical effectiveness of mirabegron versus tolterodine in the specified subgroups were not reported.

In terms of indirect evidence, the manufacturer identified 40 trials that were used to carry out an MTC evaluating oral treatments for OAB. The ERG notes that the seven trials evaluating mirabegron, including four trials excluded by the manufacturer from the direct clinical evidence, were included in the trials deemed eligible for inclusion in the MTC. The ERG had concerns around the manufacturer's choice of studies to include in the MTC based on potential clinical and methodological heterogeneity, the inconsistency identified in one or more treatment comparisons for multiple outcomes, and the number of iterations used for sampling the posterior distributions (which may be an indicator of poor mixing of data within the model). Based on these concerns, the ERG considers that the results of the manufacturer's MTC should be interpreted with caution. For the outcome of urinary frequency (micturition) per 24 hours, most differences between mirabegron 50 mg and other active treatments were not statistically significant, with only solifenacin 10 mg found to significantly reduce the number of micturition episodes per 24 hours compared with mirabegron 50 mg (MD –0.583; 95% Credible Interval [CrI]: –0.832 to –0.333). Mirabegron 50 mg was also found to be clinically less effective than solifenacin 10 mg at reducing the frequency of urgency urinary incontinence per 24 hours (MD –0.420; 95% CrI: –0.786 to –0.056). No statistically significant difference was found between mirabegron 50 mg and any other active treatment for the outcome of urgency urinary incontinence episodes. Mirabegron 50 mg was associated with a significantly lower risk (evaluated as an odds ratio [OR], where OR >1 favours mirabegron) of the adverse effect of dry mouth compared with all other antimuscarinics evaluated. The risk of constipation was significantly lower with mirabegron 50 mg compared with solifenacin 5 (OR 2.50; 95% CI: 1.41 to 4.13) and 10 mg (OR 4.37; 95% CI: 2.54 to

7.07), fesoterodine 8 mg (OR 1.93; 95% CI: 1.14 to 3.06), and trospium 60 mg (OR 7.60; 95% CI: 2.08 to 22.59). However, for most comparisons, there was no statistically significant difference between mirabegron 50 mg and other active treatments in the risk of experiencing constipationTable 30. No statistically significant differences in the risk of developing blurred vision were found between mirabegron 50 mg and other active treatments.

1.3 Summary of cost-effectiveness submitted evidence by the manufacturer

Within the published literature, the manufacturer identified 16 economic evaluations considering currently available pharmaceutical interventions for OAB. However, none of the economic evaluations identified by the manufacturer served to answer the decision problem regarding the cost-effectiveness of mirabegron in a UK OAB population. Therefore, the manufacturer constructed a *de novo* economic evaluation to assess the incremental cost-effectiveness of:

- mirabegron 50 mg versus tolterodine ER 4 mg (primary base case analysis);
- mirabegron 50 mg versus tolterodine ER 4 mg, solifenacin 5 mg and 10 mg, fesoterodine 4 mg, trospium chloride modified-release (MR) 60 mg and oxybutynin ER and IR 10 mg (secondary base case analyses).

The manufacturer's economic evaluation was carried out within a Markov cohort model which assessed the therapeutic management of patients (including complications), and the severity and progression of disease. The time horizon of the model was 5 years, in which a hypothetical cohort of OAB patients simultaneously transitioned, in monthly cycles, through severity levels representing frequency of micturition and frequency of incontinence (in line with the co-primary outcome measures of the pivotal mirabegron trials [SCORPIO, ARIES and CAPRICORN]) experienced by the patient. In addition to disease progression, captured by transitions between levels of symptom severity, the manufacturer's model assessed the therapeutic management of patients; including treatment discontinuation; treatment switch and the management of adverse events (AEs). Treatment discontinuation was assumed to be a result of either AEs or other causes (e.g. lack of efficacy). The AEs considered in the manufacturer's model were limited to dry mouth and constipation and 90% of patients experiencing an adverse event were assumed to discontinue. Discontinuation as a result of other causes was estimated based on real world persistence data for each treatment; however, as no real world persistence data is available for mirabegron, the manufacturer assumed that the persistence rate of mirabegron would be equal to the persistence rate of the treatment mirabegron was compared with (i.e. the other-cause discontinuation rate of mirabegron varied depending on the comparison made).

The manufacturer's primary base case analysis compared mirabegron 50 mg with tolterodine ER 4 mg based on clinical effectiveness data from SCORPIO. The manufacturer carried out a mixed treatment comparison "to estimate the relative efficacy and safety of mirabegron compared with all treatments

of interest"; data from the manufacturer's MTC was used to inform the secondary base case economic evaluation of mirabegron versus all comparators of interest.

The manufacturer's primary base case comparison of mirabegron 50 mg versus tolterodine ER 4 mg (based on clinical effectiveness data from SCORPIO) resulted in an estimated deterministic ICER of £4,386. Results of the secondary base case comparisons (based on clinical effectiveness data from the manufacturer's MTC) were presented within the MS individually (mirabegron versus solifenacin 10 mg [£340], fesoterodine 4 mg [£3,607], tolterodine ER 4 mg [3,715], oxybutynin ER 10 mg [3,878], trospium chloride MR 60 mg [8,881], solifenacin 5 mg [12,493], and oxybutynin IR 10 mg [14,234]) and incrementally. The manufacturer's incremental results were generated by assuming that mirabegron 50 mg was associated with an other cause discontinuation rate equal to that of patients treated with solifenacin (5 or 10 mg) and indicated that treatment with:

- oxybutynin ER 10 mg is strictly dominated (more costly and less effective) by treatment with trospium chloride MR 60 mg;
- fesoterodine 4 mg is strictly dominated by treatment with solifenacin 5 mg;
- tolterodine ER 4 mg and solifenacin 10 mg are extendedly dominated (less effective yet with a higher ICER) by treatment with mirabegron 50 mg.

1.4 ERG commentary on the robustness of evidence submitted by the manufacturer

Following detailed examination of the MS and the manufacturer's primary and secondary base case models, the ERG identified several areas of inaccuracy or uncertainty. Where possible, the ERG carried out sensitivity analyses to investigate the impact of alternative assumptions or parameters on the manufacturer's base case cost-effectiveness results. The sensitivity analyses carried out by the ERG indicate that the manufacturer's primary base case cost-effectiveness result was generally robust; cumulative impact of ERG sensitivity analyses increased the ICER by £886 (from £4,386 to £5,272). However, the manufacturer's incremental secondary base case cost-effectiveness results were substantially altered by application of the ERG's sensitivity analyses. In particular, following simultaneous application of the ERG's sensitivity analyses, tolterodine ER 4 mg and solifenacin 10 mg move from being extendedly dominated by mirabegron 50 mg to being strictly dominated by solifenacin 5 mg. Moreover, the ICER of mirabegron 50 mg versus solifenacin 5 mg increases from £12,493 to £32,712. Furthermore, the ERG was unable to quantify the impact of using alternative assumptions or parameters in all areas of uncertainty identified in the ERG's critique. Particularly, the use of clinical effectiveness estimates from the ERG's revised MTC (using a more homogeneous data set than that used to inform the manufacturer's MTC). However, the ERG notes that estimates obtained from the ERG's MTC indicate that solifenacin 5 mg is statistically significantly more effective at reducing incontinence episodes than mirabegron 50 mg (mean difference –0.386; 95%

CrI: –0.717 to –0.055). By contrast, estimates obtained from the manufacturer's MTC detected no statistically significant difference between solifenacin 5 mg and mirabegron 50 mg in reducing the number of incontinence episodes experienced per 24 hours (mean difference –0.237; 95% CrI –0.482 to 0.007). Based on this, the ERG considers that the ERG's revised ICER for the comparison of mirabegron 50 mg with solifenacin 5 mg (£32,712) is likely to be conservative; i.e. an ICER estimated using ERG MTC data is likely to be higher than £32,712.

1.4.1 Strengths

Clinical

The ERG considers the ARIES, CAPRICORN, and SCORPIO trials to be well-designed trials, and considers that the results on the effectiveness of mirabegron 50 mg and 25 mg are consistent across the trials.

Economic

The manufacturer submitted well constructed, transparent and accurate economic models. The ERG considers the manufacturer's primary base case to be robust, with respect to parameter uncertainty. In addition, the ERG considers the use of calibration techniques to incorporate MTC data to improve the accuracy of estimates from the economic model.

1.4.2 Weaknesses and areas of uncertainty

Clinical

In addition to ARIES, CAPRICORN, and SCOPRIO, the manufacturer identified four other trials evaluating the use of mirabegron in the treatment of OAB, which were described as supporting evidence and did not form part of the submitted evidence (TAURUS, DRAGON, 178-CL-045, and 178-CL-048). The manufacturer cited various reasons for exclusion of these trials from the analysis, and indicated that these trials were included only as supporting evidence. On reviewing the four RCTs, the ERG considers the RCTs to be relevant to the decision problem. Of the four additional trials, TAURUS was designed as a long-term follow-up safety study, evaluating mirabegron over a period of 12 months, compared with treatment duration of 3 months in the other trials. The remaining three RCTs evaluated mirabegron at various doses (25 mg, 50 mg, 100 mg, and 200 mg), and two of the trials (DRAGON and 178-CL-048) included a tolterodine group as an active control. The ERG considers that DRAGON, 178-CL-045, and 178-CL-048 were sufficiently similar to ARIES, CAPRICORN, and SCOPRIO to warrant inclusion in the submission and to synthesise data for mirabegron 50 mg (recommended dose).

SCORPIO, ARIES, and CAPRICORN were powered to evaluate the effectiveness of mirabegron compared with placebo. The ERG considers that, although the results of effectiveness of mirabegron

versus placebo are key as it is important to demonstrate that a treatment is more effective than placebo, analysis of mirabegron versus active treatment(s) currently used in the NHS is of more relevance to the decision problem. Thus, the ERG considers that inclusion of comparative results from the tolterodine group of SCORPIO, together with those from DRAGON and 178-CL-048, would have been appropriate and informative.

The ERG considers that limitation of the MTC to oral formulations could potentially have led to exclusion of studies relevant to the decision problem. Due to time constraints, the ERG was unable to carry out an independent systematic review of the literature or to validate the 40 studies identified by the manufacturer as relevant to the MTC. The ERG has concerns that there is considerable heterogeneity (in terms of population and quality of trials) in the MTC submitted by manufacturer and, as such, results should be interpreted with caution.

Economic

The ERG considers the manufacturer's approach to implementing discontinuation into the economic model to be a weakness; the approach taken hindered comparison of modelled results with real-world data and assumed a variable rate for mirabegron.

In addition, the ERG considers that the uncertainty inherent in an MTC exhibiting large amounts of heterogeneity is propagated through the manufacturer's economic model.

1.5 Summary of exploratory and sensitivity analyses undertaken by the ERG

Clinical

In terms of direct evidence, the ERG's meta-analysis found that, compared with tolterodine ER 4 mg, treatment with mirabegron 50 mg led to significantly fewer micturitions per 24 hours (MD –0.27; 95% CI: –0.48 to –0.06; p-value = 0.01), and significantly fewer incontinence episodes per 24 hours (MD –0.21; 95% CI –0.41 to –0.01; p-value = 0.04).

With the goal of producing a more homogeneous dataset for the MTC, trials determined by the manufacturer to be of poor methodological quality and RCTs that included a population other than OAB were excluded from the dataset used to inform the ERG's MTC. The outcomes considered in the ERG's MTC were micturition, incontinence, constipation, dry mouth, and additionally all-cause discontinuation (as the outcome of all-cause discontinuation was believed to be a potential key driver in the economic model and was not included in the manufacturer's MTC). The results of the ERG's MTC were predominantly in agreement with those of the manufacturer's MTC. For the outcome of micturition per 24 hours, whereas the manufacturer's MTC identified that solifenacin 10 mg significantly reduced micturition per 24 hours compared with mirabegron 50 mg, the ERG's MTC

found no statistically significant difference between mirabegron 50 mg and any active treatment evaluated. In terms of reduction of incontinence episodes, the results of the ERG's MTC for the comparison of solifenacin (5 and 10 mg) versus mirabegron differed from the manufacturer's MTC, with the ERG's MTC identifying a significant difference between the solifenacin at both doses and mirabegron 50 mg favouring solifenacin in each analysis; the manufacturer's MTC identified no statistically significant difference between the treatments for this outcome. For the adverse effects of dry mouth and constipation, which are associated with treatment with anticholinergics, mirabegron 50 mg was found to significantly lower the risk of dry mouth compared with all other antimuscarinics assessed, and to significantly lower the risk of constipation compared with fesoterodine 8 mg, solifenacin (5 mg and 10 mg), and trospium 60 mg. Results of the ERG's MTC for dry mouth and constipation are analogous to the results from the manufacturer's MTC.

Economic

The sensitivity analyses carried out by the ERG indicated that the manufacturer's primary base case cost-effectiveness result was generally robust with respect to the areas of uncertainty identified in the ERG's critique. However, following individual and cumulative application of the ERG's sensitivity analyses, the ICERs estimated by the manufacturer for the individual comparisons considered in the secondary base case remained relatively consistent; with the exception of comparisons between mirabegron 50 mg and solifenacin 5 and 10 mg. The impact of each sensitivity analysis on the ICERs for mirabegron 50 mg versus solifenacin 5 and 10 mg was highly variable, with ICERs ranging from £573 to the dominance of solifenacin in the comparison of mirabegron 50 mg versus solifenacin 10 mg. Similarly, ICERs ranged from £11,778 to £32,572 upon application of the ERG's sensitivity analyses to the comparison of mirabegron 50 mg versus solifenacin 5 mg. Moreover, the cumulative impact of the ERG's sensitivity analyses on the comparison of mirabegron 50 mg with solifenacin 5 and 10 mg, resulted in ICER increases of £20,218 (from £12,493 to £32,711.50) and £1,573 (from £340 to £1,913), respectively. Furthermore, following simultaneous application of the ERG's sensitivity analyses, tolterodine ER 4 mg and solifenacin 10 mg move from being extendedly dominated by mirabegron 50 mg to being strictly dominated by solifenacin 5 mg. Moreover, the ICER of mirabegron 50 mg versus solifenacin 5 mg increases from £12,493 to £32,712.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problems

In the manufacturer's submission (MS), the manufacturer provides an overview of overactive bladder syndrome (OAB), including its pathophysiology (Box 1) and prevalence (Box 2), together with the clinical consequences of OAB (Box 3). The information presented in boxes is taken directly from the MS unless otherwise stated and the references have been renumbered.

OAB is a common condition characterised by specific symptoms that usually include:

- urgency – an urgent desire to pass urine and being unable to put off going to the toilet;
- frequency – going to the toilet more than eight times a day;
- nocturia – waking to go to the toilet more than once at night;

With or without:

- urgency incontinence – when having a feeling of urgency, leaking urine before being able to get to the toilet.

The storage and expulsion of urine by the bladder is controlled by the bladder wall muscle, which is the detrusor muscle.[(1)] The relaxation and contraction of the detrusor muscle determines bladder function during filling and micturition (expulsion of urine), respectively. Several neural pathways and neurotransmitters are involved in the regulation of urine storage and subsequent emptying of the bladder, and the disruption of these signalling pathways can lead to bladder dysfunction. OAB is thought to be the result of overactivity of the detrusor muscle, causing the bladder to contract suddenly, even though the bladder may not be full.[(1)]

Box 1. Pathophysiology of overactive bladder

Overactive bladder (OAB) syndrome (or urge syndrome or urgency-frequency syndrome) has been described by the International Continence Society (ICS) as urgency, with or without urge incontinence, usually with frequency and nocturia.[(2)]

The exact aetiology of OAB is unknown, but may have a neurological[(3)] or myogenic[(4)] basis. Damage to central inhibitory pathways, leads to sensitisation of afferent nerves and increased afferent activity. Furthermore, decreased inhibitory control and increased sensitivity of the detrusor muscle can occur with triggers involuntary overactive detrusor contractions. Alterations in the functional properties of detrusor myocytes, including hypersensitivity can result in excessive spontaneous excitation and propagation.

However, symptoms such as urinary frequency, urinary incontinence and nocturia may be indications of other conditions, such as urinary tract infection, polyuria or other types of incontinence.[(5)] It is also important to differentiate urgency incontinence, which may be a symptom of OAB, from stress

urinary incontinence in which the pelvic floor muscles are too weak to prevent urination, causing urine to leak when the bladder is under pressure, for example, when coughing, laughing or sneezing.[2] However, urge and stress incontinence can occur concurrently, and this condition is referred to as mixed incontinence. An important part of the diagnosis of OAB is therefore linked to the absence of other pathological or metabolic conditions.[2]

Box 2. Prevalence of overactive bladder

The prevalence of OAB in the UK has been estimated at approximately 5 million people aged 40 years and older,[6] with prevalence increasing with advancing age.[7;8] The prevalence of OAB is similar between men and women.[6-8] A community-based survey of 2,063 adult men and women aged 40 years or older in the UK (as part of a larger European study) revealed that 19% had symptoms of OAB.[6]
Abbreviation used in box: OAB, overactive bladder.

In addition to age, several other factors are associated with risk of developing OAB. People with insulin-dependent diabetes, depression, arthritis, or increased body mass index, and those taking oral hormone-replacement therapy, have a higher risk of developing OAB.[5] Although the prevalence of OAB is similar between men and women, women are more likely to have OAB with incontinence, and men more likely to have OAB without incontinence.[7]

Box 3. Clinical consequences of OAB

Within the full European dataset, frequency was the most commonly reported symptom (85% of patients), followed by urgency (54%) and urgency incontinence (36%). Overall, 65% of men and 67% of women with OAB indicated that their symptoms had an effect on their daily lives and 60% had consulted a medical practitioner about their symptoms, although only 27% of patients were currently receiving treatment. In a European case-control study, participants (19.3% from the UK) with OAB with or without additional lower urinary tract symptoms (LUTS) reported significantly less work productivity and sexual satisfaction, higher rates of depressive symptoms and erectile dysfunction, and lower levels of overall health.[9] In the US, patients with OAB and nocturia have reported significantly higher symptom bother and decreased health-related quality of life (HRQoL) due to disrupted sleep patterns.[10] The negative impact on HRQoL increases with the number of night-time voids. OAB is associated with a variety of co-morbidities. In a US retrospective claims database analysis, the prevalence of falls and fractures (25.3% vs 16.1%), depression (10.5% vs 4.9%), urinary tract infections (UTIs) (28.0% vs 8.4%) and skin infections (3.9% vs 2.3%) was significantly higher ($p<0.0001$) for patients with OAB than for controls.[11]
Abbreviation used in box: OAB, overactive bladder.

Based on expert clinical advice, the Evidence Review Group (ERG) considers the manufacturer's overview of the underlying health problem to be accurate.

2.2 Critique of manufacturer's overview of current service provision

The manufacturer's overview of the current service provision included an overview of the treatment options for OAB (Box 4), signalling pathways and mechanism of action of mirabegron and antimuscarinic pharmacotherapy (Box 5), the National Institute for Health and Clinical Excellence (NICE) guidelines for the management of urinary incontinence in women (CG40)[12] and the management of lower urinary tract symptoms in men (CG97),[13] both of which cover the management of OAB (Box 6). The manufacturer also outlined the proposed position of mirabegron in the current treatment pathway for OAB (Box 7), and estimated the number of patients in England and Wales who would be eligible for treatment with mirabegron.

Box 4. Manufacturer's overview of treatment options for overactive bladder

Treatment options for OAB include conservative management (e.g. bladder training and electrical stimulation), pharmacotherapy and surgical intervention. In the UK, bladder training and lifestyle advice is recommended for OAB in both men and women, followed by pharmacotherapy.[12;13] The primary pharmacotherapy option is currently muscarinic receptor treatments (antimuscarinics), although the market share of the drugs prescribed has changed over recent years. In the year to March 2009 tolterodine, oxybutynin and solifenacin were the most commonly prescribed and dispensed antimuscarinics in England, accounting for 39%, 34%, 20% of the market, respectively.[14] More recent data from July 2012 has shown that whilst these drugs are still the most commonly prescribed, the share has changed with solifenacin being the most common followed by oxybutynin and tolterodine with shares of 36%, 29% and 22%, respectively, of prescriptions issued in the UK.[15]
Abbreviation used in box: OAB, overactive bladder.

The ERG considers the manufacturer's overview of the treatment options for OAB to be accurate, but considers that expansion of some points may be informative.

Conservative management includes lifestyle advice, such as caffeine reduction, modification of fluid intake and weight loss interventions.[12;13] It also covers bladder training, which, according to NICE clinical guidelines, should be offered as first-line treatment to patients with OAB. If bladder training and lifestyle advice have been ineffective, patients should be offered pharmacotherapy. As the manufacturer points out, pharmacotherapy treatments for OAB are not curative and therefore patients have to take the treatments continuously and in the long-term. Surgical procedures may be a relevant option for patients with severe symptoms, for whom conservative management and drug therapy have been unsuccessful. Surgical procedures include sacral nerve stimulation, cystoplasty, urinary diversion, and bladder wall injection with botulinum toxin A.[12;13]

Box 5. Manufacturer's overview of the signalling pathways and mechanism of action of antimuscarinic pharmacotherapy and of mirabegron

During the urine storage phase, when urine accumulates in the bladder, sympathetic nerve stimulation predominates. Noradrenaline is released from nerve terminals, leading predominantly to beta adrenoceptor activation in the bladder musculature, and hence bladder smooth muscle relaxation. During the urine voiding phase, the bladder is predominantly under parasympathetic nervous system control. Acetylcholine, released from pelvic nerve terminals, stimulates cholinergic M2 and M3 receptors, inducing bladder contraction. The activation of the M2 pathway also inhibits beta 3-adrenoceptor induced increases in cAMP. Therefore beta 3-adrenoceptor stimulation should not interfere with the voiding process.

Antimuscarinics block the muscarinic receptors in the bladder wall and therefore inhibit abnormal detrusor contractions in the bladder. The effects of these agents are not selective for the bladder but also affect the salivary gland, intestine and eye, resulting in unwanted side-effects such as dry mouth, blurred vision and constipation.[16;17]

Mirabegron is a potent and selective beta 3-adrenoceptor agonist ... that ... enhances urine storage function by stimulating beta 3-adrenoceptors in the bladder.

Abbreviations used in box: cAMP, cyclic adenosine monophosphate; M_2, muscarinic receptor 2; M_3, muscarinic receptor 3.

Box 6. Manufacturer's overview of relevant NICE clinical guidelines (CG40 and CG97)

NICE has issued separate guidelines including the management of OAB for men and women:
NICE Clinical Guideline Number 40, October 2006 'Urinary incontinence: The management of urinary incontinence in women'.[12] The guidelines (currently under review) state that 'Immediate-release non-proprietary oxybutynin should be offered to women with OAB or mixed urinary incontinence as first-line drug treatment if bladder training has been ineffective. If immediate-release oxybutynin is not well tolerated, darifenacin, solifenacin, tolterodine, trospium or an extended-release or transdermal formulation of oxybutynin should be considered as alternatives.

NICE Clinical Guideline Number 97, May 2010 'The management of lower urinary tract symptoms in men'.[13] Anticholinergics should be offered as first-line pharmacotherapy to men with storage LUTS suggestive of OAB if bladder training, lifestyle and behavioural advice and containment devices have failed.

Abbreviations used in box: CG, clinical guideline; LUTS, lower urinary tract symptoms; NICE, National Institute for Health and Clinical Excellence; OAB, overactive bladder.

Box 7. Manufacturer's proposed position of mirabegron in the treatment pathway of overactive bladder

Current treatment regimens for OAB are limited because of a lack of well-tolerated non-surgical treatment options. Mirabegron is a first-in-class pharmacotherapy with a new mechanism of action resulting in a differing side-effect profile to the currently available antimuscarinics, particularly low rates of dry mouth, similar to placebo. The addition of mirabegron to the prescribing schedule in England and Wales will provide patients with an alternative treatment for OAB with an approved efficacy and tolerability balance. Mirabegron has the potential to greatly improve patient compliance and outcomes, and may avoid the need for more invasive surgical treatments. Currently available antimuscarinics have been shown to fail to achieve a balance between efficacy and tolerability in many patients and this is reflected by the general low persistence with treatment.[18] It is anticipated that mirabegron would offer an alternative pharmacotherapy to antimuscarinics within the existing pathway for both treatment naive patients and previously treated patients – for example, for patients in whom the desired efficacy has not been achieved with antimuscarinic treatment, or for those patients who have been unable to tolerate antimuscarinic treatment. Currently these patients may progress to surgery or symptom management using incontinence pads.
Abbreviation used in box: OAB, overactive bladder.

The manufacturer states that there are no additional resources required for selection, monitoring or administration of mirabegron, and no concomitant OAB therapy is needed for patients on mirabegron, although some patients may persist with bladder training. The manufacturer indicates that the marketing authorisation for mirabegron is anticipated to cover the general OAB population, which was estimated to be approximately 5 million people aged 40 years and older in England and Wales.[6] The ERG notes that the number of patients eligible for treatment with mirabegron will likely be less than 5 million people as, according the NICE clinical guidelines, patients should only be offered pharmacotherapy if conservative treatment has been ineffective.[12;13] According to a European study, which included the UK, a large proportion of patients with OAB do not consult a medical practitioner about their symptoms.[6] Furthermore, of those seeking help, only 27% were currently taking pharmacotherapy.[6]

3 CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

The manufacturer provided a summary of the final decision problem issued by the National Institute for Health and Clinical Excellence (NICE; Table 1).[19] The manufacturer predominantly adhered to the scope issued by NICE, but, in addition to the requested outcomes, also presented data on the additional outcome of number of incontinence episodes.

Table 1. Summary of decision problem as outlined in the manufacturer's submission (reproduced from MS; pg 32)

Key parameter	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
Population	Adults with symptoms of OAB	As per NICE scope	N/A
Intervention	Mirabegron	As per NICE scope	N/A
Comparator(s)	Antimuscarinic drugs including: • oxybutynin (including modified-release preparations); • tolterodine; • fesoterodine; • solifenacin; • trospium.	As per NICE scope	N/A
Outcomes	• urinary frequency; • frequency of urge urinary incontinence; • symptoms of urgency; • nocturia; • adverse effects of treatments; • HRQoL.	As per NICE scope, and additionally: • number of incontinence episodes	N/A
Economic analysis	The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of incremental cost per QALY. The reference case stipulates that the time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	As per NICE scope	N/A
Subgroups to be considered	If the evidence allows: • men and women; • previously untreated and previously treated OAB.	As per NICE scope	N/A
Special considerations, including issues related to equity or equality	None	N/A	N/A

Abbreviations used in table: HRQoL, health-related quality of life; MS, manufacturer's submission; NICE, National Institute for Health and Clinical Excellence; OAB, overactive bladder; N/A, not applicable; NHS, National Health Service; QALYs, quality adjusted life years.

3.1 Population

The trials SCORPIO,[20] ARIES,[21] and CAPRICORN,[22] which formed the basis of the direct clinical evidence submitted by the manufacturer, enrolled adults with symptomatic overactive bladder (OAB). To be eligible for randomisation, patients were required to:

At screening:

- be of ≥18 years of age;
- have symptoms of OAB (urinary frequency and urgency with or without incontinence) for ≥3 months.

At baseline:

- have a frequency of micturition of ≥8 times per 24-hour period during the 3-day micturition diary period;
- have ≥3 episodes of urgency (Grade 3 or 4) with or without incontinence during the 3-day micturition diary period.

In addition to SCORPIO, ARIES, and CAPRICORN, the manufacturer identified four other trials evaluating treatments in patients with OAB.[23-26] The manufacturer decided against including the four additional trials as direct clinical evidence, and instead reported the trials as only supporting evidence. This issue is discussed further in Section 4. The ERG considers it important to note that, at the time of writing of the report, data from the identified RCTs have yet to be published in a peer-reviewed journal; all data are taken from the Clinical Study Reports (CSRs) for the individual trials, which the manufacturer helpfully provided.

SCORPIO, ARIES, and CAPRICORN were carried out at sites in Europe, North America, and Australia.[20-22] The population in the three trials was primarily Caucasian (88-99%). The ERG's clinical expert stated that ethnicity is unlikely to influence the development or symptoms of OAB and therefore trials involving patients with OAB carried out in any country and any population with OAB may be considered representative of patients with OAB in England and Wales.

In summary, the ERG considers the population in the three RCTs to be in agreement with the population defined in the final scope issued by NICE,[19] and to be representative of patients with OAB in the UK.

3.2 Intervention

Mirabegron is a selective beta$_3$-adrenoceptor agonist. It activates beta$_3$-adrenoceptors in the bladder wall muscle (detrusor) and a triangular region of the urinary bladder called the trigone area, causing relaxation of the bladder and facilitating urine storage.[27]

Marketing authorisation for mirabegron was granted in Japan in July 2011 (under the trade name Betanis®) and in the USA in June 2012 (under the trade name Myrbetriq®).[28] The manufacturer does not currently have a UK marketing authorisation for mirabegron. However, in October 2012, the Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion on the use of mirabegron at doses of 25 and 50 mg for the symptomatic treatment of urgency, increased micturition frequency and/or urgency incontinence as may occur in adult patients with OAB.[29] The manufacturer anticipates that the European Medicines Agency (EMA) will issue marketing authorisation in late January 2013.

In the manufacturer's submission (MS), the manufacturer presents direct evidence on the clinical effectiveness of mirabegron at doses of 25 mg,[22] 50 mg,[20-22] and 100 mg.[20;21] The licensed dose of mirabegron is anticipated to be 50 mg (recommended dose) and 25 mg for patients with renal or hepatic impairment (MS; Table 4, pg 22). The ERG notes that, based on the expected licence for mirabegron, data on a dose of 100 mg are not relevant to the decision problem that is the focus of this Single Technology Appraisal (STA).

To summarise, the ERG considers the intervention in the MS to be consistent with the anticipated licence and the final scope issued by NICE for this STA.[19]

3.3 Comparators

The comparators listed in the final scope issued by NICE were:[19]

- oxybutynin (including modified-release preparations);
- tolterodine;
- fesoterodine;
- solifenacin;
- trospium.

In the MS, the manufacturer presented direct clinical data from three trials comparing mirabegron with placebo. One of the trials (SCORPIO) had an active control arm of tolterodine, which is one of the comparators of interest.[20] However, the manufacturer stated that the trial was not a direct comparison of the two treatments as it was not powered to detect superiority or non-inferiority of mirabegron versus tolterodine, and therefore data from the active control arm were presented only as a non-statistical comparison. The ERG agrees with the manufacturer regarding the statistical power of the comparison, but considers the evidence to be relevant to the decision problem of this STA.

Of the four trials excluded from the direct clinical evidence, three trials had an active control arm of tolterodine.[23;25;26] The ERG considers that the three trials address a comparison relevant to this STA,

and that the inclusion of direct clinical evidence from these trials could provide an important contribution to the clinical evidence of the effectiveness of mirabegron compared with tolterodine.

The manufacturer did not identify any head-to-head trials comparing mirabegron with any of the other comparators listed in the NICE final scope.[19] Consequently, to assess the relative effect of mirabegron versus the comparators listed in the scope, the manufacturer carried out a mixed treatment comparison (MTC). Based on the absence of head-to-head trial data of mirabegron versus the comparators listed in the scope, the ERG considers the manufacturer's decision to synthesise relative treatment effects using an MTC to be appropriate. The ERG's evaluation of the manufacturer's MTC is discussed in Section 4.

In summary, the ERG considers that the comparators specified in the final scope issued by NICE[19] have been addressed within the MTC in the MS.

3.4 Outcomes

The outcomes listed in the final scope issued by NICE for this STA were:[19]

- urinary frequency;
- frequency of urge urinary incontinence;
- symptoms of urgency;
- nocturia;
- adverse effects of treatments;
- health-related quality of life (HRQoL).

Within the MS, the manufacturer presents direct clinical evidence from three trials: SCORPIO;[20] ARIES;[21] and CAPRICORN[22]. The two pre-specified primary outcomes in the three trials were change from baseline (CFB) to end of treatment (final visit) in:

- mean number of incontinence episodes per 24 hours based on a 3-day micturition diary;
- mean number of micturitions per 24 hours based on a 3-day micturition diary.

The manufacturer reported data from the three trials for the outcomes of urinary frequency, frequency of urge urinary incontinence, frequency of urgency episodes, level of urgency, nocturia, and HRQoL. HRQoL was measured using multiple scales and questionnaires:

- European quality of life-5 dimensions (EQ-5D);
- EQ-5D Visual analogue scale (EQ-5D VAS);
- Treatment satisfaction visual analogue scale (TS-VAS);
- Work productivity and activity impairment: specific health problem (WPAI:SHP);
- Overactive bladder questionnaire (OAB-q), disease-specific scale;
- Patient perception of bladder condition scale (PPBC), disease-specific scale.

The manufacturer also reported results for other outcomes not listed in the NICE scope:

- frequency of incontinence episodes;
- volume voided per micturition;
- mean number of pads used.

The ERG notes that the manufacturer did not specify which, if any, of the outcomes reported corresponds to the outcome of symptoms of urgency, listed in the scope.[19] Due to time constraints, the large number of included trials and large number outcomes, the ERG will focus on the outcomes listed in the final scope issued by NICE, and the additional outcome, frequency of incontinence episodes, as this informs the manufacturer's economic model. The ERG's clinical expert indicated that frequency of urgency urinary incontinence is analogous to number of incontinence episodes as incontinence is typically preceded by urgency. In the key trials evaluating mirabegron, incontinence was defined as any involuntary leakage of urine, with or without accompanying or preceding sense of urgency.

The manufacturer also reported efficacy data from the TAURUS study, which is a long-term follow-up study of the effects of mirabegron.[26] However, data were limited to 12 months and no direct statistical comparisons of efficacy between treatment groups were performed. TAURUS was the main study investigating adverse effects of mirabegron treatment; however, safety data from SCORPIO, ARIES, and CAPRICORN were also presented in the MS.

The manufacturer also presented data from an MTC for the outcomes: urinary frequency; frequency of incontinence episodes; frequency of urge urinary incontinence; and the adverse events dry mouth, constipation, and blurred vision. Considering the adverse effects evaluated, the ERG's clinical expert confirmed that dry mouth, constipation, and blurred vision are the most relevant to treatment with anticholinergics.

To summarise, the ERG considers that the outcomes listed in the NICE final scope have been addressed in the MS. However, the ERG notes that not all clinical data that could inform the different outcomes have been included.

3.5 Other relevant factors

Six of the seven identified trials had a trial duration of 3 months,[20-25] with the safety study TAURUS having a longer duration of follow-up of 1-year.[26] According to the ERG's clinical expert, it is recommended within NHS clinical practice to assess pharmaceutical treatments for OAB after three months of treatment. Based on this, the ERG considers the duration of treatment and follow-up in the included trials to be sufficient for assessing the efficacy and safety of treatment with mirabegron.

The manufacturer presented subgroup data for the two pre-specified primary outcomes in the three mirabegron trials SCORPIO, ARIES, and CAPRICORN based on gender and prior treatment (i.e., previously treated and treatment naïve). The subgroups were in line with the scope issued by NICE. The results of these analyses are discussed in further detail in Section 4.

4 CLINICAL EFFECTIVENESS

4.1 Critique of the methods of reviews

4.1.1 Searches

The manufacturer conducted two systematic reviews of the literature to identify relevant clinical data to inform on the efficacy and safety of mirabegron for the treatment of adults with symptoms of overactive bladder (OAB). One review was limited to randomised controlled trials (RCTs), and studies directly comparing mirabegron with the treatments of interest listed in the final scope issued by the National Institute for Health and Clinical Excellence (NICE).[19] In addition, this literature was used to inform the mixed treatment comparison (MTC), where head-to-head RCTs proved to be insufficient. The second review was designed to identify non-RCT evidence on the efficacy and safety of mirabegron.

The manufacturer lists the databases and trial registers searched, conference proceedings that were hand searched, and the time spans of the searches. The manufacturer also searched reference lists of identified trials and systematic reviews. The literature was searched on 13th June 2012 for both the systematic review of RCTs and non-randomised studies.

The Evidence Review Group (ERG) notes that, in the manufacturer's submission (MS), the manufacturer states that MEDLINE was searched from 1946 (MS Section 10.2.3), and, in a subsequent section, from 1950 (MS Section 10.2.4). Also, in the MS, the manufacturer states that EMBASE was searched from 1980 up to 2010 week 23 (MS Section 10.2.3) and from 1980 to 2010 week 36 (MS Section 10.2.4). In the search for non-randomised studies, the manufacturer states that EMBASE was searched from 1980 up to 2012 week 23 (MS Section 10.6.3). The discrepancy in the span of the search dates was not discussed by the manufacturer. However, the ERG considers that the variation in the cited search dates for EMBASE could be a typographical error and that the search date for both reviews was 2012 week 23. Moreover, the ERG considers that the variation in the date from which the literature was searched is unlikely to have resulted in omission of key studies.

The search terms included commonly used words to describe the disease, drug names and brand names for mirabegron and the comparators listed in the scope. It also included appropriate search terms for study design. The manufacturer did not specify a separate search strategy for identifying adverse events.

Due to time constraints, the ERG has been unable to validate fully the manufacturer's searches and verify the number of studies identified. However, the ERG considers the manufacturer's searches to be comprehensive and the search strategies used for each systematic review to be appropriate. In

addition, the ERG is unaware of any relevant studies that have been missed by the manufacturer's search.

4.1.2 Inclusion/exclusion criteria

For both the systematic review for RCTs and non-randomised studies, two reviewers independently assessed identified references for inclusion/exclusion and any discrepancies were resolved by a third reviewer.

Inclusion and exclusion criteria and appropriate flow diagrams were provided for the literature searches for RCTs and non-randomised studies (presented in Appendix 1).

Direct clinical evidence

The inclusion/exclusion criteria of RCTs for direct comparisons and non-randomised studies predominantly aligned with the final scope.[19] However, the ERG notes that the inclusion criteria relating to the intervention of interest, as stated by the manufacturer, included mirabegron or oxybutynin (including modified-release preparations) (Appendix 1). No studies evaluating oxybutynin were included by the manufacturer as direct clinical evidence on the effects of mirabegron. In addition, the manufacturer's review excluded studies evaluating transdermal oxybutynin as a comparator. The manufacturer stated that the decision to exclude studies evaluating transdermal oxybutynin was based on the differences in placebo administration (i.e., placebo patch vs oral tablet). The ERG considers the exclusion of non-oral formulations of oxybutynin to be inappropriate.

No separate search was conducted to identify studies looking at the safety of mirabegron. However, the manufacturer presents safety data from all the trials for which they also presented direct clinical effectiveness data.

The manufacturer identified seven trials in the population of interest, which evaluated mirabegron at the anticipated licensed doses (50 mg [recommended dose], and 25 mg for those with hepatic or renal failure):

- 178-CL-044 (DRAGON);[23]
- 178-CL-045;[24]
- 178-CL-046 (SCORPIO);[20]
- 178-CL-047 (ARIES);[21]
- 178-CL-048;[25]
- 178-CL-049 (TAURUS);[26]
- 178-CL-074 (CAPRICORN).[22]

However, in the MS, the direct clinical evidence is based on a pooled analysis of only three of the trials (SCORPIO, ARIES, CAPRICORN), and the safety data are based on the long-term study TAURUS. The reasons provided by the manufacturer for not including the four remaining trials in the pooled analysis for the direct clinical evidence are listed in Table 2.

Table 2. Rationale for excluding direct clinical evidence of randomised controlled trials including mirabegron from pooled analysis (reproduced from MS; Table 46, pg 114)

Study no. (acronym)	Rationale for exclusion
178-CL-044 (DRAGON)	• This study was intended to be supportive • The endpoint of mean number of incontinence episodes per 24 hours was not considered a primary endpoint in this study • This study has differences from the primary studies in the derivations of micturition-based endpoints
178-CL-045	• This study was conducted only in Japan and is supportive for this submission • The endpoint of mean number of incontinence episodes per 24 hours was not considered a primary endpoint in this study • The inclusion and exclusion criteria for defining the OAB population differ from the primary studies • Urinary urgency was captured based on whether a patient had urgency or not with an episode. Therefore, the key secondary endpoints of mean level of urgency, mean number of urgency incontinence episodes (Grade 3 or 4)/24 hr and mean number of urgency episodes (Grade 3 or 4)/24 hr) in the primary studies cannot be summarised • Collection of data for mean number of nocturia episodes/24 hr differs between this study and the primary studies
178-CL-048	• This study was conducted only in Japan and are supportive for this submission • The endpoint of mean number of incontinence episodes per 24 hours was not considered a primary endpoint in this study • The inclusion and exclusion criteria for defining the OAB population differ from the primary studies • Urinary urgency was captured based on whether a patient had urgency or not with an episode. Therefore, the key secondary endpoints of mean level of urgency, mean number of urgency incontinence episodes (Grade 3 or 4)/24 hr and mean number of urgency episodes (Grade 3 or 4)/24 hr) in the primary studies cannot be summarised • Collection of data for mean number of nocturia episodes/24 hr differs between this study and the primary studies
178-CL-049 (TAURUS)	• Differences in duration of treatment • Lack of placebo control
Abbreviation used in table: OAB, overactive bladder.	

The ERG considers the four excluded trials to be relevant to the decision problem that is the focus of this Single Technology Appraisal (STA). Therefore, these four trials are discussed throughout the report alongside the three trials included by the manufacturer in the pooled analysis of direct clinical evidence. The eligibility of inclusion of the four excluded trials is discussed in greater detail in the section that follows.

Indirect comparison

The manufacturer expanded the inclusion and exclusion criteria for the identification of trials eligible for the MTC, which was carried out to provide comparative clinical effectiveness data for mirabegron versus the comparators listed in the final scope.

Inclusion criteria:

- study duration of 4–16 weeks for safety analysis;
- study duration of 8–16 weeks for efficacy;
- appropriate measures of variability;
- primary analysis;
- outcomes as per the final scope.

Exclusion criteria:

- sub-analysis;
- pooled analysis;
- inappropriate study duration;
- no relevant outcome reported;
- not a major publication;
- inadequate reporting;
- not relevant doses/treatments compared;
- non-RCT/ not relevant study design;
- not appropriate population for analysis.

The manufacturer did not provide a rationale for, or definitions of, the additional inclusion and exclusion criteria. It is therefore difficult for the ERG to validate the appropriateness of the study selection. However, considering the inclusion criterion of minimum study duration of 8 weeks for evaluation of efficacy, the ERG considers it important to note that the ERG's clinical expert highlighted that 12 weeks' follow-up would be considered sufficient to evaluate the effects of a treatment for OAB. Furthermore, typical UK clinical practice would be to treat patients for 3 months initially, followed by switching to another treatment if no clinical improvement is observed.

4.1.3 Description of included trials, direct clinical evidence

Randomised controlled trials

The manufacturer identified seven RCTs relevant to the decision problem. However, as stated in Section 4.1.2, the manufacturer focuses on direct clinical evidence from only three of the identified trials (CAPRICORN, SCORPIO, ARIES).[20-22] Clinical data from the long-term study TAURUS were

also presented, although this study was predominantly included as a safety study on the long-term effects of mirabegron.[26] The remaining three trials were presented as only supporting evidence.[23-25] Based on the final scope issued by NICE,[19] the ERG considers all seven trials to be relevant to the decision problem that is the focus of this STA. Key characteristics of the seven RCTs are summarised in Table 3. All seven trials were double-blind, parallel group, randomised controlled studies, and were either Phase II or Phase III studies.

Table 3. Comparative summary of methodology of randomised controlled trials evaluating mirabegron

Study	Trial design	Population	Intervention and comparators	Number randomised	Trial length (weeks)	Primary outcomes
DRAGON[23]	Phase IIb, RCT, double-blind, multicentre (Europe)	OAB, ≥18 years	• Mirabegron 25 mg • Mirabegron 50 mg • Mirabegron 100 mg • Mirabegron 200 mg • Tolterodine ER 4 mg • Placebo	928	12	Urinary frequency
SCORPIO[20]	Phase III, RCT, double-blind, multicentre (Europe, Australia)	OAB, ≥18 years	• Mirabegron 50 mg • Mirabegron 100 mg • Tolterodine ER 4 mg • Placebo	1,987	12	Urinary frequency Incontinence frequency
ARIES[21]	Phase III, RCT, double-blind, multicentre (US and Canada)	OAB, ≥18 years	• Mirabegron 50 mg • Mirabegron 100 mg • Placebo	1,329	12	Urinary frequency Incontinence frequency
CAPRICORN[22]	Phase III, RCT, double-blind, multicentre (Europe, US, Canada)	OAB, ≥18 years	• Mirabegron 25 mg • Mirabegron 50 mg • Placebo	1,306	12	Urinary frequency Incontinence frequency
178-CL-045[24]	Phase II, RCT, double-blind, multicentre (Japan)	OAB, 20–80 years	• Mirabegron 25 mg • Mirabegron 50 mg • Mirabegron 100 mg • Placebo	842	12	Urinary frequency
178-CL-048[25]	Phase III, RCT, double-	OAB, ≥20 years	• Mirabegron 50 mg • Tolterodine tartrate 4 mg	1,139	12	Urinary frequency

	blind, multicentre (Japan)		• Placebo			
TAURUS[26]	Phase III, RCT, double-blind, multicentre (Europe, USA, Canada, Australia, South Africa)	OAB, ≥18 years	• Mirabegron 50 mg • Mirabegron 100 mg • Tolterodine ER 4 mg	2,452	52	Incidence and severity of TEAEs
Abbreviations used in table: ER, extended release; mg, milligram; OAB, overactive bladder; RCT, randomised controlled trial; TEAE, treatment-emergent adverse event.						

The objectives of SCORPIO, ARIES, CAPRICORN, and TAURUS were to assess the efficacy and safety of mirabegron in patients with symptoms of OAB. The study objective in DRAGON and 178-CL-045 was to evaluate the dose-response relationship of mirabegron efficacy in patients with OAB. 178-CL-048 was designed to evaluate the efficacy (superiority vs placebo), safety and pharmacokinetics of mirabegron in patients with OAB, and to evaluate the efficacy and safety of mirabegron compared with tolterodine. However, according to the manufacturer, none of the studies that included an active control arm of tolterodine (SCORPIO, TAURUS, DRAGON, 178-CL-048) was powered to detect superiority or non-inferiority of mirabegron compared with tolterodine.

All seven trials were multicentre, and all but the two studies carried out in Japan (178-CL-045, and 178-CL-048) were multinational. As mentioned in Section 3.1, the ERG's clinical expert has stated that it is unlikely that ethnicity is a determining factor in the development of OAB and therefore OAB trials comprising multiple ethnicities are likely to be representative of patients in England and Wales with OAB.

The trials all had a 2-week single-blind placebo run in, during which the patients were blinded to the identity of study drug. Randomisation was stratified by country in SCORPIO, and by study centre in ARIES, CAPRICORN, and TAURUS. It is unclear from the MS whether randomisation was stratified in DRAGON, 178-CL-045, and 178-CL-048. The duration of all trials was three months, with the exception of the safety study TAURUS, which had a follow-up period of 12 months.

Three studies used both placebo and active control (tolterodine),[20;23;25] one study used only active control,[26] and the remaining three used placebo control.[21;22;24]

As mentioned in Section 3.2, the ERG notes that based on the anticipated licence for mirabegron, only the 25 and 50 mg doses are relevant to the decision problem. Data on mirabegron used at a dose of 100 mg are therefore not discussed in the report.

Patient inclusion criteria were similar in the seven trials (summarised in Table 4).

Table 4. Patient inclusion criteria for the randomised controlled trials evaluating mirabegron

Study	Inclusion criteria
SCORPIO, ARIES, CAPRICORN, TAURUS, DRAGON	At screening: • be of ≥18 years of age; • have symptoms of OAB (urinary frequency and urgency with or without incontinence) for ≥3 months. At baseline: • frequency of micturition ≥8 times per 24-hour period during the 3-day micturition diary period; • have ≥3 episodes of urgency (Grade 3 or 4) with or without incontinence during the 3-day micturition diary period.
178-CL-045	At screening: • be aged between 20 and 80 years; • have symptoms of OAB for ≥24 weeks (6 months). At baseline: • frequency of micturition ≥8 times per 24-hour period during the 3-day run-in period; • have ≥1 episode of urgency per 24 hours or ≥1 episode of urge incontinence per 24 hours during the 3-day run-in period.
178-CL-048	At screening: • be of ≥20 years of age; • have symptoms of OAB for ≥24 weeks (6 months). At baseline: • frequency of micturition ≥8 times per 24-hour period during the 3-day run-in period; • have ≥1 episode of urgency per 24 hours or ≥1 episode of urge incontinence per 24 hours during the 3-day run-in period
Abbreviation used in table: OAB, overactive bladder.	

TAURUS primarily included patients who had completed the 12-week treatment and safety follow-up periods of SCORPIO and ARIES. However, the patients were randomised again when entering TAURUS, and patients had to have been off study medication for at least 30 days.

Patient baseline characteristics were provided for the five multinational RCTs: SCORPIO; ARIES; CAPRICORN; TAURUS; and DRAGON. Baseline characteristics for the patients enrolled in 178-CL-045 and 178-CL-048 were not presented in the MS, but the manufacturer states that baseline characteristics and OAB history were similar across treatment groups with no significant differences between groups reported in either trial.

Patient baseline characteristics were consistent between treatment groups in SCORPIO; ARIES; CAPRICORN; TAURUS; and DRAGON. Details were provided on patient demographics, OAB history, and OAB-related baseline characteristics (Appendix 2). The trials enrolled more women than men. The mean age across the trials was around 60 years. Patient OAB history and characteristics were similar across the treatment groups in the five trials. The manufacturer divided the type of OAB into:

- urgency incontinence (urge incontinence only);
- mixed (mixed stress/urge incontinence with urge as a predominant factor);
- frequency (frequency/urgency without incontinence).

The type of OAB was similar across the treatment groups but varied slightly between studies. In SCORPIO, there was imbalance between the treatment arms in the number of patients who had undergone surgery for OAB prior to enrolment, with almost double the number of patients in the mirabegron 50 mg group having surgery as in the placebo or tolterodine group. However, the ERG notes that the proportion in each arm was small (4.6% in the placebo group, 7.0% in the mirabegron 50 mg group, and 3.6% in the tolterodine group). The proportion of patients with previous OAB surgery was similarly low in the ARIES, CAPRICORN, and TAURUS (<10%). The number of patients who had previously had OAB surgery was not reported for the DRAGON trial. Around 50% of included patients in the five trials for which data are available had previously taken pharmacotherapy for OAB. Patients had a median duration of symptoms of OAB of approximately 4 years. OAB-related baseline characteristics were also consistent across treatment groups. The mean number of micturitions per 24 hours was between 11 and 12 in the five RCTs. Moreover, the overall mean level of urgency was approximately 2.4 and the mean number of urgency episodes (Grade 3 or 4) was between 5 and 6. Incontinence-related characteristics were comparable across treatment groups. The mean number of incontinence episodes per 24 hours ranged from 2.41 to 3.03 and the mean number of urgency incontinence episodes per 24 hours ranged from 2.21 to 2.88.

Patients were assessed at weeks 4, 8, and 12 in all trials with treatment duration of 12 weeks. Post treatment, patients were followed up until 30 days after end of treatment phase in SCORPIO, ARIES, and CAPRICORN, and for 2 weeks after the last visit in 178-CL-048. In TAURUS patients were assessed at 1, 3, 6, 9, and 12 months, and were not contacted after the last visit.

In SCORPIO, ARIES, and CAPRICORN the primary outcomes were frequency of incontinence episodes and micturitions. The primary outcome in DRAGON, 178-CL-045, and 178-CL-048 was urinary frequency, and, in the safety study TAURUS, the primary outcome was the incidence and severity of treatment-emergent adverse events (TEAEs). The manufacturer states that the outcomes were selected based on Committee for Medicinal Products for Human Use (CHMP) guidance on clinical investigation of medicinal products for urinary incontinence. Data for most primary and secondary outcomes were captured in a micturition diary:

- urinary frequency – measured as the mean number of micturitions per 24 hours;
- frequency of incontinence episodes (incontinence defined as any involuntary leakage of urine) – measured as the mean number of incontinence episodes per 24 hours;
- mean volume voided per micturition;
- frequency of urge urinary incontinence (urge incontinence or urge urinary incontinence defined as involuntary leakage accompanied by or immediately preceded by urgency) – measured as the mean number of urge incontinence episodes per 24 hours;
- frequency of nocturia – measured as mean number of nocturia episodes per 24 hours;
- mean number of pads used per 24 hours;
- frequency of urgency episodes and level of urgency – assessed using the 5-point categorical Patient Perception of Intensity of Urgency Scale (PPIUS).

For each micturition and/or incontinence episode, patients were asked to rate the degree of associated urgency according to the 5-point PPIUS:

0. No urgency, I felt no need to empty my bladder, but did so for other reasons.
1. Mild urgency, I could postpone voiding as long as necessary, without fear of wetting myself.
2. Moderate urgency, I could postpone voiding for a short while, without fear of wetting myself.
3. Severe urgency, I could not postpone voiding, but had to rush to the toilet in order not to wet myself.
4. Urge incontinence, I leaked before arriving to the toilet.

The manufacturer stated that urinary urgency and nocturia were captured differently in 178-CL-045 and 178-CL-048 compared with the other trials. Because of the differences, the manufacturer indicated that the outcomes of level of urgency, urgency incontinence episodes, mean number of urgency episodes (Grade 3 or 4), and nocturia were not comparable across the studies. The manufacturer also stated that there were differences in the derivations of micturition-based endpoints between DRAGON and SCORPIO, ARIES, and CAPRICORN, rendering these outcomes not comparable across the studies. However, for studies 178-CL-045 and 178-CL-048, the ERG was unable to ascertain from the relevant Clinical Study Reports (CSRs) how urgency and nocturia were assessed. In addition, the ERG identified no significant differences in derivations of micturition-based endpoints. Therefore, the ERG considers that data for urgency incontinence episodes, urgency episodes, and nocturia reported in 178-CL-045 and 178-CL-048 are of clinical relevance to the decision problem, and relevant data are therefore presented in Section 4.2.3, together with data from the other trials.

Health-related quality of life (HRQoL) and treatment satisfaction were captured using several different scales:

- European quality of life-5 dimensions (EQ-5D);
- EQ-5D Visual analogue scale (EQ-5D VAS);
- treatment satisfaction visual analogue scale (TS-VAS);
- work productivity and activity impairment: specific health problem (WPAI:SHP);
- overactive bladder questionnaire (OAB-q), disease specific scale;
- patient perception of bladder condition scale (PPBC), disease specific scale;
- King's Health questionnaire (KHQ, QoL domain) (only 178-CL-045, 178-CL-048).

The manufacturer's description of these can be found in Appendix 3. Adverse events were a secondary outcome in all seven trials.

Non-randomised trial

The manufacturer identified one non-RCT as being relevant to the decision problem, study 178-CL-051. Key characteristics of the trial are summarised in Appendix 4.

4.1.4 Description of trials included in the MTC

A summary of the trials used to inform the MTC is provided in Table 5. In total, 40 trials were included in the MTC. The trials included an OAB population; however, some studies narrowed or widened the population either by limiting it to include, for example, only women, or patients with a specific symptom of OAB. As a whole, the included trials evaluated all the relevant interventions and comparators listed in the final scope (mirabegron, oxybutynin, tolterodine, fesoterodine, solifenacin, and trospium).

Table 5. Summary of trials used to create the network for the mixed treatment comparison (adapted from MS; Table 47, pg117)

Study (primary reference)	Intervention	Trial design	Patient population	Number randomised patients	Trial length (weeks)
Mirabegron studies					
DRAGON[23]	• Mirabegron 25 mg • Mirabegron 50 mg • Mirabegron 100 mg • Mirabegron 200 mg • Tolterodine ER 4 mg • Placebo	Phase IIb, RCT, double-blind, double dummy, multicentre (Europe)	OAB, ≥18 years	928	12
SCORPIO[20]	• Mirabegron 50 mg • Mirabegron 100 mg • Tolterodine ER 4 mg • Placebo	Phase III, RCT, double-blind, multicentre (Europe, Australia)	OAB, ≥18 years	1,987	12

ARIES[21]	• Mirabegron 50 mg • Mirabegron 100 mg • Placebo	Phase III, RCT, double-blind, double dummy, multicentre (US and Canada)	OAB, ≥18 years	1,329	12
CAPRICORN[22]	• Mirabegron 25 mg • Mirabegron 50 mg • Placebo	Phase III, RCT, double-blind, double dummy, multicentre (Europe, US, Canada)	OAB, ≥18 years	1,306	12
178-CL-045[24]	• Mirabegron 25 mg • Mirabegron 50 mg • Mirabegron 100 mg • Placebo	Phase II, RCT, double-blind, multicentre (Japan)	OAB, 20–80 years	842	12
178-CL-048[25]	• Mirabegron 50 mg • Tolterodine tartrate 4 mg • Placebo	RCT, double-blind, multicentre (Japan)	OAB , ≥20 years	1,139	12
Comparator studies					
BLOSSOM[30]a	• Mirabegron 150 mg • Mirabegron 100 mg • Tolterodine ER 4 mg • Placebo	Phase II, RCT, double-blind, multicentre (Europe)	OAB, ≥18 years	262	4
Abrams 2006[31]	• Oxybutynin 5 mg TDS • Propiverine 20 mg • Propiverine 45 mg • Placebo	RCT, double-blind, multicentre (UK)	Idiopathic OAB, >18 years	77	4
Appell 2001[32]	• Tolterodine IR 2 mg BD • Oxybutynin ER 10 mg	RCT, double-blind, multicentre (USA)	OAB	378	12
Birns 2000[33]	• Oxybutynin ER 10 mg • Oxybutynin IR 5 mg BD	RCT, double-blind, double dummy, multicentre (UK)	Patients with voiding problems, 18–76 years	130	6
Cardozo 2004[34]	• Solifenacin 5 mg • Solifenacin 10 mg • Solifenacin 20 mg	RCT, double-blind, multicentre	OAB, ≥18 years	907	12
Chapple 2007[35]	• Tolterodine ER 4 mg • Fesoterodine 4 mg • Fesoterodine 8 mg • Placebo	Phase III RCT, double-blind, double dummy, multicentre	OAB, ≥18 years	1,135	12

Chapple 2004[36]	• Solifenacin 2.5 mg • Solifenacin 5 mg • Solifenacin 10 mg • Solifenacin 20 mg • Tolterodine IR 2 mg BD • Placebo	Phase II, RCT, double-blind, multicentre (Europe)	Idiopathic detrusor overactivity, 18–80 years	225	4
Chapple 2004[37]	• Solifenacin 5 mg • Solifenacin 10 mg • Tolterodine 2 mg BD • Placebo	Phase IIIa RCT, double-blind, multicentre (North America & Europe)	OAB, ≥18 years	1,081	12
Choo 2008[38]	• Solifenacin 5 mg • Solifenacin 10 mg • Tolterodine IR 2 mg BD	Phase III RCT, double-blind, multicentre (Korea)	OAB, ≥18 years	329	12
Chu 2009[39]	• Solifenacin 10 mg • Placebo	Phase III RCT, double-blind, multicentre (USA)	OAB, ≥18 years	672	12
Corcos 2006[40]	• Oxybutynin ER 5 mg • Oxybutynin ER 10 mg • Oxybutynin ER 15 mg	RCT, double-blind, multicentre (Canada)	Urge urinary incontinence ≥18 years	237	4
Diokno 2003[41]	• Oxybutynin ER 10 mg • Tolterodine ER 4 mg	RCT, double-blind, multicentre (USA)	Women with OAB, ≥18 years	790	12
Dmochowski 2003[42]	• Oxybutynin TDS • Long-acting tolterodine • Placebo	RCT, double blind, double dummy	≥18 years, taking pharmacologic treatment for OAB	361	12
Herschorn 2008[43]	• Tolterodine ER 4 mg • Placebo	RCT, double-blind, multicentre (Canada, Europe)	OAB, ≥18 years	617	12
Herschorn 2010[44]	• Solifenacin 5 mg • Oxybutynin IR 5 mg TDS	RCT, double-blind, double dummy, multicentre (Canada)	OAB, ≥18 years	132	8
Herschorn 2010[45]	• Fesoterodine 4/8 mg • Tolterodine ER 4 mg • Placebo	RCT, double-blind, double dummy, multicentre (USA)	OAB, ≥18 years	1,712	12

Ho 2010[46]	• Solifenacin 5 mg • Tolterodine ER 4 mg	Randomised, open-label, single centre (Taiwan)	OAB, ≥18 years	75	12
Homma 2003[47]	• Tolterodine ER 4 mg • Oxybutynin 3 mg TDS • Placebo	RCT, double-blind, multicentre (Japan)	OAB, ≥20 years	608	12
Jacquetin 2001[48]	• Tolterodine IR 1 mg BD • Tolterodine IR 2 mg BD • Placebo	Phase III RCT, double-blind, multicentre (France & Belgium)	OAB, ≥18 years	251	4
Kaplan 2011[49]	• Fesoterodine 4/8 mg • Tolterodine ER 4 mg • Placebo	RCT, double-blind, double dummy, multicentre (North & South America, Europe, Asia, Africa)	OAB, ≥18 years	2,417	12
Khullar 2004[50]	• Tolterodine ER 4 mg • Placebo	RCT, double-blind, multicentre (Europe)	Women ≥18 years with urge-predominant mixed incontinence	854	8
Lackner 2008[51]	• Oxybutynin ER 5 mg • Placebo	RCT, double-blind	Women ≥65 years with urge-incontinence and cognitive impairment	50	4
Lee 2002[52]	• Tolterodine IR 2 mg BD • Oxybutynin 5 mg BD	RCT, double-blind, multicentre (Korea)	OAB, ≥18 years	228	8
Malone-Lee 2001[53]	• Tolterodine IR 1 mg BD • Tolterodine IR 2 mg BD • Placebo	RCT, double-blind, multicentre (UK, France, Ireland)	≥65 years with symptoms of urinary urgency, increased frequency of micturition and/or urge incontinence	177	4
Nitti 2007[54]	• Fesoterodine 4 mg • Fesoterodine 8 mg • Placebo	RCT, double-blind, multicentre (USA)	OAB, ≥18 years	836	12

Nitti 2010[55]	• Fesoterodine 4 mg • Fesoterodine 8 mg • Fesoterodine 12 mg • Placebo	Phase II RCT, double-blind, multicentre	OAB, 18–78 years	173	8
Rackley 2006[56]	• Tolterodine ER 4 mg • Placebo	RCT, double-blind, multicentre (USA)	OAB, ≥18 years	850	12
Rogers 2008[57]	• Tolterodine ER 4 mg • Placebo	RCT, double-blind, multicentre (USA)	Sexually active women, ≥18 years with OAB	413	12
Rudy 2006[58]	• Trospium chloride 20 mg BD • Placebo	RCT, double-blind, multicentre (USA)	OAB, ≥18 years	658	12
Staskin 2007[59]	• Trospium chloride 60 mg • Placebo	RCT, double-blind, multicentre (USA)	Subjects with OAB	601	12
Van Kerrebroeck 2001[60]	• Tolterodine IR 2 mg BD • Tolterodine ER 4 mg • Placebo	RCT, double-blind, multicentre (Australasia, Europe, North America)	OAB, ≥18 years	1,529	12
Yamaguchi 2007[61]	• Solifenacin 5 mg • Solifenacin 10 mg • Propiverine 20 mg • Placebo	Phase III RCT, double-blind, multicentre (Japan)	OAB, ≥20 years	1,584	12
Yamaguchi 2011[62]	• Fesoterodine 4 mg • Fesoterodine 8 mg • Placebo	Phase III RCT, double-blind, multicentre (Asia)	OAB, ≥20 years	951	12
Zinner 2002[63]	• Tolterodine ER 4 mg • Placebo	RCT, multicentre (Europe, USA, Canada, Australia, New Zealand)	OAB, ≥18 years	1,015	12

Abbreviations used in table: BD, twice daily; ER, extended-release; IR, immediate-release; mg, milligram; OAB, overactive bladder; RCT, randomised controlled trial; TDS, three times daily; wks, weeks.

[a] The Blossom study has been included as a comparator study only as the mirabegron doses analysed (50 mg BD and 100 mg BD) are not of interest for the submission.

4.1.5 Quality assessment

The manufacturer assessed the trials included in the direct clinical evidence and the MTC against criteria adapted from guidance for undertaking reviews in health care issued by the Centre for Reviews and Dissemination,[64] as provided in the NICE template for manufacturer/sponsor submission of evidence to the Single Technology Appraisal (STA) process.[65]

Direct clinical evidence

The manufacturer's quality assessments for the RCTs informing the direct clinical evidence are summarised in Table 6.

Table 6. Quality assessment results for randomised controlled trials evaluating mirabegron

Study no. (acronym)	**DRAGON, 178-CL-044**	**178-CL-045**	**178-CL-046 SCORPIO**	**178-CL-047 ARIES**	**178-CL-048**	**TAURUS, 178-CL-049**	**178-CL-074 CAPRICORN**
Was randomisation carried out appropriately?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the concealment of treatment allocation adequate?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were there any unexpected imbalances in drop-outs between groups?	No	No	No	No	Yes	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	No	No	No	No	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The ERG notes that the manufacturer's descriptions of how the quality issues were addressed in the studies were reproduced from the individual CSRs, but the ERG considers that, in some cases, insufficient details were provided on the methods used to minimise bias and ensure methodological rigour. The manufacturer states that all seven trials were randomised. However, the ERG considers that, for some trials, it is unclear how the randomisation sequence was generated (DRAGON, 178-CL-045, 178-CL-048), and how the random allocation was concealed (SCORPIO, ARIES, TAURUS, CAPRICORN, 178-CL-045). All the trials were described as being adequately blinded. However, the ERG notes that the methods implemented to achieve and maintain blinding were not described. The ERG agrees with the manufacturer that the baseline characteristics of the treatment groups in the different trials were well balanced.

All seven trials based the primary efficacy analysis on the full analysis set (FAS) rather than an intention-to-treat (ITT) population. The FAS population comprised all randomised patients who took ≥1 dose of double-blind study drug and who had a micturition measurement in the baseline diary and ≥1 post-baseline visit diary with a micturition measurement. The manufacturer asserts that use of the FAS population is consistent with other OAB trials, and the ERG considers the use of the FAS population appropriate. In SCORPIO, ARIES, CAPRICORN, and TAURUS missing data were handled using last observation carried forward (LOCF) methodology. The proportion of patients who discontinued was small in individual trials and rates were comparable across the studies with duration of treatment of 3 months (Table 7). The ERG notes that the rate of discontinuation was markedly higher in the long term study TAURUS, with about 23% discontinuing treatment at 12 months. The manufacturer's quality assessment, with accompanying comments from the ERG, is presented in Appendix 5.

Table 7. Discontinuation in the randomised controlled trials evaluating mirabegron

	Mirabegron 50 mg	Mirabegron 25 mg	Tolterodine	Placebo
178-CL-046 (SCORPIO)				
N randomised	497	N/A	495	497
N discontinued (%)	57 (11.5)	N/A	50 (10.1)	44 (8.9)
Adverse event	25	N/A	24	13
Lack of efficacy	6	N/A	3	5
Other reasons	26	N/A	23	26
178-CL-047 (ARIES)				
N randomised	442	N/A	N/A	454
N discontinued (%)	59 (13.3)	N/A	N/A	69 (15.2)
Adverse event	18	N/A	N/A	17
Lack of efficacy	1	N/A	N/A	9
Other reasons	40	N/A	N/A	43

178-CL-074 (CAPRICORN)				
N randomised	440	433	N/A	433
N discontinued (%)	54 (12.3)	45[a] (10.6)	N/A	66 (15.2)
Adverse event	12	17	N/A	15
Lack of efficacy	3	4	N/A	11
Other reasons	39	24	N/A	40
178-CL-049 (TAURUS)				
N randomised	815	N/A	813	N/A
N discontinued (%)	186 (22.8)	N/A	192 (23.6)	N/A
Adverse event	52	N/A	49	N/A
Lack of efficacy	34	N/A	45	N/A
Other reasons	100	N/A	98	N/A
178-CL-044 (DRAGON)				
N randomised	169	169	85	169
N discontinued (%)	NR	NR	NR	NR
Adverse event	3	10	2	5
Lack of efficacy	NR	NR	NR	NR
Other reasons	NR	NR	NR	NR
178-CL-045				
N randomised	208	211	N/A	214
N discontinued (%)	13 (6.3)	11 (5.2)	N/A	16 (7.5)
Adverse event	8	6	N/A	6
Lack of efficacy	NR	NR	N/A	NR
Other reasons	NR	NR	N/A	NR
178-CL-048				
N randomised	380	N/A	378	381
N discontinued (%)	31 (8.2)	N/A	23 (6.1)	31 (8.1)
Adverse event	15	N/A	13	9
Lack of efficacy	NR	N/A	NR	NR
Other reasons	NR	N/A	NR	NR

[a] Reported as 46 in the manufacturer's submission.

Abbreviations used in table: mg, milligram; N/A, not applicable; NR, not reported.

The manufacturer's quality assessment of the non-randomised trial can be found in Appendix 5.

Indirect clinical evidence

A summary of the manufacturer's quality assessments for the additional RCTs used to populate the network for the MTC can be found in Appendix 5. The manufacturer states that there are no doubts about the relevance of these trials when performing the MTC analyses. However, according to the manufacturer's quality assessment, a number of the studies had a high risk of bias for several of the quality questions, or it was unclear how the quality issues had been addressed. Due to time constraints, the ERG was unable to validate the quality assessment for each individual trial. However, based on the manufacturer's quality assessment of the trials, the ERG has concerns that it might be

inappropriate to include all the identified trials in the MTC. Issues regarding the quality of the trials included in the MTC are discussed in more detail in Section 4.4.

4.1.6 Description and critique of statistical approach and data synthesis, direct clinical evidence

Individual RCTs

SCORPIO, ARIES, CAPRICORN, and 178-CL-048 were powered to show superiority of mirabegron over placebo. The planned sample size for TAURUS was not based on a formal sample size calculation, but rather on an estimate of the number of patients who would enrol in this study on completion of either SCORPIO or ARIES. As noted in Section 4.2.1, the manufacturer states that the studies that included an active control arm (tolterodine) were not powered to detect superiority or non-inferiority of mirabegron over tolterodine.

In all seven trials, efficacy data were analysed based on the FAS rather than the ITT population. The rationale provided by the manufacturer, for the use of FAS, was that "in non-fatal conditions such as OAB at least one post-dose assessment is required for meaningful data about the study drug….In addition, the FAS population was required in order to conduct an MTC with the currently available antimuscarinics as the ITT population was not reported across all trials" (MS; pg 47).

The ERG considers the manufacturer's use of the FAS population to be appropriate and to be consistent with other studies of antimuscarinics. The ERG acknowledges the manufacturer's point that the ITT population was not reported across all trials of currently available antimuscarinics. The manufacturer performed sensitivity analysis on the primary outcome data from SCORPIO, ARIES, and CAPRICORN using the ITT population (urinary frequency and frequency of incontinence). The ITT population comprised all randomised patients who took ≥1 dose of double-blind study drug and who had a baseline diary with micturition measurements. Safety and adverse events were based on the safety analysis set (SAS), which was defined as all randomised patients who took ≥1 dose of double-blind study drug.

In SCORPIO, ARIES, and CAPRICORN all outcomes were analysed as change from baseline (CFB) to endpoint or final visit, and reported with standard error and 95% confidence interval (CI) for the FAS and ITT populations, respectively. On request, the manufacturer clarified that "endpoint" and "final visit" both refer to data acquired at the 12-week assessment, with missing values handled using LOCF. The manufacturer also reported the difference in CFB in the outcomes assessed between mirabegron and placebo, with standard error and 95% CI. However, for trials including a tolterodine active control group, the manufacturer did not present data for the difference in CFB between mirabegron and tolterodine. The manufacturer indicates that the individual trials were not powered to

detect superiority or non-inferiority of mirabegron versus tolterodine. The LOCF approach was also used in TAURUS to account for missing data.

In SCORPIO, ARIES, and CAPRICORN a multiplicity adjustment was used to account for the multiple outcomes and the resulting increased probability of type I errors. The ERG considers this reasonable.

Meta-analysis

The manufacturer did not perform a meta-analysis of the identified RCTs, with no rationale for this decision provided by the manufacturer. However, the manufacturer reported the results of a pre-specified pooled analysis of SCORPIO, ARIES, and CAPRICORN using the FAS populations, including subgroup analyses based on gender and previous treatment with antimuscarinics (i.e., yes versus no). The manufacturer pooled data for the placebo and mirabegron 50 mg treatment groups from the three RCTs; data from the mirabegron 25 mg treatment group in CAPRICORN were not pooled as CAPRICORN was the only study reported as direct clinical evidence evaluating mirabegron at this dose. Although the results of the manufacturer's pooled analysis augment the results for the individual trials on the clinical effectiveness of mirabegron compared with placebo, the ERG considers that the analysis does not fully inform the decision problem that is the focus of this STA as placebo is not a comparator of interest in the final scope. The ERG carried out a meta-analysis of data from RCTs evaluating mirabegron and tolterodine, which is listed as a comparator of interest; the results of the ERG's meta-analysis are described in Section 4.4.1.

The pooled analysis involved a multiplicity adjustment based on the included trials. The outcomes were analysed using analysis of covariance (ANCOVA) with treatment group, gender and study as factors and baseline values as a covariate. Stratified rank ANCOVA was used for hypothesis testing. For each endpoint variable the stratified rank ANCOVA was performed for the pairwise comparisons of mirabegron 50 mg versus placebo. No statistical assessment of heterogeneity was performed on the pooled analysis. Subgroup analyses were performed based on gender, and previously treated versus treatment-naïve patients, in accordance with the NICE final scope.

4.1.7 Description and critique of statistical approach and data synthesis, MTC

The manufacturer conducted a Bayesian MTC using a Markov Chain Monte Carlo (MCMC) simulation to estimate the relative efficacy and safety of mirabegron compared with all comparators listed in the final scope for this STA, and versus placebo. As the manufacturer states, oxybutynin and tolterodine are available in extended release (ER) and immediate release (IR) formulations. The two formulations were assumed to have similar efficacy (supported by the manufacturer's clinical expert),

and were therefore not separated for analyses on efficacy. However, they were separated for analyses on safety.

For each population, a fixed effect and a random effect model were used with a non-informative prior distribution allowing for correlation between different arms within multi-arm studies. The model with the best fit, as assessed by the deviance information criterion (DIC), was selected (i.e., the model with the lowest DIC). Tolterodine 4 mg was selected as the reference treatment for analyses of efficacy outcomes, as this treatment was the comparator in the health economic model of mirabegron and was also the most widely reported active comparator in the trials included in the MTC. For the analyses of safety outcomes, tolterodine 4 mg was selected as the reference treatment.

Input data

For continuous data, the mean changes and associated standard errors (SEs) reported in included trials were used in the MTC. If the SE was not reported, it was derived from the standard deviation of change, variance or CI around the mean, where available. When the mean change was not reported, it was calculated as the difference between mean at 12 weeks and mean at baseline, where available.

For the MTC of safety outcomes, dichotomous data were extracted, that is, the number of patients experiencing the specific event and the total number of patients in each group in the trial. If the number of patients experiencing the event was not reported in the trial, the number of patients with the event was estimated by using the reported percentage and the total number of patients in the group.

Model specifications

A normal likelihood with identity link was assumed for continuous outcomes (mean changes) and binomial likelihood with logit link was associated to the binary data (adverse events).

Vague priors were used for all the parameters in the MTC. In the analysis of efficacy outcomes and safety outcomes, a non-informative prior of N (0, 10^4) was used for the treatment effect and the study at baseline, with the exception of the analysis of dry mouth for which a non-informative prior of N (0, 100) was used.

The parameters in the distributions of random effects for between study correlation have vague prior distributions with Uniform (0, 5) for continuous data and Uniform (0, 2) for binary data.

The effect of treatment compared with mirabegron 50 mg was calculated directly in the model, as the difference between the effect of treatment and the effect of mirabegron 50 mg.

Having evaluated the WinBUGS code used by the manufacturer, the ERG is confident that an appropriate model structure has been used.

Assessment of model convergence

The convergence of models was assessed based on three diagnostics tools – Brooks-Gelman-Rubin diagnostic tool in WinBUGS, together with the inspection of the auto-correlation and history plots. The results of the diagnostic tools were not presented in the MS and the ERG has not had the opportunity to validate model convergence.

Test of inconsistency

Inconsistency was assessed using the node splitting method developed by Dias.[66] This method facilitates checking for consistency by comparing the direct and indirect evidence on each pairwise comparison (node) and shows how these combine in the MTC analysis.

The ERG is concerned about the results of the test for inconsistency (MS, Appendix 17) as significant inconsistency is identified in one or more treatment comparisons for: micturitions; incontinence episodes; and dry mouth.

Implementation of statistical analyses

Analyses were performed using WinBUGS version 1.4 statistical software (MRC Biostatistics Unit, Cambridge, UK). WinBUGS codes provided in the NICE Digital Support Unit Technical Support Document 2[67] were used for both fixed effect and random effect models. The code takes into account that data might come from multi-arm studies (studies with three or more arms).

In all MTC analyses, an initial burn-in of 100,000 iterations was discarded and all the results were based on a further sample of 350,000 iterations, with the exception of the analyses of adverse events using the random effects model, which were based on a further sample of 500,000 iterations.

The ERG has concerns around the number of iterations used for sampling the posterior distributions as this may be an indicator of poor mixing of data within the model. The high number of iterations may have been required to ensure convergence.

Assessment of heterogeneity

Heterogeneity was assessed by determining the DIC for each MTC analyses using both fixed and random effect models. The model with the lower DIC was chosen as it indicates the best fit. The ERG is confident that the DIC has been used appropriately by the manufacturer to identify the best fitting model. However, for the random effects models used, there is no estimate of the between pairwise comparisons heterogeneity (tau) presented by the manufacturer. At clarification, the ERG requested further information on the number of effective parameter estimates used within each analysis and the residual deviance (i.e., how well the model fitted the underlying data) – for a good fit the two values should be broadly similar. The data provided by the manufacturer are presented in Table 8.

Table 8. Overview of the number of effective parameters and residual deviance for the outcomes assessed in the MTC

Outcome	Number of effective parameters[a]	Residual deviance (by model)[b]	
		Fixed effects	Random effects
Micturition	39	49	51
Incontinence	61	29	33
Urgency incontinence	46	38	53
Dry mouth	103	590	580
Constipation	96	437	437
Blurred vision	63	227	227

[a] Estimated by the Evidence Review Group.

[b] As provided in the manufacturer's clarification response.

Overall, based on the potential for clinical and methodological heterogeneity, in addition to the evidence presented on inconsistency and model fit, the ERG believes that the results calculated by the manufacturer's MTC should be interpreted with caution.

4.1.8 Summary statement

The ERG considers the manufacturer's search strategies for RCT and non-RCT evidence to be appropriate and is satisfied that all relevant trials were identified. The inclusion and exclusion criteria used to identify relevant RCTs are also considered to be reasonable. However, in the ERG's opinion, the manufacturer excluded four eligible RCTs from the pooled analysis of direct clinical evidence, citing various reasons for exclusion. In brief, the manufacturer asserts that three of the RCTs were included as supporting evidence for the direct evidence submitted from ARIES, SCORPIO, and CAPRICORN; the manufacturer's rationale for exclusion of the four studies is summarised in Table 2. Details and data from one RCT excluded from the pooled analysis were reported in the main body of the MS but were considered separately from the three RCTs included in the pooled analysis. Details for the remaining three RCTs were presented in only the appendix of the MS. On reviewing the four RCTs, the ERG considers the RCTs to be relevant to the decision problem, and that it would have been appropriate to synthesise the reported data with that from the three key trials described by the manufacturer. The ERG has, therefore, summarised information and results from all seven trials in the report. The manufacturer also implemented additional inclusion and exclusion criteria for the selection of trials to inform the MTC. Several of the reasons were not clearly defined, making it difficult for the ERG to validate the appropriateness of the study selection.

The ERG considers the seven trials informing the direct evidence on the efficacy and safety of mirabegron to generally be well designed, although in several cases the methodology regarding randomisation, allocation concealment and blinding was not clear. Three trials compared mirabegron versus both placebo and tolterodine, three compared mirabegron versus only placebo and the remaining trial compared mirabegron versus tolterodine. In terms of patient baseline characteristics,

the groups were reasonably well balanced within trials and characteristics were comparable across the trials. The trials were conducted in different countries and ethnic populations, but the OAB characteristics of the included patients were comparable across the trials. The outcomes assessed in the trials and presented in the MS are clinically relevant and address the decision problem as outlined in the final scope issued by NICE. The manufacturer did not perform a meta-analysis of the included trials, but presented the results of a pre-specified pooled analysis of three of the trials.

The trials informing the direct clinical evidence only studied mirabegron compared with one of the comparators of interest listed in the scope, tolterodine. The manufacturer identified 40 trials that could inform an MTC, including all the comparators of interest as listed in the scope. The ERG notes that the seven trials evaluating mirabegron, including the four trials excluded by the manufacturer from the direct clinical evidence, were included in the trials deemed eligible for inclusion in the MTC. Most of the trials included in the MTC evaluated treatment in patients with OAB population, as in the key trials evaluating mirabegron. However, the ERG considers it important to note that some studies had a narrower population, including only women or investigating treatment of a specific symptom of OAB. Additionally, the manufacturer's quality assessment indicated considerable variability in study quality among the trials. The ERG considers it likely that variation in study quality and population could introduce considerable heterogeneity into the MTC carried out by the manufacturer. The ERG carried out an MTC excluding trials of poor methodological quality and those that included a population other than OAB; the results of the ERG's analysis are discussed in Section 4.4.2.

4.2 Summary and critique of the direct clinical evidence

4.2.1 Efficacy results

Within the MS, clinical evidence data are derived from three trials evaluating mirabegron in OAB: SCORPIO;[20] ARIES;[21] and CAPRICORN.[22] Long-term efficacy data from a 1-year study (TAURUS) are also presented.[26] Clinical evidence data from three additional trials, which the manufacturer indicates are included as supporting evidence (DRAGON,[23] 178-CL-045,[24] and 178-CL-048[25]), are included in Appendix 14 of the MS. For SCORPIO, ARIES, and CAPRICORN the manufacturer presents data on outcomes assessed based on CFB for the individual treatment groups within each trial, and the difference between mirabegron and placebo at the end point. However, for SCORPIO, analyses of the difference between mirabegron and the active control tolterodine for the outcomes assessed are not reported. The ERG considers that, although the results of effectiveness of mirabegron versus placebo are key as it is important to demonstrate that a treatment is more effective than placebo, analysis of mirabegron versus active treatment currently used in the NHS is of more relevance to the decision problem To facilitate meta-analysis, at clarification, the ERG requested additional data for the comparison of mirabegron versus tolterodine for the relevant trials (SCORPIO, TAURUS, DRAGON, and 178-CL-048); TAURUS had a longer duration of follow-up compared

with the other trials and, thus, to give an indication of the longer term efficacy of mirabegron, the ERG requested data at both the 3 and 12 month time point. In response to the ERG's request, the manufacturer provided data for the comparison of mirabegron versus tolterodine from DRAGON and 178-CL-048, but not from TAURUS or SCORPIO. The manufacturer maintained that no comparisons of mirabegron 50 mg versus tolterodine ER 4 mg were conducted in SCORPIO as the study was not powered to test superiority or non-inferiority of mirabegron versus tolterodine.

Due to the large number of included studies and multiple outcomes, the ERG has limited the summary of the clinical evidence to the outcomes stated in the scope of this STA, and the additional outcome of incontinence frequency, as data from this outcome inform the economic model. The ERG reports data for the comparison of mirabegron versus placebo for information, but, as noted earlier, considers that results of mirabegron versus tolterodine are more relevant to the decision problem.

Urinary frequency

Urinary frequency was measured as the mean number of micturitions per 24 hours. In the three trials from which the direct clinical evidence is derived, mirabegron 50 mg was associated with a statistically significant reduction in number of micturitions per 24 hours at 3 months' follow-up compared with placebo (p-value <0.05; Table 9). The results from the individual trials are in accordance with the manufacturer's pooled analysis of SCORPIO, ARIES, and CAPRICORN (MD –0.55; 95% CI: –0.75 to –0.36; p-value <0.001). However, two trials (DRAGON, and 178-CL-045) that compared mirabegron 50 mg versus tolterodine, and for which data are available, found no statistically significant difference between the treatments in urinary frequency (Table 9). When evaluating the lower dose of mirabegron (25 mg), in CAPRICORN, mirabegron was associated with a statistically significant reduction in urinary frequency compared with placebo. Results from 178-CL-045 support the finding from CAPRICORN. However, the DRAGON trial found no statistically significant difference between mirabegron 25 mg and placebo or between mirabegron 25 mg and tolterodine for this outcome.

Table 9. Change from baseline to final visit in mean number of micturitions per 24 hours

Study	Mirabegron 50 mg vs tolterodine		Mirabegron 50 mg vs placebo		Mirabegron 25 mg vs tolterodine		Mirabegron 25 mg vs placebo	
	MD	95% CI	MD	95% CI	MD	95% CI	MD	95% CI
178-CL-046 (SCORPIO)	NR	NR	–0.60	–0.90 to –0.29	N/A	N/A	N/A	N/A
178-CL-047 (ARIES)	N/A	N/A	–0.61	–0.98 to –0.24	N/A	N/A	N/A	N/A
178-CL-074 (CAPRICORN)	N/A	N/A	–0.42	–0.76 to –0.08	N/A	N/A	–0.47	–0.82 to –0.13
178-CL-044 (DRAGON)	–0.14	–0.80 to 0.53	–0.64	–1.19 to –0.10	0.06	–0.60 to 0.72	–0.45	–0.99 to 0.10
178-CL-045	N/A	N/A	–0.74	–1.12 to –0.36	N/A	N/A	–0.66	–1.04 to –0.28
178-CL-048	–0.25	–0.55 to 0.04	–0.86	–1.16 to –0.57	N/A	N/A	N/A	N/A
Manufacturer's pooled analysis			–0.55	–0.75 to –0.36				
			Mirabegron 50 mg, N=1,324 Placebo, N=1,328					
Abbreviations used in table: CI, confidence interval; MD, mean difference; mg, milligram; NA, not applicable; NR, not reported.								

Frequency of incontinence

Frequency of incontinence was recorded as the mean number of incontinence episodes per 24 hours, and was assessed in FAS patients who had ≥1 incontinence episode at baseline (FAS-I population). On request, the manufacturer clarified that incontinence included any involuntary leakage of urine. Treatment with mirabegron 50 and 25 mg led to a statistically significant reduction in the number of incontinence episodes per 24 hours at 3 months' follow-up compared with placebo in each of the six trials individually (p-value <0.05; Table 10) and in the manufacturer's pooled analysis of ARIES, and CAPRICORN (MD –0.40; 95% CI: –0.58 to –0.21; p-value <0.001). However, in two RCTs (DRAGON and 178-CL-048), the difference in frequency of incontinence between mirabegron 50 or 25 mg and tolterodine was not statistically significant (Table 10).

Frequency of urge urinary incontinence

Frequency of urge urinary incontinence was recorded as the mean number of urgency incontinence episodes per 24 hours, and was assessed in the FAS-I population. On request, the manufacturer clarified that urgency incontinence included involuntary leakage of urine accompanied or preceded by urgency. With similar results to the frequency of incontinence episodes, all six trials comparing mirabegron 50 mg versus placebo found mirabegron was associated with a statistically significant decrease in frequency of urgency incontinence per 24 hours (p-value <0.05; Table 11). The manufacturer's pooled analysis of data from SCORPIO, ARIES, and CAPRICORN is in agreement with the results from the individual trials (MD –0.40; 95% CI: –0.57 to –0.23; p-value <0.001). Results of analysis of data from three RCTs (CAPRICORN, and DRAGON, 178-CL-045) comparing mirabegron 25 mg with placebo favour mirabegron, but the difference reaches statistical significance in only two RCTs (CAPRICORN, and DRAGON). Again, in two RCTs (DRAGON, and 178-CL-048), the difference between mirabegron 25 mg or 50 mg and tolterodine was found to be not statistically significant.

Level of urgency

After 12 weeks' treatment, the level of urgency was lower for patients on mirabegron 50 mg compared with placebo, and the difference was statistically significant in three out of four studies from which data are available for this outcome (Table 12); mean level of urgency was not assessed in the 178-CL-045 or 178-CL-048 trial. The manufacturer's pooled analysis of SCORPIO, ARIES, and CAPRICORN showed a statistically significant decrease in the level of urgency between mirabegron and placebo (MD –0.11; 95% CI: –0.16 to –0.07; p-value <0.001). There was no statistically significant difference between mirabegron 50 or 25 mg and tolterodine (DRAGON, and 178-CL-048), or between mirabegron 25 mg and placebo (CAPRICORN, and DRAGON).

Table 10. Change from baseline to final visit in mean number of incontinence episodes per 24 hours

Study	Mirabegron 50 mg vs tolterodine		Mirabegron 50 mg vs placebo		Mirabegron 25 mg vs tolterodine		Mirabegron 25 mg vs placebo	
	MD	95% CI	MD	95% CI	MD	95% CI	MD	95% CI
178-CL-046 (SCORPIO)	NR	NR	–0.41	–0.72 to –0.09	N/A	N/A	N/A	N/A
178-CL-047 (ARIES)	N/A	N/A	–0.34	–0.66 to –0.03	N/A	N/A	N/A	N/A
178-CL-074 (CAPRICORN)	N/A	N/A	–0.42	–0.76 to –0.08	N/A	N/A	–0.40	–0.74 to –0.06
178-CL-044 (DRAGON)	–0.34	–1.06 to 0.39	–0.62	–1.22 to –0.02	–0.56	–1.29 to 0.18	–0.84	–1.45 to –0.23
178-CL-045	N/A	N/A	–0.40	–0.67 to –0.13	N/A	N/A	–0.39	–0.67 to –0.11
178-CL-048	–0.10	-0.36 to 0.15	-0.42	-0.67 to -0.17	N/A	N/A	N/A	N/A
Manufacturer's pooled analysis			–0.40	–0.58 to –0.21				
			Mirabegron 50 mg, N=862 Placebo, N=878					
Abbreviations used in table: CI, confidence interval; MD, mean difference; mg, milligram; N/A, not applicable; NR, not reported.								

Table 11. Change from baseline to final visit in mean number of urgency incontinence episodes per 24 hours

Study	Mirabegron 50 mg vs tolterodine		Mirabegron 50 mg vs placebo		Mirabegron 25 mg vs tolterodine		Mirabegron 25 mg vs placebo	
	MD	95% CI	MD	95% CI	MD	95% CI	MD	95% CI
178-CL-046 (SCORPIO)	NR	NR	–0.35	–0.65 to –0.05	N/A	N/A	N/A	N/A
178-CL-047 (ARIES)	N/A	N/A	–0.43	–0.72 to –0.15	N/A	N/A	N/A	N/A
178-CL-074 (CAPRICORN)	N/A	N/A	–0.39	–0.69 to –0.08	N/A	N/A	–0.36	–0.67 to –0.05
178-CL-044 (DRAGON)	–0.37	–0.99 to 0.24	–0.69	–1.18 to –0.19	–0.55	–1.18 to 0.07	–0.86	–1.38 to –0.35
178-CL-045	N/A	N/A	–0.27	–0.53 to –0.01	N/A	N/A	–0.24	–0.51 to 0.02
178-CL-048	–0.04	–0.28 to 0.21	–0.36	–0.59 to –0.12	N/A	N/A	N/A	NA
Manufacturer's pooled analysis			–0.40	–0.57 to –0.23				
			Mirabegron 50 mg, N=862 Placebo, N=878					
Abbreviations used in table: CI, confidence interval; MD, mean difference; mg, milligram; NA, not applicable; NR, not reported.								

Table 12. Change from baseline to final visit in mean level of urgency

Study	Mirabegron 50 mg vs tolterodine		Mirabegron 50 mg vs placebo		Mirabegron 25 mg vs tolterodine		Mirabegron 25 mg vs placebo	
	MD	**95% CI**	**MD**	**95% CI**	**MD**	**95% CI**	**MD**	**95% CI**
178-CL-046 (SCORPIO)	NR	NR	–0.09	–0.16 to –0.01	N/A	N/A	N/A	N/A
178-CL-047 (ARIES)	N/A	N/A	–0.11	–0.18 to –0.04	N/A	N/A	N/A	N/A
178-CL-074 (CAPRICORN)	N/A	N/A	–0.14	–0.22 to –0.06	N/A	N/A	–0.07	–0.15 to 0.01
178-CL-044 (DRAGON)	–0.08	–0.22 to 0.05	–0.04	–0.20 to 0.13	–0.07	–0.23 to 0.10	–0.12	–0.25 to 0.02
Manufacturer's pooled analysis			–0.11	–0.16 to –0.07				
			Mirabegron 50 mg, N=1,324 Placebo, N=1,328					
Abbreviations used in table: CI, confidence interval; MD, mean difference; mg, milligram; N/A, not applicable; NR, not reported.								

Number of urgency episodes

The number of urgency episodes (Grade 3 or 4) per 24 hours was lower in the mirabegron 25 and 50 mg groups compared with the placebo group at 12 weeks' follow-up in all trials, although the difference did not reach statistical significance in all trials (Table 13). However, in the manufacturer's pooled analysis of SCORPIO, ARIES, and CAPRICORN, the difference between mirabegron 50 mg and placebo favoured mirabegron and was statistically significant (MD –0.64; 95% CI: –0.89 to –0.39; p-value <0.001). However, in the two RCTs evaluating mirabegron and tolterodine (DRAGON, and 178-CL-048), there was no statistically significant difference in the number of urgency episodes per 24 hours between mirabegron (25 and 50 mg) and tolterodine.

Nocturia

Treatment with mirabegron 50 mg and mirabegron 25 mg was associated with reduction in nocturia (i.e., waking at night one or more times to void) episodes per 24 hours compared placebo at 12 weeks in all trials (Table 14). However, the difference between groups reached statistical significance in only three (out of six) trials and one trial (out of three) for mirabegron 50 mg and mirabegron 25 mg, respectively. In the manufacturer's pooled analysis of SCORPIO, ARIES, and CAPRICORN, mirabegron 50 mg was associated with a statistically significant reduction in nocturia episodes per 24 hours compared with placebo (MD –0.14; 95% CI: –0.23 to –0.05; p-value = 0.003). However, there was no statistically significant difference in the number of nocturia episodes per 24 hours between mirabegron (50 mg or 25 mg) and tolterodine in two RCTs (DRAGON, and 178-CL-048; Table 14).

Table 13. Change from baseline to final visit in mean number of urgency episodes (Grade 3 or 4) per 24 hours

Study	Mirabegron 50 mg vs tolterodine		Mirabegron 50 mg vs placebo		Mirabegron 25 mg vs tolterodine		Mirabegron 25 mg vs placebo	
	MD	95% CI	MD	95% CI	MD	95% CI	MD	95% CI
178-CL-046 (SCORPIO)	NR	NR	–0.60	–1.02 to –0.18	N/A	N/A	N/A	N/A
178-CL-047 (ARIES)	N/A	N/A	–0.75	–1.20 to –0.30	N/A	N/A	N/A	N/A
178-CL-074 (CAPRICORN)	N/A	N/A	–0.59	–1.01 to –0.16	N/A	N/A	–0.33	–0.76 to 0.10
178-CL-044 (DRAGON)	–0.22	–1.06 to 0.62	–0.60	–1.29 to 0.08	–0.31	–1.14 to 0.52	–0.70	–1.38 to –0.01
178-CL-045	N/A	N/A	–0.29	–0.77 to 0.19	N/A	N/A	–0.27	–0.75 to 0.21
178-CL-048	–0.13	–0.49 to 0.23	–0.54	–0.90 to –0.18	N/A	N/A	N/A	N/A
Manufacturer's pooled analysis			–0.64	–0.89 to –0.39				
			Mirabegron 50 mg, N=1,324 Placebo, N=1,328					
Abbreviations used in table: CI, confidence interval; MD, mean difference; mg, milligram; N/A, not applicable; NR, not reported.								

Table 14. Change from baseline to final visit in mean number of nocturia episodes per 24 hours

Study	Mirabegron 50 mg vs tolterodine		Mirabegron 50 mg vs placebo		Mirabegron 25 mg vs tolterodine		Mirabegron 25 mg vs placebo	
	MD	95% CI	MD	95% CI	MD	95% CI	MD	95% CI
178-CL-046 (SCORPIO)	NR	NR	–0.15	–0.28 to –0.02	N/A	N/A	N/A	N/A
178-CL-047 (ARIES)	N/A	N/A	–0.18	–0.36 to –0.01	N/A	N/A	N/A	N/A
178-CL-074 (CAPRICORN)	N/A	N/A	–0.04	–0.20 to 0.12	N/A	N/A	–0.01	–0.17 to 0.15
178-CL-044 (DRAGON)	–0.01	–0.27 to 0.25	–0.22	–0.44 to –0.01	0.06	–0.20 to 0.32	–0.15	–0.36 to 0.07
178-CL-045	N/A	N/A	–0.16	–0.33 to 0.00	N/A	N/A	–0.20	–0.36 to –0.04
178-CL-048	–0.02	–0.15 to 0.11	–0.12	–0.25 to 0.01	N/A	N/A	N/A	N/A
Manufacturer's pooled analysis			–0.14	–0.23 to –0.05				
			Mirabegron 50 mg, N=1,324 Placebo, N=1,328					
Abbreviations used in table: CI, confidence interval; MD, mean difference; mg, milligram; N/A, not applicable; NR, not reported.								

Health-related quality of life

The manufacturer presents assessments of HRQoL based on multiple questionnaires, including both generic (e.g., EQ-5D) and disease-specific questionnaires (e.g., OAB-q). The ERG considers it important to note that NICE has specified the EQ-5D as its preferred method of utility measurement. However, the ERG recognises that the use of generic questionnaires across a wide range of health conditions results in the loss of sensitivity to clinically important changes in health when applied to a specific patient population, such as patients with OAB. For completeness, the ERG presents and discusses the results for both the EQ-5D and the OAB-q as reported by the manufacturer.

EQ-5D

Quality of life was assessed using EQ-5D in the modified-intention-to-treat (m-ITT) population (3,741 patients). The mITT population comprised all study patients who were randomised, received at least one dose of double-blind study medication and completed the EQ-5D questionnaire at baseline and at least once post-baseline, excluding any patients who presented serious deviations from the protocol or for whom the EQ-5D questionnaire data was not available at 12 weeks.

The manufacturer states that there were no significant differences between treatment groups in SCORPIO, ARIES, and CAPRICORN. No EQ-5D data from the individual trials are presented in the MS; however, a post-hoc pooled analysis of SCORPIO, ARIES, and CAPRICORN showed no statistically significant difference in change from baseline utility score between mirabegron and placebo at 12 weeks' follow-up (MD 0.007; p-value = 0.30). By contrast, treatment with mirabegron 50 mg was found to be superior to tolterodine ER 4 mg in terms of change from baseline utility score at 12 weeks (MD 0.019; p-value ≤0.05; MS, pg 104).

OAB-q HRQoL

In SCORPIO, ARIES, and CAPRICORN, treatment with mirabegron was found to improve a patients' quality of life more than placebo, based on the disease-specific quality of life measure OAB-q, with the difference reaching statistical significance in SCORPIO and ARIES (summarised in Table 15). However, the ERG considers it important to note that a difference of 10 points from baseline has been suggested to represent a minimally important difference on the OAB-q.[68] Data reported in the MS (summarised in the footnote to Table 15) indicate that OAB-q score improved by at least 10 points from baseline in all groups relevant to the decision problem (mirabegron 50 mg and 25 mg and tolterodine ER 4 mg), as well as in the placebo groups.

Table 15. OAB-q HRQoL total score for SCORPIO, ARIES, and CAPRICORN

Study	Mirabegron 50 mg vs tolterodine ER 4 mg		Mirabegron 50 mg vs placebo		Mirabegron 25 mg vs tolterodine		Mirabegron 25 mg vs placebo	
	MD	95% CI	MD	95% CI	MD	95% CI	MD	95% CI
178-CL-046 (SCORPIO)	NR	NR	2.3	0.2 to 4.5[a]	N/A	N/A	N/A	N/A
178-CL-047 (ARIES)	N/A	N/A	4.1	1.6 to 6.6[b]	N/A	N/A	N/A	N/A
178-CL-074 (CAPRICORN)	N/A	N/A	1.2	–1.0 to 3.4[c]	N/A	N/A	1.3	–0.9 to 3.5[d]

[a] Adjusted mean CFB to final visit in OAB-q: 16.1 in the mirabegron 50 mg group versus 13.7 in the placebo group.

[b] Adjusted mean CFB to final visit in OAB-q: 14.8 in the mirabegron 50 mg group versus 10.7 in the placebo group.

[c] Adjusted mean CFB to final visit in OAB-q: 14.3 in the mirabegron 50 mg group versus 13.0 in the placebo group.

[d] Adjusted mean CFB to final visit in OAB-q: 16.1 in the mirabegron 50 mg group versus 14.8 in the tolterodine group.

Abbreviations used in table: CI, confidence interval; ER, extended release; HRQoL, health-related quality of life; MD, mean difference; mg, milligram; NA, not applicable; NR, not reported; OAB-q, overactive bladder questionnaire.

Non-randomised trial

Key efficacy and safety results for 178-CL-051 are provided in Appendix 6. The trial, in which 202 patients were treated with mirabegron 50 mg for one year (with an optional dose increase to 100 mg at week eight), showed that improvements in number of micturitions, urgency episodes, incontinence episodes, urge incontinence episodes, and nocturia episodes, were maintained until the end of study. There were no major differences in the incidence of adverse events or treatment-related adverse events between the subjects maintained at 50 mg and those increased to 100 mg, and most adverse events were of mild severity.

4.2.2 Safety results

Safety data for mirabegron were based primarily on the long term (1 year) TAURUS study,[(26)] which evaluated the safety of mirabegron at doses of 50 mg (recommended dose) and 100 mg versus tolterodine ER 4 mg. The ERG considers it appropriate to focus on safety data from TAURUS as the primary objective of the study was to assess the safety of mirabegron. However, the ERG notes that no relative risk or risk difference with associated 95% CIs was presented for any of the safety data.

In TAURUS, the proportion of patients experiencing an adverse event was similar between mirabegron 50 mg (60%) and tolterodine ER 4 mg (63%) (Table 16). Similarly, the proportion of patients discontinuing treatment as the result of an adverse event was similar in the two groups (5.9% in the mirabegron 50 mg group vs 5.7 % in the tolterodine ER 4 mg group).

Table 16. Overview of treatment-emergent adverse events, TAURUS (adapted from MS; Table 64, pg 160)

Adverse event	Mirabegron 50 mg N=812	Tolterodine ER 4 mg N=812
TEAEs	485 (59.7)	508 (62.6)
Mild	222 (27.3)	251 (30.9)
Moderate	212 (26.1)	218 (26.8)
Severe	51 (6.3)	39 (4.8)
Treatment-related TEAEs	213 (26.2)	224 (27.6)
Deaths	2 (0.2)	2 (0.2)
SAEs	42 (5.2)	44 (5.4)
Treatment-related SAEs	10 (1.2)	5 (0.6)
TEAEs leading to study drug discontinuation	48 (5.9)	46 (5.7)
Treatment-related TEAEs leading to study drug discontinuation	35 (4.3)	31 (3.8)
Abbreviations used in table: AE, adverse event; ER, extended-release; mg, milligram; SAE, serious adverse event; TEAE, treatment-emergent adverse event.		

Among the adverse events the manufacturer considered treatment related, the most common adverse events occurring in ≥2% of patients in any group included hypertension, dry mouth, constipation and headache (Table 17). The proportion of patients experiencing the commonly occurring adverse events was similar for the mirabegron 50 mg and tolterodine ER 4 mg groups, with the exception of dry mouth which occurred more frequently in the tolterodine ER 4 mg group than in the mirabegron 50 mg group (Table 17).

Table 17. Common treatment-related treatment-emergent adverse events occurring in ≥2% of patients in any treatment group, TAURUS (adapted from MS; Table 66, pg 161)

MedDRA (v9.1) preferred term	Mirabegron 50 mg N=812	Tolterodine ER 4 mg N=812
Hypertension	43 (5.3)	42 (5.2)
Dry mouth	20 (2.5)	67 (8.3)
Constipation	18 (2.2)	19 (2.3)
Headache	18 (2.2)	14 (1.7)
Abbreviations used in table: ER, extended-release; mg, milligram.		

Safety data from other key RCTs evaluating mirabegron

Safety data from the six trials with a treatment duration of 3 months (SCORPIO, ARIES, CAPRICORN, DRAGON, 178-CL-045, and 178-CL-048) indicated a similar frequency of occurrence of adverse events to those observed in TAURUS (summarised in Table 18). However, a larger proportion of patients in 178-CL-045 and 178-CL-048 experienced adverse events than in the other studies. The number of serious adverse events was low in all trials.

Table 18. Overview of safety data from randomised controlled trials evaluating mirabegron with a treatment duration of 3 months

Study	Mirabegron 50 mg	Mirabegron 25 mg	Tolterodine ER 4 mg	Placebo
Adverse events				
178-CL-046 (SCORPIO)	211 (42.8)	N/A	231 (46.7)	214 (43.3)
178-CL-047 (ARIES)	228 (51.6)	N/A	N/A	227 (50.1)
178-CL-074 (CAPRICORN)	208 (47.3)	210 (48.6)	N/A	217 (50.1)
178-CL-044 (DRAGON)	74 (43.8)	74 (43.8)	41 (48.2)	73 (43.2)
178-CL-045	171 (82.2)	169 (80.5)	N/A	157 (74.1)
178-CL-048	281 (74.1)	N/A	305 (81.3)	292 (77.0)
Treatment-related adverse event				
178-CL-046 (SCORPIO)	100 (20.3)	N/A	131 (26.5)	89 (18.0)
178-CL-047 (ARIES)	80 (18.1)	N/A	N/A	66 (14.6)
178-CL-074 (CAPRICORN)	76 (17.3)	87 (20.1)	N/A	77 (17.8)
178-CL-044 (DRAGON)	38 (22.5)	34 (20.1)	13 (15.3)	26 (15.4)
178-CL-045	51 (24.5)	49 (23.3)	N/A	40 (18.9)
178-CL-048	93 (24.5)	N/A	131 (34.9)	91 (24.0)
SAE				
178-CL-046 (SCORPIO)	14 (2.8)	N/A	11 (2.2)	8 (1.6)
178-CL-047 (ARIES)	11 (2.5)	N/A	N/A	9 (2.0)
178-CL-074 (CAPRICORN)	4 (0.9)	7 (1.6)	N/A	12 (2.8)
178-CL-044 (DRAGON)	1 (0.6)	1 (0.6)	1 (1.2)	1 (0.6)
178-CL-045	1 (0.5)	3 (1.4)	N/A	4 (1.9)
178-CL-048	3 (0.8)	N/A	4 (1.1)	4 (1.1)
Discontinuations due to adverse event				
178-CL-046 (SCORPIO)	24 (4.9)	N/A	22 (4.4)	13 (2.6)
178-CL-047 (ARIES)	18 (4.1)	N/A	N/A	17 (3.8)
178-CL-074 (CAPRICORN)	11 (2.5)	17 (3.9)	N/A	16 (3.7)
178-CL-044 (DRAGON)	4 (2.4)	9 (5.3)	1 (1.2)	5 (3.0)
178-CL-045	7 (3.4)	5 (2.4)	N/A	4 (1.9)
178-CL-048	12 (3.2)	N/A	12 (3.2)	8 (2.1)
Abbreviations used in table: ER, extended release; mg, milligram; N/A, not applicable; SAE, serious adverse event.				

Considering adverse events occurring in ≥2% of patients, frequency of treatment-related adverse events in SCORPIO, ARIES, and CAPRICORN was similar to those observed in TAURUS (summarised in Table 19). However, of the three trials submitted by the manufacturer as direct evidence on clinical effectiveness, incidence of dry mouth occurred in ≥2% of patients (any treatment group) in only SCORPIO. In SCORPIO, the incidence of dry mouth was similar in the mirabegron 50 mg group and the placebo group, but considerably higher in the tolterodine ER 4 mg group, which mirrors the results observed in TAURUS.

Table 19. Common treatment-related treatment-emergent adverse events in ≥2% of patients in any treatment group, SCORPIO, ARIES, and CAPRICORN

Study	Mirabegron 50 mg	Mirabegron 25 mg	Tolterodine ER 4 mg	Placebo
Hypertension				
178-CL-046 (SCORPIO)	20 (4.1)	N/A	30 (6.1)	23 (4.7)
178-CL-047 (ARIES)	14 (3.2)	N/A	N/A	17 (3.8)
178-CL-074 (CAPRICORN)	31 (7.0)	30 (6.9)	N/A	23 (5.3)
Dry mouth				
178-CL-046 (SCORPIO)	9 (1.8)	N/A	47 (9.5)	9 (1.8)
Headache				
178-CL-046 (SCORPIO)	13 (2.6)	N/A	11 (2.2)	6 (1.2)
178-CL-047 (ARIES)	11 (2.5)	N/A	N/A	3 (0.7)
178-CL-074 (CAPRICORN)	4 (0.9)	4 (0.9)	N/A	9 (2.1)
Abbreviation used in table: ER, extended release; mg, milligram; N/A, not applicable.				

The manufacturer also provides a summary of treatment-emergent adverse events (TEAEs) of interest, which included cardiovascular type events (hypertension, Torsades de Pointes/QTc prolongation events, cardiac arrhythmias), urinary retention type events, hypersensitivity type events, syncope/seizure type events and hepatic type events. Within the MS, the manufacturer did not indicate how these effects had been determined to be of special interest. However, the CSRs of ARIES,[21] CAPRICORN,[22] and SCORPIO[20] state that TEAEs of interest were identified based on observations from nonclinical and clinical studies of mirabegron. In patients with OAB, adverse events identified by the sponsor as TEAEs of interest those presented by the manufacturer in the submission. TEAEs of interest associated with mirabegron are presented in Table 20. The ERG notes that the proportion of patients experiencing the individual TTEAEs of interest is similar for mirabegron 50 mg and tolterodine ER 4 mg.

Table 20. Treatment-emergent adverse events of interest in SCORPIO, ARIES, and CAPRICORN (adapted from MS; Tables 73 [pg 168], 76 [pg 170], and 79 [pg 172])

Study	Mirabegron 50 mg	Mirabegron 25 mg	Tolterodine ER 4 mg	Placebo
Hypertension type				
178-CL-046 (SCORPIO)	38 (7.7)	N/A	47 (9.5)	46 (9.3)
178-CL-047 (ARIES)	33 (7.5)	N/A	N/A	32 (7.1)
178-CL-074 (CAPRICORN)	49 (11.1)	52 (12.0)	N/A	37 (8.5)
Torsades de Pointes/QTc prolongation type				
178-CL-046 (SCORPIO)	0 (0)	N/A	2 (0.4)	0 (0)
178-CL-047 (ARIES)	0 (0)	N/A	N/A	0 (0)
178-CL-074 (CAPRICORN)	0 (0)	0 (0)	N/A	0 (0)
Cardiac arrhythmia				
178-CL-046 (SCORPIO)	11 (2.2)	N/A	16 (3.2)	5 (1.0)
178-CL-047 (ARIES)	9 (2.0)	N/A	N/A	4 (0.9)
178-CL-074 (CAPRICORN)	13 (3.0)	13 (3.0)	N/A	11 (2.5)

Urinary retention				
178-CL-046 (SCORPIO)	1 (0.2)	N/A	3 (0.6)	3 (0.6)
178-CL-047 (ARIES)	0 (0)	N/A	N/A	3 (0.7)
178-CL-074 (CAPRICORN)	0 (0)	0 (0)	N/A	1 (0.2)
Hypersensitivity				
178-CL-046 (SCORPIO)	22 (4.5)	N/A	20 (4.0)	16 (3.2)
178-CL-047 (ARIES)	16 (3.6)	N/A	N/A	23 (5.1)
178-CL-074 (CAPRICORN)	13 (3.0)	15 (3.5)	N/A	15 (3.5)
Syncope/seizure				
178-CL-046 (SCORPIO)	0 (0)	N/A	1 (0.2)	0 (0)
178-CL-047 (ARIES)	0 (0)	N/A	N/A	0 (0)
178-CL-074 (CAPRICORN)	0 (0)	0 (0)	N/A	2 (0.5)
Hepatic disorders				
178-CL-046 (SCORPIO)	11 (2.2)	N/A	10 (2.0)	7 (1.4)
178-CL-047 (ARIES)	6 (1.4)	N/A	N/A	5 (1.1)
178-CL-074 (CAPRICORN)	4 (0.9)	6 (1.4)	N/A	5 (1.2)
Abbreviations used in table: ER, extended release; mg, milligram; N/A, not applicable.				

4.2.3 Subgroup analysis results

Based on the final scope,[19] the manufacturer reported subgroup analyses of:

- men and women;
- previously untreated and previously treated OAB.

The manufacturer reported data for primary outcomes evaluated in the three key trials submitted as direct clinical evidence, that is, frequency of micturition per 24 hours and of incontinence episodes per 24 hours, for the comparison of mirabegron versus placebo. Subgroup analyses of mirabegron versus tolterodine were not reported.

Gender

Compared with placebo, treatment with mirabegron resulted in a statistically significant reduction in the mean number of micturitions per 24 hours, from baseline to final visit, for both male and female patients (Table 21). Among women, there were significantly fewer episodes of incontinence in the mirabegron group than in the placebo group (Table 22). However, in men, there was so statistically significant difference between mirabegron and placebo for this outcome. There was no statistically significant difference between the subgroups for either outcome (gender interaction p-value for micturitions was 0.16, and for incontinence episodes was 0.22).

Table 21. Change from baseline to final visit in mean number of micturitions per 24 hours, by gender, pre-specified pooled analysis of SCORPIO, ARIES, and CAPRICORN (adapted from MS; Table 43, pg 109)

	Placebo	Mirabegron 50 mg
Males		
n at baseline	362	382
Adjusted mean CFB (SE) 95% CI	–0.92 (0.135) –1.18 to –0.66	–1.29 (0.131) 1.55[a] to –1.04
Adjusted mean difference vs placebo (SE) 95% CI	N/A N/A	–0.37 (0.187) –0.74 to –0.01
Females		
n at baseline	966	942
Adjusted mean CFB (SE) 95% CI	–1.31 (0.082) –1.47 to –1.15	–1.93 (0.084) –2.09 to –1.77
Adjusted mean difference vs placebo (SE) 95% CI	N/A N/A	–0.62 (0.117) –0.85 to –0.39
Gender interaction p-value	**0.16**	

[a] As reported in the manufacturer's submission. The Evidence Review Group considers that this is potentially a typographical error and should perhaps read –1.55.

Abbreviations used in table: CFB, change from baseline; CI, confidence interval; mg, milligram; N/A, not applicable; SE, standard error.

Table 22. Change from baseline to final visit in mean number of incontinence episodes per 24 hours, by gender, pre-specified pooled analysis of SCORPIO, ARIES, and CAPRICORN (adapted from MS; Table 42, pg 107)

	Placebo	Mirabegron 50 mg
Males		
n at baseline	154	168
Adjusted mean CFB (SE) 95% CI	–1.41 (0.159) –1.72 to –1.10	–1.48 (0.152) –1.78 to –1.18
Adjusted mean difference vs placebo (SE) 95% CI	N/A N/A	–0.07 (0.220) –0.50 to 0.36
Females		
n at baseline	724	694
Adjusted mean CFB (SE) 95% CI	–1.03 (0.074) –1.17 to –0.89	–1.50 (0.075) –1.65 to –1.35
Adjusted mean difference vs placebo (SE) 95% CI	N/A N/A	–0.47 (0.105) –0.67 to –0.26
Gender interaction p-value	**0.22**	

Abbreviations used in table: CFB, change from baseline; CI, confidence interval; mg, milligram; N/A, not applicable; SE, standard error.

Previously treated versus treatment naïve patients

In the subgroup analysis of previously treated versus treatment-naïve patients, mirabegron 50 mg significantly reduced the mean number of micturitions per 24 hours in both populations compared with placebo (Table 2). In patients who had received prior treatment for OAB, mirabegron 50 mg was also associated with a statistically significant decrease in the number of incontinence

episodes (Table 24). However, in treatment naïve patients, the difference between mirabegron 50 mg and placebo was not statistically significant for this outcome (Table 24). There was no statistically significant difference between the subgroups for either outcome (previously treated vs treatment naïve interaction p-value for micturitions was 0.10 and for incontinence episodes was 0.095).

Table 23. Change from baseline to final visit in mean number of micturitions per 24 hours, previously treated versus treatment-naïve patients, pre-specified pooled analysis of SCORPIO, ARIES, and CAPRICORN (adapted from MS; Table 45, pg 113)

	Placebo	**Mirabegron 50 mg**
Previously treated		
n	704	688
Adjusted mean CFB (SE) 95% CI	–0.93 (0.097) –1.12 to –0.74	–1.67 (0.098) –1.86 to –1.48
Adjusted mean difference vs placebo (SE) 95% CI	N/A N/A	–0.74 –1.01 to –0.47
Treatment-naïve		
n	624	636
Adjusted mean CFB (SE) 95% CI	–1.51 (0.103) –1.71 to –1.31	–1.84 (0.102) –2.04 to –1.64
Adjusted mean difference vs placebo (SE) 95% CI	N/A N/A	–0.33 –0.62 to –0.05
Population interaction p-value	**0.10**	
Abbreviations used in table: CFB, change from baseline; CI, confidence interval; mg, milligram; N/A, not applicable; SE, standard error.		

Table 24. Change from baseline to final visit in mean number of incontinence episodes per 24 hours, previously treated versus treatment-naïve patients, pre-specified pooled analysis of SCORPIO, ARIES, and CAPRICORN (adapted from MS; Table 44, pg 111)

	Placebo	**Mirabegron 50 mg**
Previously treated		
n	518	506
Adjusted mean CFB (SE) 95% CI	–0.92 (0.087) –1.09 to –0.75	–1.49 (0.088) –1.66 to –1.32
Adjusted mean difference vs placebo (SE) 95% CI	N/A N/A	–0.57 –0.81 to –0.33
Treatment-naïve		
n	360	356
Adjusted mean CFB (SE) 95% CI	–1.35 (0.104) –1.55 to –1.14	–1.50 (0.105) –1.71 to –1.29
Adjusted mean difference vs placebo (SE) 95% CI	N/A N/A	–0.15 –0.44 to 0.14
Population interaction p-value	**0.095**	
Abbreviations used in table: CFB, change from baseline; CI, confidence interval; mg, milligram; N/A, not applicable; SE, standard error.		

4.3 Summary and critique of the mixed treatment comparison

4.3.1 Results MTC

An overview of the comparators and the number of studies evaluating the outcomes assessed in the MTCs carried out by the manufacturer is provided in Table 25. Figures showing the network diagram of the different comparators and the direct comparisons between them for each are provided in Appendix 7. The manufacturer presented MTC results for the outcomes of:

- micturition (urinary frequency);
- urgency urinary incontinence;
- dry mouth;
- constipation;
- blurred vision.
- frequency of incontinence (one of the primary outcomes assessed in the trials evaluating mirabegron).

However, as for the direct clinical evidence, the ERG has chosen to focus on the outcomes specified in the NICE final scope for this STA (micturition, frequency of urgency urinary incontinence and adverse effects of treatment),[19] together with the additional outcome, frequency of incontinence episodes, as this informs the manufacturer's economic model.

The ERG notes that the manufacturer did not provide a rationale for the choice of outcomes explored using the MTC. The ERG considers it important to reiterate concerns noted earlier around the potential for clinical and methodological heterogeneity in the manufacturer's choice of studies for analysis, the identification of inconsistency by the manufacturer in one or more treatment comparisons for multiple outcomes, and the number of iterations used for sampling the posterior distributions (which may be an indicator of poor mixing of data within the model). Based on these concerns, the ERG considers that the results of the manufacturer's MTC should be interpreted with caution.

Table 25. Overview of the number of trials evaluating the different outcomes

Outcome	Placebo	Mirabegron 50 mg	Tolterodine 4 mg	Oxybutynin 5 mg	Oxybutynin 10 mg	Oxybutynin 15 mg	Solifenacin 5 mg	Solifenacin 10 mg	Fesoterodine 4 mg	Fesoterodine 8 mg	Trospium 60 mg
Micturition	20	6	14	2	-	-	4	5	3	5	-
Incontinence	13	6	8	2	-	-	3	4	1	1	-
Urge incontinence	15	6	10	1	-	-	3	3	3	4	-
Dry mouth	32	6	23	2	6	3	7	6	4	6	3
Constipation	31	6	21	2	4	3	7	6	4	5	3
Blurred vision	18	6	11	2	4	2	6	6	1	1	1
Abbreviation used in table: mg, milligram.											

Micturition (urinary frequency)

The results for the MTC on urinary frequency are presented by the manufacturer using the fixed effects model (as it had a lower DIC than the random effects model). As discussed previously, the ERG agrees that the fixed effects model is the better fitting of the two models but is still a poor fit for the underlying data (see Section 4.1.7).

Of the active treatments, only solifenacin 10 mg was found to have a statistically significant benefit over mirabegron 50 mg in reducing the number of micturition episodes per 24 hours (Table 26). No statistically significant difference was found between mirabegron 50 mg and any other active treatment.

Table 26. Estimate of the treatment effect versus mirabegron 50 mg for micturition (a positive value favours mirabegron; adapted from manufacturer's clarification response; Table 8)

Comparator	Mean difference	95% Credible Interval
Tolterodine 4 mg	0.157	–0.0002 to 0.3154
Fesoterodine 4 mg	0.137	–0.1613 to 0.4345
Fesoterodine 8 mg	–0.048	–0.2489 to 0.1524
Oxybutynin 10 mg	0.139	–0.5290 to 0.8058
Placebo	0.696	0.5544 to 0.8378
Solifenacin 10 mg	–0.583	–0.8324 to –0.3326
Solifenacin 5 mg	–0.240	–0.4921 to 0.0132
Trospium 60 mg	–0.124	–0.5767 to 0.3261
Abbreviation used in table: mg, milligram.		

Urge incontinence

The results for the MTC on urge incontinence are presented by the manufacturer using the fixed effects model (as it had a lower DIC than the random effects model). As discussed previously, the ERG agrees that the fixed effects model is the better fitting of the two models but is still a poor fit for the underlying data.

Seventeen studies were included in the MTC assessing the relative efficacy of OAB treatments on the frequency of urge incontinence episodes (Table 27). Of the active treatments, only solifenacin 10 mg demonstrated a statistically significant benefit over mirabegron 50 mg in reducing the frequency of urgency incontinence. No statistically significant difference was found between mirabegron 50 mg and any other active treatment.

Table 27. Estimate of the effect of treatment versus mirabegron 50 mg for urgency incontinence (a positive value favours mirabegron; adapted from manufacturer's clarification response; Table 8)

Comparator	Mean difference	95% Credible Interval
Tolterodine 4 mg	0.095	–0.123 to 0.307
Fesoterodine 4 mg	–0.034	–0.384 to 0.303
Fesoterodine 8 mg	–0.225	–0.535 to 0.048
Oxybutynin 10 mg	–0.279	–0.945 to 0.385
Placebo	0.437	0.255 to 0.624
Solifenacin 10 mg	–0.420	–0.786 to –0.056
Solifenacin 5 mg	–0.288	–0.642 to 0.071
Trospium 60 mg	–0.112	–0.707 to 0.485
Abbreviation used in table: mg, milligram.		

Incontinence

The results for the MTC on frequency of incontinence episodes are presented by the manufacturer using the fixed effects model (as it had a lower DIC than the random effects model). As discussed previously, the ERG agrees that the fixed effects model is the better fitting of the two models but is still a poor fit for the underlying data.

Fifteen studies reported data on the change from baseline to end of study in incontinence episodes. No statistically significant differences in the change in frequency of incontinence episodes from baseline were found between any of the active treatments. However, the results favoured solifenacin (5 mg and 10 mg) compared with mirabegron 50 mg, with solifenacin being associated with a greater reduction in incontinence episodes compared with mirabegron 50 mg, although the difference did not reach statistical significance (Table 28).

Table 28. Estimate of the treatment effect versus mirabegron 50 mg for incontinence (a positive value favours mirabegron; adapted from manufacturer's clarification response; Table 8)

Comparator	Mean difference	95% Credible Interval
Tolterodine 4 mg	0.082	–0.0649 to 0.2286
Fesoterodine 4 mg	0.107	–0.3911 to 0.6033
Fesoterodine 8 mg	0.226	–0.2770 to 0.7299
Oxybutynin 10 mg	0.137	–0.3986 to 0.6752
Placebo	0.497	0.3724 to 0.6225
Solifenacin 10 mg	–0.240	–0.4875 to 0.0066
Solifenacin 5 mg	–0.237	–0.4824 to 0.0073
Abbreviation used in table: mg, milligram.		

Dry mouth

The results for the MTC on dry mouth are presented by the manufacturer using the random effects model (as it had a lower DIC than the fixed effects model). As discussed previously, the ERG agrees that the random effects model is the better fitting of the two models but is still a poor fit for the underlying data.

All forty studies identified by the manufacturer were included in the MTC assessing the adverse event of dry mouth. All antimuscarinics were associated with a significantly higher risk (evaluated as an odds ratio [OR]) of dry mouth compared with mirabegron 50 mg (Table 29), whereas placebo had a probability of dry mouth similar to that of mirabegron 50 mg, with no statistically significant difference found between placebo and mirabegron 50 mg.

Table 29. Estimate of the effect of treatment versus mirabegron 50 mg for dry mouth (an odds ratio above 1 favours mirabegron; adapted from manufacturer's clarification response; Table 8)

Comparator	OR	95% Credible Interval
Tolterodine ER 4 mg	4.17	2.73 to 6.12
Fesoterodine 4 mg	4.44	2.69 to 6.97
Fesoterodine 8 mg	9.70	6.11 to 14.69
Oxybutynin ER 10 mg	6.80	3.89 to 11.25
Oxybutynin ER 15 mg	7.86	2.91 to 17.48
Oxybutynin ER 5 mg	4.13	1.56 to 9.02
Oxybutynin IR 10 mg	14.07	6.57 to 26.4
Oxybutynin IR 15 mg	39.21	14.98 to 85.64
Oxybutynin IR 9 mg	10.78	5.59 to 18.92
Placebo	1.30	0.86 to 1.92
Solifenacin 10 mg	10.08	6.03 to 15.97
Solifenacin 5 mg	4.23	2.48 to 6.83
Tolterodine IR 4 mg	7.04	4.31 to 11.03
Trospium 40 mg	5.67	2.96 to 9.98
Trospium 60 mg	4.48	1.60 to 10.46
Abbreviations used in table: ER, extended release; IR, immediate release; mg, milligram; OR, odds ratio.		

Constipation

The results for the MTC on constipation are presented by the manufacturer using the fixed effects model (as it had a lower DIC than the random effects model). As discussed previously, the ERG agrees that the fixed effects model is the better fitting of the two models but is still a poor fit for the underlying data.

Thirty seven studies reporting the number of patients suffering from constipation were included in the MTC. The risk of constipation was significantly lower with mirabegron 50 mg compared with solifenacin 5 and 10 mg, fesoterodine 8 mg, and trospium 60 mg. However, for most comparisons,

there was no statistically significant difference between mirabegron 50 mg and other active treatments in the risk of experiencing constipation (Table 30Table 30).

Table 30. Estimate of the effect of treatment versus mirabegron 50 mg for constipation (an odds ratio above 1 favours mirabegron; adapted from manufacturer's clarification response; Table 8)

Comparator	OR	95% Credible Interval
Tolterodine ER 4 mg	1.11	0.72 to 1.65
Fesoterodine 4 mg	1.07	0.58 to 1.81
Fesoterodine 8 mg	1.93	1.14 to 3.06
Oxybutynin ER 10 mg	1.02	0.53 to 1.79
Oxybutynin ER 15 mg	2.16	0.27 to 8.28
Oxybutynin ER 5 mg	2.46	0.42 to 8.71
Oxybutynin IR 15 mg	1.61	0.42 to 4.38
Oxybutynin IR 9 mg	0.99	0.41 to 1.99
Placebo	0.73	0.48 to 1.07
Solifenacin 10 mg	4.37	2.54 to 7.07
Solifenacin 5 mg	2.50	1.41 to 4.13
Tolterodine IR 4 mg	1.03	0.59 to 1.67
Trospium 40 mg	1.69	0.88 to 2.98
Trospium 60 mg	7.60	2.08 to 22.59
Abbreviations used in table: ER, extended release; IR, immediate release; mg, milligram; vs, versus.		

Blurred vision

The results for the MTC on blurred vision are presented by the manufacturer using the random effects model (as it had a lower DIC than the fixed effects model). As discussed previously, the ERG agrees that the random effects model is the better fitting of the two models but is still a poor fit for the underlying data.

Twenty three studies were included in the MTC. No statistically significant differences in the risk of developing blurred vision were found between mirabegron 50 mg and other active treatments or placebo (Table 31).

Table 31. Estimate of the effect of treatment versus mirabegron 50 mg for blurred vision (an odds ratio above 1 favours mirabegron; adapted from manufacturer's clarification response; Table 8)

Comparator	OR	95% Credible Interval
Tolterodine ER 4 mg	1.44	0.56 to 3.13
Fesoterodine 4 mg	0.80	0.04 to 3.71
Fesoterodine 8 mg	0.73	0.04 to 3.40
Oxybutynin ER 10 mg	2.61	0.21 to 12.12
Oxybutynin ER 15 mg	7.07	0.02 to 41.75
Oxybutynin ER 5 mg	5.12	0.05 to 28.87

Oxybutynin IR 15 mg	2.45	0.07 to 13.72
Oxybutynin IR 9 mg	0.40	0 to 2.58
Placebo	5.27	0.90 to 18.47
Solifenacin 10 mg	0.79	0.31 to 1.71
Solifenacin 5 mg	1.94	0.67 to 4.50
Tolterodine IR 4 mg	1.15	0.38 to 2.71
Trospium 40 mg	0.75	0.23 to 1.83
Trospium 60 mg	2.44	0.15 to 11.93
Abbreviations used in table: ER, extended release; IR, immediate release; mg, milligram; OR, odds ratio.		

4.4 Additional work on clinical effectiveness undertaken by the ERG

To inform the decision problem that is the focus of the STA, the ERG performed additional analyses of three of the RCTs evaluating mirabegron 50 mg and that had an active control arm of tolterodine (SCORPIO,[20] DRAGON,[23] and 178-CL-048[25]). Data from the long-term study TAURUS[26] are presented separately and are not included in any meta-analysis, as patients in TAURUS were mainly recruited from SCORPIO and ARIES. The ERG believes that this may potentially introduce selection bias. Because of time constraints, the ERG's additional analyses focused on outcomes used to inform the manufacturer's economic model; frequency of micturition; frequency of incontinence episodes; adverse events of dry mouth and constipation; and discontinuations. For RCTs evaluating mirabegron, data were derived from individual trial CSRs.[20;23];[25]

4.4.1 Direct clinical evidence

Statistical approach and data synthesis

Data were analysed using Review Manager 5. Continuous outcomes were analysed as a mean difference (MD) with a 95% CI using Generic Inverse Variance (GIV). Dichotomous outcomes were analysed using Mantel-Haenzsel odds ratios (OR) with a 95% CI. Statistical heterogeneity was assessed with the I^2 measurement.

Results of the ERG's meta-analysis

Clinical effectiveness outcomes

Meta-analysis of SCORPIO, DRAGON and 178-CL-048 showed that treatment with mirabegron 50 mg led to significantly fewer micturitions per 24 hours compared with treatment with tolterodine 4 mg (MD –0.27; 95% CI: –0.48 to –0.06; p-value = 0.01; I^2 = 0%; Table 32). However, data from TAURUS favoured tolterodine 4 mg, although the difference was not statistically significant (Table 32). As in the analysis of micturition, in the meta-analysis, treatment with mirabegron 50 mg led to significantly fewer incontinence episodes per 24 hours compared with tolterodine 4 mg (MD –0.21; 95% CI –0.41 to –0.01; p-value = 0.04; I^2 = 0%). However, data from TAURUS found that

mirabegron 50 mg was associated with a statistically significant increase in the number of incontinence episodes compared with tolterodine (p-value = 0.04; Table 32).

Table 32. Results of the ERG's meta-analysis: outcomes evaluating clinical effectiveness (a negative mean difference favours mirabegron)

Outcome	Mirabegron 50 mg vs tolterodine	
	MD	95% CI
Micturitions per 24 hours		
SCORPIO	–0.34	–0.65 to –0.03
178-CL-048	–0.27	–0.59 to 0.05
DRAGON	0.15	–0.60 to 0.90
TAURUS	0.12	–0.11 to 0.35
ERG's meta-analysis (SCORPIO, DRAGON, 178-CL-048)[a]	–0.27	–0.48 to –0.06
Frequency of incontinence episodes per 24 hours		
SCORPIO	–0.30	–0.61 to 0.01
178-CL-048	–0.15	–0.42 to 0.12
DRAGON	–0.09	–0.91 to 0.73
TAURUS	0.25	0.01 to 0.49
ERG's meta-analysis (SCORPIO, DRAGON, 178-CL-048)[b]	–0.21	–0.41 to –0.01

[a] Micturitions: mirabegron 50 mg, N=1,009, and tolterodine 4 mg, N=928.
[b] Incontinence episodes: mirabegron 50 mg, N=667, and tolterodine 4 mg, N=593.
Abbreviations used in table: CI, confidence interval; MD, mean difference; mg, milligram.

Adverse event outcomes

For the outcome of dry mouth, in all four trials, mirabegron 50 mg was associated with a lower risk of experiencing dry mouth compared with tolterodine (Table 33). In the ERG's meta-analysis, the difference between mirabegron and tolterodine was found to be statistically significant, and favoured mirabegron (OR 0.22; 95% CI: 0.14 to 0.34; p-value <0.00001; I^2 = 0%). Considering constipation and discontinuation, the ERG's meta-analysis found no significant difference between mirabegron 50 mg and tolterodine in the risk of either experiencing constipation or discontinuing treatment (constipation: OR 0.98; 95% CI: 0.56 to 1.72; I^2 = 0%; discontinuation: OR 1.31; 95% CI: 0.96 to 1.79; I^2 = 0%; Table 33). The results from the long-term study TAURUS are in agreement with the ERG's meta-analysis for the outcomes of constipation and discontinuation.

Table 33. Results of the ERG's meta-analysis: outcomes evaluating adverse events (an odds ratio <1 favours mirabegron)

Outcome	Mirabegron 50 mg vs tolterodine	
	OR	95% CI
Dry mouth		
SCORPIO	0.26	0.09 to 0.37
178-CL-048	0.16	0.17 to 0.47
DRAGON	0.49	0.10 to 2.50
TAURUS	0.31	0.19 to 0.50
ERG's meta-analysis (SCORPIO, DRAGON, 178-CL-048)[a]	0.22	0.14 to 0.34
Constipation		
SCORPIO	0.80	0.31 to 2.04
178-CL-048	0.92	0.42 to 1.98
DRAGON	3.09	0.37 to 26.11
TAURUS	1.05	0.58 to 1.89
ERG's meta-analysis (SCORPIO, DRAGON, 178-CL-048)[b]	0.98	0.56 to 1.72
Discontinuation		
SCORPIO	1.15	0.77 to 1.72
178-CL-048	1.37	0.78 to 2.40
DRAGON	2.86	0.81 to 10.10
TAURUS	0.96	0.76 to 1.20
ERG's meta-analysis (SCORPIO, DRAGON, 178-CL-048)[c]	1.31	0.96 to 1.79

[a] Dry mouth: mirabegron 50 mg, N=1,041, and tolterodine 4 mg, N=955.
[b] Constipation: mirabegron 50 mg, N=1,041, and tolterodine 4 mg, N=955.
[c] Discontinuation: mirabegron 50 mg, N=1,046, and tolterodine 4 mg, N=958.
Abbreviations used in table: CI, confidence interval; mg, milligram; OR, odds ratio.

4.4.2 Mixed treatment comparison

Statistical approach and data synthesis

As the ERG has concerns about the MTC conducted by the manufacturer based on clinical, methodological and statistical heterogeneity, the decision was taken to carry out an independent MTC with the goal of using a more homogeneous dataset. Unfortunately, due to capacity constraints, the ERG was unable to conduct its own systematic review of the available evidence, or appraise the RCTs identified by the manufacturer. The ERG, therefore, used the 40 RCTs identified by the manufacturer for the MTC but applied the restrictions that follow:

- excluded RCTs that included patients other than those with OAB, that were carried out in a single gender population, or that reported on outcomes available at a time point other than 12 weeks (based on the manufacturer's summary; Table 5 of the ERG report);

- excluded RCTs that were deemed to be of poor methodological quality based on the manufacturer's summary (Appendix 3 of the ERG report, with an RCT with a "Yes" in all of the first four categories assessed as being of acceptable quality);
- included only outcomes and treatment formulations and doses used in the economic model supplied by the manufacturer.

Applying the restrictions listed above provided a pool of 22 RCTs for use in the ERG's MTC, [32;34;35;38;39;43;44;47;49;54;56;59-63] including the six RCTs evaluating mirabegron directly (SCORPIO, ARIES, CAPRICORN, DRAGON, 178-CL-045, and 178-CL-048).

The outcomes considered in the ERG's MTC were micturition, incontinence, constipation, dry mouth, and additionally all-cause discontinuation (as the outcome of all-cause discontinuation was believed to be a potential key driver in the economic model and was not included in the manufacturer's MTC). In addition, only results comparing placebo and active treatments versus mirabegron 50 mg are presented (i.e., results of active interventions versus each other are not reported).

The methodology used by the ERG was similar to that used by the manufacturer and as critiqued in Section 4.1.7. Comparison of the number of unconstrained data points with the residual deviance in the preferred model from the manufacturer's MTC and from the ERG's MTC indicates that there is a greater degree of concordance with the values obtained from the ERG's analyses (Table 34). This suggests that, for each outcome, the ERG's MTC would be considered a better fit of the underlying data set and would therefore produce potentially more reliable results.

Table 34. Overview of the number of unconstrained data points and residual deviance in the preferred model for the manufacturer's MTC and the preferred model from the ERG's MTC (where model selection was based on lowest DIC)

Outcome	Manufacturer's MTC		ERG's MTC	
	Number of unconstrained data points[a]	Residual deviance in preferred model[b]	Number of unconstrained data points	Residual deviance in preferred model
Micturition	39	49	51	46
Incontinence	61	29	32	33
Constipation	103	580	59	61
Dry mouth	96	437	59	43
Discontinuation	N/A	N/A	57	50

[a] As assessed by the Evidence Review Group.

[b] As provided in the manufacturer's clarification response.

Abbreviations used in table: ERG, Evidence Review Group; MTC, mixed treatment comparison; N/A, not applicable.

Results

Micturition

Data were extracted from 19 RCTs,[20-25;32;34;35;38;39;43;47;54;59-63] with one trial providing data from two populations (young and old).[63] The DIC was lower for the random effects model than for the fixed effects model (20 vs 21, respectively).

The ERG found no significant difference between mirabegron 50 mg and any of the other active treatments assessed in the outcome of micturition per 24 hours. The results of the manufacturer's MTC and the ERG's MTC are compared in Table 35.

Table 35. Results of the manufacturer's MTC for micturition compared with the ERG's MTC using mirabegron 50 mg as the baseline treatment (a positive value favours mirabegron; a negative value favours the alternative treatment)

Micturition	Manufacturer's MTC		ERG's MTC	
	MD	95% CrI	MD	95% CrI
Fesoterodine 4 mg	0.137	–0.161 to 0.435	0.381	–0.398 to 1.154
Fesoterodine 8 mg	–0.048	–0.249 to 0.152	0.138	–0.636 to 0.913
Oxybutynin 10 mg	0.139	–0.529 to 0.806	–0.536	–1.849 to 0.782
Placebo	0.696	0.554 to 0.838	0.700	0.254 to 1.141
Solifenacin 5 mg	–0.240	–0.492 to 0.013	–0.193	–1.066 to 0.672
Solifenacin 10 mg	–0.583	–0.832 to –0.333	–0.560	–1.346 to 0.225
Tolterodine 4 mg	0.157	–0.0002 to 0.315	0.087	–0.421 to 0.591
Trospium 60 mg	–0.124	–0.577 to 0.326	–0.121	–1.351 to 1.091
Abbreviations used in table: CrI, credible interval; ERG, Evidence Review Group; MD, mean difference; mg, milligram; MTC, mixed treatment comparison.				

Incontinence

Data were extracted from 12 RCTs,[20-25;32;39;60-63] with one trial providing data from two different populations (young and old).[63] The DIC was found to be lower for the fixed effects model when compared to the random effects model (–15 vs –7, respectively).

Of the active treatments assessed, mirabegron 50 mg was found to be significantly less effective than only solifenacin (5 mg and 10 mg) at reducing frequency of incontinence episodes. The results of the manufacturer's MTC and the ERG's MTC are compared in Table 36.

Table 36. Results of the manufacturer's MTC for incontinence compared with the ERG's MTC using mirabegron 50 mg as the baseline treatment (a positive value favours mirabegron; a negative value favours the alternative treatment)

Comparator	Manufacturer's MTC		ERG's MTC	
	MD	95% CrI	MD	95% CrI
Fesoterodine 4 mg	0.107	–0.391 to 0.603	0.108	–0.383 to 0.597
Fesoterodine 8 mg	0.226	–0.277 to 0.730	0.231	–0.277 to 0.731
Oxybutynin 10 mg	0.137	–0.399 to 0.675	–0.476	–1.011 to 0.054
Placebo	0.497	0.372 to 0.623	0.499	0.370 to 0.627
Solifenacin 5 mg	–0.237	–0.482 to 0.007	–0.386	–0.717 to –0.055
Solifenacin 10 mg	–0.240	–0.488 to 0.007	–0.380	–0.694 to –0.067

Tolterodine 4 mg	0.082	–0.065 to 0.2296	0.066	–0.089 to 0.221
Abbreviations used in table: CrI, credible interval; ERG, Evidence Review Group; MD, mean difference; mg, milligram; MTC, mixed treatment comparison; OR, odds ratio.				

Constipation

Data were extracted from 22 RCTs,[20-25;32;34;35;38;39;43;45;47;49;54;56;59-63] with one trial providing data from two different populations (young and old).[63] The DIC was found to be lower for the fixed effects model compared with the random effects model (325 vs 327, respectively).

In the ERG's MTC, compared with the active treatments evaluated, mirabegron 50 mg was found to be associated with a significantly lower risk of constipation than fesoterodine 8 mg, solifenacin (5 mg and 10 mg), and trospium 60 mg. No other statistically significant differences were found between mirabegron 50 mg and the other active treatments. The results of the manufacturer's MTC and the ERG's MTC are compared in Table 37.

Table 37. Results of the manufacturer's MTC for constipation compared with the ERG's MTC using mirabegron 50 mg as the baseline treatment (an odds ratio >1 favours mirabegron; an odds ratio <1 favours the alternative treatment)

Comparator	Manufacturer's MTC		ERG's MTC	
	OR	95% CrI	OR	95% CrI
Fesoterodine 4 mg	1.066	0.576 to 1.808	1.181	0.610 to 2.093
Fesoterodine 8 mg	1.926	1.142 to 3.059	2.115	1.134 to 3.639
Fesoterodine 4/8 mg	N/A	N/A	1.744	0.913 to 3.036
Oxybutynin ER 10 mg	1.021	0.527 to 1.793	1.363	0.477 to 3.086
Oxybutynin IR 15 mg	1.614	0.416 to 4.375	0.954	0.433 to 1.826
Placebo	0.732	0.483 to 1.066	0.738	0.485 to 1.082
Solifenacin 5 mg	2.501	1.410 to 4.127	2.114	1.159 to 3.590
Solifenacin 10 mg	4.369	2.540 to 7.071	4.522	2.598 to 7.471
Tolterodine ER 4 mg	1.109	0.716 to 1.647	1.085	0.701 to 1.620
Tolterodine IR 4 mg	1.034	0.594 to 1.673	1.259	0.661 to 2.180
Trospium 60 mg	7.604	2.08 to 22.59	7.629	2.116 to 22.950
Abbreviations used in table: CrI, credible interval; ER, extended release; ERG, Evidence Review Group; IR, immediate release; mg, milligram; MTC, mixed treatment comparison; OR, odds ratio.				

Dry mouth

Data were extracted 22 RCTs,[20-25;32;34;35;38;39;43;45;47;49;54;56;59-63] with one trial providing data from two different populations (young and old).[63] The DIC was found to be lower for the fixed effects model when compared to the random effects model (384 vs 385, respectively).

Mirabegron 50 mg was found to be associated with a significantly lower risk of dry mouth compared with all other antimuscarinics assessed. The results of the manufacturer's MTC and the ERG's MTC are compared in Table 38.

Table 38. Results of the manufacturer's MTC for dry mouth compared with the ERG's MTC using mirabegron 50 mg as the baseline treatment (an odds ratio >1 favours mirabegron; an odds ratio <1 favours the alternative treatment)

Comparator	Manufacturer's MTC		ERG's MTC	
	OR	95% CrI	OR	95% CrI
Fesoterodine 4 mg	4.436	2.693 to 6.974	4.695	2.921 to 7.279
Fesoterodine 8 mg	9.7	6.109 to 14.686	11.240	7.072 to 17.279
Fesoterodine 4/8 mg	N/A	N/A	10.890	6.960 to 16.580
Oxybutynin ER 10 mg	6.795	3.894 to 11.25	3.715	1.997 to 6.402
Oxybutynin IR 15 mg	39.208	14.98 to 85.64	11.840	6.969 to 19.140
Placebo	1.303	0.859 to 1.916	1.254	0.849 to 1.829
Solifenacin 5 mg	4.229	2.484 to 6.825	3.328	1.969 to 5.338
Solifenacin 10 mg	10.078	6.027 to 15.97	9.772	5.977 to 15.300
Tolterodine ER 4 mg	4.168	2.733 to 6.117	4.603	3.134 to 6.700
Tolterodine IR 4 mg	7.042	4.311 to 11.03	6.670	4.166 to 10.300
Trospium 60 mg	4.481	1.598 to 10.46	4.418	1.707 to 9.903
Abbreviations used in table: CrI, credible interval; ER, extended release; ERG, Evidence Review Group; IR, immediate release; mg, milligram; MTC, mixed treatment comparison; N/A, not applicable; OR, odds ratio.				

Discontinuation

Data were extracted from 22 RCTs.[20-25;32;34;35;38;39;43;45;47;49;54;56;59-63] The DIC was found to be lower for the fixed effects model when compared to the random effects model (373 vs 375, respectively).

Of the active treatments assessed, only oxybutynin IR 15 mg was found to significantly increase the risk of discontinuation compared with mirabegron 50 mg. All other differences did not reach statistical significance. The results of the manufacturer's MTC and the ERG's MTC are compared in Table 39.

Table 39. Results of the manufacturer's MTC for discontinuation compared with the ERG's MTC using mirabegron 50 mg as the baseline treatment (an odds ratio >1 favours mirabegron; an odds ratio <1 favours the alternative treatment)

Comparator	Manufacturer's MTC		ERG's MTC	
	OR	95% CrI	OR	95% CrI
Fesoterodine 4 mg	N/A	N/A	1.288	0.888 to 1.815
Fesoterodine 8 mg	N/A	N/A	1.350	0.927 to 1.905
Oxybutynin ER 10 mg	N/A	N/A	1.101	0.542 to 2.005
Oxybutynin IR 15 mg	N/A	N/A	2.671	1.597 to 4.220
Placebo	N/A	N/A	1.039	0.860 to 1.245
Solifenacin 5 mg	N/A	N/A	1.031	0.718 to 1.435
Solifenacin 10 mg	N/A	N/A	1.057	0.769 to 1.422
Tolterodine ER 4 mg	N/A	N/A	0.865	0.692 to 1.069
Tolterodine IR 4 mg	N/A	N/A	0.889	0.517 to 1.411
Trospium 60 mg	N/A	N/A	1.310	0.729 to 2.180

Abbreviations used in table: CrI, credible interval; ER, extended release; ERG, Evidence Review Group; IR, immediate release; mg, milligram; MTC, mixed treatment comparison; N/A, not applicable; OR, odds ratio.

Summary

The results from the ERG's MTC are in general agreement with those of the manufacturer's MTC, with few differences between the two analyses. Considering micturition (one of the primary outcomes in the key trials evaluating mirabegron), the manufacturer's MTC found mirabegron 50 mg to be more effective than tolterodine 4 mg and less effective than solifenacin 10 mg at reducing micturition, whereas the ERG's MTC identified no significant difference between mirabegron 50 mg and any other active treatments for this outcome. In terms of reduction of incontinence episodes, the results of the ERG's MTC for the comparison of solifenacin (5 and 10 mg) versus mirabegron differed from the manufacturer's MTC, with the ERG's MTC identifying a significant difference between the solifenacin at both doses and mirabegron 50 mg favouring solifenacin in each analysis; the manufacturer's MTC identified no statistically significant difference between the treatments for this outcome. In the ERG's MTC, only oxybutynin IR 15 mg was found to significantly increase the risk of discontinuation compared with mirabegron 50 mg. All of the ERG's MTCs were found to be a better fit of the underlying data than the manufacturer's MTCs.

4.5 Conclusions of the clinical effectiveness section

4.5.1 Summary of clinical results

- The submitted direct clinical evidence was derived from the RCTs SCORPIO, ARIES, CAPRICORN, and TAURUS. However, the manufacturer also recognised the relevance of the RCTs DRAGON, 178-CL-045, and 178-CL-048, for which some efficacy and safety data were presented. The ERG is confident that no relevant clinical trials were overlooked.
- The objectives in the trials from which the direct clinical evidence is derived were to assess the efficacy and safety of mirabegron versus placebo or tolterodine in patients with OAB. Based on ERG clinical expert opinion, the OAB patients included in the trials are representative of the OAB population in England and Wales. All trials included a group evaluating mirabegron 50 mg, which is the anticipated licensed dose of mirabegron (25 mg dose of mirabegron for patients with renal or hepatic failure).
- Mirabegron 50 mg was more effective than placebo across all clinical efficacy outcomes, and was associated with improvements in health-related quality of life (HRQoL) as measured using the disease specific QoL measure OAB-q. However, there was no statistically significant difference between mirabegron and placebo in EQ-5D utility score.
- In the ERG's opinion, direct evidence for the clinical efficacy of mirabegron 50 mg compared with tolterodine ER 4 mg is of more relevance to the decision problem. Two RCTs found no statistically significant difference between mirabegron 50 mg and tolterodine 4 mg in micturitions, incontinence episodes, urgency incontinence episodes, level of urgency, number of urgency episodes, or nocturia. However, as the manufacturer highlights, the two RCTs were not powered to assess superiority or non-inferiority of mirabegron compared with tolterodine.

- Mirabegron 50 mg treatment was associated with a larger improvement in EQ-5D utility score compared with tolterodine ER 4 mg. The difference in disease-specific QoL measure OAB-q was not reported.
- Additional meta-analyses by the ERG of three RCTs that included a mirabegron 50 mg and a tolterodine 4 mg treatment group showed a small, but statistically significant, improvement in micturitions and incontinence episodes with mirabegron 50 mg compared with tolterodine 4 mg.
- The overall adverse event and safety profile of mirabegron 50 mg was comparable to tolterodine ER 4 mg and placebo, including treatment-related adverse events, serious adverse events, and adverse events leading to discontinuations. However, of adverse events occurring in ≥2 % of patients in any group (common adverse events), the incidence of dry mouth was significantly lower with mirabegron 50 mg compared with tolterodine ER 4 mg.
- There were no statistically significant differences between the subgroups of female versus male OAB patients, or OAB patients previously treated for the condition versus treatment naïve patients.
- The manufacturer's MTC identified statistically significant benefits for mirabegron 50 mg compared with other active treatments over a range of outcomes. Mirabegron 50 mg was more found to be more effective at reducing micturition than tolterodine 4 mg. However, solifenacin 10 mg was found to be significantly more effective than mirabegron 50 mg at reducing micturition and urge incontinence. Considering adverse events, mirabegron was found to be associated with a lower risk of developing dry mouth than all other antimuscarinics evaluated and a lower risk of developing constipation than solifenacin (5 mg and 10mg), fesoterodine 8 mg, and trospium 60 mg. The ERG has reservations about the manufacturer's MTC based on clinical, methodological, and statistical heterogeneity.
- The ERG's MTC identified no significant difference between mirabegron 50 mg and any other active treatments for the outcome of micturition. Solifenacin (5 mg and 10mg) was found to be significantly more effective than mirabegron 50mg at reducing incontinence. In agreement with the manufacturer's MTC, the ERG found that mirabegron was associated with a lower risk of developing constipation compared with fesoterodine 8 mg, solifenacin (5 mg and 10 mg), and trospium 60 mg, and that mirabegron was associated with a lower risk of developing dry mouth compared with all the other antimuscarinics evaluated. Only oxybutynin IR 15 mg was found to significantly increase the risk of discontinuation compared with mirabegron 50 mg. All of the ERG's MTCs were found to be a better fit of the underlying data than the manufacturer's MTCs.

4.5.2 Clinical issues

- Direct clinical evidence submitted by the manufacturer was derived from three RCTs – SCORPIO, ARIES, and CAPRICORN. Comparisons evaluated in the trials were limited to mirabegron (25 mg, 50 mg, and 100 mg) versus placebo and/or tolterodine 4 mg but no other active comparators listed in the scope. As highlighted by the manufacturer, the three studies were not powered to evaluate superiority or non-inferiority of mirabegron versus tolterodine, which is the comparison with the submitted clinical evidence that the ERG considers to be more relevant to the decision problem.
- The manufacturer omitted from the direct evidence three additional trials that, on evaluation, the ERG consider are relevant to the decision problem.
- Data were not provided from the SCORPIO RCT to inform the comparison of mirabegron versus tolterodine. Although the ERG appreciates the manufacturer's point that SCORPIO was not powered to detect a difference between the two active treatments, the ERG considers that the results from the tolterodine group could have contributed to the meta-analysis and MTC.

- The ERG considers insufficient details were provided on the methods used within the trials to minimise bias and ensure methodological rigour.
- The manufacturer limited the MTC to oral formulations. The ERG considers that it would have been appropriate to evaluate non-oral preparations, for example, oxybutynin is available as a transdermal patch. The ERG, therefore, considers that there could potentially be additional studies available to inform an MTC. Due to time constraints, the ERG was unable to carry out an independent systematic review of the literature or to validate the 40 studies identified by the manufacturer as relevant to the MTC.
- The ERG considers that there is considerable heterogeneity (in terms of population and quality of trials) in the MTC submitted by manufacturer and as such results should be interpreted with caution.

5 COST EFFECTIVENESS

This section provides a structured description and critique of the systematic literature review and *de novo* economic evaluation submitted by the manufacturer. The manufacturer provided a written submission of the economic evidence together with four electronic versions of the Microsoft© EXCEL-based economic model. The location of the key economic information within the manufacturer's submission (MS) is summarised in Table 40.

Table 40. Summary of key information within the manufacturer's submission

Information	Section (MS)
Details of the systematic review of the economic literature	7.1
Model structure	7.2.2 to 7.2.6
Technology	7.2.7 to 7.2.8
Clinical parameters and variables	7.3
Measurement and valuation of health effects and adverse events	7.4
Resource identification, valuation and measurement	7.5
Sensitivity analysis	7.6
Results	7.7
Validation	7.8
Subgroup analysis	7.9
Interpretation of economic evidence	7.10
Abbreviation used in table: MS, manufacturer's submission.	

5.1 Summary and critique of the manufacturer's review of cost-effectiveness evidence

The manufacturer carried out a systematic review of the literature with the aim of identifying economic evaluations and costing studies considering treatments for overactive bladder (OAB). Searches of the following databases: Medline, Embase, Medline (R) In-Process, EconLIT and NHS Economic Evaluation Database (NHS EED) were carried out on 26th November 2011; no date restrictions were applied to the search. The evidence review group (ERG) notes that the search terms used were reasonable and both inclusion and exclusion criteria were explicitly stated. However, the ERG notes that the manufacturer did not supplement the database search with hand-searching of review bibliographies, conference abstracts or manufacturer's databases. Although, based on supplementary searches, the ERG considers it unlikely that any relevant publications were excluded.

The manufacturer's review identified seven costing studies (discussed further in Section 5.2.8) and 16 economic evaluations. Of the 16 economic evaluations, 10 were cost-utility analyses,[69-78] one a cost-consequence analysis[79] and the remaining five were cost-effectiveness analyses.[80-84] Eight of the analyses were carried out using the analytical framework of a Markov model;[71;72;75-78;80;81] seven were carried out using a decision tree[69;70;73;74;79;82;84] and one used an empirically derived algorithm based model.[83] Eight studies did not provide details of the patient population considered,[69;70;72-74;78;81;84] six

were reported to consider patients with OAB,[71;76;77;79;80;83] one considered patients with "urge or mixed incontinence with a primary-urge component"[82] and the remaining study considered a population of patients with urge incontinence.[75] Of the identified studies, six focused on the UK,[70;71;77;81-83] three on Sweden,[72-74] three on Canada,[75;78;80] two on the USA[69;84] and the remaining two studies considered Spain[79] and Italy,[76] respectively. All of the identified studies considered time horizons of 1 year or less (range 3 months to 1 year), with the majority (13 studies) considering costs and consequences over 1 year.[70-81;83] Of those studies that used a shorter time horizon, one[69] used a 6 month time horizon, based on the rationale that 6 months appropriately reflects the initial phase of OAB treatment. A further study used a 6 month time horizon with the justification that extrapolation of 3 month data beyond 6 months would introduce substantial uncertainty.[82] The remaining study used a 3 month time horizon; although, the authors did not consider the rationale or impact on the value of the analysis of the time horizon used.[84] All of the identified economic evaluations considered currently available pharmaceutical interventions for OAB; however, none considered the cost-effectiveness of mirabegron. Table 41 summarises the economic evaluations identified by the manufacturer's systematic review.

Table 41. Summary of relevant economic evaluation (adapted from MS; Table 80; pg 179)

Study, Country, year	Aim	Model structure, perspective, time horizon	Patient population	QALYs (intervention, comparator)	Costs (intervention, comparator)	ICER (cost/QALY)
Arlandis-Guzman,[79] Spain, 2011	Assess the economic value of OAB treatment with fesoterodine relative to tolterodine ER and solifenacin	• Decision tree • Societal, healthcare payer • 1 year	≥18 years, ≥8 micturitions/24 hrs, OAB symptoms with urinary urgency and ≥1 urge urinary incontinence episode/24 hrs	QALYs gained vs baseline after 52 Wks: • Fesoterodine: 0.01014 • Tolterodine: 0.00846 • Solifenacin: 0.00957	Total cost (52 wks), in €: • Fesoterodine 1,937 • Tolterodine 2,089 • Solifenacin 1,960	Only reported as: • Fesoterodine was cost saving vs tolterodine • Fesoterodine was cost saving vs solifenacin
Arikian,[69] USA, 2000	Evaluate the relative treatment costs and cost effectiveness of oxybutynin IR, oxybutynin CR and tolterodine IR	• Decision tree • US payer • 6 months	NR	• QALYs not reported • Cost per success and cost per continent day	Cost Oxybutynin CR/ oxybutynin IR/ tolterodine IR; over 6 months, in US$: • Surgery as 2nd line:1402/1395/1650 • Surgery as 3rd line: 893/818/918	Cost per success and cost per continent day were reported. Overall result: increased use of oxybutynin CR first-line would lead to cost savings for payers. Oxybutynin CR/ oxybutynin IR/ tolterodine IR; over 6 months, in US$ **Cost/success** • Surgery as 2nd line: 2682/3022/5176 • Surgery as 3rd line:1708/1774/ 2881 **Cost/continent days** • Surgery as 2nd line:

						18.70/21.60/37.20 Surgery as 3rd line: 11.90/12.60/20.70
Cardozo,[70] UK, 2010	Assess the cost-effectiveness of solifenacin vs other antimuscarinic strategies commonly used in UK clinical practice.	• Decision tree • UK NHS payer • 1 year	NR	QALYs for urgency/ frequency/ incontinence for the 1000-patient cohort • Fesoterodine 4 mg/8 mg†: 709.6/718.3/ 692.5 • Oxybutynin IR 15 mg: NA/719.6/ 691.7 • Propiverine ER 20 mg: 708.9/718.0/ 688.0 • Solifenacin 5mg/10mg†: 712.3/723.1/ 695.0 • Tolterodine ER 4mg: 709.7/718.1/ 688.0 • Tolterodine IR 2 mg/4 mg†: NA/718.5/688.1	Total cost for 1000 patients, by symptoms (urgency/frequency/incontinence), in £: • Fesoterodine 4 mg or 8 mg: 484,553/462,230/469,062 • Oxybutynin IR 15mg: NA/159,896/171,891 • Propiverine ER 20mg: 443,455/420,377/437,683 • Solifenacin 5 mg/10 mg†: 470,840/443,282/456,048 • Tolterodine ER 4mg: 480,090/458,720/476/167 • Tolterodine IR 2 mg /4 mg†: NA/472,183/490,554	**ICERs for solifenacin** 5 mg/10 mg† compared with: • Fesoterodine 4 mg/8 mg†: dominant for all symptoms • Oxybutynin IR 15 mg: NA for urgency, £80,009 for frequency and £87,162 for incontinence • Propiverine ER 20mg: £8087 for urgency, £4457 for frequency and £2639 for incontinence • Tolterodine ER 4 mg: dominant for all symptoms • Tolterodine IR 2 mg /4 mg: NA for urgency, dominant frequency and incontinence
Getsios,[80] Canada, 2004	Describe a model comparing health-economic outcomes for the ER formulation of oxybutynin and tolterodine IR in a population of community-dwelling Canadian adults with OAB	• Markov • Healthcare payer • 1 year	Community-dwelling Canadian adults with OAB	• QALYs not reported • Cost per incontinent episode avoided	Annual costs per patient, in Can$: • Oxybutynin ER: 688 • Tolterodine IR: 656 Saving of Can$32 per year per patient, increasing to Can$42 when comorbidities and surgery are included.	• Oxybutynin ER dominated tolterodine
Getsios,[81]	Evaluate the cost-	• Markov		• QALYs not reported	1-year total costs, in £:	• Oxybutynin dominates

UK, 2004	effectiveness of oxybutynin ER relative to tolterodine IR, for OAB	• Healthcare payer • 1 year		• Difference in QALYs is minimal <0.01 per patient in favour of Oxybutynin	• Oxybutynin ER 10 mg: 332 • Tolterodine IR 2 mg: 418	tolterodine
Guest,[82] Austria, France, UK, 2004	Estimate the cost effectiveness of oxybutynin CR compared with oxybutynin IR and tolterodine in the treatment of OAB	• Decision tree • UK NHS payer, France Social Security, Austria Sick Funds, and patient-perspective • 6 months	≥18 years, with urge or mixed incontinence with a primary-urge component	• QALYs not reported • Cost in reducing the frequency of incontinence at 6 months • Cost in reducing micturition frequency at 6 months	6-monthly total costs per patient: UK/France/ Austria, in € • Oxybutynin CR 10 mg: 1078/872/912 • Oxybutynin IR 10 mg: 1097/834/986 Tolterodine 40 mg: 1359/861/1108	• Starting treatment with oxybutynin CR dominant in the UK and Austria, and cost-effective in France.
Hakkaart,[71] UK, 2009	Estimate the cost per QALY of solifenacin at two doses (vs placebo), over a time horizon of 12-months	• Markov • Healthcare payer • 1 year	≥18 with symptoms of OAB (including urinary frequency, urgency, or urge incontinence) for more than three months	Mean QALY/patient: Solifenacin 5 mg and 10 mg: 0.711 Placebo: 0.697	Total cost/patient, 6 months, in £: • Placebo: 253 • Solifenacin 5 mg: 484 • Solifenacin 10 mg: 597	• Solifenacin 5 mg vs placebo: £17,602 • Solifenacin 10 mg vs placebo: £24,464
Herschorn,[78] Canada, 2010	Estimate the cost effectiveness of solifenacin 5mg/day compared with oxybutynin IR 15mg/day in patients with OAB	• Markov • Canadian healthcare payer • 1 year		• Solifenacin 5 mg: 0.696 • Oxybutynin IR 5mg: 0.686	Total costs for 1 year, in Can$: • Solifenacin 5 mg: 695 • Oxybutynin IR 5 mg: 550	**ICER** • Without incontinence pads: solifenacin vs oxybutynin: $14,092 • With incontinence pads: solifenacin dominant

Hughes,[83] UK, 2004	Calculate and compare the cost-effectiveness of oxybutynin ER, tolterodine ER, tolterodine IR and oxybutynin IR	• Algorithm based model • UK NHS payer • 1 year	Hypothetical cohort of patients with urge incontinence associated with OAB	• QALYs not reported • Cost per incontinence-free week	Total annual cost, in £: • Oxybutynin ER: 76.77 • Oxybutynin IR: 39.61 • Tolterodine IR: 74.21 • Tolterodine ER: 63.91	**ICER (cost/incontinence-free week)** • Oxybutynin IR (vs NR): £5.26 • Oxybutynin ER vs tolterodine IR: £84.82 • Tolterodine IR: dominated • Tolterodine ER vs oxybutynin IR: £7.14
Ko,[84] USA, 2006	Compare the cost-effectiveness of various antimuscarinic agents for the treatment of OAB	• Decision tree • USA payer • 3 months	NR	• QALYs not reported • Average cost/patient with continue and successful treatment	• Average 3-month cost/per patient, in $: • Solifenacin 3373 • Oxybutynin TD 3603 • Darifenacin 3633 • Oxybutynin ER 3646 • Tolterodine ER 3659 • Trospium 3722 • Tolterodine IR 3750 • Oxybutynin IR 3769	Solifenacin dominated all other comparators
Kobelt, [72] Sweden, 1998	Develop a simulation model to calculate the incremental cost-effectiveness and cost-utility of new treatments for OAB (tolterodine vs no treatment)	• Markov • Perspective: NR • 1 year	NR	Mean cumulative utility with tolterodine is 0.6977 vs 0.6728 with no treatment (for 1 year)	Cost, in $: Tolterodine: 59.2/month (dose NR, price based on anticipated sales price in Sweden) Incremental cost per patient and year with tolterodine is SEK5309 ($699) vs no treatment	Tolterodine vs no treatment: SEK213,042 (US$28,032)
Milsom,[73] Denmark, Finland, Norway, Sweden,	Compare the cost-effectiveness of solifenacin flexible dosing (5-10 mg) with tolterodine	• Decision tree • Societal and healthcare payer	NR	NR. There were only minor differences in QoL between the three treatment options.	Total yearly costs/patient (Sweden/Norway/Finland/Denmark), in €: • Placebo: 712/869/626/806 • Solifenacin flexible:	Sweden/Denmark/Norway/Finl and **ICER** Total cost • Solifenacin vs placebo:

2009	SR 4 mg or placebo for patients with OAB symptoms	• 1 year			1142/11091076/1149 • Tolterodine SR 4 mg: 1216/1205/1122/1277	€27,603/€14,318/ €26,817/€20,457 Solifenacin vs tolterodine: Dominance in all country settings
Nilsson,[74] Sweden, 2012	Analyse the cost-effectiveness of newer anticholinergic drugs in relation to oxybutynin IR and no treatment for patients with urgency urinary incontinence	• Decision tree • Healthcare payer • 1 year	NR	• Oxybutynin: 0.9376 • Newer drugs (solifenacin, tolterodine, fesoterodine, darifenacin, oxybutynin patch): 0.9435 • No treatment (no effect): 0.9301 • No treatment (placebo effect): 0.9389	Total cost (1 year), in €: • Oxybutynin 1038 • Newer drugs: 1229 • No treatment (no effect): 1012 • No treatment (placebo effect): 951	• Oxybutynin vs no treatment (no effect): €8,400 • Oxybutynin vs no treatment (placebo effect): Dominated • Newer anticholinergic drugs vs no treatment (no effect): €21,045 • Newer anticholinergic drugs vs no treatment (placebo effect): €65,435 • Newer anticholinergic drugs vs oxybutynin: €37,119
O'Brien,[75] Canada, 2001	Examine the cost-effectiveness of tolterodine for patients with urge incontinence who discontinue initial therapy with oxybutynin	• Markov • Societal • 1 year	Adult patients with urge incontinence	QALYs by disease state; normal/mild/moderate/severe: • Switch to no therapy: 0.03/0.17/0.29/0.18 • Switch to tolterodine: 0.07/0.27/0.24/0.11 No therapy: 0.67 per patient; tolterodine: 0.69 per patient	Total cost for 1 year, in Can$: • No switch: 367 • Switch to tolterodine: 530	• Switch to tolterodine is cost-effective with an **ICER** of Can$9982
Pradelli,[76] Italy, 2009	Investigate the pharmacoeconomic performance of treatment with solifenacin, when compared with tolterodine and	• Markov • Societal, Italian Health Service • 1 year	A patient cohort representative of the Italian patient population with OAB	QALY/patient • Solifenacin:0.810 • Tolterodine 0.800 • Placebo: 0.776 • No treatment: 0.740	Total cost per year, in €: • Solifenacin 5 mg: 834 • Tolterodine ER 4 mg: 988 • Placebo: 204 • No treatment: 305	**ICER (€/QALY)** • Solifenacin vs placebo: €18,612 • Solifenacin vs no treatment: €7634 • Tolterodine vs placebo:

	placebo, in patients with OAB					€33,309 • Tolterodine vs no treatment: €11,457
Speakman,[77] UK, 2008	Evaluate the cost-utility of solifenacin, compared with tolterodine in the treatment of OAB	• Markov • UK NHS payer • 1 year	Adults with OAB	• Solifenacin: 0.709 • Tolterodine: 0.705	Total cost (1 year), in £: • Solifenacin 5 mg/day: 509 • Tolterodine: 526	Solifenacin is dominant compared with tolterodine

†Where two dosages are quoted, results were given in the publication for the pooled dosage cohorts only.

Abbreviations used in table: CR, controlled-release; ER, extended-release; HS, health state; ICER, incremental cost-effectiveness ratio; ICUR, incremental cost-utility ratio; IR, immediate-release; NA, not applicable; NR, not reported; OAB, overactive bladder; QALY, quality adjusted life year; SR, sustained-release; TD, transdermal; wk, week.

5.2 Summary and ERG critique of the de novo *economic evaluation submitted by the manufacturer*

In support of this STA, the manufacturer submitted four electronic versions of the Microsoft© EXCEL-based economic model, as follows:

- a primary base case model, based on efficacy data from SCORPIO which considered the comparison of mirabegron 50 mg with tolterodine extended release (ER) 4 mg;
- a secondary base case model, based on efficacy data from the manufacturer's MTC, considering mirabegron 50 mg versus all comparators (except oxybutynin immediate response [IR] 10 mg) listed in the NICE scope [((19))];
- a version of the secondary base case model including oxybutynin IR 10 mg;
- a version of the secondary base case model including the impact of co-morbidity.

The ERG considers the manufacturer's models to be generally well constructed and largely transparent. In addition, the ERG considers that disaggregating the submitted economic analyses into distinct versions of the model facilitated examination of each analysis.

5.2.1 NICE reference case checklist (TABLE ONLY)

Tables 42 and 43 summarise the ERG's assessment of the manufacturer's economic evaluation against the NICE reference case and Philips checklists,[(85)] respectively.

Table 42. NICE reference case

Attribute	**Reference case**	**Does the *de novo* economic evaluation match the reference case?**
Decision problem	The scope developed by NICE	Yes
Comparator(s)	Alternative therapies routinely used in the NHS	Yes
Perspective costs	NHS and Personal Social Services	Yes
Perspective benefits	All health effects on individuals	Yes
Form of economic evaluation	Cost-utility analysis	Yes
Time horizon	Sufficient to capture differences in costs and outcomes	Yes.
Synthesis of evidence on outcomes	Systematic review	Yes
Outcome measure	QALYs	Yes
Health states for QALY	Described using a standardised and validated instrument	Utility values were obtained from trial based EQ-5D and OAB-q data.
Benefit valuation	Time-trade off or standard gamble	Yes
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	Yes
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes

Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes
Sensitivity analysis	Probabilistic sensitivity analysis	Yes. The manufacturer carried out deterministic sensitivity analysis, scenario analysis and probabilistic sensitivity analysis.
Abbreviations used in table: EQ-5D, EuroQol 5 dimensions questionnaire; NICE, National Institute for Health and Clinical Excellence; NHS, National Health Service; OAB, overactive bladder; QALY, quality adjusted life year.		

Table 43. Philips checklist

Dimension of quality	Comments
Structure	
S1: Statement of decision problem/objective	Clearly stated.
S2:Statement of scope/perspective	The scope and perspective of the model were clearly stated; the manufacturer fully followed the NICE scope.
S3: Rationale for structure	The ERG notes that the manufacturer's rationale for the structure of the model was based on previous publications of related economic evaluations. With the exception of the implementation of discontinuation, the ERG considers the model structure to be appropriate and well-constructed.
S4: Structural assumptions	The structural assumptions were transparent. In addition, a number of scenario and sensitivity analyses were undertaken to test the robustness of the different assumptions.
S5: Strategies/comparators	All relevant comparators were evaluated.
S6: Model type	Correct, cost-utility analysis.
S7: Time horizon	The ERG considers that the chosen time horizon of 5 years was sufficient to capture treatment-relate4d costs and consequences in a condition where treatment offers no survival benefit.
S8: Disease states/pathways	The ERG agrees with the pathways/health states modelled.
S9: Cycle length	The ERG considers 1 month a reasonable cycle length to capture the consequences of model events. Half-cycle correction was not used as it was anticipated that the full cost of prescription would be incurred regardless of patient transitions.
Data	
D1: Data identification	Data were taken from SCORPIO to inform the primary base case analysis. The manufacturer also carried out an MTC which was used to inform the secondary base case analyses. Utility data from SCORPIO were used to inform all analyses.
D2: Pre-model data analysis	Correct formulae were used in all pre-model data analyses validated by the ERG. However, the ERG was unable to validate the derivation of calibrated beta coefficients.
D2a: Baseline data	Baseline data were taken from SCORPIO.
D2b: Treatment effects	Treatment effects for each outcome were estimated from SCORPIO or the manufacturer's MTC. The model used multinomial logistic regression models to inform transition probabilities. Extrapolation of treatment effects is clearly described and justified.
D2d: Quality of life weights (utilities)	Utility data from SCORPIO were used in the submitted models.
D3: Data incorporation	The manufacturer clearly described how data were used in the model, all sources were referenced and copies of referenced papers were provided. Standard distributions were used for different outcomes (e.g. the gamma distribution for

	costs).
D4: Assessment of uncertainty	The assessment of sensitivity was thorough and generally robust. Probabilistic, one-way sensitivity analysis and various scenario analyses were reported satisfactorily.
D4a: Methodological	Appropriate analytical methods were used, and were supported with sensitivity and scenario analyses to test the robustness of the chosen base case approach.
D4b: Structural	The manufacturer described deterministic sensitivity analysis and scenario analysis in detail.
D4c: Heterogeneity	Heterogeneity was partially addressed by the analysis of different subgroups of patients. However, the secondary base case analyses were based on an MTC which exhibited high levels of heterogeneity.
D4d: Parameter	The ERG notes that the manufacturer's assessment of uncertainty did not account for correlation of regression parameters. Therefore, it is possible that the magnitude of parameter variation has not been fully accounted for.
Consistency	
C1: Internal consistency	The model seems to be mathematically sound with no obvious inconsistencies. The manufacturer reported that the model underwent validation and verification.
C2: External consistency	The model results are intuitive and conclusions are valid given the data presented.
Abbreviations used in table: ERG, Evidence Review Group; MTC, mixed treatment comparison; NICE, National Institute for Health and Clinical Excellence.	

5.2.2 Population

As per the scope issued by NICE for this STA,[19] the manufacturer's base case analysis considered a patient population of adults with OAB; the expected licensed indication for mirabegron. Baseline characteristics were taken from SCORPIO (pooled data from the three treatment arms) and were used to inform the manufacturer's economic evaluation. In addition, based on data from SCORPIO, the manufacturer carried out the following subgroup analyses:

- male patients;
- female patients;
- previously treated patients;
- treatment-naïve patients.

5.2.3 Interventions and comparators

The manufacturer's primary base case analysis compared mirabegron 50 mg with tolterodine ER 4 mg. Clinical effectiveness data from SCORPIO were used to inform the primary base case economic evaluation. As stipulated in the final scope issued by NICE, the manufacturer also compared mirabegron 50 mg with solifenacin 5 mg and 10 mg, fesoterodine 4 mg, trospium chloride 60 mg modified-release (MR) and oxybutynin (ER and IR) 10 mg. All antimuscarinics were assumed to be used within their marketing authorisation; mirabegron was assumed to be used within the anticipated

marketing authorisation. The manufacturer carried out a mixed treatment comparison (MTC; see Sections 4.1.7 and 4.3.1) "to estimate the relative efficacy and safety of mirabegron compared with all treatments of interest" (MS; pg 128) Data from the manufacturer's MTC was used to inform the secondary base case economic evaluation of mirabegron versus all comparators of interest and results were presented individually and as a fully incremental analyses (see Section 5.2.9). However, as highlighted in Section 4.1.7, the ERG has concerns regarding the level of heterogeneity identified in the manufacturer's MTC. Therefore, the ERG carried out a revised MTC, using a subset of the trials presented in the MS. The ERG selected a subset of trials that were classified as good quality (by the manufacturer) and that provided a homogeneous patient population for analysis. The methods and results of this analysis are presented in full in Section 4.4.2. To summarise, the ERG's MTC resulted in treatment effects with respect to frequency of micturition that, compared with the manufacturer's MTC results, were:

- more favourable for: tolterodine ER 4 mg, oxybutynin 10 mg (ER and IR assumed to be of equal efficacy), solifenacin 10 mg and solifenacin 5 mg;
- less favourable for: fesoterodine 4 mg, fesoterodine 8 mg and placebo;
- equal (within 0.01) for trospium chloride MR 60 mg.

Furthermore, compared to the manufacturer's MTC results, treatment effect, with respect to frequency of incontinence were:

- more favourable for: tolterodine ER 4 mg and oxybutynin 10 mg (ER and IR assumed to be of equal efficacy);
- less favourable for: fesoterodine 8 mg, solifenacin 10 mg, solifenacin 5 mg and placebo;
- equal (within 0.01) for fesoterodine 4mg

The impact of the findings from the ERG's MTC on the manufacturer's secondary base case cost-effectiveness results are discussed further in Section 6. However, it is important to note that no statistically significant differences, between mirabegron 50 mg and each active comparator considered, were identified for the outcome of frequency of micturition. Furthermore, the only comparisons that showed a statistically significant difference for the outcome of frequency of incontinence episodes were mirabegron 50 mg versus solifenacin 5 mg and versus solifenacin 10 mg; in both comparisons mirabegron was significantly less effective.

5.2.4 Model structure

The manufacturer's *de novo* Markov model considered the costs and consequences of mirabegron versus currently available antimuscarinics for overactive bladder (OAB). The therapeutic management of patients (including complications), severity and progression of disease were assessed in a hypothetical cohort of OAB patients in monthly cycles over a 5 year time horizon. The model was

constructed to assess costs and consequences from a societal or NHS Payer perspective. However, in the MS, only results from an NHS payer perspective were reported. Costs and benefits were discounted at a rate of 3.5% per annum in line with NICE reference case.[86]

Disease severity and progression

Disease severity was assumed to be a combination of the mean number of micturitions and the mean number of incontinence (rather than urge incontinence) episodes per day; which is in line with the co-primary outcome measures of the pivotal mirabegron trials (SCORPIO, ARIES and CAPRICORN). However, despite being included in the International Continence Society's definition of OAB ("urgency, with or without urge incontinence, usually with frequency and nocturia",[2] nocturia and symptoms of urgency were excluded from the manufacturer's model. The manufacturer's rationale for the exclusion of these aspects of OAB is presented in Box 8.

Box 8. The manufacturer's rationale for exclusion of urgency and nocturia from the modelled definition of OAB severity

Urgency is subjective in nature, and within clinical trials it is measured using varying instruments, and with alternative different severity thresholds, making comparisons difficult and potentially adding considerable uncertainty to the analyses. Therefore it was considered appropriate to exclude urgency from the model. Nocturia has multiple aetiologies and is multi-factorial in nature and therefore may not just be related to OAB. It has therefore been excluded from the model, consistent with previously published models.[69;70;72-77]
Abbreviation used in box: OAB, overactive bladder.

The ERG received expert clinical advice which indicated that the manufacturer's rationale for excluding nocturia was appropriate. Regarding urgency, expert clinical opinion is that urgency is an important aspect of OAB and likely to be a key component of disease severity. However, clinical advice also highlighted that there are few validated instruments used in the measurement of urgency and that definitions are likely to differ among trials. Therefore, based on expert clinical opinion, the ERG agrees with the manufacturer that the exclusion of urgency and nocturia from the model is reasonable. Furthermore, the ERG notes that, in SCORPIO, mirabegron demonstrated an improvement in the mean number of urgency (Grade 3/4 per 24 hours) episodes (change from baseline[CFB]: –2.25) and level of urgency (CFB: –0.31 [p=0.018]) that was of higher than the improvement in these outcomes observed for tolterodine (CFB: –2.07 [p=0.050] and –0.29 [p=0.085] for mean number urgency episodes and mean level of urgency, respectively). Therefore, the ERG considers that, in the primary base case, the direction of any potential bias from the exclusion of symptoms of urgency would be against mirabegron. However, the ERG notes that with respect to symptoms of urgency, there is an absence of evidence on the relative effectiveness of mirabegron

versus antimuscarinics other than tolterodine. Therefore, an assessment of bias in the comparison of mirabegron with other antimuscarinics (secondary base case) was not possible.

Within the manufacturer's model, patients were simultaneously categorized into five groups for micturition and five groups for incontinence (Table 44). The categories represent the quintiles of frequency of micturition and incontinence observed in a pooled analysis of the pivotal mirabegron trials (ARIES, SCORPIO and CAPRICORN, described in Section 4.1.6). Level 1 of each symptom represents the threshold usually used to define OAB; i.e. patients in level 1 for frequency of micturition and incontinence are not considered to have OAB.

Table 44. Micturition and incontinence categories (reproduced from MS; pg 192; Table 81)

Symptom	Level 1	Level 2	Level 3	Level 4	Level 5
Mean number of micturitions per day	≤8	>8 to ≤10	>10 to ≤12	>12 to ≤14	>14
Mean number of incontinence episodes per day	0	>0 to ≤1	>1 to ≤2	>2 to ≤3	>3

The manufacturer considered the Pearson's correlation coefficient to assess any potential relationship between the frequency of micturition and incontinence. A small, positive ($r=0.19094$: $-1<r<1$ where $r<0$ indicates negative correlation, $r>0$ indicates positive correlation), yet statistically significant (p-value <0.0001) correlation was detected; i.e. patients who experience greater frequency of micturition are more likely to suffer from incontinence and vice versa. However, within the model, the manufacturer assumed that the frequency of micturition was independent of the frequency of incontinence experienced by a patient. The ERG considers that exclusion of the statistically significant correlation between these two outcomes may have compromised the accuracy of the model with respect to the distribution of patients across different symptom severity levels. However, the ERG notes that the correlation between these outcomes is unlikely to be affected by treatment and therefore may not result in model bias either towards or against mirabegron.

The overall severity of a patient's OAB was determined by a linear combination of the severity level for each symptom; resulting in a matrix of 25 severity profiles (health states) depicted in Table 45.

Table 45. OAB severity profiles (health states) included in the manufacturer's model

Incontinence level	Micturition level				
	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I	J
3	K	L	M	N	O
4	P	Q	R	S	T
5	U	V	W	X	Y
For example, patients in severity level N would experience 12 to 14 micturitions per day and 1 or 2 episodes of incontinence					

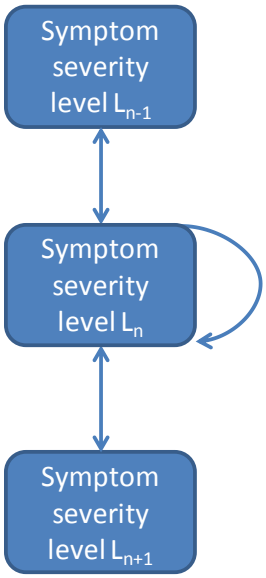
daily (see Table 44).

At model entry, patients were distributed across the 25 severity profiles (Table 46) and assigned to treatment with either mirabegron or a comparator antimuscarinic. Then in monthly cycles, patients could simultaneously transition through the five severity levels of micturition and the five severity levels of incontinence; i.e. a patient's severity profile was reassessed each month according to improvement, deterioration or stabilisation of the individual symptoms of micturition frequency and incontinence (Figure 1).

Table 46. Initial distribution of patients across symptom severity levels (adapted from MS; pg 195; Table 83)

Severity level	General OAB population	
	Micturition	Incontinence
1	6.30%	38.87%
2	30.69%	18.84%
3	27.18%	14.64%
4	19.46%	9.18%
5	16.37%	18.47%
Abbreviation used in table: OAB, overactive bladder.		

Figure 1. Symptom specific transitions between severity levels (occur simultaneously for each considered symptom)



Therapeutic management

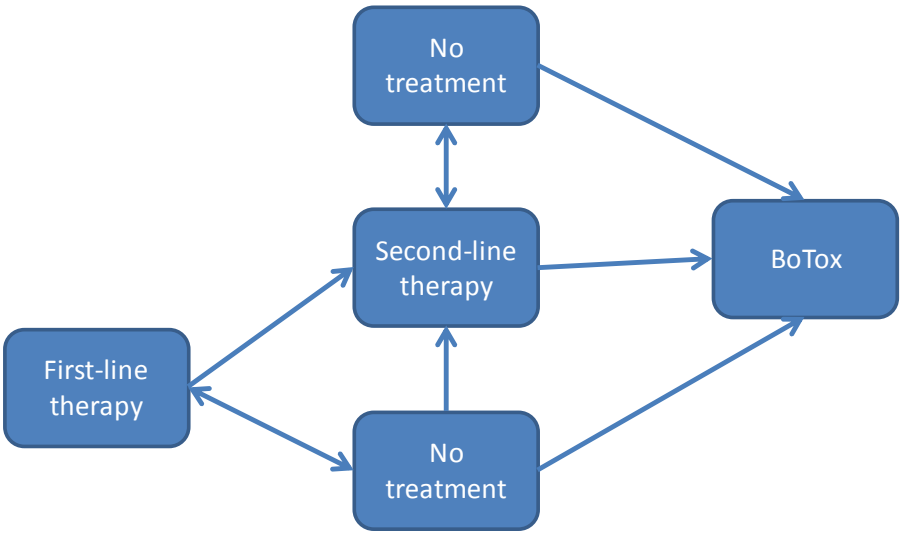
In addition to disease progression, captured by transitions between levels of symptom severity, the manufacturer's model assessed the therapeutic management of patients; including treatment discontinuation; treatment switch and the management of adverse events (AEs).

Treatment discontinuation was assumed to be a result of either AEs or other causes (e.g. lack of efficacy). The AEs considered in the manufacturer's model were limited to dry mouth and constipation. The manufacturer's rationale for excluding other AEs (such as blurred vision) was based on evidence from a European cross-sectional survey of physicians and OAB patients by Compion et al.[87] Compion et al. reported that dry mouth and constipation were the two most frequently reported side effects causing treatment switch. Based on expert clinical advice, the ERG accepts the manufacturer's assumption that dry mouth and constipation would be the main drivers of AE related discontinuation. Furthermore, the ERG notes that the relative differences between mirabegron and tolterodine with respect to other treatment-related AEs, e.g. hypertension) observed in SCORPIO and TAURUS were not substantial. Therefore, the ERG considers it unlikely that the exclusion of other AE would bias the model either towards or against mirabegron.

Figure 2 summarises the treatment switch and discontinuation pathways permitted within the manufacturer's model. Patients who discontinued their original therapy (mirabegron 50 mg or comparator) could either switch to a second antimuscarinic (assumed to be solifenacin 5 mg) or opt to receive no treatment; patients who discontinued from their second antimuscarinic therapy were assumed to receive either botulinum toxin (BoTox) or no treatment. At any point in the model time horizon, patients receiving no treatment were able to restart pharmacological therapy with:

- their original antimuscarinic, a second antimuscarinic or BoTox – for patients who discontinued from their originally assigned therapy;
- their second antimuscarinic or BoTox – for patients who discontinued from their second antimuscarinic therapy.

Figure 2. Treatment switch and discontinuation pathways permitted in the manufacturer's model.



Within the manufacturer's model, the reason for discontinuation (i.e. as a result of AEs or other causes) affected the timing and probability of discontinuation. Patients who discontinued as a result of

other causes did so immediately; i.e. within the same cycle. Patients who experienced an AE either discontinued immediately or remained on treatment. Patients who remained on treatment (original or second treatment) following the occurrence of an AE moved into a separate health state in which they accrued the associated AE disutility and were subject to a higher probability of discontinuation. Figure 3 displays the full model structure, including disease progression, treatment switch and AEs. Table 47 summarises the estimates used within the manufacturer's model to inform treatment discontinuation and switch.

Figure 3. Model structure

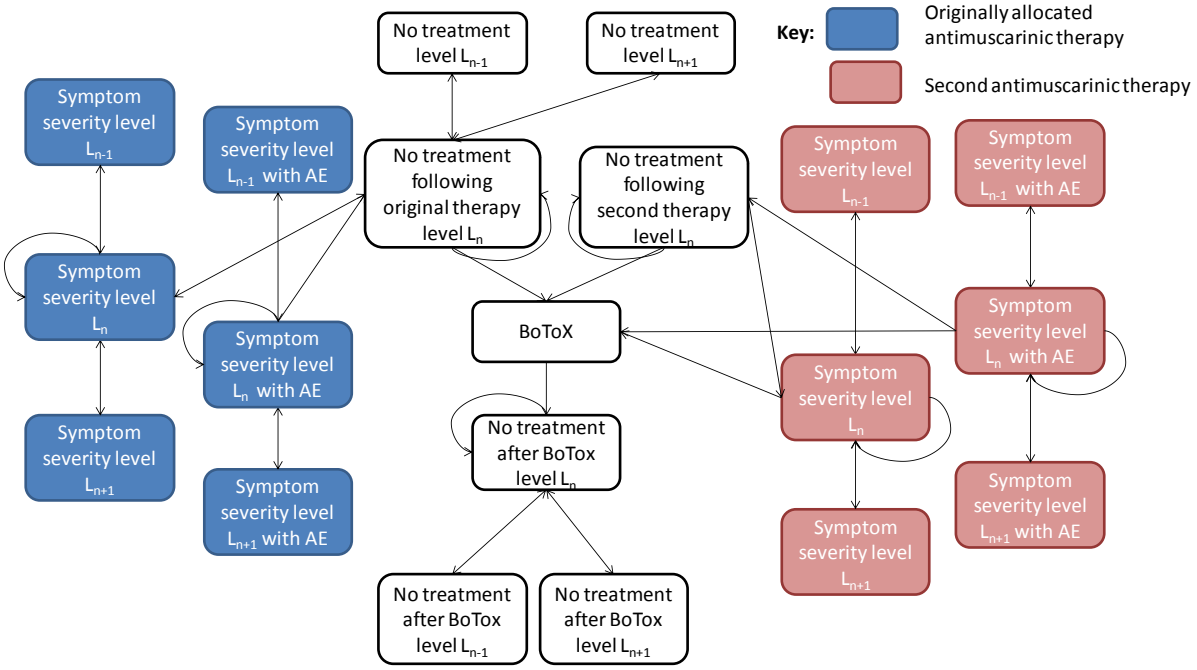


Table 47. Treatment discontinuation and switch parameters used in the manufacturer's model

Parameter	Reason for discontinuation		Source
	Adverse events	Other causes	
Probability of immediate discontinuation (per month)	Treatment specific discontinuation rate x Probability of an AE	Treatment specific discontinuation rate	Wagg et al. 2012,[18] Castro-diaz 2001[88]
Probability of discontinuation in subsequent cycles (per month)	90%	N/A	Expert opinion
For patients discontinuing from original therapy			
Probability of receiving second antimuscarinic	26.1%		Odeyemi et al.[89]
Probability of receiving no treatment	73.9%		Calculated[a]

Probability of restarting pharmacological therapy following discontinuation to no treatment	5.6%	Expert opinion
Probability that treatment restarted is original therapy	33.3%	Expert opinion
Probability that treatment restarted is second antimuscarinic	66.6%	Expert opinion
Probability of receiving BoTox following discontinuation to no treatment	0.08%	Expert opinion
For patients discontinuing from second antimuscarinic therapy		
Probability of receiving no treatment	99.92%	Calculated[b]
Probability of receiving BoTox	0.08%	Expert opinion
Probability of restarting pharmacological therapy following discontinuation to no treatment	5.6%	Expert opinion
Probability of receiving BoTox following discontinuation to no treatment	0.08%	Expert opinion

[a] 1-probability of receiving second antimuscarinic.

[b] 1-probability of BoTox.

Abbreviations used in table: AE, adverse event; BoTox; Botulinum toxin.

To summarise, the ERG considers that the structure of the manufacturer's model was reasonable and appropriately captured the consequences (costs and benefits) of treatments for OAB. However, the ERG notes that no rationale was provided for assuming that patients who fail on conservative antimuscarinic therapy would receive BoTox, rather than other invasive procedures recommended by NICE.[(12)] In addition, the ERG notes that a majority of parameters informing treatment discontinuation and switch were based on expert clinical opinion. Furthermore, the ERG notes that these values were estimated through open discussion, rather than through the use of elicitation techniques.[(90)] Therefore, the ERG considers that this aspect of the manufacturer's model will be subject to additional parameter uncertainty. In particular, the probability of discontinuation in subsequent cycles for patients experiencing an AE; expert clinical advice received by the ERG indicated that a 90% probability of discontinuation in patients who are experiencing an AE may be too high. However, the ERG notes that the manufacturer has included a sensitivity analysis examining the impact of assuming a much lower (50%) probability of AE related discontinuation on the cost-effectiveness results (see Section 5.2.10).

Furthermore, the ERG notes that although the manufacturer has disaggregated discontinuation as a result of an AE and discontinuation as a result of other-causes, in some instances the same probability is used for both. In particular, the probability of other-cause discontinuation is used to inform the probability of immediate discontinuation (i.e. within the same cycle) as a result of an AE. The probability of other-cause discontinuation was assumed to be treatment specific and in the manufacturer's base case (primary and secondary), was derived from the published literature to exclude discontinuation as a result of AEs. Therefore, the ERG considers the application of the probability of other cause discontinuation to the probability of discontinuation as a result of an AE to be inappropriate. Full details of the literature sources and calculations used to derive these treatment specific probabilities and ERG sensitivity analyses around these parameters are presented in Section 5.2.6.

Finally, the ERG notes that, within the manufacturer's model, the number of times a patient may discontinue to no treatment (from either first- or second-antimuscarinic treatment) and then reinitiate previous therapy was unlimited. However, the ERG acknowledges that this is a consequence of the 'lack of memory' attributed to the Markov model structure used. Furthermore, the ERG investigated the potential impact of this assumption in a sensitivity analysis which assumed that no patients would restart their previous therapy. The impact of this on the manufacturer's primary base case cost-effectiveness results was to increase the ICER by ~3% (from £4,386 to £4,516). The impact of this sensitivity analysis on the incremental results (secondary base case) is presented in Section 6 and Appendix 11.

5.2.5 Summary of model parameters

Table 48 summarises all parameters used in the manufacturer's primary base case model. However, with the exception of β-parameters (provided in Appendix 8), the manufacturer's secondary base case model uses the same parameters as the primary base case.

Table 48: Summary of base case model parameters

Parameter	Base case value	DSA values	PSA
Statistical distributions for proportions of patients by severity level at baseline			
Micturition 1	6.30%	0% - 0%	Dirichlet distribution (α1, α2, α3, α4, α5)=(120,585,518,371,312)
Micturition 2	30.69%	100% - 0%	
Micturition 3	27.18%	0% - 0%	
Micturition 4	19.46%	0% - 0%	
Micturition 5	16.37%	0% - 100%	
Incontinence 1	38.87%	100% - 0%	Dirichlet distribution (α1, α2, α3, α4, α5)=(741,359,279,175 ,352)
Incontinence 2	18.84%	0% - 0%	
Incontinence 3	14.64%	0% - 0%	
Incontinence 4	9.18%	0% - 0%	
Incontinence 5	18.47%	0% - 100%	

Probabilities of transition between different severity levels, by treatment			
Beta coefficients for Mirabegron 50 mg			
Micturition 1 (5 as reference)	0.6037	0.2239 – 0.9835	Normal distribution (μ , σ)=(0.6037,0.1938)
Micturition 2 (5 as reference)	0.3803	0.0295 – 0.7311	Normal distribution (μ , σ)=(0.3803,0.1790)
Micturition 3 (5 as reference)	0.1454	-0.1876 – 0.4784	Normal distribution (μ , σ)=(0.1454,0.1699)
Micturition 4 (5 as reference)	0.0665	-0.2738 – 0.4068	Normal distribution (μ , σ)=(0.0665,0.1736)
Incontinence 1 (5 as reference)	0.3617	0.0054 – 0.7180	Normal distribution (μ , σ)=(0.3617,0.1818)
Incontinence 2 (5 as reference)	0.4634	0.1043 – 0.8225	Normal distribution (μ , σ)=(0.4634,0.1832)
Incontinence 3 (5 as reference)	-0.0251	-0.4042 – 0.3540	Normal distribution (μ , σ)=(-0.0251,0.1934)
Incontinence 4 (5 as reference)	0.2040	-0.2119 – 0.6199	Normal distribution (μ , σ)=(0.2040,0.2122)
Beta coefficients for Tolterodine ER 4 mg			
Micturition 1 (5 as reference)	0.3667	-0.0073 – 0.7407	Normal distribution (μ , σ)=(0.3667,0.1908)
Micturition 2 (5 as reference)	0.1826	-0.1610 – 0.5262	Normal distribution (μ , σ)=(0.1826,0.1753)
Micturition 3 (5 as reference)	-0.0609	-0.3867 – 0.2649	Normal distribution (μ , σ)=(-0.0609,0.1662)
Micturition 4 (5 as reference)	0.0550	-0.2739 – 0.3839	Normal distribution (μ , σ)=(0.0550,0.1678)
Incontinence 1 (5 as reference)	0.1431	-0.2028 – 0.4890	Normal distribution (μ , σ)=(0.1431,0.1765)
Incontinence 2 (5 as reference)	0.1768	-0.1735 – 0.5271	Normal distribution (μ , σ)=(0.1768,0.1787)
Incontinence 3 (5 as reference)	-0.3271	-0.7009 – 0.0467	Normal distribution (μ , σ)=(-0.3271,0.1907)
Incontinence 4 (5 as reference)	-0.0298	-0.4385 – 0.3789	Normal distribution (μ , σ)=(-0.0298,0.2085)
Beta coefficients for Solifenacin 5 mg			
Micturition 1 (5 as reference)	0,9977	0,6237 – 1.3717	Normal distribution (μ , σ)=(0,9977,0,1908)
Micturition 2 (5 as reference)	0,4933	0,1497 – 0.8639	Normal distribution (μ , σ)=(0,4933,0,1753)
Micturition 3 (5 as reference)	0,0384	0,3641 - -0.2874	Normal distribution (μ , σ)=(0,0384,0,1662)
Micturition 4 (5 as reference)	-0,0729	0,2560 - -0.4017	Normal distribution (μ , σ)=(-0,0729,0,1678)
Incontinence 1 (5 as reference)	1,1403	0,7944 – 1.4863	Normal distribution (μ , σ)=(1,1403,0,1765)
Incontinence 2 (5 as reference)	0,7343	0,3840 – 1.0845	Normal distribution (μ , σ)=(0,7343,0,1787)
Incontinence 3 (5 as reference)	0,0347	0,4084 - -0.3391	Normal distribution (μ , σ)=(0,0347,0,1907)

Incontinence 4 (5 as reference)	0,1136	0,5223 - -0.2950	Normal distribution (μ , σ)=(0,1136,0,2085)
Probability of having a dry mouth AE			
Mirabegron 50 mg	2.80%	2.1% - 3.5%	Beta distribution (α,ß)=(47.60,1652.40)
Tolterodine ER 4 mg	10.10%	8.7% - 11.5%	Beta distribution (α,ß)=(113.12,1006.86)
No treatment	0%	NA	NA
Probability of having a constipation AE			
Mirabegron 50 mg	1.60%	1% - 2.20%	NA
Tolterodine ER 4 mg	2%	1.40% - 2.60%	NA
No treatment	0%	NA	NA
Probability of success of botulinum toxin (all patients)			
	79%	60% - 92%	NA
Utilities according to symptom severity – EQ-5D (coefficients of regression equation)			
Micturition 1 (5 as reference)	0.0632	0.0453 – 0.0811	Normal distribution (μ , σ)=(0.0632,0.0091)
Micturition 2 (5 as reference)	0.0422	0.0258 – 0.0587	Normal distribution (μ , σ)=(0.0422,0.0084)
Micturition 3 (5 as reference)	0.0204	0.0045 – 0.0363	Normal distribution (μ , σ)=(0.0204,0.0081)
Micturition 4 (5 as reference)	0.0104	-0.0054 – 0.0262	Normal distribution (μ , σ)=(0.0104,0.0081)
Incontinence 1 (5 as reference)	0.0586	0.0422 – 0.0749	Normal distribution (μ , σ)=(0.0586,0.0083)
Incontinence 2 (5 as reference)	0.0437	0.0271 – 0.0602	Normal distribution (μ , σ)=(0.0437,0.0084)
Incontinence 3 (5 as reference)	0.0314	0.0142 – 0.0486	Normal distribution (μ , σ)=(0.0314,0.0088)
Incontinence 4 (5 as reference)	0.0128	-0.0056 – 0.0313	Normal distribution (μ , σ)=(0.0128,0.0094)
Utility decrement associated with AE			
All AE	-0.0357	0 - -0.1	NA
Pad use per day by level of incontinence (coefficients of linear regression equation)			
Incontinence 1	0.17	0.150 – 0.198	NA
Incontinence 2	0.75	0.687 – 0.817	NA
Incontinence 3	1.38	1.282 – 1.486	NA
Incontinence 4	1.89	1.745 – 2.039	NA
Incontinence 5	3.34	3.167 – 3.511	NA
Monthly probability of discontinuation of OAB therapy			
Without AEs	6.40%	0% - 14.5%	NA
With AEs	90%	50% - 100%	Beta distribution (α,ß)=(6.92,0.77)
Monthly probability of switch after discontinuation of OAB therapy			
Probability of switch, among all patients discontinuing OAB treatment	26.06%	15.32% - 50%	Beta distribution
Monthly probabilities of restarting OAB therapy among patients without treatment			

Monthly probability of restarting treatment	10%	0.05% - 20%	Beta distribution (α,ß)=(1.74,15.63)
Split between different medications, for general OAB population*			
- Initial treatment (mirabegron or tolterodine)	33.33%	0% - 50%	NA
- Next line A	33.33%	0% - 50%	NA
- Next line B	33.33%	0% - 50%	NA
Monthly probability of transition to botulinum toxin			
Monthly probability of having botulinum toxin injection in the general OAB population	0.01%	0% - 0.05%	Beta distribution (α,ß)= (0.70,834.78)
Resource utilisation (physician visits and botulinum toxin reinjections)			
Number of GP consultations	1 visit at the start and at every switch	0 - 2	Lognormal distribution (μ , σ)=(1,0.20)
Number of specialist consultations	1.5 visits at the start and at every switch	1 - 3	Lognormal distribution (μ , σ)=(1.5,0.95)
Number of Botulinum toxin reinjections, following success of first injection	0.17 per month	0	NA
Model inputs: Monthly OAB medication costs			
Mirabegron 50 mg	£28.00	NA	NA
Tolterodine ER 4 mg	£28.01	£8.4	NA
Model inputs: unit costs of health care resources			
GP consultation	£36	NA	NA
Specialist visit: Follow-up visit	£96	NA	NA
Botulinum toxin injection: Initial / Reinjections	£ 1158 / £964	NA	NA
Incontinence pad (per pad)	£0.16	NA	NA
Model inputs: cost of absenteeism			
Proportion of workers	NA	46.28%	NA
Labour cost per month	NA	£2,923	NA
Discount rates			
Costs	3.5%	3.5%-6%	NA
Outcomes (QALYs)	3.5%	0%-6%	NA
Abbreviations used in table: AE, adverse event; DSA, deterministic sensitivity analysis; ER, extended release; GP, general practitioner; NA, not applicable; OAB, overactive bladder; PSA, probabilistic sensitivity analysis; QALY, quality adjusted life year.			

5.2.6 Treatment effectiveness and extrapolation

The manufacturer's model captures the effect of antimuscarinic treatment on: disease severity; AEs; and discontinuation. As discussed in Section 5.2.3, clinical effectiveness data from SCORPIO were used to inform the manufacturer's primary base case economic evaluation (of mirabegron 50 mg versus tolterodine ER 4 mg); in particular, disease severity and incidence of AEs. Whereas, data from the manufacturer's MTC were used to inform the clinical effectiveness (disease severity and AE rates) of other relevant comparators listed in the final scope issued by NICE.[19] However, in the manufacturer's base case models (primary and secondary), the effect of treatment on discontinuation was obtained from the literature and expert clinical opinion.

Disease severity – primary base case model

As discussed in Section 5.2.4, a patient's disease severity profile was assumed to be a linear combination of severity with respect to frequency of micturition and incontinence episodes per day; with transitions between severity levels occurring simultaneously for each symptom. In the primary base case, the probabilities of transition between severity levels for each symptom were derived from multinomial logistic regression models estimated from SCORPIO data.

Multinomial logistic regression models are a technique used to assess the impact of covariates on categorical response data, with more than two categories. Response categories may be ordinal (consisting of ordered categories) or nominal (consisting of unordered categories) depending on the nature of the response. In the case of ordinal response categories, separate logistic regression models may be developed to assess the impact of covariates on each category; the results of each analysis are then pooled to give the overall result.[91]

However, the use of separate logistic regression models relies on the proportional odds assumption that "the effects of any explanatory variables are consistent across different pairs of symptom levels"; i.e. that the explanatory variables have the same effect on the odds of transition between levels of symptom severity regardless of the cut-off level used" (MS; pg 195). The manufacturer assessed the validity of the proportional odds assumption using a "multinomial logit model estimated from SCORPIO for each symptom variable" (Manufacturer's clarification response; pg 28). The null hypothesis of equal coefficients associated with different severity levels was tested and rejected at the 0.01% significance level (i.e. $p<0.0001$). Therefore, for each symptom, a single multinomial logistic regression model was used to determine the probability of transition between levels of symptom severity.

Each regression model incorporated covariates for: treatment; symptom severity in previous month (t); gender; and age. In addition, the manufacturer stated that "The interaction between the covariates of treatment and symptom severity was also tested and appeared not significant" (MS; pg 195). As

part of the clarification process, the ERG requested further details of the manufacturer's rationale for using regression analysis to estimate transition probabilities, the identification and selection of covariates and the test used to assess interaction between treatment and symptom severity. The manufacturer's rationale for using regression analysis to obtain transition probabilities is displayed is Box 9. In addition, the manufacturer described the process undertaken for covariate selection (Box 10).

Box 9. Rationale for using regression analysis to inform transition probabilities (reproduced from manufacturer's clarification response; pg 27)

We used a regression analysis to obtain the probability (from Astellas trials) of being in a given severity level in the next cycle according to treatment, severity level in current cycle, time of follow-up visit, and adjust on known confounding variables such as gender and age.

There are three advantages of using this regression model rather than simply using proportions of patients moving from one level to another:

1. The probabilities are adjusted on potential confounders, such as age and gender
2. This reduces the number of parameters in the model: the transition matrix for each symptom and each treatment has 25 parameters, or 20 parameters considering that the sum of probabilities in each row must be 1. As transition probabilities are different in the first, second and subsequent months, this would represent 60 parameters for each symptom and treatment. With the regression model, the number of parameters is reduced to four coefficients for each treatment, and 36 parameters not related to treatment. This facilitated the sensitivity analysis on the treatment effect, and most importantly, this made it possible to use a calibration approach to obtain transition probabilities for treatment not included in SCORPIO. It would not be possible to obtain three matrices with 60 independent parameters by calibration.
3. If we had used the matrices based on proportions of patients, patients lost to follow-up would be ignored. Using the regression analysis, accounting for repeated observations, data for patients lost to follow-up are implicitly imputed.

Box 10. Manufacturer's description of covariate selection (reproduced from manufacturer's clarification response; pg 28-29)

> Treatment was included in the model because the objective was to estimate the effect of treatment. Symptom severity in previous month was also entered in the model because we expected that the severity level in a given month would be dependent on the severity level in the previous month (e.g. patients were more likely to improve than worsen during the trial, thus it would be rather unlikely for a patient at Level 1 to worsen to Level 5 in the next month, whereas patients at Level 5 could stay in the same level). Results confirmed that symptom severity in previous month was a good predictor of severity in current month. Age and gender were also entered as adjustment factors, because we expected that probabilities of improvement or worsening (independent of treatment, i.e. according to natural disease history) might vary with age and gender. The effect of gender was significant, the effect of age was not, but the variable was left in the model nevertheless; removing that variable had little impact on other coefficients.
>
> [In addition] a logistic model that contained the interaction between treatment and severity in the previous month showed that this interaction was not significant.

The ERG notes that covariate selection was pragmatic rather than systematic and it is not clear whether clinical experts were involved in the process. However, the ERG accepts that the manufacturer's rationale for using regression analysis to determine transition probabilities and considers the manufacturer's regression models to have face validity. In addition, the ERG considers it unlikely that any important covariates have been overlooked.

Therefore, the final regression equations used to determine the probability of transition to symptom level *j* (for frequency of micturition and number of incontinence episodes per day) were constructed as follows:

$$\Pr ob(Severity_{t+1} = j|x) = \frac{e^{\beta_0^j + \beta_1^j .Treatment + \beta_2^j .Severity_t + \beta_3^j .Sex + \beta_4^j .Age}}{1 + \sum_{k=1}^{j-1} e^{\beta_0^k + \beta_1^k .Treatment + \beta_2^k .Severity_t + \beta_3^k .Sex + \beta_4^k .Age}}$$

Furthermore, the manufacturer highlights that "the model accounts for the fact that probabilities of improvement or worsening of symptoms may differ between the short-term and the long-term" (MS; pg 189) That is, for each considered symptom, three transition probability matrices were developed based on data from SCORPIO for baseline to month 1, month 1 to month 2 and month 2 to month 3. These are displayed in Table 49 and Table 50 for micturition and incontinence, respectively.

Table 49. Treatment-specific transition probabilities between micturition levels (adapted from MS; pg 348; Tables 156 and 157)

Treatment	**Mirabegron 50 mg**					**Tolterodine ER 4 mg**				
To:	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
From:	***Severity level at 1 month***									
1	0.805	0.180	0.013	0.002	0.000	0.799	0.186	0.013	0.002	0.000
2	0.408	0.465	0.113	0.012	0.002	0.397	0.472	0.113	0.015	0.003
3	0.160	0.387	0.343	0.084	0.026	0.152	0.381	0.335	0.100	0.031
4	0.055	0.202	0.368	0.251	0.124	0.050	0.188	0.340	0.281	0.141
5	0.030	0.074	0.156	0.241	0.500	0.025	0.064	0.133	0.251	0.527
From:	***Severity level at 2 months***									
1	0.761	0.213	0.021	0.004	0.001	0.754	0.219	0.021	0.005	0.001
2	0.334	0.476	0.162	0.023	0.004	0.324	0.480	0.162	0.028	0.005
3	0.107	0.321	0.399	0.132	0.040	0.100	0.312	0.385	0.155	0.048
4	0.030	0.138	0.352	0.323	0.157	0.027	0.126	0.319	0.355	0.175
5	0.014	0.043	0.128	0.268	0.546	0.011	0.037	0.109	0.275	0.568
From:	***Severity level at 3 months***									
1	0.734	0.237	0.024	0.004	0.001	0.726	0.243	0.024	0.005	0.001
2	0.302	0.497	0.175	0.021	0.005	0.293	0.501	0.175	0.025	0.006
3	0.094	0.326	0.420	0.115	0.046	0.088	0.317	0.405	0.135	0.055
4	0.027	0.140	0.372	0.282	0.179	0.024	0.128	0.337	0.311	0.200
5	0.012	0.042	0.129	0.223	0.594	0.004	0.020	0.086	0.243	0.646
Abbreviation used in the table: ER, extended release.										

Table 50. Treatment specific transition probabilities between levels of incontinence (adapted from MS; pg 349; Tables 159-160)

Treatment	**Mirabegron 50 mg**					**Tolterodine ER 4 mg**				
To:	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
From:	***Severity level at 1 month***									
1	0.879	0.100	0.012	0.005	0.005	0.884	0.094	0.011	0.005	0.006
2	0.518	0.364	0.078	0.022	0.018	0.532	0.349	0.074	0.022	0.023
3	0.348	0.354	0.184	0.076	0.037	0.359	0.341	0.175	0.077	0.048
4	0.209	0.290	0.219	0.158	0.125	0.211	0.273	0.203	0.157	0.157
5	0.123	0.134	0.135	0.144	0.463	0.113	0.115	0.114	0.130	0.528
From:	***Severity level at 2 months***									
1	0.866	0.105	0.015	0.007	0.007	0.871	0.098	0.014	0.007	0.009
2	0.484	0.361	0.096	0.033	0.026	0.497	0.346	0.091	0.033	0.033
3	0.305	0.329	0.212	0.105	0.050	0.313	0.316	0.201	0.106	0.064
4	0.168	0.247	0.231	0.199	0.154	0.168	0.231	0.213	0.196	0.192
5	0.089	0.103	0.129	0.164	0.515	0.080	0.087	0.107	0.146	0.580
From:	***Severity level at 3 months***									
1	0.850	0.120	0.015	0.008	0.008	0.856	0.113	0.014	0.008	0.010
2	0.454	0.394	0.091	0.034	0.026	0.467	0.379	0.086	0.035	0.033
3	0.284	0.357	0.201	0.109	0.050	0.293	0.343	0.190	0.110	0.064
4	0.156	0.267	0.218	0.206	0.152	0.156	0.250	0.201	0.203	0.190
5	0.083	0.112	0.122	0.170	0.512	0.052	0.070	0.093	0.167	0.618
Abbreviation used in table: ER, extended release.										

For patients who remained on treatment for longer than 3 months, transition matrices from month 2 to month 3 were reapplied for each cycle until discontinuation. The ERG notes that the reapplication of these transition probabilities assumes that the treatment effect seen in month 3 would be maintained for the duration of treatment. The manufacturer's rationale for the assumption of constant treatment effect from month 3 was based on evidence from a long-term (24 to 32 weeks) open-label extension study of patients treated with fesoterodine.[92] The results of the study by Van Kerrebroeck et al. suggested that treatment effect at 4 months was maintained through to month 24. In addition, the ERG notes that evidence from the manufacturer's long-term safety study (TAURUS) suggested that treatment effect with mirabegron is likely to be sustained for up to 1 year. Therefore, the ERG considers the manufacturer's assumption of sustained treatment effect to be reasonable.

Disease severity – other treatments of interest

Based on clinical effectiveness data from the manufacturer's MTC, the effect of treatment on disease severity was considered within the model for the following treatment regimens:

- solifenacin 5 mg;
- solifenacin 10 mg;
- fesoterodine 4 mg;
- trospium chloride MR 60 mg;
- oxybutynin ER 10 mg;
- oxybutynin IR 10 mg (assumed to have the same efficacy as oxybutynin ER 10 mg).

However, rather than direct implementation of the MTC results within the model (e.g. by the application of hazard ratios [HRs] to a selected baseline), the manufacturer used a calibration method based on the following methodology proposed by Vanni et al.[93]

1. choose parameters to be varied in the calibration approach.
2. choose the data target for calibration.
3. choose the measure of goodness-of-fit to be used.
4. choose parameter search strategy.
5. set convergence criteria for estimate of goodness-of-fit.
6. set stopping rule for calibration.
7. implement model calibration results.

The manufacturer chose, to base the calibration procedure, around the β-parameters (β_1-β_4) of the multinomial logistic regression models, used to inform patient transitions through levels of symptom severity. Clinical effectiveness data from the manufacturer's MTC, in particular, mean change in symptoms from baseline (to 3 months) were chosen as the targets for calibration. The manufacturer states that "The aim of the calibration method was to determine the β_1,…, β_4 estimates in the logistic

model for a given treatment by minimizing the distance between the mean change in symptoms from baseline [to 3 months] predicted by the [economic] model and the mean change [in symptoms] determined from a MTC" (MS; pg 196). The manufacturer highlighted that the mean change in symptoms from baseline to 3 months could not be directly obtained from the economic model programmed in Excel© (MS; pg 197). Therefore, a sub-model (not submitted) was programmed in Scilab© to facilitate the estimation of modelled mean change in symptoms. As a result of time constraints and the absence of the Scilab© sub-model in the MS, the ERG were unable to validate the manufacturer's estimates.

Optimisation (i.e. searching for β-parameters such that the difference between the mean change in symptoms predicted by the economic model and the mean change in symptoms determined from the MTC was equal to zero) was carried out in Scilab© (code not submitted). The manufacturer did not specify the: goodness-of-fit measure; strategy for β-parameter selection; convergence criteria; or stopping rule used. However, the manufacturer did highlight that "there was potentially an infinity of solutions. Therefore three solutions, i.e. three series of beta coefficients, were generated for each symptom by the calibration procedure." (MS; pg 197) The following sets of β-parameters were used to initiate the calibration procedure:

1. coefficients (β-parameters) for mirabegron 50 mg from the logistic regression based on the SCORPIO study;
2. coefficients for tolterodine ER 4 mg from the logistic regression based on the SCORPIO study;
3. coefficients for solifenacin 5 mg from a logistic regression based on data from the study 905-CL-015.

Within the MS, the manufacturer presented β-parameters derived from optimisation of the coefficients for mirabegron 50 mg (MS; Table 170; pg 359). However, these coefficients did not match those used in the submitted economic models. As part of the clarification process, the ERG requested the β-parameters obtained from optimisation on the coefficients obtained from tolterodine and solifenacin data; from SCORPIO and study 905-CL-015, respectively. In the clarification response, the manufacturer provided the requested β-coefficients. In addition, the manufacturer highlighted that the β-coefficients presented in the MS were in error and provided the actual coefficients used in the secondary base case model (Appendix 8).

The manufacturer also highlighted that in the primary base case model, the coefficients used to inform solifenacin efficacy (assumed to be second antimuscarinic used) were derived from calibration on mirabegron data, rather than from solifenacin data from study 905-CL-015. The manufacturer's rationale for this was that use of calibrated β-coefficients rather than those derived from a single trial optimised the use of available data (manufacturer's clarification response, pg 33). However, the ERG

notes that β-coefficients informing mirabegron and tolterodine efficacy were trial rather than calibrated data based. Therefore, the ERG carried out a sensitivity analysis to assess the impact of using trial rather than calibrated data to inform the efficacy of solifenacin on the manufacturer's primary base case cost-effectiveness results. The primary base case ICER did not change.

In addition, the ERG carried out sensitivity analyses to establish the impact of using β-coefficients derived from calibration on tolterodine ER 4 mg data and solifenacin 5 mg data. The individual (mirabegron versus comparator) results of these sensitivity analyses are presented in Appendix 9. To summarise, using β-coefficients derived from calibration on tolterodine data resulted in ICERs that were less favourable for mirabegron. Whereas, using β-parameters derived from calibration on solifenacin data resulted in ICERs that were more favourable for mirabegron. Based on this, the ERG considers the manufacturer's use of β-parameters derived from calibration on mirabegron data in the secondary base case model to be reasonable.

Adverse events

As discussed in Section 5.2.4, the AEs considered in the manufacturer's model were limited to dry mouth and constipation. Monthly probabilities of dry mouth and constipation (Table 51) were derived from SCORPIO to inform the primary base case analysis (mirabegron 50 mg versus tolterodine ER 4 mg). For all treatments considered in the secondary base case analyses, log odds ratios for each treatment compared with mirabegron were estimated from the manufacturer's MTC, and used to adjust the probability of each AE (Table 52).

Table 51. Probability of adverse events used in the manufacturer's primary base case analysis

Treatment	Adverse event		Source
	Dry mouth: mean (95% CI)	Constipation: mean (95% CI)	
3-month probabilities			
Mirabegron 50 mg	2.8% (2.1% to 3.5%)	1.6% (1% to 2.2%)	SCORPIO
Tolterodine ER 4 mg	10.1% (8.7% to 11.5%)	2.0% (1.4% to 2.6%)	
Monthly probabilities			
Mirabegron 50 mg	0.9%	0.5%	Calculated[a]
Tolterodine ER 4 mg	3.5%	0.7%	

[a] Monthly probability = 1-((1-3-month probability)^(4/12)).
Abbreviation used in table: ER, extended release.

Table 52. Probability of adverse events used in the manufacturer's secondary base case analyses

Treatment	Odds ratio (OR)[a]	3-month probability[b]	Monthly probability[c]	Source
Dry mouth				
Mirabegron 50 mg	N/A	2.80%	0.94%	Reference / SCORPIO
Tolterodine ER 4 mg	4.17	10.72%	3.71%	MTC results: Random effect model
Solifenacin 5 mg	4.23	10.86%	3.76%	MTC results: Random effect model
Solifenacin 10 mg	10.08	22.50%	8.15%	MTC results: Random effect model
Fesoterodine 4 mg	4.44	11.33%	3.93%	MTC results: Random effect model
Oxybutynin ER 10 mg	6.80	16.37%	5.78%	MTC results: Random effect model
Oxybutynin IR 10 mg	14.07	28.84%	10.72%	MTC results: Random effect model
Trospium MR 60 mg	4.48	11.43%	3.97%	MTC results: Random effect model
Constipation				
Mirabegron 50 mg	N/A	1.60%	0.54%	Reference / SCORPIO
Tolterodine ER 4 mg	1.11	1.77%	0.59%	MTC results: Fixed effect model
Solifenacin 5 mg	2.50	3.91%	1.32%	MTC results: Fixed effect model
Solifenacin 10 mg	4.37	6.63%	2.26%	MTC results: Fixed effect model
Fesoterodine 4 mg	1.07	1.70%	0.57%	MTC results: Fixed effect model
Oxybutynin ER 10 mg	1.02	1.63%	0.55%	MTC results: Fixed effect model
Oxybutynin IR 10 mg	1.02	1.63%	0.55%	Assumption (same as Oxybutynin ER 10 mg)

[a] Versus mirabegron 50 mg – OR <1 favours comparator, OR ≥1 favours mirabegron

[b] OR*baseline probability/((1–baseline probability)+OR*baseline probability)

[c] Monthly probability = 1–((1–3-month probability)^(4/12))

The ERG notes that AE data from the 12-week SCORPIO study, rather than the 12 month TAURUS study have been used to inform the manufacturer's model. The use of data from a trial of shorter duration may result in unnecessary extrapolation of AE rates over time. Furthermore, the ERG notes that the relative difference (between mirabegron and tolterodine) in AE rates was higher in SCORPIO than TAURUS (Table 53).

Table 53. Monthly probability of TEAEs from SCORPIO and TAURUS

Adverse event	SCORPIO		TAURUS	
	Mirabegron 50 mg	Tolterodine ER 4 mg	Mirabegron 50 mg	Tolterodine ER 4 mg
Dry mouth				
12-week probability	2.8%	10.1%	–	–
12-month probability	–	–	2.8%	8.6%
Monthly probability	0.9%	3.5%	0.2%	0.7%
Constipation				
12-week probability	1.6%	2.0%	-	-
12-month probability	–	–	2.8%	2.7%
Monthly probability	0.5%	0.7%	0.2%	0.2%

Abbreviations used in table: ER, extended release; TEAE, treatment emergent adverse events.

Therefore, the ERG carried out a sensitivity analysis on the manufacturer's primary base case which used AE data from TAURUS. The primary base case ICER decreased by £72 (from £4,386 to £4,314) suggesting that regarding the comparison of mirabegron and tolterodine, the use of SCORPIO data is conservative.

Discontinuation

As discussed in Section 5.2.4, within the manufacturer's base case models, discontinuation was disaggregated into discontinuation as a result of an AE and discontinuation as a result of other causes. Discontinuation as a result of an AE was assumed to occur either immediately (i.e., within the same cycle) or in subsequent cycles. The probability of discontinuation as a result of other causes was treatment specific and in the manufacturer's base case, was determined from the published literature. A retrospective longitudinal study of a UK prescriptions database by Wagg et al. together with a 12-week observational study of OAB patients conducted in Spain (Castro-Diaz et al.[88] were used to inform the probability of other cause discontinuation. Wagg et al. reported that "at 12 months, the proportions of patients still on their original treatment were: solifenacin 35%, tolterodine ER 28%,…., oxybutynin ER 26%, trospium 26%,…., oxybutynin IR 22%".[18] Whilst Castro-Diaz et al. reported that 24% of patients switched treatment as a result of side effects.[88] From this information, the manufacturer assumed that the 12-month probability of other cause discontinuation would be:

$$P_{OC\_12} = (1 - persistance_{12}) * (1 - \%discontinue\_AE)$$

Where: P_{OC_12} is the probability of discontinuation as a result of other causes; $persistance_{12}$ is the treatment specific persistence rate at 12 months (from Wagg et al.); %discontinue_AE is the percentage of patients who switch antimuscarinic therapy as a result of AEs (Castro-Diaz et al.). Table 54 summarises the annual and monthly probabilities of other cause discontinuation used in the manufacturer's base case models.

Table 54. Probability of other cause discontinuation used in the manufacturer's base case models

Treatment	Persistence rate at 12 months	Probability of other-cause discontinuation	
		Annual	Monthly
Mirabegron	N/A	N/A	N/A
Tolterodine ER	28.20%	54.6%	6.36%
Solifenacin	35.00%	49.4%	5.52%
Trospium chloride	25.90%	56.3%	6.67%
Fesoterodine	28.20%	54.6%	6.36%
Oxybutynin ER	21.70%	59.5%	7.26%
Oxybutynin IR	21.70%	59.5%	7.26%

Abbreviations used in table: ER, extended release; IR, immediate release; N/A, not applicable.

The manufacturer highlighted that no real world persistence data were available for mirabegron; therefore for the purposes of the model, the manufacturer assumed that the persistence of mirabegron was equal to that of the comparator (i.e., varied according to the comparison being made). In addition, the manufacturer carried out a sensitivity analysis which used the mean duration of treatment (5.2 months) observed in SCORPIO; this resulted in a monthly other cause discontinuation rate of 14.5% (MS; pg 199). However, the manufacturer states that as 28% of patients remained on treatment at 12 months the true mean could not be calculated (MS; pg 200).

The ERG considers it inappropriate to assume that discontinuation associated with mirabegron varies depending upon the comparison made; particularly, when trial data is available to inform the relative level of discontinuation associated with each treatment. Moreover, the ERG notes that the manufacturer reported that 28% of mirabegron patients were observed to persist with treatment after 12 months; although, it is unclear which data the manufacturer used to inform this statement. Therefore, the ERG carried out a sensitivity analysis using a 12-month persistence rate of 28% for mirabegron. This resulted in a £3 decrease in the primary base case ICER (from £4,386 to £4,383). The impact of this sensitivity analysis on the secondary base case results is presented in Section 6 and Appendix 11. In addition, the ERG consider it important to note that with the exception of oxybutynin IR 15 mg, no statistically significant differences in all cause discontinuation were identified in the ERG's MTC (see Section 4.4.2). Furthermore, mean estimates of the relative difference in discontinuation between mirabegron 50 mg and tolterodine ER 4 mg suggest that patients receiving treatment with tolterodine are less likely to discontinue treatment than patients treated with mirabegron 50 mg. Based on this, the ERG considers that the assumption of a 28% persistence rate for mirabegron 50 mg (i.e. a persistence rate equal to that of tolterodine) is likely to be favourable (i.e. any bias likely to be towards mirabegron).

Furthermore, the ERG notes that the manufacturer assumed that the probability of immediate discontinuation as a result of an AE was equal to the probability of other cause discontinuation multiplied by the probability of an AE. The ERG notes that the probability of other cause discontinuation was calculated to explicitly remove discontinuation as a result of an AE; therefore, the ERG considers it inappropriate to use this probability to inform the probability of immediate discontinuation as a result of an AE event.

5.2.7 Health-related quality of life

The manufacturer's models incorporated utility values derived from health related quality of life (HRQoL) data collected in SCORPIO. As discussed in Section 4.1.3, SCORPIO patients completed EuroQoL and OAB-q questionnaires, each of which were completed every 4 weeks. EuroQoL and

OAB-q questionnaire data were used to estimate EQ-5D (based on UK time trade-off data[94] and OAB-5D (using an algorithm developed by Yang et al.,[95] utility scores; applied in the manufacturer's base case and sensitivity analyses (see Section 5.2.10), respectively. In addition, the manufacturer carried out a systematic review of the literature which facilitated assessment of the consistency of utility scores derived from SCORPIO with published values.

Further to the derivation of utility scores from SCORPIO, the manufacturer developed utility scores based on data from the three key mirabegron trials (ARIES, CAPRICORN and SCORPIO). These data were used to assess the sensitivity of the EQ-5D and OAB-5D instruments according to patient response.

EQ-5D data derived from SCORPIO

As previously stated, the manufacturer's base case analyses were informed by EQ-5D utility scores estimated from SCORPIO. In particular, EQ-5D data from the three arms of SCORPIO were used to formulate a linear regression model; which, in turn was used to estimate utility according to a patient's disease severity profile (dependant on severity level of individual symptoms see Section 5.2.4 for full details).

As part of the clarification process, the ERG requested an explanation of the manufacturer's rationale for using linear regression models to estimate utility. The manufacturer clarified the rationale for the approach taken; stating that age and gender are known to be important factors in QoL. Therefore, regression analysis was used, so that these potentially confounding factors could be taken account of. In addition, the manufacturer stated that a regression modelling approach enabled the effect of each symptom (micturition and incontinence) to be considered separately (Manufacturer's clarification response; pg 29). The ERG agrees with the manufacturer that the use of regression analysis is appropriate and could minimise the risk of overestimating any utility benefit associated with mirabegron.

The following linear regression model was used to estimate a patient's utility based on symptom severity, age, gender and country (as a random effect).

$$Utility = \beta_0 + \beta_1 ClassMict + \beta_2 ClassInco + \beta_3 Age + \beta_4 Sex + \varepsilon_{patient} + \varepsilon_{country}$$

Where: β_0-β_4 are the regression coefficients; ClassMict is the level of severity with respect to frequency of micturition experienced by the patient; ClassInco is the level of severity with respect to number of incontinence episodes experienced by the patient; $\varepsilon_{patient}$ is a random error term included to account for repeated measures by patient and $\varepsilon_{country}$ is a random error term included to account for between country variation. The manufacturer states that "there was no significant treatment effect, independent of symptom severity" (MS; pg 207). As part of the clarification process, the ERG

requested further details on the selection of covariates used in the utility regression model. The manufacturer states that age, gender and country were the only covariates considered, based on prior knowledge of their potential confounding influence. In addition, the manufacturer states that the interaction between the numbers of micturition and incontinence episodes was tested (Wald test: p-value 0.0566) and found not to be significant (Manufacturer's clarification response; pg 30). Table 55 displays the coefficients obtained from the manufacturer's regression analysis of SCORPIO EQ-5D data. The utility values associated with each of the 25 disease severity profiles considered in the manufacturer's base case models are summarised in Table 56.

Table 55. The linear regression model for utility based on EQ-5D (reproduced from MS; Table 93; pg 207)

Effect	Level	Estimate
Intercept (β_0)	–	0.7838
Age (β_3)	–	–0.00041
Micturition severity level (β_1)	1	0.06321
	2	0.04224
	3	0.02042
	4	0.01039
	5	0
Incontinence severity level (β_2)	1	0.05859
	2	0.04367
	3	0.03141
	4	0.01282
	5	0
Gender (β_4)	F	–0.04412
	M	0

Table 56. EQ-5D utility values by disease severity profile (adapted from MS; Table 94; pg 207)

Incontinence frequency level	Micturitions frequency level				
	1	2	3	4	5
1	0.85	0.83	0.81	0.80	0.79
2	0.83	0.81	0.79	0.78	0.77
3	0.82	0.80	0.78	0.77	0.76
4	0.80	0.78	0.76	0.75	0.74
5	0.79	0.77	0.75	0.74	0.73

The ERG notes that covariate selection was neither systematic nor rigorous and that expert clinical advice was not sought in the formulation of the linear regression models. However, based on comparison with the published literature, the ERG considers the utility values estimated by the manufacturer's regression model to be reasonable.

In addition to disease severity in the estimation of patient utility, disutilities associated with AEs were included in the manufacturer's overall utility calculations. Patients, who experienced an AE (dry

mouth or constipation) and remained on treatment, accrued an adverse-event related disutility, for as long as they remained on treatment. However, patients who experienced an AE, yet immediately (i.e. within the same cycle) discontinued treatment were not assumed to accrue any AE related disutility.

AE-related disutility was derived from a repeated regression model based on the same data used to inform the linear regression models of utility; i.e. EQ-5D data collected in SCORPIO. Patients were categorised according to AE occurrence since their last visit (0/1). These data were then adjusted for age, gender, symptom severity (micturition, incontinence and urgency) and between country variations and were used to predict patient utility. In the MS, the manufacturer reported that "significant differences in utility between patients who reported dry mouth or constipation AE, and those who did not report such AE was found" (MS; pg 218). Moreover, based on the repeated regression analysis, the manufacturer estimated the utility decrement associated with an AE (dry mouth or constipation) to be -0.0357.

Within the MS, the manufacturer highlighted that "it was felt appropriate to calculate AE utilities using this repeated regression model, however it should be noted that utility decrements for AEs derived from the regression model used to calculate health state utilities elicited a near identical figure of -0.03558" (MS; pg 218). In addition, as part of the clarification process, the ERG requested further details on the rationale for using a repeated regression model and justification of the covariates selected. In the clarification response, the manufacturer provided the requested details (Boxes 11 and 12).

Box 11. Manufacturer's rationale for selecting a repeated regression model to estimate AE-related disutilities (manufacturer's clarification response; pg 32)

> A repeated observations model was used to account for the dependence between different utility assessments within individuals. Assuming that utilities at different visits are independent would have led to underestimate the variability around estimates of disutilities. The Akaike Information Criteria confirmed that the model accounting for repeated measures was better.

Box 12. Manufacturer's justification of covariates selected for the repeated regression model (manufacturer's clarification response; pg 32)

> We entered gender, age and geographical regions for the same reason as mentioned in B3. Age in particular was thought to be a potential important confounding factor as EQ-5D utilities are strongly related to age, and adverse events are more frequent amongst the elderly. The proportion of patients experiencing AEs was estimated at 16.00% amongst patients aged 65 years or over, and 13.82% amongst those under 65 years, in the pooled analysis of SCORPIO, ARIES and CAPRICORN (independence chi-square test, p=0.0423).
>
> In addition, we expected symptoms could also be a confounding factor. In the pooled analysis of SCORPIO, ARIES and CAPRICORN, patients experiencing an AE had significantly more episodes of incontinence, compared with patients without an AE (1.24 incontinence episodes per day vs 0.94, respectively, p=0.0004). Thus failing to adjust for incontinence severity would have led to an overestimate of the disutility associated with adverse events.
>
> We entered urgency, as well as micturitions and incontinence, but this had little impact on results: the disutility associated with AEs was estimated at 0.357 with urgency in the model and 0.356 without urgency in the model.

The ERG considers the manufacturer's rationale for using a repeated regression model to be acceptable, particularly in light of the sensitivity analysis (using the same linear regression model used to derive disease severity utilities) carried out by the manufacturer. Furthermore, the ERG considers the selection of covariates to be evidence-based and reasonable.

Manufacturer's systematic review

In addition to utility scores derived from clinical trial data, the manufacturer carried out a systematic literature review. The review was carried out in August 2012 and aimed to identify health state utility value (HSUV) studies relevant to patients with OAB. The following electronic databases were searched; Medline and Medline in-process (R), Embase, EconLit, NHS EED and the Cochrane library. Supplementary hand-searches of bibliographies (included studies and systematic reviews that were no more than 3 years old), related NICE technology appraisal, the cost-effectiveness analyses (CEA) registry, and the websites of EQ-5D and Research Papers in Economics (RePEc) were also carried out. The ERG notes that the manufacturer's searches were comprehensive, with reasonable and explicitly stated inclusion and exclusion criteria applied. Therefore, the ERG considers it unlikely that any relevant studies have been missed.

The review identified 10 HSUV studies deemed relevant to the scope of this STA.[9;96-104] Of these, seven studies used EQ-5D questionnaires to elicit utilities,[10;96;98-102] one study converted SF-12 to EQ-5D[97] and one study did not clearly report elicitation or valuation methods used.[103] The remaining study was a mapping study, which carried out regression analysis to "investigate the association between patient characteristics and disease-specific and generic quality of life (QOL) as

well as the degree of bother in women seeking treatment for urinary incontinence".[104] Of the studies using EQ-5D elicitation, one used US weights[101] to value EQ-5D scores, two reported the use of UK weights,[9, 96] three were assumed to use UK weights,[98-100] and the source of weights used in the remaining study could not be inferred or determined.[102] All studies were published between 2003 and 2012. Two of the studies were carried out in the UK,[98;99] two in the USA,[100;101] one in Sweden[102] and one in Turkey;[103] the remaining four studies were multinational.[9;96;97;104] Table 57 summarises the included studies.

Table 57. Summary of the studies included from the manufacturer's QoL systematic review (adapted from MS; Table 97; pg 211)

Author & year	Country	Elicitation	Valuation
Coyne 2008[9]	Multinational	EQ-5D	Health states valued using UK weights
Tincello 2010[96]	Multinational	EQ-5D	Health states valued using UK weights
Verheggen 2012[97]	Multinational	SF-12 converted to EQ-5D using published algorithms.	Health states valued using UK weights
Monz 2007[104]	Multinational	Regression analysis mapping study	
Haywood 2008[98]	UK	EQ-5D	Source of weights for health states not reported (assumed to be UK weights)
Currie 2006[99]	UK	EQ-5D	Source of weights for health states not reported (assumed to be UK weights)
Harvie 2010[100]	USA	EQ-5D	Source of weights for health states not reported (assumed to be UK weights)
Patterson 2011[101]	USA	EQ-5D	Health states valued using US weights
Kobelt 2003[102]	Sweden	EQ-5D.	Not stated
Sut 2012[103]	Turkey	Not stated	Not stated

The manufacturer states that the utility values estimated from SCORPIO were generally comparable with those reported in the published literature; although, the trial based utility values provided a higher level of granularity. In addition, the manufacturer highlights that patient populations assessed in the published literature were often mixed, limiting the comparability with clinical trial data. The ERG agrees with the manufacturer's assertion of comparability and considers the use of trial based rather than published utility values in the model to be appropriate.

Utility scores based on data from ARIES, CAPRICORN and SCORPIO

Based on pooled HRQoL data from all treatment arms of ARIES, CAPRICORN and SCORPIO, the manufacturer used linear regression models to estimate mean EQ-5D and OAB-5D utility scores, for symptoms of micturition, incontinence and urgency, by level of severity (Table 58). No rationale for inclusion of urgency within these regression analyses was provided in the MS. For each symptom, the regression models used were adjusted for age, gender, geographical region and random patient effects (i.e. to account for utility scores from one individual at different assessment visits). In addition, the manufacturer states that the Pearson correlation between EQ-5D and OAB-5D utilities was estimated;

however, no details of the results of this assessment were presented in the MS and so the ERG is unable to comment on the importance of this. Furthermore, the manufacturer reported that "Using a similar method, differences between OAB-5D and EQ-5D utilities were estimated by symptom level, and tested for the null hypothesis of equal mean OAB-5D and EQ-5D utilities" (MS; pg 208). Similarly, results of this analysis were not presented; therefore, the ERG is unable to comment on the relevance of this assessment to the manufacturer's base case or sensitivity analyses results.

Table 58. Mean utility scores estimated from regression analyses of pooled ARIES, CAPRICORN and SCORPIO HRQoL data (reproduced from MS; Table 96; pg 208)

Clinical symptom	Symptom levels	Level definition	% patients	Utility mean (±SD)	Utility mean (±SD)
				EQ-5D	OAB-5D
Micturitions	1	< 8	21.2	0.85 (±0.21)	0.90 (±0.08)
	2	8 – <=10	30.7	0.84 (±0.20)	0.87 (±0.09)
	3	10 – <=12	22.7	0.82 (±0.21)	0.85 (±0.09)
	4	12 – <=14	13.2	0.80 (±0.22)	0.82 (±0.09)
	5	>14	12.3	0.78 (±0.23)	0.80 (±0.09)
Incontinence episodes	1	0	50.3	0.85 (±0.19)	0.89 (±0.08)
	2	>0 – <=1	19.7	0.82 (±0.20)	0.85 (±0.09)
	3	1 – <=2	11.0	0.80 (±0.22)	0.83 (±0.09)
	4	2 – <=3	6.9	0.78 (±0.23)	0.81 (±0.09)
	5	>3	12.2	0.76 (±0.26)	0.79 (±0.09)
Urgency Grade 3 episodes	1	<1	23.7	0.86 (±0.19)	0.90 (±0.08)
	2	1 – <=3	30.3	0.83 (±0.21)	0.87 (±0.09)
	3	3 – <=5	21.6	0.81 (±0.22)	0.84 (±0.09)
	4	5 – <=7	11.7	0.81 (±0.22)	0.82 (±0.09)
	5	>7	12.7	0.78 (±0.24)	0.80 (±0.09)
Abbreviations used in table: OAB, overactive bladder; SD, standard deviation.					

For each level of severity of micturition and incontinence, the ERG compared the mean EQ-5D scores reported in Table 58 with average EQ-5D scores calculated from Table 56. The results of which are presented in Table 59. The ERG notes that the mean utilities estimated from the pooled analysis are generally higher than those estimated from SCORPIO data alone. Therefore, the ERG considers the use of utility data from SCORPIO in the model to be likely to bias against the more effective treatment.

Table 59. Results of ERG comparison of average EQ-5D scores estimated from SCORPIO data with mean EQ-5D scores estimated from pooled ARIES, CAPRICORN and SCORPIO data

Symptom	Level of severity	Average EQ-5D estimated from SCORPIO[a]	Mean EQ-5D estimated from pooled analysis	Difference (SCORPIO – pooled data)
Micturition	1	0.82	0.85	–0.03
	2	0.80	0.84	–0.04
	3	0.78	0.82	–0.04
	4	0.77	0.80	–0.03
	5	0.76	0.78	–0.02
Incontinence	1	0.82	0.85	–0.03
	2	0.80	0.82	–0.02
	3	0.79	0.80	–0.01
	4	0.77	0.78	–0.01
	5	0.76	0.76	0.00
[a] Calculated from Table 56.				

In addition to the analyses described above, the manufacturer states that "linear models, predicting mean OAB-5D and EQ-5D utilities according to severity levels of the three symptoms, with adjustment on gender and age, were created" (MS; pg 208). From which the ERG infers that two linear models simultaneously accounting for the severity of each symptom (micturition, incontinence and urgency) were developed to estimate EQ-5D and OAB-5D scores, respectively. The manufacturer states that the purpose of these models was to provide "a way to derive utilities from the micturition diary data, which were collected in the clinical studies" (MS; pg 208). However, neither further details nor results of the linear models used were presented in the MS; therefore, the ERG is unable to comment on the relevance or otherwise of these analyses.

Finally, the manufacturer briefly described an assessment of the sensitivity of the instruments (EQ-5D and OAB-5D) using "a linear model, with the co-variables of gender, age and response as fixed effects, and geographical region as random effect……..to provide adjusted means (SD) of utility changes from baseline to week 12 by response level" (MS; pg 209). As part of the clarification process, the ERG requested the results of this assessment, which the manufacturer provided (Table 60).

Table 60. Responsiveness of EQ-5D and OAB-5D to changes in clinical symptoms (reproduced from manufacturer's clarification response; Table 20; pg 31)

Clinical symptoms (per 24 hours)	Response in symptoms	N (%)	CFB in utility EQ-5D, adjusted mean (±SD)	P value vs 'Stable'†	CFB in utility OAB-5D, adjusted mean (±SD)	P value vs 'Stable'†
Micturitions	Improvement ≥1 level	2,286 (56.4%)	0.046 (±0.213)	0.0058	0.076 (±0.092)	<0.0001
	Stable	1,260 (31.1%)	0.027 (±0.202)	N/A	0.039 (±0.087)	N/A
	Worsening ≥1 level	506 (12.5%)	0.023 (±0.195)	0.7291	0.021 (±0.085)	0.0001
Incontinence episodes	Improvement ≥1 level	1,874 (46.3%)	0.053 (±0.223)	0.0002	0.081 (±0.096)	<0.0001
	Stable	1849 (45.6%)	0.028 (±0.206)	N/A	0.042 (±0.089)	N/A
	Worsening ≥1 level	329 (8.1%)	0.007 (±0.197)	0.0662	0.022 (±0.086)	<0.0001
Abbreviations used in table: CFB, change from baseline; N/A, not applicable; OAB, overactive bladder, SD, standard deviation.						

Within the clarification response, the manufacturer states that "Changes in OAB-5D utilities are significantly different between worsening and stable patients, but changes in EQ-5D utilities are not significantly different" (Manufacturer's clarification response; pg 31). The ERG agrees with the manufacturer's assertion and based on this analysis considers the use of EQ-5D utility data in the model to be conservative (i.e. any bias is likely to be against the most effective treatment).

5.2.8 Resources and costs

As discussed in Section 5.1, the manufacturer carried out a systematic literature review to identify economic evaluations and costing studies relevant to the use of mirabegron in a patient population with OAB. Seven costing studies were identified in the review.[105-111] Of these, one study was carried out in Sweden[105] and the remaining studies were carried out in the USA. Table 61 summarises the identified costing studies.

Table 61. Costing studies identified in the manufacturer systematic literature review (reproduced from MS; Table 99; pg 219)

Reference	Country	Study objective
Altman 2009[105]	Sweden	National analysis of utilisation and costs associated with the pharmacological treatment for OAB
Jumadilova 2006[106]	USA	Costs related to comorbidities associated with OAB
Nitz 2005[107]	USA	To compare post-treatment medical costs for OAB patients when treatment is one of the following: oxybutynin IR, tolterodine ER and oxybutynin ER
Noe 2002[108]	USA	To compare the estimated first-line treatment costs of tolterodine ER vs oxybutynin CR in patients with OAB

Perfetto 2005[109]	USA	To compare 1 year healthcare costs for OAB patients treated with oxybutynin ER vs tolterodine ER in a cost minimisation model
Varadharajan 2005[110]	USA	To examine the economic impact of oxybutynin IR, tolterodine ER and oxybutynin ER among commercially insured
Zinner 2008[111]	USA	Resource use and work productivity for patients switching from tolterodine ER to solifenacin
Abbreviations used in table: CR, controlled-release; ER, extended-release; IR, immediate-release; OAB, overactive bladder.		

As highlighted by the manufacturer, none of the identified studies were carried out in the UK; therefore, resource use data relevant to the UK setting were obtained from the following sources:

- British National Formulary (BNF) Volume 63;[112]
- Unit Costs of Health and Social Care 2011;[113]
- AgeUK incontinence website;[114]
- Nottingham Urology Group;[115]

Within the manufacturer's economic evaluation, three types of cost were accounted for, these were the cost of: interventions; healthcare professionals (HCP) and incontinence pad utilisation. The manufacturer assumed that AEs would not be associated with any additional cost; over and above the cost of specialist referrals for treatment switch. Given the small proportion of patients experiencing AEs and the low cost of any associated treatments for AEs, the ERG considers the exclusion of specifically AE related costs to be reasonable.

Intervention costs

The monthly acquisition cost of each considered intervention is summarised in Table 62. The manufacturer used the list price reported in BNF 63 for all antimuscarinic interventions (except mirabegron, which was disclosed by the manufacturer); the ERG notes that, as highlighted by the manufacturer, the list price for all considered interventions was the same in BNF 63 and 64.[112;116] Antimuscarinic costs were calculated assuming that patients used one tablet a day per month. The cost of BoTox was sourced from private care costs reported by the Nottingham Urology Group.[115] In addition, based on expert clinical opinion, patients that were successfully treated with BoTox were assumed to receive 2 reinjections every year.

Table 62; Intervention costs used in the manufacturer's model (adapted from MS; Tables 100 and 103; pgs 220 and 221)

OAB medication	Cost per pack (£)	No. of tablets per pack	Cost per day (£)	Cost per month[a] (£)	Source
Antimuscarinics					
Mirabegron 50 mg	29.00	30	0.97	29.40	Astellas
Tolterodine ER 4 mg	25.78	28	0.92	28.01	BNF63[(112)]
Solifenacin 5mg	27.62	30	0.92	28.00	
Solifenacin 10 mg	35.91	30	1.20	36.41	
Trospium chloride MR 60 mg	23.05	28	0.82	25.04	
Fesoterodine 4 or 8 mg	25.78	28	0.92	28.01	
Oxybutynin ER 10 mg	27.54	30	0.92	27.92	
Oxybutynin IR 10 mg (cost of 5mg)	11.60	84	0.14	8.40	
Injection	**Cost of injection (£)**	**Visits per year[b]**	**Visits per month**	**Cost per month (£)**	**Source**
Botulinum toxin injection					
Initial	1,158	–		1,158	Nottingham Urology Group[(115)]
Reinjections	964	2	0.17	163.88	

[a] Considering (365/12) days per month.

[b] Based on clinical expert opinion.

Abbreviations used in table: BNF, British National Formulary; ER, extended-release; IR, immediate-release; mg, milligram; MR, modified-release.

The ERG considers that appropriate methodology was used to calculate monthly antimuscarinic drug and BoTox injection costs. However, the ERG notes that the manufacturer used private care costs to inform the cost of BoTox injections. The ERG identified NHS reference costs for inpatient and outpatient BoTox injections of £321 and £212, respectively (HRG code XD09Z; Torsion Dystonias and other involuntary movements drugs band 1) (ref). Therefore, the ERG carried out a sensitivity analysis where initial BoTox injections were assumed to cost £321 and reinjections were assumed to cost £212. Following implementation of these costs in the manufacturer's model, the primary base case ICER increased by £810 (from £4,386 to £5,196). The impact of altering the cost of BoTox on the manufacturer's incremental results is presented in Section 6 and Appendix 11.

Healthcare professional costs

Regarding patient care by healthcare professionals, within the economic models, the manufacturer assumed that patients would require:

- one GP consultation at treatment initiation and treatment switch;
- one and a half specialist consultations at treatment initiation and treatment switch.

These assumptions were based on expert clinical opinion reported in a study by Cardozo et al.[(70)] Cardozo et al. assessed the cost-effectiveness of solifenacin versus other antimuscarinic strategies

commonly used in UK clinical practice. The cost associated with GP (£36) and specialist consultations (£96) were sourced from the Unit Costs of Health and Social Care 2011[113] and NHS payment by results (PbR) tariff, 2010-2011, respectively. As highlighted by the manufacturer, the cost of an outpatient specialist urology visit reported in NHS reference costs 2010-2011 (£91) is marginally lower than the cost reported in NHS PbR tariff 2010-2011 (MS; pg 219). However, contrary to the manufacturer's assertion that the "use of the PbR tariff provides a conservative estimation of costs for comparators" (MS; pg 219), the ERG notes that the use of a lower cost for specialist outpatient visits favours mirabegron; i.e. the manufacturer's primary base case ICER increased by £107 (from £4,386 to £4,493) when the NHS reference cost was used. The impact of using the NHS reference cost on the manufacturer's incremental results is presented in Section 6 and Appendix 11.

Incontinence pad costs

The number of incontinence pads used per month varied with respect to the severity of incontinence a patient experienced. The mean number of pads used by level of incontinence severity was calculated using data from all three arms of SCORPIO (Table 63). These values were then assumed to apply to all patients regardless of treatment received (including patients not receiving any treatment). The unit cost of an incontinence pad was assumed to be £0.16, with AgeUK incontinence cited as the reference source. No further details of assumptions used to calculate this cost (e.g. products included or pad size assumed) were provided in the MS; therefore, the ERG were unable to verify this cost.

Table 63. Incontinence pad use and associated monthly cost used in the manufacturer's model (adapted from MS; Table 102; pg 221)

Incontinence severity level	Pad use per day	Pad use per month[a]	Monthly cost[b] (£)
1	0.17	5.29	0.85
2	0.75	22.87	3.66
3	1.38	42.10	6.74
4	1.89	57.55	9.21
5	3.34	101.56	16.25

[a] Pad use per day multiplied by 365/12.

[b] Pad use per month multiplied by cost per pad £0.16.

To summarise, the costs associated with each treatment arm were primarily comprised of medication and incontinence pad costs. Treatments that reduced the severity of incontinence accrued less costs as patients required fewer incontinence pads. Treatments that were associated with higher levels of discontinuation or AEs implicitly accrued higher costs as a result of treatment switching.

5.2.9 Cost effectiveness results

The manufacturer presented two sets of base case results, the primary base case (mirabegron 50 mg versus tolterodine ER 4 mg) based on efficacy data from SCORPIO, and the secondary base case (mirabegron 50 mg versus all comparators of interest) based on efficacy data from the manufacturer's MTC. The secondary base case results were presented both individually and incrementally. Table 64 displays the primary base case results. Tables 65 and 66 display the individual and incremental secondary base case results, respectively. The manufacturer presented both deterministic (using mean parameter values only) and probabilistic (assessing the simultaneous effect of parameter uncertainty) primary base case results with ICERs of £4,386 and £4,886, respectively. Based on these, the manufacturer asserted that mirabegron 50 mg is cost-effective when compared to tolterodine ER 4 mg.

Table 64. Primary base case results based on efficacy data from SCORPIO (adapted from MS; Tables 115 and 118, pgs 234 and 238)

Treatment	Total		Incremental		ICER (£/QALY) versus tolterodine
	Costs (£)	QALYs	Costs (£)	QALYs	
Deterministic results					
Tolterodine ER 4 mg	1,607.75	3.755	–	–	–
Mirabegron 50 mg	1,645.62	3.764	37.88	0.009	4,385.65
Probabilistic results (only incremental values are presented)					
Mirabegron vs tolterodine			49.86	0.010	4,886.00
Abbreviations used in table: ICER, incremental cost-effectiveness ratio; mg, milligram; QALY, quality adjusted life year.					

Table 65. Individual secondary base case results, mirabegron versus antimuscarinics, based on MTC results (reproduced from MS; Table 116; pg 234)

Treatment	Total			Incremental			ICER (£/QALY) versus mirabegron
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Solifenacin 10 mg	1,647.60	4.666	3.762	3.53	0	0.010	340.15
Fesoterodine 4 mg	1,601.40	4.666	3.758	38.09	0	0.011	3,606.71
Tolterodine ER 4 mg	1,601.64	4.666	3.759	37.85	0	0.010	3,714.98
Oxybutynin ER 10mg	1,587.06	4.666	3.755	42.12	0	0.011	3,877.57
Trospium chloride MR 60 mg	1,551.86	4.666	3.759	83.89	0	0.010	8,881.48
Solifenacin 5 mg	1,592.94	4.666	3.768	58.19	0	0.005	12,493.21
Oxybutynin IR 10 mg	1,421.00	4.666	3.752	208.18	0	0.015	14,233.83
Abbreviations used in table: ER, extended-release; ICER, incremental cost-effectiveness ratio; IR, immediate-release; LYG, life year gained; mg, milligram; MR, modified-release; QALY, quality adjusted life year.							

Table 66. Incremental secondary base case results, assuming mirabegron persistence is equal to solifenacin persistence (reproduced from manufacturer's model)

Intervention	Total		Incremental (versus previous treatment)		ICER (£/QALY) versus oxybutynin IR 10 mg	Incremental ICER (£/QALY)
	Cost (£)	QALYs	Cost (£)	QALYs		
Oxybutynin IR 10 mg	1,421.00	3.752	-	-	-	-
Trospium chloride MR 60 mg	1,551.86	3.759	130.86	0.007	18,816.13	18,816.13[a]
Oxybutynin 10mg ER	1,587.06	3.755	35.20	-0.003	44,127.74	Strictly dominated[b]
Solifenacin 5 mg	1,592.94	3.768	5.89	0.012	10,812.95	4,591.75[c]
Fesoterodine 4 mg	1,601.40	3.758	8.46	-0.009	26,336.26	Strictly dominated[d]
Tolterodine ER 4 mg	1,601.64	3.759	0.23	0.000370	25,017.43	Extendedly dominated[e]
Solifenacin 10 mg	1,647.60	3.762	45.97	0.003	22,261.85	Extendedly dominated[f]
Mirabegron 50 mg	1,651.14	3.772	3.53	0.010	11,193.61	12,493.21[g]

[a] versus oxybutynin IR 10 mg.

[b] by trospium chloride 60 mg MR.

[c] versus trospium chloride 60 mg MR.

[d] by solifenacin 5 mg.

[e] by mirabegron 50 mg.

[f] by mirabegron 50 mg.

[g] versus solifenacin 5 mg.

Abbreviations used in table: ER, extended-release; ICER, incremental cost-effectiveness ratio; IR, immediate-release; MR, modified-release; QALY, quality adjusted life year.

As discussed in Section 5.2.6, the manufacturer highlighted that no real world persistence data were available for mirabegron. Therefore, the manufacturer assumed that mirabegron would be associated with the same persistence rate as that of the comparator (i.e. the persistence rate associated with mirabegron varied depending on the comparison made). In order to compute the incremental results presented in the MS, the manufacturer assumed that the persistence rate of mirabegron was equal to that of solifenacin. The ERG notes that solifenacin is the treatment associated with the highest persistence rates. The impact of the persistence rate associated with mirabegron on the incremental results is discussed further in Section 6 and Appendix 11.

In addition to the cost-effectiveness results, the manufacturer presented a comparison of disease severity as estimated by the primary base case model with that observed in SCORPIO (Table 67).

Table 67. Primary base case model outcomes compared with the clinical results of SCORPIO (adapted from MS; Table 112, pg 226)

Outcome		Clinical trial result	Model result	error in prediction	Magnitude of error (error/clinical trial result)
Treatment	**Severity level**				
Micturition					
Mirabegron	1	33.40%	31.70%	–1.70%	5.09%
	2	31.40%	30.20%	–1.20%	3.82%
	3	18.80%	19.90%	1.10%	5.85%
	4	9.20%	9.10%	–0.10%	1.09%
	5	7.30%	9.10%	1.80%	24.66%
Tolterodine	1	32.40%	29.60%	–2.80%	8.64%
	2	29.70%	29.40%	–0.30%	1.01%
	3	18.50%	19.30%	0.80%	4.32%
	4	9.40%	10.80%	1.40%	14.89%
	5	10.10%	11.00%	0.90%	8.91%
Incontinence					
Mirabegron	1	62.70%	61.90%	–0.80%	1.28%
	2	19.00%	19.30%	0.30%	1.58%
	3	7.60%	6.90%	–0.70%	9.21%
	4	4.40%	4.70%	0.30%	6.82%
	5	6.40%	7.20%	0.80%	12.50%
Tolterodine	1	63.70%	61.40%	–2.30%	3.61%
	2	17.10%	17.60%	0.50%	2.92%
	3	5.90%	6.50%	0.60%	10.17%
	4	4.60%	4.90%	0.30%	6.52%
	5	8.70%	9.60%	0.90%	10.34%

The proportions of patients by severity level, for micturition and incontinence, predicted by the manufacturer's primary base case model are largely consistent with the distribution of patients (by severity level) observed in SCORPIO; i.e. the majority of errors were within 10% of the clinical trial result. However, the ERG notes that the manufacturer's primary base case model appears to overestimate the benefit of mirabegron 50 mg compared with tolterodine ER 4 mg. In particular, the proportions of patients predicted to be in lower levels (1 or 2) of severity (micturition and incontinence), whilst smaller than the proportions observed in SCORPIO for both treatments, are underestimated to a greater degree in the tolterodine model arm. Similarly, the proportions of patients predicted to be in higher levels (4 or 5) of severity (micturition and incontinence) are overestimated to a greater degree in the tolterodine arm.

Further to the comparison of model results with clinical trial data, the manufacturer presented details of disaggregated QALYs and costs; these are displayed in Tables 68 and 69, respectively. The manufacturer also provided a graphical depiction of mean utility by treatment over time, along with useful plots (Figure 4) depicting the proportion of patients over time who:

- continued to receive originally allocated treatment;
- were receiving BoTox;
- experienced AEs (dry mouth or constipation).

Table 68. Summary of QALYs gained for each treatment by health state (adapted from MS; Table 113, pg 233)

Health state	QALY mirabegron	QALY tolterodine	Increment	Absolute increment	% absolute increment
Baseline	3.4005	3.4010	–0.0005	0.0005	3.82%
Micturition severity level 1	0.0735	0.0679	0.0055	0.0055	46.91%
Micturition severity level 2	0.0591	0.0584	0.0007	0.0007	5.88%
Micturition severity level 3	0.0198	0.0200	–0.0002	0.0002	1.95%
Micturition severity level 4	0.0060	0.0065	–0.0004	0.0004	3.73%
Micturition severity level 5	0.0000	0.0000	0.0000	0.0000	0.05%
Incontinence severity level 1	0.1527	0.1497	0.0030	0.0030	25.36%
Incontinence severity level 2	0.0350	0.0341	0.0010	0.0010	8.31%
Incontinence severity level 3	0.0137	0.0140	–0.0003	0.0003	2.72%
Incontinence severity level 4	0.0035	0.0036	–0.0001	0.0001	1.27%
Incontinence severity level 5	0.0000	0.0000	0.0000	0.0000	0.00%
Total	3.7638	3.7552	0.0086	0.0118	100.00%
Abbreviation used in table: QALY, quality adjusted life year.					

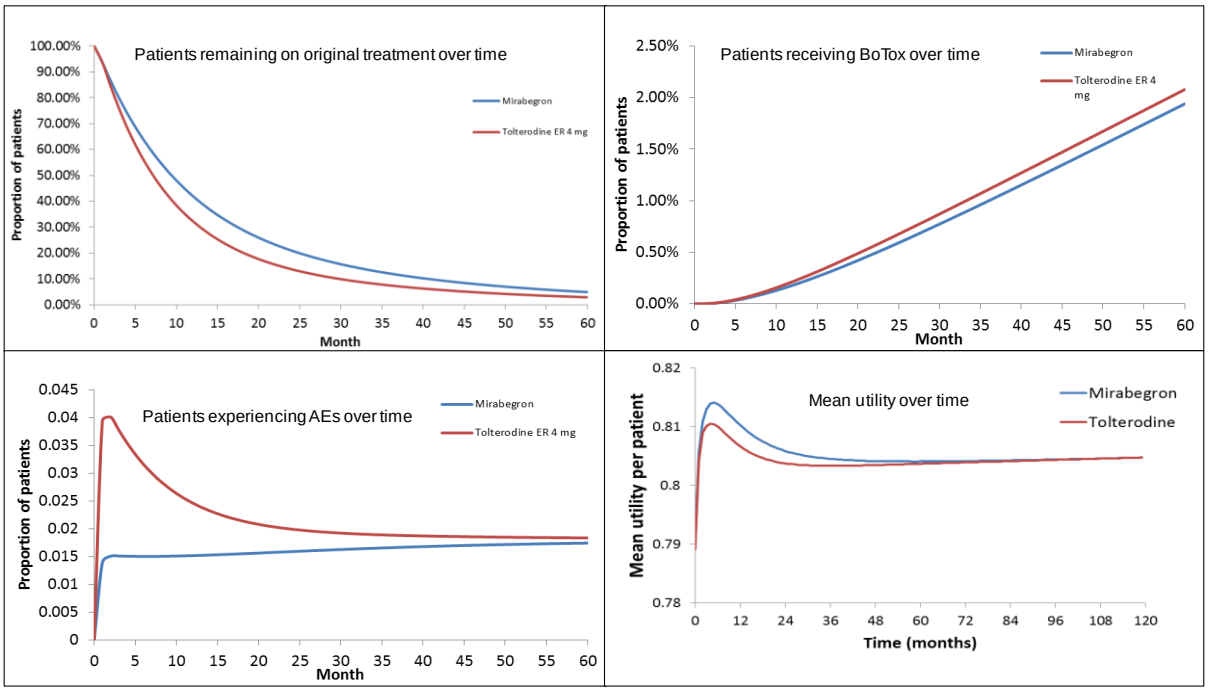
Table 69. Summary of predicted resource use by category of cost (reproduced from MS; Table 114, pg 233)

Item	Cost mirabegron (£)	Cost tolterodine (£)	Increment	Absolute increment	% absolute increment
Drug cost	451.43	343.70	107.72	107.72	46.6%
Other OAB medication	364.92	393.42	–28.50	28.50	12.3%
Primary care visit	101.38	105.83	–4.45	4.45	1.9%
Specialist (urology) follow-up visit	405.53	423.31	–17.78	17.78	7.7%
Initial botulinum toxin injection	25.50	27.42	–1.92	1.92	0.8%
Repeat botulinum toxin injection	68.16	75.36	–7.19	7.19	3.1%
Incontinence pads	228.70	238.71	–10.00	10.00	4.3%
Total	1,645.62	1,607.75	37.88	231.10	100%
Abbreviation used in table: OAB, overactive bladder.					

The ERG notes that mirabegron accrued less costs as a result of other OAB medication and specialist follow-up visits (i.e. following treatment switch). In addition, the ERG notes that over half of the QALY gain associated with mirabegron compared with tolterodine was accrued by patients in the lower micturition and incontinence severity levels (1 and 2). Moreover, the benefit associated with mirabegron versus tolterodine with respect to the severity of incontinence and micturition has been observed to be overestimated (Table 67). Therefore, based on the fact that treatment efficacy with

respect to the outcomes of micturition and incontinence drive the QALY gain associated with mirabegron, the ERG considers the manufacturer's model to be biased in favour of mirabegron.

Figure 4. Treatment, adverse events and utility accrual over time (adapted from MS; Figures 39–41 and Figure 46; pgs 227–229 and 232)



5.2.10 Sensitivity analyses

The manufacturer carried out several analyses assessing the sensitivity of the base case models to uncertainty associated with model parameters and structural assumptions. Within the MS, the results of sensitivity analyses were only presented for the primary base case (mirabegron 50 mg versus tolterodine ER 4 mg) model; however, the results of all sensitivity analyses were available in the manufacturer's submitted models. This section focuses on sensitivity analyses carried out on the manufacturer's primary base case; however, the results of deterministic sensitivity analyses on the manufacturer's secondary base case model are presented in Appendix 10. To summarise, the impact of the manufacturer's one-way sensitivity analyses on the secondary base case results was similar to the impact of one-way sensitivity analyses on the primary base case cost-effectiveness results. As a result of time constraints it was not feasible to consider the probabilistic results of each of the comparisons made in the manufacturer's secondary base case.

Parameter uncertainty

The impact of uncertainty around parameter estimates was assessed using one way (individual assessment of the impact of variation in each parameter) and probabilistic (simultaneous assessment of impact of parameter variation) sensitivity analyses.

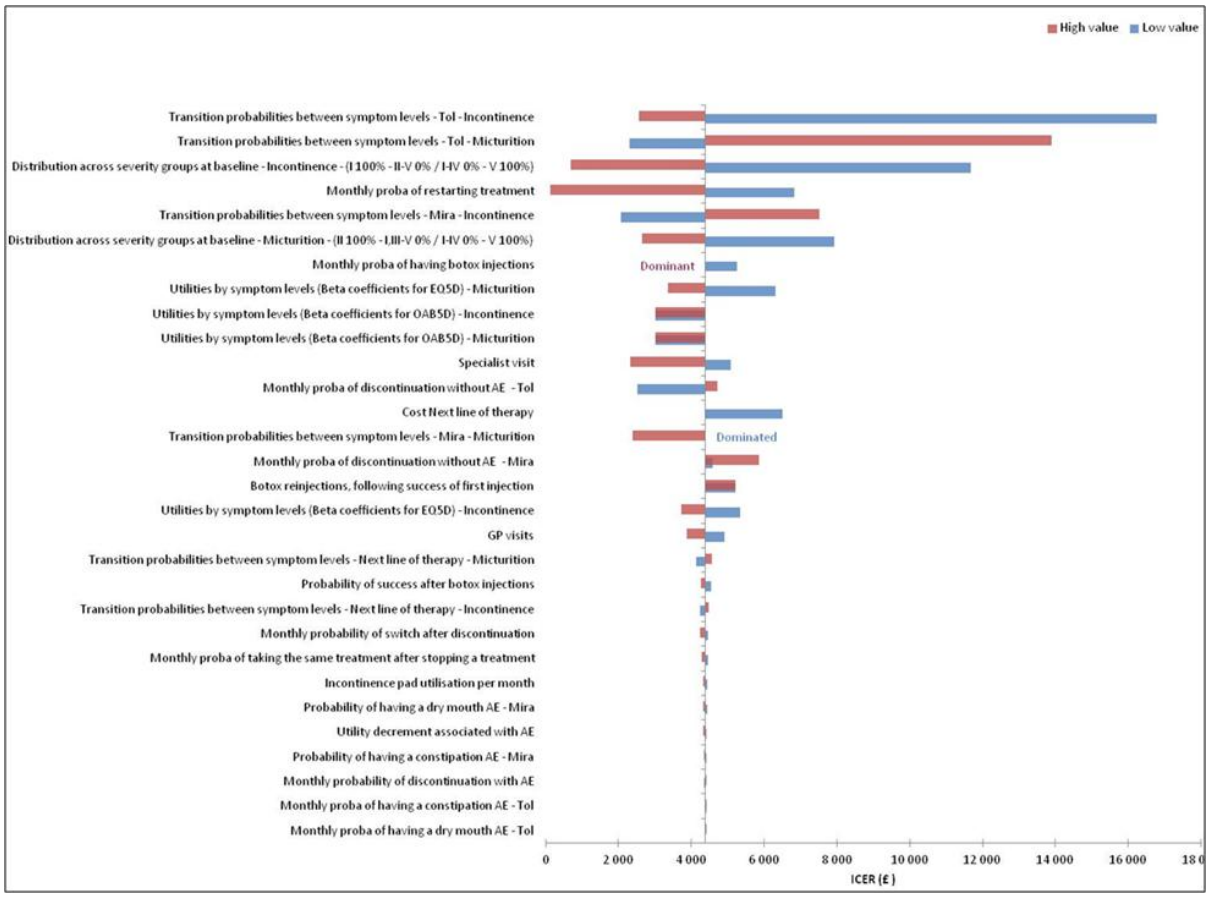
One-way sensitivity analyses

The manufacturer carried out one-way sensitivity analyses on all model parameters considered to be associated with uncertainty, namely the:

- distribution of patients across symptom severity levels at baseline;
- probabilities of transition between symptom levels (β-coefficients of regression equation);
- utility values by symptom levels (β-coefficients of regression equation);
- probabilities of treatment-related events (discontinuation, switch),
- probabilities related to BoTox injections, (probability of having the injection in the first month and probability of success after the injections);
- probabilities of AEs;
- utility decrement associated with AEs;
- resource use parameters (GP visits, specialist visits, medication costs, incontinence pad use).

The manufacturer used deterministic results as the base for one way sensitivity analysis. However, given the high level of consistency observed (Table 64), between the deterministic and probabilistic ICERs, the ERG considers that the use of deterministic rather than probabilistic results was appropriate. One way sensitivity analyses were carried out by consecutively recalculating the ICER (mirabegron 50 mg versus tolterodine ER 4 mg) using lower and higher values for each parameter considered. Where possible, the upper and lower bounds of a parameter's 95% confidence interval were used; however, where confidence intervals were unobtainable, extreme low and high values were assumed (see Table 48 for full details of values used). Figure 5 displays the tornado diagram generated by the manufacturer to display the impact of individual variations in the considered parameters on the primary base case cost-effectiveness results.

Figure 5. Results of manufacturer's one-way sensitivity analysis on primary (mirabegron 50 mg versus tolterodine ER 4 mg) base case cost-effectiveness results (reproduced from MS; Figure 48; pg 237)



The ERG considers that the manufacturer's one-way sensitivity analyses were thorough. In addition, the ERG notes that these analyses indicate that the manufacturer's primary base case result is relatively robust; i.e. insensitive to individual variation in parameter estimates. The parameter's that had the highest impact on the ICER were transition probabilities between symptom levels and the baseline distribution of patients across symptom severity levels; suggesting that efficacy was a key driver of the manufacturer's models. However, the ERG notes that most of the analyses returned ICER estimates of less than £17,000 per QALY gained.

Probabilistic sensitivity analysis

The simultaneous impact of parameter uncertainty on the manufacturer's cost-effectiveness results was assessed through probabilistic sensitivity analysis (PSA). Statistical distributions were assigned to each parameter included in the PSA, from which estimates were simultaneously sampled (1,000 times) to inform the manufacturer's model. Table 70 summarises the parameters included and distributions used.

Table 70. Distributions used to inform the manufacturer's probabilistic sensitivity analysis

Parameter	Distribution
Baseline proportions of patients by level of symptom severity	Dirichlet
Regression coefficients (β-coefficients) for: • multinomial logistic regression (informing transition probabilities); • linear regression model (informing utility values).	Normal
Probabilities: • dry mouth; • BoTox injection) • discontinuation as a result of an AE; • other cause discontinuation; • treatment switch after discontinuation; • restarting treatment	Beta
Resource use: • Number of GP consultations; • Number of specialist visits.	Lognormal
Abbreviations used in table: AE, adverse events; BoTox, botulinum toxin; GP, general practitioner.	

Generally, the ERG considers that the manufacturer used appropriate statistical distributions from which to sample model parameters. However, the ERG considers it important to note that, in the PSA, the manufacturer may not have accounted for correlation between β-coefficients of the regression equations used to inform the economic model. The manufacturer sampled β-coefficients from assigned normal distributions; however, the ERG did not identify the use of Cholesky decomposition and it is unclear whether the manufacturer accounted for correlation between β-coefficients used within the same regression model. The exclusion of any assessment of correlation from the manufacturer's PSA, is likely to decrease the accuracy around estimates of uncertainty; however, it is unclear whether uncertainty would be over- or under-estimated.

Figure 6 displays the results of the manufacturer's PSA in the cost-effectiveness plane. The cost-effectiveness acceptability curve (CEAC) is presented in Figure 7. The manufacturer concluded that mirabegron 50 mg has an 89.4% probability of being cost-effective versus tolterodine ER 4 mg at a threshold of £20,000 per QALY.

Figure 6. Results of the manufacturer's probabilistic sensitivity analysis (mirabegron 50 mg versus tolterodine ER 4 mg) in the cost-effectiveness plane (reproduced from MS; Figure 49, pg 238)

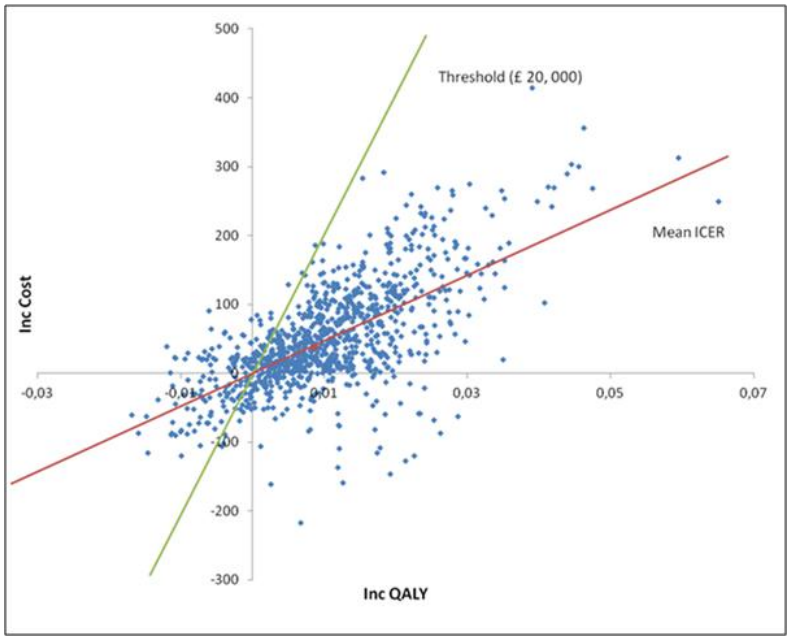
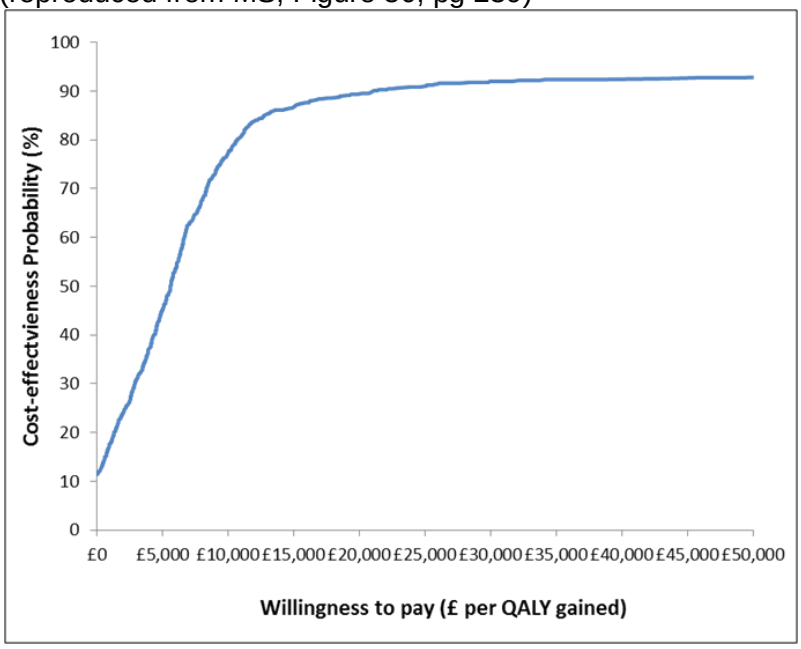


Figure 7. Cost-effectiveness acceptability curve: mirabegron 50 mg vs tolterodine ER 4 mg (reproduced from MS; Figure 50; pg 239)



Structural uncertainty

Assessment of uncertainty associated with the manufacturer's structural assumptions included sensitivity analyses around:

- model time horizon;
- the impact of OAB-related co-morbidities;
- the use of disease specific (OAB-5D) rather than generic (EQ-5D) utility values.

The manufacturer's base case models used 5 year time horizons, generic utility data and excluded the impact of co-morbidities. In structural sensitivity analyses, the manufacturer considered:

- time horizons of 1, 2 and 10 years;
- utility values derived from the disease specific OAB-5D measure;
- co-morbidities of depression, fractures (as a result of falls), urinary tract infections (UTIs) and skin infections.

As outlined in Section 5.2, the manufacturer submitted a version of the secondary base case model which included co-morbidity data. However, the manufacturer did not submit a version of the primary base case model which included co-morbidity data; therefore, the ERG was unable to validate this aspect of the manufacturer's sensitivity analyses. In addition, the ERG notes that assessment of the impact of time horizon and the use of OAB-5D based utility values on the secondary base case results was not presented in the MS.

Disease specific (OAB-5D) utility values were estimated using a similar linear regression model to that used to estimate EQ-5D utility values (see Section 5.2.7). Table 71 summarises the OAB-5D utility values associated with each of the 25 disease severity profiles considered in the manufacturer's base case models.

Table 71. OAB-5D utility values by disease severity profile (adapted from MS; Table 95; pg 208)

Incontinence frequency level	**Micturitions frequency level**				
	1	2	3	4	5
1	0.92	0.88	0.85	0.84	0.82
2	0.89	0.85	0.83	0.81	0.79
3	0.87	0.83	0.80	0.78	0.77
4	0.85	0.81	0.79	0.77	0.75
5	0.84	0.80	0.78	0.76	0.74

To assess the impact of co-morbidity on the cost-effectiveness results, the manufacturer applied monthly probabilities of depression, fractures, UTIs and skin infections to all patients regardless of health state or treatment. The probability associated with each co-morbidity was assumed to vary between patients who were continent and those who were not (Table 72). Patients who experienced

co-morbidities accrued additional costs and disutilities according to the co-morbidities experienced (Table 73); costs and disutilities were applied for one cycle only.

Table 72. Monthly probability of co-morbidities used in manufacturer's sensitivity analyses (adapted from MS; Table 104; pg 223)

Co-morbidity	Continent patients (level 1)	Incontinent patients (level 2 or above)	Source
Falls with fractures	0.42%	0.90%	Arlandis-Guzman, 2011[79]
Depression	0.70%	1.72%	
Skin infection	1.78%	1.55%	
UTI	3.17%	5.12%	
Abbreviation used in table: UTI, urinary tract infection.			

Table 73. Cost and disutility associated with each considered co-morbidity

Co-morbidity	Disutility	Source	Cost	Source
Falls with fractures	–0.239	Peasgood, 2009[117] First year utility loss	£5,048.00	NHS 2011-12 tariff information
Depression	–0.248	NICE CG90 Oct 2009[118] Moderate depression patients on citalopram, over 12 months Mean QALY for a moderate depressive patients on citalopram:0.602 Mean QALY for a healthy UK person :0.85	£1,522.00	Moderate depression patients on citalopram (NICE CG90, Oct 2009)[118]
Skin infection	–0.017	Assumption: Utility loss of 0.2 over one cycle.	£96.00	Assumption: unit cost of a specialist visit
UTI	–0.024	Barry, 1997[119] Utility loss of 0.2894 for a day with UTI symptoms, over one cycle	£96.00	Assumption: unit cost of a specialist visit
Abbreviations used in table: CG, clinical guideline; NHS, National Health Service; NICE, National Institute for Health and Clinical Excellence; QALY, quality adjusted life year; UTI, urinary tract infection.				

Tables 74 and 75 summarise the impact of the manufacturer's structural sensitivity analyses on the primary and secondary base case cost-effectiveness results, respectively.

Table 74: Results of structural sensitivity analysis on manufacturer's primary base case (mirabegron 50 mg versus tolterodine ER 4 mg) cost-effectiveness results (adapted from MS; Tables 120-121; pgs 239-240)

Structural sensitivity analysis		Incr. costs (£)	Incr. QALYs	ICER (£/QALY)
Time horizon	1 year	12.37	0.00315	3,925
	2 years	25.50	0.00588	4,338
	5 years[a]	**37.88**	**0.00864**	**4,386**
	10 years	33.60	0.00914	3,675
Inclusion of comorbidities[b]		37.88	0.01943	1,950
Use of OAB-5D utilities		37.88	0.01259	3,008

[a] Primary base case result

[b] ERG unable to validate.

Abbreviations used in table: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OAB-5D, Overactive bladder-5D (based on OAB-q questionnaire); QALYs, quality adjusted life years.

Table 75. Results of co-morbidity scenario analysis on manufacturer's secondary base case (comparator versus mirabegron 50 mg) cost-effectiveness results (adapted from MS; Table 122; pg 240)

Population	Incr. costs	Incr. QALYs	ICER (versus mirabegron 50 mg)
Tolterodine ER 4 mg	–£11.29	0.01060	Dominant
Solifenacin 5 mg	£57.74	0.00466	£12,386
Solifenacin 10 mg	–£40.67	0.01075	Dominant
Trospium chloride 60 mg	£49.05	0.00973	£5,038
Fesoterodine 4 mg	–£17.80	0.01102	Dominant
Oxybutynin ER 10 mg	£3.42	0.01118	£306
Oxybutynin IR 10 mg	£150.11	0.01511	£9,937

Abbreviations used in table: ER, extended release; IR, immediate release.

The ERG notes that the primary base case cost effectiveness results were relatively insensitive to changes in time horizon. However, the primary base case ICER decreased by £2,436 when co-morbidity was taken into account; although, as mentioned previously the ERG were unable to validate this analysis. Furthermore, the ERG notes that the inclusion of co-morbidity has a varied impact on the individual comparisons (versus mirabegron 50 mg) of the secondary base case results (Table 75). In particular, the ERG notes that contrary to the primary base case model, the inclusion of co-morbidity data in the secondary base case model results in mirabegron being dominated by tolterodine ER 4 mg.

5.2.11 Subgroup analyses

As specified in the final scope issued by NICE,[19] the manufacturer carried out the following subgroup analyses:

- male patients;
- female patients;
- previously treated patients;
- treatment-naïve patients.

Each subgroup analysis was based on data from a "pre-specified pooled analysis of the three primary studies, SCORPIO, ARIES and CAPRICORN" (MS; pg 93; see Section 4.1.6 for further details of pooled analysis). These data were used to inform the following model parameters:

- probability of switch after discontinuation (assumed to be zero in previously treated patients);
- probability of having BoTox injections at 1 month (assumed to be zero in all but the previously treated patient population, i.e. no other second line treatment assumed for previously treated patients);
- transition probabilities between symptom levels (derived from data for subgroup under consideration);
- utilities by symptom levels (derived from data for subgroup under consideration).

The primary (mirabegron 50 mg versus tolterodine ER 4 mg) results of the manufacturer's subgroup analyses are presented in Table 76.

Table 76. Results of the manufacturer's primary (mirabegron 50 mg versus tolterodine ER 4 mg) subgroup analyses (reproduced from MS; Table 124; pg 242)

Treatment	Total			Incremental			ICER (£/QALY) incremental
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Previously treated subgroup							
Tolterodine ER 4 mg	1,643.26	4.666	3.640	–	–	–	–
Mirabegron 50 mg	1,681.32	4.666	3.650	38.07	0	0.0099	3,836
Treatment-naïve subgroup							
Tolterodine ER 4 mg	1,535.76	4.666	3.847	–	–	–	–
Mirabegron 50 mg	1,576.03	4.666	3.855	40.27	0	0.0076	5,315
Male subgroup							
Tolterodine ER 4 mg	1,411.85	4.666	3.888	–	–	–	–
Mirabegron 50 mg	1,455.80	4.666	3.889	43.96	0	0.0011	38,708
Female subgroup							
Tolterodine ER 4 mg	1,694.47	4.666	3.684	–	–	–	–
Mirabegron 50 mg	1,732.20	4.666	3.697	37.73	0	0.0122	3,091
Abbreviations used in table: ER, extended release; ICER, incremental cost-effectiveness ratio; IR, immediate release; LYG, life year gained; QALY, quality adjusted life year.							

The ERG notes that subgroup analyses were only available for the manufacturer's primary comparison of mirabegron 50 mg versus tolterodine ER 4 mg. Moreover, the ERG considers that these subgroup analyses indicate that the manufacturer's primary base case ICER is robust (i.e. relatively insensitive to changes) with respect to the subgroups considered, with the exception of male gender. Furthermore, the ERG notes that the number of male patients recruited to the manufacturer's key trials was lower; therefore, reducing statistical power to detect differences in efficacy. In addition, male patients displayed lower baseline severity of OAB and experienced a higher placebo response (MS; pg 105).

5.2.12 Model validation and face validity check

Box 13 summarises the model validation and face validity checks carried out by the manufacturer.

Box 13. Manufacturer's model validation and face validity assessments (reproduced from MS; pg 241; references have been renumbered)

> The model underwent verification and validation consistent with recommendations by Philips et al.[85] and the ISPOR Task Force.[120] Verification is defined as the process of determining the model is implemented correctly and accurately. Validation refers to the process of evaluating the degree to which the model represents the real world data. Within the verification process the model was checked for internal consistency, accurate data inputs, and logical and mathematically correct calculations. Calculation checks were carried out in order to identify errors (such as probabilities not summing to 1) and to ensure symmetry was present, i.e. outcomes were the same for treatments in different sections of the model. This included testing the model using null or extreme values and comparing the results to expected results. The validation exercise comprised comparison of model outputs to clinical trial data used in the model, face validity and checking of key assumptions by clinical experts.

The ERG considers that the manufacturer's validation process was reasonable; although, it is not clear by whom verification was carried out. However, the ERG did not identify any programming errors during model examination and critique.

5.3 Conclusions of the cost-effectiveness section

The manufacturer submitted four versions of an EXCEL-based economic model, which presented the cost-effectiveness of mirabegron 50 mg:

- versus tolterodine ER 4 mg (primary base case),
- versus all comparators (except oxybutynin IR 10 mg) listed in the NICE final scope for this STA (secondary base case),
- versus oxybutynin IR 10 mg (included in the secondary base case);
- versus all comparators (except oxybutynin IR 10 mg) listed in the NICE final scope for this STA including the impact of selected co-morbidities.

In addition, the manufacturer carried out subgroup analyses on:

- previously treated patients;
- patients who were treatment naïve;
- male patients;
- female patients.

Therefore, the ERG considers the manufacturer's submission to have fully addressed all aspects of the NICE scope. In addition, the ERG considers the manufacturer's economic models to be accurately programmed and largely transparent. However, the ERG identified some areas of uncertainty or inaccuracy in the parameter estimates and structural assumptions used in the manufacturer's models. These are detailed below, along with their potential impact on the manufacturer's cost-effectiveness results.

6 IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

Following detailed examination of the MS and the manufacturer's primary and secondary base case models, the ERG identified some areas of inaccuracy or uncertainty, namely:

- the uncertainty resulting from heterogeneity associated with estimates from the manufacturer's MTC (see Section 5.2.3);
- the assumption of variable other cause discontinuation for mirabegron patients (see Section 5.2.6);
- the assumption that immediate (i.e. within the same cycle) discontinuation as a result of an AE would be equivalent to the rate of other cause discontinuation (see Section 5.2.4);
- the possibility of infinite treatment discontinuation and reinitiation, a factor of the "lack of memory" associated with the Markov model (see Section 5.2.4);
- the use of AE rates from SCORPIO rather than the manufacturer's safety study TAURUS (see Section 5.2.6);
- the cost associated with BoTox injections (see Section 5.2.8);
- the use of NHS pbR tariffs rather than reference costs to inform the cost of outpatient specialist visits (see Section 5.2.8);
- the exclusion of correlation from the PSA (see Section 5.2.10).

Where possible, the ERG carried out sensitivity analyses to investigate the impact of alternative assumptions or parameters on the manufacturer's deterministic base case cost-effectiveness results; deterministic, rather than probabilistic values were used as a result of time constraints. The individual and cumulative impact, of the ERG's sensitivity analyses, on the manufacturer's primary base case cost-effectiveness result is displayed in Table 77. The impact (individual and cumulative) of these sensitivity analyses on the cost-effectiveness results of each comparison (mirabegron versus comparator) is summarised in Table 78. In addition, the combined impact of the ERG's sensitivity analyses on the manufacturer's incremental secondary base case results has been computed and is displayed in Table 79 (Full details are provided in Appendix 11).

Table 77. Individual and cumulative results of ERG's sensitivity analysis on manufacturer's primary base case cost-effectiveness results

Sensitivity analysis	Treatment	Cost	QALY	Inc cost	Inc QALY	ICER (individual)	ICER (Cumulative)
Manufacturer's base case	**Tolterodine ER 4 mg**	**1,607.75**	**3.755**	**–**	**–**	**–**	**–**
	Mirabegron 50 mg	**1,645.62**	**3.764**	**37.88**	**0.009**	**4,385.65**	**4,385.65**
Uncertainty associated with manufacturer's MTC data	N/A, primary base case model is based on trial rather than MTC data						
Assuming mirabegron is associated with a 28% persistence rate	Tolterodine ER 4 mg	1,607.75	3.755	–	–	–	–
	Mirabegron 50 mg	1,645.28	3.764	37.53	0.009	4,382.63	4,382.63
The use of AE-specific (rather than other cause) immediate discontinuation rates	The ERG was unable to assess the impact of this sensitivity analysis as a result of the lack of AE-specific discontinuation rates						
Probability of reinitiating original therapy set to 0	Tolterodine ER 4 mg	1,591.67	3.756	–	–	–	–
	Mirabegron 50 mg	1,624.30	3.763	32.64	0.007	4,516.14	4,512.21
Use of adverse event rates from TAURUS	Tolterodine ER 4 mg	1,603.03	3.753	–	–	–	–
	Mirabegron 50 mg	1,634.91	3.761	31.88	0.007	4,313.61	4,409.69
The use of NHS reference costs for BoTox injections	Tolterodine ER 4 mg	1,529.14	3.755	–	–	–	–
	Mirabegron 50 mg	1,574.02	3.764	44.87	0.009	5,195.88	5,180.84
The use of NHS reference costs for outpatient specialist visits	Tolterodine ER 4 mg	1,585.70	3.755	–	–	–	–
	Mirabegron 50 mg	1,624.50	3.764	38.80	0.009	4,492.89	5,271.75
Assessment of the impact of correlation in the manufacturer's PSA	N/A as sensitivity analysis is based on the deterministic primary base case results						

Abbreviations used in table: AE, adverse event; BoTox, botulinum toxin; ER, extended release; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; Inc, incremental; NHS, National Health Service; PSA, probabilistic sensitivity analysis; QALY, quality adjusted life year.

Table 78. Individual and cumulative ICERs (£/QALY) from ERG's sensitivity analyses on the cost-effectiveness results of each comparison (mirabegron versus comparator) made in the manufacturer's secondary base case

Treatment	ERG's sensitivity analyses								
	Manufacturer's base case	**Uncertainty associated with manufacturer's MTC data**	**Assuming mirabegron is associated with a 28% persistence rate**	**The use of AE-specific (rather than other cause) immediate discontinuation rates**	**Probability of reinitiating original therapy set to 0**	**Use of AE rates from TAURUS**	**The use of NHS reference costs for BoTox injections**	**The use of NHS reference costs for outpatient specialist visits**	**Assessment of the impact of correlation in the manufacturer's PSA**
ICER for each sensitivity analysis (i.e., showing individual impact of each sensitivity analysis)									
Solifenacin 10 mg	340.15	The ERG were not explicitly able to quantify the impact of uncertainty associated with the heterogeneity present in the manufacturer's MTC.	Dominant	The ERG were unable to assess the impact of this sensitivity analysis as a result of the lack of AE-specific discontinuation rates	806.92	Relative AE events versus mirabegron using TAURUS data were not available. Therefore the ERG were unable to assess this uncertainty in the secondary base case CE results	2,164.84	572.79	N/A as sensitivity analysis is based on the deterministic primary base case results
Fesoterodine 4 mg	3,606.71		3,605.91		3,688.36		4,335.77	3,702.75	
Tolterodine ER 4 mg	3,714.98		3,715.11		3,791.96		4,430.84	3,809.55	
Oxybutynin ER 10mg	3,877.57		3,843.00		3951.23		4,754.83	3,987.39	
Trospium chloride MR 60 mg	8,881.48		8,421.70		8,745.49		10,242.61	9,052.29	
Solifenacin 5 mg	12,493.21		32,571.50		11,778.18		14,714.70	12,792.91	
Oxybutynin IR 10 mg	14,233.83		12,595.41		13,979.58		15,260.24	14,357.67	

Cumulative ICERs (i.e., showing cumulative impact of sensitivity analyses)									
Solifenacin 10 mg	340.15	The ERG were not explicitly able to quantify the impact of uncertainty associated with the heterogeneity present in the manufacturer's MTC.	Dominant	The ERG were unable to assess the impact of this sensitivity analysis as a result of the lack of AE-specific discontinuation rates	Dominant	Relative AE events versus mirabegron using TAURUS data were not available. Therefore the ERG were unable to assess this uncertainty in the secondary base case CE results	1,662.65	1,913.33	N/A as sensitivity analysis is based on the deterministic primary base case results
Fesoterodine 4 mg	3,606.71		3,605.91		3,687.04		4,478.33	4,571.38	
Tolterodine ER 4 mg	3,714.98		3,715.11		3,791.54		4,568.29	4,659.90	
Oxybutynin ER 10mg	3,877.57		3,843.00		3,927.49		4,930.03	5,043.61	
Trospium chloride MR 60 mg	8,881.48		8,421.70		8,298.32		9,728.65	9,889.48	
Solifenacin 5 mg	12,493.21		32,571.50		27,727.62		32,181.63	32,711.50	
Oxybutynin IR 10 mg	14,233.83		12,595.41		12,332.34		13,458.81	13,582.07	
Abbreviations used in table: AE, adverse event; BoTox, botulinum toxin; ER, extended release; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; IR, immediate release; MR, modified release; MTC, mixed treatment comparison; NA, not applicable; NHS, National Health Service; PSA, probabilistic sensitivity analysis; QALY, quality adjusted life year.									

Table 79. Combined impact of ERG's sensitivity analyses on manufacturer's incremental secondary base case cost-effectiveness results

Treatment	Manufacturer's base case ICER (£/QALY)	Total		Incremental (versus previous therapy)		ICER (£/QALY) versus oxybutynin IR 10 mg	Incremental ICER (£/QALY)
		Cost (£)	QALY	Cost (£)	QALY		
Oxybutynin IR 10 mg	–	1,329.37	3.753	–	–	–	–
Trospium chloride MR 60 mg	18,816.13[a]	1,437.46	3.759	108.09	0.006	19,401.56	19,401.56[a]
Oxybutynin 10mg ER	Strictly dominated[b]	1,467.09	3.756	29.63	–0.003	45,763.53	Strictly dominated[b]
Solifenacin 5 mg	4,591.75[c]	1,477.88	3.766	10.80	0.010	11,483.59	£5,491.22[c]
Fesoterodine 4 mg	Strictly dominated[d]	1,484.00	3.759	6.12	–0.007	27,923.18	Strictly dominated[d]
Tolterodine ER 4 mg	Extendedly dominated[e]	1,484.65	3.759	0.65	0.000	26,571.76	Strictly dominated[d]
Solifenacin 10 mg	Extendedly dominated[e]	1,512.41	3.761	27.77	0.002	22,478.28	Strictly dominated[d]
Mirabegron 50 mg	12,493.21[f]	1,524.29	3.768	11.88	0.006	13,582.07	32,711.50[f]

[a] versus Oxybutynin IR 10 mg.

[b] by trospium chloride 60 mg MR.

[c] versus trospium chloride 60 mg MR.

[d] by solifenacin 5 mg.

[e] by mirabegron 50 mg.

[f] versus solifenacin 5 mg.

Abbreviations used in table: ER, extended release; ICER, incremental cost-effectiveness ratio; IR, immediate release; MR, modified release; QALY, quality adjusted life year.

The sensitivity analyses carried out by the ERG indicated that the manufacturer's primary base case cost-effectiveness result was generally robust with respect to the areas of uncertainty identified in the ERG's critique (Table 77). However, as a result of a paucity of data, the ERG were unable to assess the impact of using AE-specific (rather than other cause) immediate discontinuation rates. Although, the primary cost-effectiveness result has been demonstrated to be robust to alternative estimates of discontinuation (see Section 5.2.10).

Following individual and cumulative application of the ERG's sensitivity analyses, the ICERs estimated by the manufacturer for the individual comparisons considered in the secondary base case remained relatively consistent; with the exception of comparisons between mirabegron 50 mg and solifenacin 5 and 10 mg (Table 78). The ERG notes that the impact of each sensitivity analysis on the ICERs for mirabegron 50 mg versus solifenacin 5 and 10 mg was highly variable, with ICERs ranging from £573 to the dominance of solifenacin in the comparison of mirabegron 50 mg versus solifenacin 10 mg. Similarly, ICERs ranged from £11,778 to £32,572 upon application of the ERG's sensitivity analyses to the comparison of mirabegron 50 mg versus solifenacin 5 mg. Moreover, the cumulative impact of the ERG's sensitivity analyses on the comparison of mirabegron 50 mg with solifenacin 5 and 10 mg, resulted in ICER increases of £20,218 (from £12,493 to £32,711.50) and £1,573 (from £340 to £1,913), respectively.

Regarding the impact of the ERG's sensitivity analyses on the manufacturer's incremental secondary base case cost-effectiveness results; the ERG notes that, following simultaneous application of the ERG's sensitivity analyses, tolterodine ER 4 mg and solifenacin 10 mg move from being extendedly dominated by mirabegron 50 mg to being strictly dominated by solifenacin 5 mg (Table 79). Moreover, the ICER of mirabegron 50 mg versus solifenacin 5 mg increases from £12,493 to £32,712. Furthermore, the ERG notes that the ERG's revised ICER of mirabegron 50 mg versus solifenacin 5 mg (£32,712) is based on clinical effectiveness estimates from the manufacturer's MTC. As discussed in Section 4.4.2, as a result of ERG concern regarding the level of heterogeneity present in the manufacturer's MTC, the ERG carried out a revised MTC using a more homogeneous data set. However, the manufacturer's model structure did not facilitate implementation of data obtained from the ERG's MTC (i.e. the manufacturer's model used calibrated beta coefficients rather than relative estimates such as hazard ratios or odds ratios); therefore, the ERG were unable to quantify the impact of incorporating clinical effectiveness estimates from the ERG's MTC into the economic model. However, the ERG notes that estimates obtained from the ERG's MTC indicate that solifenacin 5 mg is statistically significantly more effective at reducing incontinence episodes than mirabegron 50 mg (mean difference [95% CrI]: -0.386[-0.717 to -0.055]). By contrast, estimates obtained from the manufacturer's MTC detected no statistically significant difference between solifenacin 5 mg and mirabegron 50 mg in reducing the number of incontinence episodes experienced per 24 hours (mean difference [95% CrI]: -0.237 [-0.482 to 0.007]). Based on this, the ERG considers that the ERG's revised ICER for the comparison of mirabegron 50 mg with solifenacin 5 mg is likely to be conservative; i.e. an ICER estimated using ERG MTC data is likely to be higher than £32,712.

7 OVERALL CONCLUSIONS

7.1 Summary of clinical effectiveness

The manufacturer presents the case for the use of mirabegron in the treatment of overactive bladder (OAB) based on data derived from three randomised controlled trials (RCTs): ARIES; CAPRICORN; and SCORPIO. All three RCTs were multiple arm trials in patients with symptoms of OAB, with SCORPIO and ARIES evaluating the clinical effectiveness of mirabegron at doses of 50 and 100 mg versus placebo, and CAPRICORN evaluating mirabegron at doses of 50 and 25 mg, again against placebo. The primary outcomes in the three trials were change from baseline (CFB) in frequency of micturition and in frequency of incontinence episodes.

At the time of writing of the ERG's report, mirabegron does not have a European licence for use in the treatment of OAB. However, the Committee for Medicinal Products for Human Use (CHMP) has adopted a positive opinion on the use of mirabegron at doses of 50 mg (recommended dose) and 25 mg (for patients with renal or hepatic failure) for the symptomatic treatment of urgency, increased micturition frequency and/or urgency incontinence as may occur in adult patients with OAB. Based on the positive opinion issued by the CHMP, the ERG does not consider data on mirabegron 100 mg to be relevant to the decision problem that is the focus of this Single Technology Appraisal (STA).

The direct evidence submitted predominantly compares the effects of mirabegron versus placebo, which, in the context of the comparisons of interest, the ERG considers does not fully address the decision problem. SCORPIO included an additional treatment group of an active control (tolterodine), which the manufacturer chose not to report as SCORPIO was not powered to evaluate the superiority or non-inferiority of mirabegron versus tolterodine. The ERG acknowledges the manufacturer's point but considers that exclusion of data from the tolterodine treatment group in SCORPIO means that there is no direct evidence relevant to the decision problem presented within the manufacturer's submission (MS).

In addition to ARIES, CAPRICORN, and SCOPRIO, the manufacturer identified four other trials evaluating the use of mirabegron in the treatment of OAB, which were described as supporting evidence and did not form part of the submitted evidence (TAURUS, DRAGON, 178-CL-045, and 178-CL-048). The manufacturer cited various reasons for exclusion of these trials from the analysis. Of the four additional trials, TAURUS was designed as a long-term follow-up safety study, with treatment duration of 12 months, compared with 3 months in the other trials. On reviewing the four RCTs, the ERG considers the RCTs to be relevant to the decision problem.

For SCORPIO, ARIES, and CAPRICORN the manufacturer presents data on outcomes assessed based on CFB for the individual treatment groups within each trial, and the difference between

mirabegron and placebo at the end point. For SCORPIO, analyses of the difference between mirabegron and the active control tolterodine for the outcomes assessed are not reported. The ERG considers that, although the results of effectiveness of mirabegron versus placebo are key as it is important to demonstrate that a treatment is more effective than placebo, analysis of mirabegron versus active treatment(s) currently used in the NHS is of more relevance to the decision problem.

Considering the six trials evaluating mirabegron 50 mg with treatment duration of 3 months (ARIES, CAPRICORN, SCORPIO, DRAGON, 178-CL-045, and 178-CL-048), mirabegron 50 mg was found to be more effective than placebo at reducing all clinical outcomes evaluated, with most differences reaching statistical significance. However, results from two trials that included tolterodine as an active control group and for which data were available suggest that mirabegron is of similar clinical effectiveness to tolterodine, with no statistically significant differences noted between the two active treatments for the outcomes reported. It should be noted that trials including tolterodine as an active control were not powered to evaluate the superiority or non-inferiority of mirabegron versus tolterodine.

The manufacturer did not perform a meta-analysis of the included trials, but presented the results of a pre-specified pooled analysis of ARIES, CAPRICORN, and SCORPIO for the comparison of mirabegron versus placebo. The results of the manufacturer's analysis support the results from the individual RCTs, with mirabegron being statistically significantly more effective than placebo at improving all clinical outcomes. However, as noted earlier, the ERG considers that comparison versus placebo does not inform the decision problem. The ERG carried out an independent meta-analysis of the available data for the comparison of mirabegron versus tolterodine, focusing on the primary clinical outcomes of frequency of urination and of incontinence episodes. The ERG's meta-analysis found that, compared with tolterodine 4 mg, treatment with mirabegron 50 mg led to significantly fewer micturitions per 24 hours (MD –0.27; 95% CI: –0.48 to –0.06; p-value = 0.01), and significantly fewer incontinence episodes per 24 hours (MD –0.21; 95% CI –0.41 to –0.01; p-value = 0.04).

In terms of indirect evidence, the manufacturer identified 40 trials that were used to carry out a mixed treatment comparison (MTC). The manufacturer focused on oral preparations of treatments. The ERG considers exclusion of studies evaluating non-oral preparations to be inappropriate and a deviation from the final scope, which specified modified-release formulations of oxybutynin (available as a transdermal patch). The ERG notes that the seven trials evaluating mirabegron, including the four trials excluded by the manufacturer from the direct clinical evidence, were included in the trials deemed eligible for inclusion in the MTC. The ERG had concerns around the potential clinical and methodological heterogeneity across the identified RCTs and as result carried out an MTC excluding trials of poor methodological quality and RCTs that included patients substantially different from

those enrolled in ARIES, CAPRICORN, and SCORPIO (e.g., population limited to women or patients with a specific symptom of OAB). The outcomes considered in the ERG's MTC were micturition, incontinence, constipation, dry mouth, and additionally all-cause discontinuation (as the outcome of all-cause discontinuation was believed to be a potential key driver in the economic model and was not included in the manufacturer's MTC). The results of the ERG's MTC are in general agreement with those of the manufacturer's MTC. For micturition per 24 hours, the ERG's MTC found no statistically significant differences between mirabegron 50 mg and all active treatments evaluated. However, mirabegron 50 mg was found to be significantly less effective than solifenacin (5 mg and 10 mg) at reducing frequency of incontinence episodes, but all other differences were not statistically significant. For the adverse effects of dry mouth and constipation, which are associated with anticholinergics, mirabegron 50 mg was found to significantly lower the risk of dry mouth compared with all other antimuscarinics assessed, and to significantly lower the risk of constipation compared with fesoterodine 8 mg, solifenacin (5 mg and 10 mg), and trospium 60 mg.

Indirect evidence on clinical effectiveness suggests that mirabegron is of comparative clinical efficacy to currently used treatments for OAB, with a lower risk of experiencing common adverse effects associated with anticholinergics, in particular dry mouth. However, the ERG notes that trials evaluating transdermal formulations of oxybutynin were omitted from the MTC, and that inclusion of additional evidence could affect the results.

7.2 Summary of cost-effectiveness

In support of this STA, the manufacturer presented two base case cost-effectiveness analyses, as follows:

- a primary base case analysis considering the comparison of mirabegron 50 mg with tolterodine ER 4 mg, using efficacy data from SCORPIO;
- secondary base case analyses considering mirabegron 50 mg versus all comparators listed in the NICE scope,[19] using efficacy data from the manufacturer's MTC.

The manufacturer's primary base case comparison of mirabegron 50 mg versus tolterodine ER 4 mg (based on clinical effectiveness data from SCORPIO) resulted in an estimated ICER of £4,386. Results of the secondary base case comparisons (based on clinical effectiveness data from the manufacturer's MTC) were presented within the MS individually (mirabegron versus solifenacin10 mg [£340], fesoterodine 4 mg [£3,607], tolterodine ER 4 mg [3,715], oxybutynin ER 10mg [3,878], trospium chloride MR 60 mg [8,881], solifenacin 5 mg [12,493], and oxybutynin IR 10 mg [14,234]) and incrementally. The manufacturer's incremental results were generated by assuming that mirabegron 50 mg was associated with an other cause discontinuation rate equal to that of patients treated with solifenacin (5 or 10 mg) and indicated that treatment with:

- oxybutynin ER 10 mg is strictly dominated (more costly and less effective) by treatment with trospium chloride MR 60 mg;
- fesoterodine 4 mg is strictly dominated by treatment with solifenacin 5 mg;
- tolterodine ER 4 mg and solifenacin 10 mg are extendedly dominated (less effective yet with a higher ICER) by treatment with mirabegron 50 mg.

Following detailed examination of the MS and the manufacturer's primary and secondary base case models, the ERG identified several areas of inaccuracy or uncertainty. Where possible, the ERG carried out sensitivity analyses to investigate the impact of alternative assumptions or parameters on the manufacturer's base case cost-effectiveness results. The sensitivity analyses carried out by the ERG indicate that the manufacturer's primary base case cost-effectiveness result was generally robust; cumulative impact of ERG sensitivity analyses increased the ICER by £886 (from £4,386 to £5,272). However, the manufacturer's incremental secondary base case cost-effectiveness results were substantially altered by application of the ERG's sensitivity analyses. In particular, following simultaneous application of the ERG's sensitivity analyses, tolterodine ER 4 mg and solifenacin 10 mg move from being extendedly dominated by mirabegron 50 mg to being strictly dominated by solifenacin 5 mg. Moreover, the ICER of mirabegron 50 mg versus solifenacin 5 mg increases from £12,493 to £32,712. Furthermore, the ERG was unable to quantify the impact of using alternative assumptions or parameters in all areas of uncertainty identified in the ERG's critique. Particularly, the use of clinical effectiveness estimates from the ERG's revised MTC (using a more homogeneous data set than that used to inform the manufacturer's MTC). However, the ERG notes that estimates obtained from the ERG's MTC indicate that solifenacin 5 mg is statistically significantly more effective at reducing incontinence episodes than mirabegron 50 mg (mean difference –0.386; 95% CrI: –0.717 to –0.055). By contrast, estimates obtained from the manufacturer's MTC detected no statistically significant difference between solifenacin 5 mg and mirabegron 50 mg in reducing the number of incontinence episodes experienced per 24 hours (mean difference –0.237; 95% CrI: –0.482 to 0.007). Based on this, the ERG considers that the ERG's revised ICER for the comparison of mirabegron 50 mg with solifenacin 5 mg is likely to be conservative; that is, an ICER estimated using ERG MTC data is likely to be higher than £32,712.

7.3 Implications for research

The results of the manufacturer's and ERG's MTCs vary in relation to the efficacy of mirabegron versus solifenacin, tolterodine 4 mg and oxybutynin. In addition, the ERG notes that sensitivity analysis carried out on the manufacturer's economic model suggested that mirabegron 50 mg versus solifenacin 5 mg is the comparison of interest with respect to fully incremental cost-effectiveness analysis. Therefore, the ERG considers that there is a need for further research into the clinical benefit of mirabegron 50 mg compared with solifenacin 5 mg, as a priority over direct comparisons with other therapies for OAB. Furthermore, the ERG considers that there is a need for a more robust synthesis of the direct and indirect evidence available in the literature (published and unpublished), including all treatments used in OAB.

8 REFERENCES

(1) Andersson KE, Arner A. Urinary bladder contraction and relaxation: physiology and pathophysiology. *Physiol Rev* 2004 Jul; **84**(3):935-86.

(2) Abrams P, Cardozo L, Fall M, Griffiths D, Rosier P, Ulmsten U, *et al*. The standardisation of terminology of lower urinary tract function: report from the Standardisation Sub-committee of the International Continence Society. *Neurourol Urodyn* 2002; **21**(2):167-78.

(3) de Groat WC. A neurologic basis for the overactive bladder. *Urology* 1997; **50**(6A Suppl):36-52.

(4) Brading AF, Turner WH. The unstable bladder: towards a common mechanism. *Br J Urol* 1994; **73**(1):3-8.

(5) Ellsworth P. Overactive bladder: etiology, diagnosis, and impact. eMedicine 2011. [last accessed 2013 January]. Available from: http://emedicine.medscape.com/article/459502-overview

(6) Milsom I, Abrams P, Cardozo L, Roberts RG, Thuroff J, Wein AJ. How widespread are the symptoms of an overactive bladder and how are they managed? A population-based prevalence study. *BJU Int* 2001; **87**(9):760-6.

(7) Stewart WF, Van Rooyen JB, Cundiff GW, Abrams P, Herzog AR, Corey R, *et al*. Prevalence and burden of overactive bladder in the United States. *World J Urol* 2003; **20**(6):327-36.

(8) Irwin DE, Milsom I, Hunskaar S, Reilly K, Kopp Z, Herschorn S, *et al*. Population-based survey of urinary incontinence, overactive bladder, and other lower urinary tract symptoms in five countries: results of the EPIC study. *Eur Urol* 2006; **50**(6):1306-14.

(9) Coyne KS, Sexton CC, Irwin DE, Kopp ZS, Kelleher CJ, Milsom I. The impact of overactive bladder, incontinence and other lower urinary tract symptoms on quality of life, work productivity, sexuality and emotional well-being in men and women: results from the EPIC study. *BJU Int* 2008; **101**(11):1388-95.

(10) Coyne KS, Zhou Z, Bhattacharyya SK, Thompson CL, Dhawan R, Versi E. The prevalence of nocturia and its effect on health-related quality of life and sleep in a community sample in the USA. *BJU Int* 2003; **92**(9):948-54.

(11) Darkow T, Fontes CL, Williamson TE. Costs associated with the management of overactive bladder and related comorbidities. *Pharmacotherapy* 2005; **25**(4):511-9.

(12) National Institute for Health and Clinical Excellence. Urinary incontinence: the management of urinary incontinence in women. NICE 2006. [last accessed 2013 January]. Available from: http://www.nice.org.uk/nicemedia/live/10996/30282/30282.pdf

(13) National Institute for Health and Clinical Excellence. Lower urinary tract symptoms: the management of lower urinary tract symptoms in men. NICE 2010. [last accessed 2013 January]. Available from: http://www.nice.org.uk/nicemedia/live/12984/48557/48557.pdf

(14) National Institute for Health and Clinical Excellence. NICE implementation uptake report: drugs used in the management of urinary incontinence. NICE 2009. [last accessed 2013 January]. Available from: http://www.nice.org.uk/media/386/B5/ImpUptakeReportCG40UrinaryIncontinence.pdf

(15) Astellas Pharma. Data on file: July MAT market share for UK OAB medications. 2012.

(16) Chapple CR, Khullar V, Gabriel Z, Muston D, Bitoun CE, Weinstein D. The effects of antimuscarinic treatments in overactive bladder: an update of a systematic review and meta-analysis. *Eur Urol* 2008; **54**(3):543-62.

(17) Nabi G, Cody JD, Ellis G, Herbison P, Hay-Smith J. Anticholinergic drugs versus placebo for overactive bladder syndrome in adults. *Cochrane Database Systematic Reviews* 2006. CD003781.

(18) Wagg A, Compion G, Fahey A, Siddiqui E. Persistence with prescribed antimuscarinic therapy for overactive bladder: a UK experience. *BJU Int* 2012; **110**(11):1767-74.

(19) National Institute for Health and Clinical Excellence. Mirabegron for the treatment of symptoms associated with overactive bladder. Final scope. NICE 2012. [last accessed 2013 January]. Available from: http://www.nice.org.uk/nicemedia/live/13828/60747/60747.pdf

(20) Astellas Pharma. Data on file: A randomised, double-blind, parallel group, placebo and active controlled, multicenter study to assess the efficacy and safety of mirabegron in subjects with symptoms of overactive bladder. 178-CL-046 (SCORPIO). Clinical Study Report. 2010.

(21) Astellas Pharma. Data on file: A Phase III, randomised, double-blind, parallel group, placebo controlled, multicenter study to assess the efficacy and safety of mirabegron in subjects with symptoms of overactive bladder. 178-CL-047 (ARIES). Clinical Study Report. 2010.

(22) Astellas Pharma. Data on file: A Phase III, randomized, double-blind, parallel group, placebo controlled, multicenter study to assess the efficacy and safety of the beta-3 agonist mirabegron (25 mg qd and 50 mg qd) in subjects with symptoms of overactive bladder. 178-CL-074 (CAPRICORN). Clinical Study Report. 2010.

(23) Astellas Pharma. Data on file: A randomized, double-blind, parallel group, placebo and active controlled, multi-center dose ranging study with the Beta-3 agonist YM178 in patients with symptomatic overactive bladder. 178-CL-044 (DRAGON). Clinical Study Report. 2010.

(24) Astellas Pharma. Data on file: A Phase 2, double-blind, placebo-controlled, parallel group study of YM178 in patients with overactive bladder. 178-CL-045. Clinical Study Report. 2010.

(25) Astellas Pharma. Data on file: Phase III Study of YM178, a double-blind group comparison study in patients with overactive bladder. 178-CL-048. Clinical Study Report. 2010.

(26) Astellas Pharma. Data on file: A randomized, double-blind, parallel group, active controlled, multi-center long-term study to assess the safety and efficiacy of the beta-3 agonist mirabegron (YM178) 50 mg qd and 100 mg qd in subjects with symptoms of overactive bladder. 178-CL-049 (TAURUS). Clinical Study Report. 2010.

(27) European Medicines Agency. Betmiga. EMA 2012. [last accessed 2013 January]. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/002388/smops/Positive/human_smop_000433.jsp&mid=WC0b01ac058001d127

(28) Food and Drug Administration. FDA Approved Drug Products: Myrbetriq. FDA 2012. [last accessed 2013 January]. Available from: http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm?fuseaction=Search.DrugDetails

(29) European Medicines Agency. Betmiga; mirabegron. EMA 2012. [last accessed 2013 January]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion_-_Initial_authorisation/human/002388/WC500134082.pdf

(30) Astellas Pharma. Data on file. A randomized, double-blind, parallel group, proof of concept study of YM178 in comparison with placebo and tolterodine in patients with symptomatic overactive bladder (BLOSSOM). 178-CL-008. Clinical Study Report. 2006.

(31) Abrams P, Cardozo L, Chapple C, Serdarevic D, Hargreaves K, Khullar V. Comparison of the efficacy, safety, and tolerability of propiverine and oxybutynin for the treatment of overactive bladder syndrome. *Int J Urol* 2006; **13**(6):692-8.

(32) Appell RA, Sand P, Dmochowski R, Anderson R, Zinner N, Lama D, *et al*. Prospective randomized controlled trial of extended-release oxybutynin chloride and tolterodine tartrate in the treatment of overactive bladder: results of the OBJECT Study. *Mayo Clin Proc* 2001; **76**(4):358-63.

(33) Birns J, Lukkari E, Malone-Lee JG. A randomized controlled trial comparing the efficacy of controlled-release oxybutynin tablets (10 mg once daily) with conventional oxybutynin tablets (5 mg twice daily) in patients whose symptoms were stabilized on 5 mg twice daily of oxybutynin. *BJU Int* 2000; **85**(7):793-8.

(34) Cardozo L, Lisec M, Millard R, Van Vierssen TO, Kuzmin I, Drogendijk TE, *et al*. Randomized, double-blind placebo controlled trial of the once daily antimuscarinic agent solifenacin succinate in patients with overactive bladder. *J Urol* 2004; **172**(5 Pt 1):1919-24.

(35) Chapple C, Van KP, Tubaro A, Haag-Molkenteller C, Forst HT, Massow U, *et al*. Clinical efficacy, safety, and tolerability of once-daily fesoterodine in subjects with overactive bladder. *Eur Urol* 2007; **52**(4):1204-12.

(36) Chapple CR, Arano P, Bosch JL, De RD, Kramer AE, Ridder AM. Solifenacin appears effective and well tolerated in patients with symptomatic idiopathic detrusor overactivity in a placebo- and tolterodine-controlled phase 2 dose-finding study. *BJU Int* 2004; **93**(1):71-7.

(37) Chapple CR, Rechberger T, Al-Shukri S, Meffan P, Everaert K, Huang M, *et al*. Randomized, double-blind placebo- and tolterodine-controlled trial of the once-daily antimuscarinic agent solifenacin in patients with symptomatic overactive bladder. *BJU Int* 2004; **93**(3):303-10.

(38) Choo MS, Lee JZ, Lee JB, Kim YH, Jung HC, Lee KS, *et al*. Efficacy and safety of solifenacin succinate in Korean patients with overactive bladder: a randomised, prospective, double-blind, multicentre study. *Int J Clin Pract* 2008; **62**(11):1675-83.

(39) Chu F, Smith N, Uchida U. Efficacy and safety of solifenacin succinate 10 mg once daily: a multicentre, phase III, randomized, double-blind, placebo-controlled, parallel-group trial in patients with overactive bladder. *Curr Ther Res Clin Exp* 2009; **70**(6): 405–420.

(40) Corcos J, Casey R, Patrick A, Andreou C, Miceli PC, Reiz JL, *et al*. A double-blind randomized dose-response study comparing daily doses of 5, 10 and 15 mg controlled-release oxybutynin: balancing efficacy with severity of dry mouth. *BJU Int* 2006; **97**(3):520-7.

(41) Diokno AC, Appell RA, Sand PK, Dmochowski RR, Gburek BM, Klimberg IW, *et al*. Prospective, randomized, double-blind study of the efficacy and tolerability of the extended-release formulations of oxybutynin and tolterodine for overactive bladder: results of the OPERA trial. *Mayo Clin Proc* 2003; **78**(6):687-95.

(42) Dmochowski RR, Sand PK, Zinner NR, Gittelman MC, Davila GW, Sanders SW. Comparative efficacy and safety of transdermal oxybutynin and oral tolterodine versus placebo in previously treated patients with urge and mixed urinary incontinence. *Urology* 2003; **62**(2):237-42.

(43) Herschorn S, Heesakkers J, Castro-Diaz D, Wang JT, Brodsky M, Guan Z. Effects of tolterodine extended release on patient perception of bladder condition and overactive bladder symptoms*. *Curr Med Res Opin* 2008; **24**(12):3513-21.

(44) Herschorn S, Stothers L, Carlson K, Egerdie B, Gajewski JB, Pommerville P, *et al*. Tolerability of 5 mg solifenacin once daily versus 5 mg oxybutynin immediate release 3 times daily: results of the VECTOR trial. *J Urol* 2010; **183**(5):1892-8.

(45) Herschorn S, Swift S, Guan Z, Carlsson M, Morrow JD, Brodsky M, *et al*. Comparison of fesoterodine and tolterodine extended release for the treatment of overactive bladder: a head-to-head placebo-controlled trial. *BJU Int* 2010; **105**(1):58-66.

(46) Ho CH, Chang TC, Lin HH, Liu SP, Huang KH, Yu HJ. Solifenacin and tolterodine are equally effective in the treatment of overactive bladder symptoms. *J Formos Med Assoc* 2010; **109**(10):702-8.

(47) Homma Y, Paick JS, Lee JG, Kawabe K. Clinical efficacy and tolerability of extended-release tolterodine and immediate-release oxybutynin in Japanese and Korean patients with an overactive bladder: a randomized, placebo-controlled trial. *BJU Int* 2003; **92**(7):741-7.

(48) Jacquetin B, Wyndaele J. Tolterodine reduces the number of urge incontinence episodes in patients with an overactive bladder. *Eur J Obstet Gynecol Reprod Biol* 2001; **98**(1):97-102.

(49) Kaplan SA, Schneider T, Foote JE, Guan Z, Carlsson M, Gong J. Superior efficacy of fesoterodine over tolterodine extended release with rapid onset: a prospective, head-to-head, placebo-controlled trial. *BJU Int* 2011; **107**(9):1432-40.

(50) Khullar V, Hill S, Laval KU, Schiotz HA, Jonas U, Versi E. Treatment of urge-predominant mixed urinary incontinence with tolterodine extended release: a randomized, placebo-controlled trial. *Urology* 2004; **64**(2):269-74.

(51) Lackner TE, Wyman JF, McCarthy TC, Monigold M, Davey C. Randomized, placebo-controlled trial of the cognitive effect, safety, and tolerability of oral extended-release oxybutynin in cognitively impaired nursing home residents with urge urinary incontinence. *J Am Geriatr Soc* 2008; **56**(5):862-70.

(52) Lee JG, Hong JY, Choo MS, Kwon HY, Chung DY, Lee KS, *et al*. Tolterodine: as effective but better tolerated than oxybutynin in Asian patients with symptoms of overactive bladder. *Int J Urol* 2002; **9**(5):247-52.

(53) Malone-Lee JG, Walsh JB, Maugourd MF. Tolterodine: a safe and effective treatment for older patients with overactive bladder. *J Am Geriatr Soc* 2001; **49**(6):700-5.

(54) Nitti VW, Dmochowski R, Sand PK, Forst HT, Haag-Molkenteller C, Massow U, *et al*. Efficacy, safety and tolerability of fesoterodine for overactive bladder syndrome. *J Urol* 2007; **178**(6):2488-94.

(55) Nitti VW, Rovner ES, Bavendam T. Response to fesoterodine in patients with an overactive bladder and urgency urinary incontinence is independent of the urodynamic finding of detrusor overactivity. *BJU Int* 2010; **105**(9):1268-75.

(56) Rackley R, Weiss JP, Rovner ES, Wang JT, Guan Z. Nighttime dosing with tolterodine reduces overactive bladder-related nocturnal micturitions in patients with overactive bladder and nocturia. *Urology* 2006; **67**(4):731-6.

(57) Rogers R, Bachmann G, Jumadilova Z, Sun F, Morrow JD, Guan Z, *et al*. Efficacy of tolterodine on overactive bladder symptoms and sexual and emotional quality of life in sexually active women. *Int Urogynecol J Pelvic Floor Dysfunct* 2008; **19**(11):1551-7.

(58) Rudy D, Cline K, Harris R, Goldberg K, Dmochowski R. Time to onset of improvement in symptoms of overactive bladder using antimuscarinic treatment. *BJU Int* 2006; **97**(3):540-6.

(59) Staskin D, Sand P, Zinner N, Dmochowski R. Once daily trospium chloride is effective and well tolerated for the treatment of overactive bladder: results from a multicenter phase III trial. *J Urol* 2007; **178**(3 Pt 1):978-83.

(60) Van KP, Kreder K, Jonas U, Zinner N, Wein A. Tolterodine once-daily: superior efficacy and tolerability in the treatment of the overactive bladder. *Urology* 2001; **57**(3):414-21.

(61) Yamaguchi O, Marui E, Kakizaki H, Itoh N, Yokota T, Okada H, *et al*. Randomized, double-blind, placebo- and propiverine-controlled trial of the once-daily antimuscarinic agent solifenacin in Japanese patients with overactive bladder. *BJU Int* 2007; **100**(3):579-87.

(62) Yamaguchi O, Nishiszawa O, Takeda M, Yoshida M, Choo MS, Gu Lee J, *et al*. Efficacy, safety and tolerability of fesoterodine in Asian patients with overactive bladder. *LUTS. Lower Urinary Tract Symptoms* 2011; **3**(1):43-50.

(63) Zinner NR, Mattiasson A, Stanton SL. Efficacy, safety, and tolerability of extended-release once-daily tolterodine treatment for overactive bladder in older versus younger patients. *J Am Geriatr Soc* 2002; **50**(5):799-807.

(64) Centre for Reviews and Dissemination. CRD's guidance for undertaking reviews in healthcare. Centre for Reviews and Dissemination 2011. [last accessed 2013 January]. Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

(65) National Institute for Health and Clinical Excellence. Single Technology Appraisal (STA): Specification for manufacturer/sponsor submission of evidence. NICE 2012. [last accessed 2013 January]. Available from: http://www.nice.org.uk/media/D6D/E2/SpecificationForManufacturerSponsorSubmissionEvidence.doc

(66) Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010; **30**(29):932-944.

(67) Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. NICE 2012. [last accessed 2013 January]. Available from: http://www.nicedsu.org.uk/TSD2%20General%20meta%20analysis.final.08.05.12.pdf

(68) Coyne KS, Matza LS, Thompson CL, Kopp ZS, Khullar V. Determining the importance of change in the overactive bladder questionnaire. *J Urol* 2006; **176**(2):627-32.

(69) Arikian SR, Casciano J, Doyle JJ, Tarride JE, Casciano RN. A pharmacoeconomic evaluation of two new products for the treatment of overactive bladder. *Manag Care Interface* 2000; **13**(2):88-94.

(70) Cardozo L, Thorpe A, Warner J, Sidhu M. The cost-effectiveness of solifenacin vs fesoterodine, oxybutynin immediate-release, propiverine, tolterodine extended-release and tolterodine immediate-release in the treatment of patients with overactive bladder in the UK National Health Service. *BJU Int* 2010; **106**(4):506-14.

(71) Hakkaart L, Verboom P, Phillips R, Al MJ. The cost utility of solifenacin in the treatment of overactive bladder. *Int Urol Nephrol* 2009; **41**(2):293-8.

(72) Kobelt G, Jonsson L, Mattiasson A. Cost-effectiveness of new treatments for overactive bladder: the example of tolterodine, a new muscarinic agent: a Markov model. *Neurourol Urodyn* 1998; **17**(6):599-611.

(73) Milsom I, Axelsen S, Kulseng-Hansen S, Mattiasson A, Nilsson CG, Wickstrom J. Cost-effectiveness analysis of solifenacin flexible dosing in patients with overactive bladder symptoms in four Nordic countries. *Acta Obstet Gynecol Scand* 2009; **88**(6):693-9.

(74) Nilsson FO, Linner L, Samuelsson E, Milsom I. Cost-effectiveness analysis of newer anticholinergic drugs for urinary incontinence vs oxybutynin and no treatment using data on persistence from the Swedish prescribed drug registry. *BJU Int* 2012; **110**(2):240-6.

(75) O'Brien BJ, Goeree R, Bernard L, Rosner A, Williamson T. Cost-Effectiveness of tolterodine for patients with urge incontinence who discontinue initial therapy with oxybutynin: a Canadian perspective. *Clin Ther* 2001; **23**(12):2038-49.

(76) Pradelli L, Iannazzo S. Solifenacin in the treatment of overactive bladder syndrome in Italian patients: pharmacoeconomic evaluation. *J Med Econ* 2009; **12**(1):25-35.

(77) Speakman M, Khullar V, Mundy A, Odeyemi I, Bolodeoku J. A cost-utility analysis of once daily solifenacin compared to tolterodine in the treatment of overactive bladder syndrome. *Curr Med Res Opin* 2008; **24**(8):2173-9.

(78) Herschorn S, Vicente C, Piwko C. Canadian cost-effectiveness analysis of solifenacin compared to oxybutynin immediate-release in patients with overactive bladder. *J Med Econ* 2010; **13**(3):508-15.

(79) Arlandis-Guzman S, Errando-Smet C, Trocio J, Arumi D, Rejas J. Cost-effectiveness analysis of antimuscarinics in the treatment of patients with overactive bladder in Spain: a decision-tree model. *BMC Urol* 2011; **11**:9.

(80) Getsios D, Caro JJ, Ishak KJ, El-Hadi W, Payne K. Canadian economic comparison of extended-release oxybutynin and immediate-release tolterodine in the treatment of overactive bladder. *Clin Ther* 2004; **26**(3):431-8.

(81) Getsios D, Caro JJ, Ishak KJ, El-Hadi W, Payne K, O'Connel M, *et al*. Oxybutynin extended release and tolterodine immediate release : a health economic comparison. *Clin Drug Investig* 2004; **24**(2):81-8.

(82) Guest JF, Abegunde D, Ruiz FJ. Cost effectiveness of controlled-release oxybutynin compared with immediate-release oxybutynin and tolterodine in the treatment of overactive bladder in the UK, France and Austria. *Clin Drug Investig* 2004; **24**(6):305-21.

(83) Hughes DA, Dubois D. Cost-effectiveness analysis of extended-release formulations of oxybutynin and tolterodine for the management of urge incontinence. *Pharmacoeconomics* 2004; **22**(16):1047-59.

(84) Ko Y, Malone DC, Armstrong EP. Pharmacoeconomic evaluation of antimuscarinic agents for the treatment of overactive bladder. *Pharmacotherapy* 2006; **26**(12):1694-702.

(85) Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, *et al*. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004; **8**(36):iii-xi, 1.

(86) National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. NICE 2008. [last accessed 2013 January]. Available from: http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf

(87) Compion G, Jackson J, Janes J. Reasons for switching antimuscarinic therapy: results from a European cross-sectional survey of physicians, and patients with OAB. 27th Annual Congress of the European Association of Urology, 24–28 February 2012, Paris, France. Poster number 691; 2012.

(88) Castro-Diaz D, Miranda P, Sanchez-Ballester F, Arumi D, Lizarraga I, Ebel-Bitoun C. Why OAB treatment change is so common? "IMPACTA" study to evaluate reasons for treatment change. Joint Annual Meeting of the International Continence Society (ICS) and International Urogynecological Association (IUGA), 23–27 August 2010, Toronto, Ontario, Canada.

(89) Odeyemi IA, Dakin HA, O'Donnell RA, Warner J, Jacobs A, Dasgupta P. Epidemiology, prescribing patterns and resource use associated with overactive bladder in UK primary care. *Int J Clin Pract* 2006; **60**(8):949-58.

(90) O'Hagan A, Buck CE, Daneshkhah A. *Uncertain Judgements: Eliciting Expert Probabilities*. 2006. John Wiley and Sons, Chichester, UK.

(91) Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd Edition. 2000. John Wiley and Sons, New York, USA.

(92) Van Kerrebroeck PE, Heesakkers J, Berriman S, Padmanabhan AL, Carlsson M, Guan Z. Long-term safety, tolerability and efficacy of fesoterodine treatment in subjects with overactive bladder symptoms. *Int J Clin Pract* 2010; **64**(5):584-93.

(93) Vanni T, Karnon J, Madan J, White RG, Edmunds WJ, Foss AM, *et al*. Calibrating models in economic evaluation: a seven-step approach. *Pharmacoeconomics* 2011; **29**(1):35-49.

(94) Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997; **35**(11):1095-108.

(95) Yang Y, Brazier J, Tsuchiya A, Coyne K. Estimating a preference-based single index from the Overactive Bladder Questionnaire. *Value Health* 2009; **12**(1):159-66.

(96) Tincello D, Sculpher M, Tunn R, Quail D, van der Vaart H, Falconer C, *et al*. Patient characteristics impacting health state index scores, measured by the EQ-5D of females with stress urinary incontinence symptoms. *Value Health* 2010; **13**(1):112-8.

(97) Verheggen BG, Lee R, Lieuw On MM, Treur MJ, Botteman MF, Kaplan SA, *et al*. Estimating the quality-of-life impact and cost-effectiveness of alpha-blocker and anti-muscarinic combination treatment in men with lower urinary tract symptoms related to benign prostatic hyperplasia and overactive bladder. *J Med Econ* 2012; **15**(3):586-600.

(98) Haywood KL, Garratt AM, Lall R, Smith JF, Lamb SE. EuroQol EQ-5D and condition-specific measures of health outcome in women with urinary incontinence: reliability, validity and responsiveness. *Qual Life Res* 2008; **17**(3):475-83.

(99) Currie CJ, McEwan P, Poole CD, Odeyemi IA, Datta SN, Morgan CL. The impact of the overactive bladder on health-related utility and quality of life. *BJU Int* 2006; **97**(6):1267-72.

(100) Harvie H, Shea J, Saks E, Schwartz JS, Arya L. Utility preference scores for urge and stress urinary incontinence. Neurourology and Urodynamics. Joint Annual Meeting of the International Continence Society (ICS) and International Urogynecological Association (IUGA). 23–27 August 2010, Toronto, Ontario, Canada.

(101) Patterson D, Geisler BP, Morse A. Determining health-related quality of life and health state utility values of urinary incontinence in women. *Female Pelvic Med Reconstr Surg* 2011; **17**(6):305-7.

(102) Kobelt G, Borgstrom F, Mattiasson A. Productivity, vitality and utility in a group of healthy professionally active individuals with nocturia. *BJU Int* 2003; **91**(3):190-5.

(103) Sut HK, Kaplan PB, Sut N, Tekbas S. The assessment of quality of life in female Turkish patients with overactive bladder. *Int J Nurs Pract* 2012; **18**(1):20-7.

(104) Monz B, Chartier-Kastler E, Hampel C, Samsioe G, Hunskaar S, Espuna-Pons M, *et al*. Patient characteristics associated with quality of life in European women seeking treatment for urinary incontinence: results from PURE. *Eur Urol* 2007; **51**(4):1073-81.

(105) Altman D, Granath F, Mattiasson A, Falconer C. Anticholinergic drug use for overactive bladder in Sweden: a nationwide pharmacoepidemiological study. *Int Urogynecol J Pelvic Floor Dysfunct* 2009; **20**(11):1285-91.

(106) Jumadilova Z, Varadharajan S, Girase P, Ollendorf DA. Retrospective evaluation of outcomes in patients with overactive bladder receiving tolterodine versus oxybutynin. *Am J Health Syst Pharm* 2006; **63**(23):2357-64.

(107) Nitz NM, Jumadilova Z, Darkow T, Frytak JR, Bavendam T. Medical costs after initiation of drug treatment for overactive bladder: effects of selection bias on cost estimates. *Am J Manag Care* 2005; **11**(4 Suppl):S130–S139.

(108) Noe L, Becker R, Williamson T, Chen D. A pharmacoeconomic model comparing two long-acting treatments for overactive bladder. *J Manag Care Pharm* 2002; **8**(5):343-52.

(109) Perfetto EM, Subedi P, Jumadilova Z. Treatment of overactive bladder: a model comparing extended-release formulations of tolterodine and oxybutynin. *Am J Manag Care* 2005; **11**(4 Suppl):S150–S157.

(110) Varadharajan S, Jumadilova Z, Girase P, Ollendorf DA. Economic impact of extended-release tolterodine versus immediate- and extended-release oxybutynin among commercially insured persons with overactive bladder. *Am J Manag Care* 2005; **11**(4 Suppl):S140–S149.

(111) Zinner N, Noe L, Rasouliyan L, Marshall T, Seifeldin R. Impact of solifenacin on resource utilization, work productivity and health utility in overactive bladder patients switching from tolterodine ER. *Curr Med Res Opin* 2008; **24**(6):1583-91.

(112) British National Formulary. BNF 63. BNF 2012. [last accessed 2013 January]. Available from: http://www.bnf.org/bnf/index.htm

(113) Curtis L. Unit Costs of Health and Social Care 2011. 2011.

(114) AgeUK incontinence. AgeUK incontinence 2013. [last accessed 2013 January]. Available from: https://www.ageukincontinence.co.uk/

(115) Nottingham Urology Group. Nottingham Urology Group 2013. [last accessed 2013 January]. Available from: http://www.nottinghamurologygroup.co.uk/treatments/bladder-botox-injections

(116) British National Formulary. BNF 64. BNF 2013. [last accessed 2013 January]. Available from: http://www.bnf.org/bnf/index.htm

(117) Peasgood T, Herrmann K, Kanis JA, Brazier JE. An updated systematic review of Health State Utility Values for osteoporosis related conditions. *Osteoporos Int* 2009; **20**(6):853-68.

(118) National Institute for Health and Clinical Excellence. CG90: Depression: the treatment and management of depression in adults (updated edition). NICE 2013. [last accessed 2013 January]. Available from: http://guidance.nice.org.uk/CG90

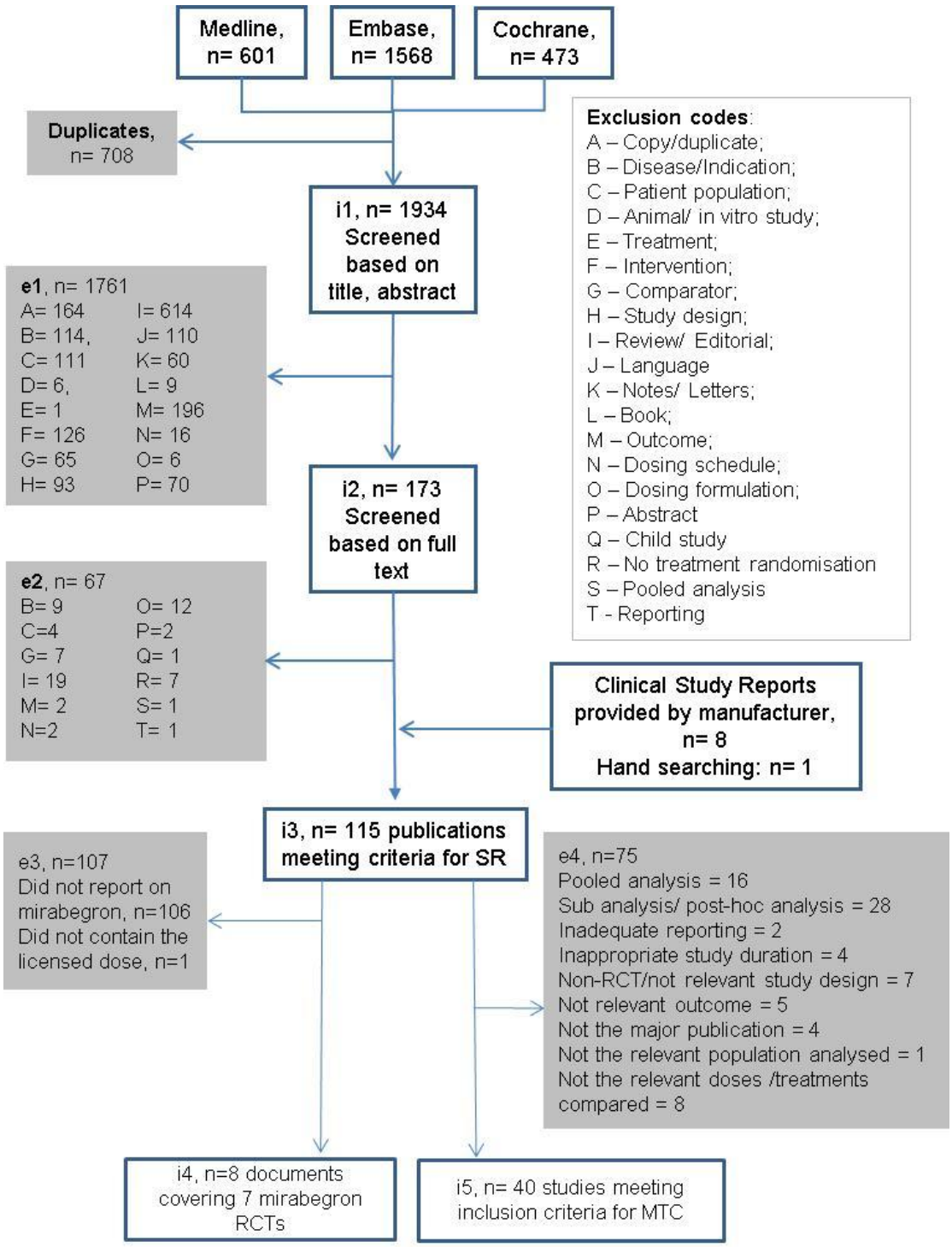
(119) Barry HC, Ebell MH, Hickner J. Evaluation of suspected urinary tract infection in ambulatory women: a cost-utility analysis of office-based strategies. *J Fam Pract* 1997; **44**(1):49-60.

(120) Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, *et al*. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices – Modeling Studies. *Value Health* 2003; **6**(1):9-17.

9 APPENDICES

Appendix 1 Flow diagram and inclusion/exclusion criteria for RCT and non-RCT search

Schematic for the systematic review of RCT clinical evidence (reproduced from MS; Figure 2, pg 39)

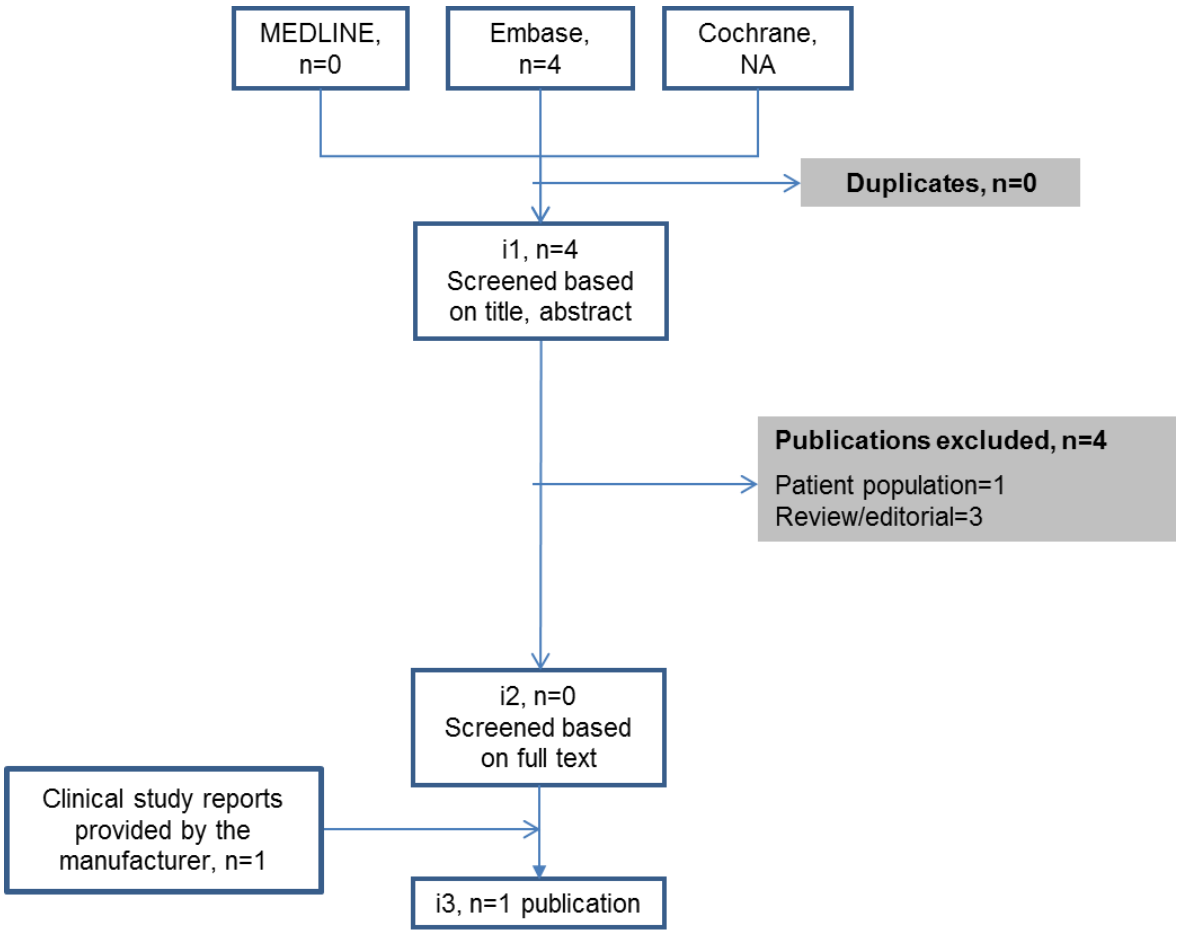


Abbreviations used in figure: MTC, mixed treatment comparison; SR, systematic review.

Inclusion/ exclusion criteria for RCT search (reproduced from MS; Section 10.2.6, pg 286)

	Description	Justification
Inclusion criteria		
Population	Adults with symptoms of OAB	As specified by final scope
Interventions	Mirabegron Oxybutynin (including modified-release preparations)	As specified by final scope
Outcomes	Symptoms of urgency Urinary frequency Frequency of urge urinary incontinence Nocturia Adverse effects of treatment	As specified by final scope
Study design	Prospective RCTs	Non-RCT studies were identified through a separate search.
Language restrictions	Non-English publications without an English abstract to be excluded at first pass stage. English abstracts of non-English publications, to be reviewed to assess eligibility.	
Exclusion criteria		
Population	Patients <18 years of age. Patients with LUTS	
Interventions	Studies not investigating mirabegron or a relevant comparator	
Outcomes	Studies not reporting the outcomes listed in the scope.	
Study design	Non-randomised controlled studies. Observational studies	Non-RCT studies were identified through a separate search.
Language restrictions	Non-English publications	
Abbreviations used in table: RCT, randomised controlled trial; OAB, overactive bladder; LUTS, lower urinary tract symptoms		

Schematic for the systematic review of non-RCT evidence for mirabegron (reproduced from MS; Figure 3, pg 144)



Inclusion/ exclusion criteria for non-RCT search (reproduced from MS; Section 10.6.6, pg 302)

	Description	Justification
Inclusion criteria		
Population	Adults with symptoms of OAB	As outlined in draft scope
Interventions	Mirabegron	As outlined in draft scope
Outcomes	Symptoms of urgency Urinary frequency Frequency of urge urinary incontinence Nocturia Adverse effects of treatment	As outlined in draft scope
Study design	Prospective observational studies	RCT studies were identified through a separate search
Language restrictions	English publications or non-English publications with an English abstract	
Exclusion criteria		
Population	Patients <18 years of age. Patients with LUTS	
Interventions	Studies not investigating Mirabegron or a relevant comparator	
Outcomes	Studies not reporting the outcomes listed in the scope.	
Study design	Randomised controlled studies Observational studies with a retrospective design.	RCT studies were identified through a separate search.
Language restrictions	Non-English publications	
Abbreviations used in table: LUTS, lower urinary tract symptoms; OAB, overactive bladder; RTC, randomised controlled trial.		

Appendix 2 Patient baseline characteristics in included RCTs (reproduced from MS; Table 10, pg 48)

178-CL-046 (SCORPIO)	**Placebo N=480**	**Mirabegron**		**Tolterodine SR 4mg N=475**	**Total N=1,906**
		50 mg N=473	**100 mg N=478**		
Sex, n (%)					
Male	134 (27.9)	133 (28.1)	138 (28.9)	129 (27.2)	534 (28.0)
Female	346 (72.1)	340 (71.9)	340 (71.1)	346 (72.8)	1372 (72.0)
Age in years,					
Mean (SD)	59.3 (12.15)	59.2 (12.15)	58.9 (12.69)	59.1 (12.75)	59.1 (12.43)
Age group in years, n (%)					
<65	302 (62.9)	302 (63.8)	306 (64.0)	291 (61.3)	1201 (63.0)
≥65	178 (37.1)	171 (36.2)	172 (36.0)	184 (38.7)	705 (37.0)
<75	437 (91.0)	430 (90.9)	435 (91.0)	442 (93.1)	1744 (91.5)
≥75	43 (9.0)	43 (9.1)	43 (9.0)	33 (6.9)	162 (8.5)
Race, n (%)					
White	477 (99.4)	468 (98.9)	474 (99.2)	472 (99.4)	1891 (99.2)
Black or African American	2 (0.4)	1 (0.2)	1 (0.2)	2 (0.4)	6 (0.3)
Asian	0	2 (0.4)	2 (0.4)	1 (0.2)	5 (0.3)
Other	1 (0.2)	2 (0.4)	1 (0.2)	0	4 (0.2)
BMI in Kg/m^2	*n=480*	*n=473*	*n=477*	*n=475*	*n=1,905*
Mean (SD)	27.8 (4.97)	27.5 (4.90)	28.0 (4.87)	27.9 (4.97)	27.8 (4.93)
Geographical region†, n (%)					
Eastern Europe	221 (46.0)	210 (44.4)	221 (46.3)	221 (46.5)	873 (45.8)
Western Europe†	259 (54.0)	263 (55.6)	257 (53.8)	254 (53.5)	1033 (54.2)

178-CL-047 (ARIES)	**Placebo N=433**	**Mirabegron**		**Total N=1,270**
		50 mg N=425	**100 mg N=412**	
Sex, n (%)				
Male	101 (23.3)	116 (27.3)	103 (25.0)	320 (25.2)
Female	332 (76.7)	309 (72.7)	309 (75.0)	950 (74.8)
Age in years				
Mean (SD)	60.1 (13.74)	59.6 (13.34)	60.8 (13.02)	NR
Age group in years, n (%)				
<65	261 (60.3)	261 (61.4)	244 (59.2)	766 (60.3)
≥65	172 (39.7)	164 (38.6)	168 (40.8)	504 (39.7)
<75	366 (84.5)	367 (86.4)	345 (83.7)	1,078 (84.9)
≥75	67 (15.5)	58 (13.6)	67 (16.3)	192 (15.0)
Race, n (%)				
White	378 (87.3)	378 (88.9)	364 (88.4)	1,120 (88.2)
Black or African American	44 (10.2)	29 (6.8)	35 (8.5)	108 (8.5)
Asian	6 (1.4)	11 (2.6)	6 (1.5)	23 (1.8)
Other	5 (1.2)	7 (1.6)	7 (1.7)	19 (1.5)

Ethnicity, n (%)				
Hispanic/Latino	23 (5.3)	22 (5.2)	31 (7.5)	76 (5.9)
Non-Hispanic/non-Latino	410 (94.7)	403 (94.8)	381 (92.5)	1,194 (94.0)
BMI in Kg/m^2	*n=432*	*n=425*	*n=412*	
Mean (SD)	30.4 (7.43)	30.0 (6.59)	30.3 (7.09)	NR
Geographical region, n (%)				
Northeastern US	75 (17.3)	72 (16.9)	77 (18.7)	224 (17.6)
Midwestern US	57 (13.2)	56 (13.2)	48 (11.7)	161 (12.7)
Southern US	150 (34.6)	140 (32.9)	139 (33.7)	429 (33.8)
Western US	110 (25.4)	113 (26.6)	106 (25.7)	329 (25.9)
Canada	41 (9.5)	44 (10.4)	42 (10.2)	127 (10.0)
178-CL-074 (CAPRICORN)	**Placebo N=415**	**Mirabegron**		**Total N=1,251**
		25 mg N=410	**50 mg N=426**	
Sex, n (%)				
Male	127 (30.6)	134 (32.7)	133 (31.2)	394 (31.5)
Female	288 (69.4)	276 (67.3)	293 (68.8)	857 (68.5)
Age in years				
Mean (SD)	58.2 (13.83)	58.8 (12.68)	60.4 (12.26)	NR
Age group in years, n (%)				
<65	261 (62.9)	263 (64.1)	262 (61.5)	786 (62.8)
≥65	154 (37.1)	147 (35.9)	164 (38.5)	465 (37.2)
<75	371 (89.4)	378 (92.2)	378 (88.7)	1,127 (90.1)
≥75	44 (10.6)	32 (7.8)	48 (11.3)	124 (9.9)
Race, n (%)				
White	372 (89.6)	373 (91.0)	389 (91.3)	1,134 (90.6)
Black or African American	34 (8.2)	31 (7.6)	31 (7.3)	96 (7.7)
Asian	7 (1.7)	5 (1.2)	4 (0.9)	16 (1.3)
Other	2 (0.5)	1 (0.2)	2 (0.5)	5 (0.4)
Ethnicity, n (%)				
Hispanic/ Latino	21 (5.1)	22 (5.4)	21 (4.9)	64 (5.1)
Non-Hispanic/ non-Latino	394 (94.9)	388 (94.6)	405 (95.1)	1,187 (94.9)
BMI in Kg/m^2	*n=415*	*n=410*	*n=426*	
Mean (SD)	29.1 (6.27)	29.6 (6.32)	29.5 (6.52)	NR
Geographical region, n (%)				
Eastern Europe	73 (17.6)	75 (18.3)	74 (17.4)	222 (17.7)
Western Europe	123 (29.6)	117 (28.5)	119 (27.9)	359 (28.7)
Northeastern US	39 (9.4)	38 (9.3)	41 (9.6)	118 (9.4)
Midwestern US	22 (5.3)	24 (5.9)	22 (5.2)	68 (5.4)
Southern US	67 (16.1)	68 (16.6)	74 (17.4)	209 (16.7)
Western US	60 (14.5)	64 (15.6)	65 (15.3)	189 (15.1)
Canada	31 (7.5)	24 (5.9)	31 (7.3)	86 (6.9)

[†]For the purposes of this summary, Australia was included within the Western Europe category.

Abbreviations in table: BMI, body mass index; cm, centimetre; Kg, kilogram; mg, milligram; NR, not reported; SD, standard deviation; SR, slow-release.

Patient demographics of participants across randomised groups in TAURUS (reproduced from MS; Table 59, pg 155)

Baseline characteristics	**Randomised groups**			**Total N=2,444**
	Mirabegron 50 mg N=812	**Mirabegron 100 mg N=820**	**Tolterodine ER 4 mg N=812**	
Sex, n (%)				
Male	210 (25.9%)	212 (25.9%)	212 (26.1%)	634 (25.9%)
Female	602 (74.1%)	608 (74.1%)	600 (73.9%)	1810 (74.1%)
Age in years, mean (SD)	59.2 (12.56)	60.1 (11.92)	59.6 (12.47)	59.6 (12.32)
Age group in years, n (%)				
<65	523 (64.4%)	504 (61.5%)	509 (62.7%)	1536 (62.8%)
≥65	289 (35.6%)	316 (38.5%)	303 (37.3%)	908 (37.2%)
<75	737 (90.8%)	739 (90.1%)	729 (89.8%)	2205 (90.2%)
≥75	75 (9.2%)	81 (9.9%)	83 (10.2%)	239 (9.8%)
Race, n (%)				
White	778 (95.8%)	774 (94.4%)	780 (96.1%)	2332 (95.4%)
Black or African American	22 (2.7%)	30 (3.7%)	20 (2.5%)	72 (2.9%)
Asian	8 (1.0%)	8 (1.0%)	5 (0.6%)	21 (0.9%)
Other	4 (0.5%)	8 (1.0%)	7 (0.9%)	19 (0.8%)
Ethnicity, n (%)				
Hispanic/Latino	23 (2.8%)	20 (2.4%)	32 (4.0%)	75 (3.1%)
Non-Hispanic/Non-Latino	789 (97.2%)	800 (97.6%)	778 (96.0%)	2367 (96.9%)
BMI in Kg/m^2	*N=811*	*N=819*	*N=809*	*N=2,439*
Mean (SD)	29.0 (6.29)	28.8 (5.99)	28.5 (5.69)	28.8 (5.99)
BMI category in Kg/m^2, n (%)				
<25	229 (28.2%)	231 (28.2%)	224 (27.7%)	684 (28.0%)
25 to <30	294 (36.3%)	319 (38.9%)	328 (40.5%)	941 (38.6%)
≥30	288 (35.5%)	269 (32.8%)	257 (31.8%)	814 (33.4%)
Geographical region, n (%)				
Eastern Europe	260 (32.0%)	270 (32.9%)	258 (31.8%)	788 (32.2%)
Western Europe	257 (31.7%)	242 (29.5%)	262 (32.3%)	761 (31.1%)
Southern hemisphere	34 (4.2%)	39 (4.8%)	37 (4.6%)	110 (4.5%)
Canada	44 (5.4%)	47 (5.7%)	45 (5.5%)	136 (5.6%)
Northeastern US	55 (6.8%)	56 (6.8%)	53 (6.5%)	164 (6.7%)
Midwestern US	29 (3.6%)	33 (4.0%)	33 (4.1%)	95 (3.9%)
Southern US	67 (8.3%)	68 (8.3%)	57 (7.0%)	192 (7.9%)
Western US	66 (8.1%)	65 (7.9%)	67 (8.3%)	198 (8.1%)

Abbreviations used in table: BMI, body mass index; ER, extended-release; kg, kilogram; m, metre; mg, milligram; SD, standard deviation; US, United States.

Patient demographics of participants across randomised groups in DRAGON (reproduced from MS; Table 132, pg 319)

178-CL-044 (DRAGON)	Placebo N=166	Mirabegron				Tolterodine SR 4mg N=85	Total N=919
		25 mg N=167	50 mg N=167	100 mg N=168	200 mg N=166		
Sex, n (%)							
Male	15 (9.0)	20 (12.0)	18 (10.8)	17 (10.1)	12 (7.2)	16 (18.8)	98 (10.7)
Female	151 (91.0)	147 (88.0)	149 (89.2)	151 (89.9)	154 (92.8)	69 (81.2)	821 (89.3)
Age in years							
Mean (SD)	57.1 (12.9)	57.2 (12.1)	56.9 (12.5)	57.1 (12.5)	58.0 (13.7)	56.6 (12.8)	57.2 (12.7)
Range	21-80	20-78	26-84	21-91	18-82	27-78	18-91
Age group in years, n (%)							
≤ 65	122 (73.5)	117 (70.1)	125 (74.9)	126 (75.0)	113 (68.1)	64 (75.3)	667 (72.6)
>65	44 (26.5)	50 (29.9)	42 (25.1)	42 (25.0)	53 (31.9)	21 (24.7)	252 (27.4)
>75	11 (6.6)	5 (3.0)	9 (5.4)	14 (8.3)	13 (7.8)	5 (5.9)	57 (6.2)
Race, n (%)							
Caucasian	166 (100)	162 (97.0)	162 (97.0)	167 (99.4)	164 (98.8)	81 (95.3)	902 (98.2)
Black	0	2 (1.2)	0	0	0	0	2 (0.2)
Asian	0	1 (0.6)	0	1 (0.6)	0	2 (2.4)	4 (0.4)
Other	0	1 (0.6)	3 (1.8)	0	0	1 (1.2)	5 (0.5)
Missing	0	1 (0.6)	2 (1.2)	0	2 (1.2)	1 (1.2)	6 (0.7)
Weight in kg							
Mean (SD)	75.1 (14.3)	75.8 (13.2)	72.9 (13.2)	73.0 (12.8)	73.7 (14.2)	73.9 (14.7)	74.1 (13.7)
Range	46-132	49-129	47-121	49-120	40-125	45-129	40-132
Height in cm							
Mean (SD)	164.5 (7.1)	165.2 (7.7)	164.7 (8.2)	164.2 (7.2)	163.2 (7.8)	165.3 (7.1)	164.5 (7.6)
Range	149-184	145-190	131-190	150-190	147-199	148-183	131-199
Abbreviations used in table: cm, centimetre; kg, kilogram; mg, milligram; SD, standard deviation; SR, slow-release.							

OAB history in participants across randomised groups in SCORPIO, ARIES, and CAPRICORN (reproduced from MS; Table 11, pg 50)

178-CL-046 (SCORPIO)	**Placebo N=480**	**Mirabegron**		**Tolterodine SR 4mg N=475**
		50 mg N=473	**100 mg N=478**	
Type of OAB[†], n (%)				
Urgency incontinence	201 (41.9)	192 (40.6)	179 (37.4)	184 (38.7)
Frequency	177 (36.9)	173 (36.6)	183 (38.3)	186 (39.2)
Mixed	102 (21.3)	108 (22.8)	116 (24.3)	105 (22.1)
Prior OAB surgery, n (%)	22 (4.6)	33 (7.0)	28 (5.9)	17 (3.6)
Previous OAB drug, n (%)	238 (49.6)	240 (50.7)	237 (49.6)	231 (48.6)
Reason for previous OAB drug discontinuation[‡], n (%)				
Insufficient effect	159 (66.8)	160 (66.7)	159 (67.1)	155 (67.1)
Poor tolerability	68 (28.6)	65 (27.1)	64 (27.0)	56 (24.2)
Duration of OAB symptoms (months)				
Mean (SD)	76.9 (92.15)	78.7 (85.68)	85.3 (95.24)	76.3 (93.40)
Median	50.5	49.9	53.4	47.2
Range	3 – 688	3 – 637	3 – 567	3 – 711

178-CL-047 (ARIES)	**Placebo N=433**	**Mirabegron**	
		50 mg N=425	**100 mg N=412**
Type of OAB[†], n (%)			
Urgency incontinence	124 (28.6)	135 (31.8)	118 (28.6)
Frequency	133 (30.7)	134 (31.5)	139 (33.7)
Mixed	176 (40.6)	156 (36.7)	155 (37.6)
Prior OAB surgery, n (%)	49 (11.3)	53 (12.5)	46 (11.2)
Previous OAB drug, n (%)	249 (57.5)	242 (56.9)	223 (54.1)
Reason for previous OAB drug discontinuation[‡], n (%)			
Insufficient effect	166 (66.7)	161 (66.5)	137 (61.4)
Poor tolerability	60 (24.1)	49 (20.2)	49 (22.0)
Duration of OAB symptoms (months)			
Mean (SD)	91.9 (108.52)	84.0 (94.61)	91.8 (108.44)
Median	52.4	51.9	52.0
Range	3 – 816	3 – 634	3 – 865

178-CL-074 (CAPRICORN)	Placebo N=415	Mirabegron	
		25 mg N=410	50 mg N=426
Type of OAB†, n (%)			
Urgency incontinence	117 (28.2)	156 (38.0)	164 (38.5)
Frequency	161 (38.8)	130 (31.7)	114 (26.8)
Mixed	137 (33.0)	124 (30.2)	148 (34.7)
Prior OAB surgery, n (%)	43 (10.4)	25 (6.1)	40 (9.4)
Previous OAB drug, n (%)	217 (52.3)	219 (53.4)	206 (48.4)
Reason for previous OAB drug discontinuation‡, n (%)			
Insufficient effect	141 (65.0)	149 (68.0)	143 (69.4)
Poor tolerability	57 (26.3)	48 (21.9)	59 (28.6)
Duration of OAB symptoms (months)			
Mean (SD)	91.4 (96.08)	97.4 (115.14)	93.7 (98.94)
Median	63.0	59.8	62.7
Range	3 - 590	3 – 759	3 – 688

†Types of OAB were defined as follows: urgency incontinence = urge incontinence only, mixed = mixed stress/urge incontinence with urge as a predominant factor, frequency = frequency/urgency without incontinence.

‡Patients could choose >1 reason for discontinuation of previous OAB drug.

Abbreviations used in table: mg, milligram; OAB, overactive bladder; SD, standard deviation; SR, slow-release.

OAB history in participants across randomised groups in TAURUS (reproduced from MS; Table 60, pg 156)

Baseline characteristics	Randomised groups		
	Mirabegron 50 mg N=812	Mirabegron 100 mg N=820	Tolterodine ER 4 mg N=812
Type of OAB[†], n (%)			
Urgency incontinence	296 (36.5)	305 (37.2)	317 (39.0)
Frequency	284 (35.0)	287 (35.0)	285 (35.1)
Mixed	232 (28.6)	228 (27.8)	210 (25.9)
Prior OAB surgery, n (%)			
Yes	75 (9.2)	70 (8.5)	68 (8.4)
No	737 (90.8)	750 (91.5)	744 (91.6)
Previous OAB drug[‡], n (%)			
Yes	446 (54.9)	419 (51.1)	447 (55.0)
No	366 (45.1)	401 (48.9)	365 (45.0)
Previous non-drug treatment for OAB[§], n (%)			
Yes	37 (4.6)	24 (2.9)	32 (3.9)
Biofeedback	0	0	0
Exercises	28 (3.4)	20 (2.4)	25 (3.1)
Electrical stimulation	1 (0.1)	0	0
Behavioural	8 (1.0)	4 (0.5)	10 (1.2)
Pessaries	3 (0.4)	1 (0.1)	0
Implants	0	0	0
Other	5 (0.6)	3 (0.4)	4 (0.5)
No	775 (95.4)	796 (97.1)	780 (96.1)
Reason for previous OAB drug discontinuation[¶], n (%)			
Insufficient effect	297 (66.6)	268 (64.0)	283 (63.3)
Poor tolerability	97 (21.7)	108 (25.8)	122 (27.3)
Duration of OAB symptoms (months)			
Mean (SD)	87.4 (96.28)	87.9 (91.52)	83.8 (87.34)
Median	55.9	56.4	55.9
Range	(3 – 653)	(3 – 692)	(3 – 642)

[†]Types of OAB were defined as follows: urgency incontinence = urge incontinence only, mixed = mixed stress/urge incontinence with urge as a predominant factor, frequency = frequency/urgency without incontinence.

[‡]'Yes' included patients who received a marketed drug with an indication for OAB. It did not include patients who only received an OAB drug as an investigational study medication in a previous clinical study.

[§]Non-drug treatment which ended ≥30 days prior to screening were not included.

[¶]Patients could chose >1 reason for discontinuation of previous OAB drug.

Abbreviations used in table: ER, extended-release; mg, milligram; OAB, overactive bladder; SD, standard deviation.

OAB history in participants across randomised groups in DRAGON (reproduced from MS; Table 133, pg 321)

178-CL-044 (DRAGON)	Placebo N=166	Mirabegron					Total N=919
		25 mg N=167	50 mg N=167	100 mg N=168	200 mg N=166	Tolterodine SR 4mg N=85	
Type of OAB, n (%)							
Urge incontinence only	74 (44.6)	79 (47.3)	67 (40.1)	67 (39.9)	63 (38.0)	38 (44.7)	388 (42.2)
Mixed incontinence (urge as predominant factor)	52 (31.3)	41 (24.6)	47 (28.1)	54 (32.1)	63 (38.0)	24 (28.2)	281 (30.6)
Without incontinence	52 (31.3)	47 (28.1)	53 (31.7)	47 (28.0)	40 (24.1)	23 (27.1)	250 (27.2)
Previous OAB drug within 1 year of study start, n (%)							
Yes, at least 1 effective	41 (24.7)	40 (24.0)	39 (23.4)	42 (25.0)	34 (20.5)	19 (22.4)	215 (23.4)
Yes, none effective	30 (18.1)	42 (25.1)	38 (22.8)	39 (23.2)	38 (22.9)	16 (18.8)	203 (22.1)
No	95 (57.2)	85 (50.9)	90 (53.9)	87 (51.8)	94 (56.6)	50 (58.8)	501 (54.5)
Duration of OAB symptoms (months)							
	N=63	*N=63*	*N=53*	*N=67*	*N=54*	*N=31*	*N=331*
Mean (SD)	54.2 (66.9)	48.0 (35.7)	45.1 (53.7)	40.6 (48.8)	43.4 (32.9)	46.5 (44.7)	46.3 (48.9)
Median	35.0	44.0	31.0	27.0	33.0	43.0	34.0
Range	6-390	3-241	6-343	6-357	4-135	3-230	3-390
Treatment other than drug, n (%)	51 (30.7)	7 (34.1)	49 (29.3)	44 (26.2)	40 (24.1)	22 (25.9)	263 (28.6)
Abbreviations used in table: mg, milligram; OAB, overactive bladder; SD, standard deviation.							

OAB-related baseline characteristics in participants across randomised groups in SCORPIO, ARIES, and CAPRICORN (reproduced from MS; Table 12, pg 52)

178-CL-046 (SCORPIO)	**Placebo N=480**	**Mirabegron**		**Tolterodine SR 4mg N=475**
		50 mg N=473	**100 mg N=478**	
Mean number of micturitions per 24 hours				
Mean (SD)	11.71 (3.138)	11.65 (2.972)	11.51 (2.703)	11.55 (2.779)
Range	5.3 – 25.0	6.7 – 25.7	6.7 – 23	6.0 – 22.7
Mean volume voided per micturition (mL)				
	n=480	*n=472*	*n=478*	*n=475*
Mean (SD)	156.7 (52.51)	161.1 (58.40)	158.2 (53.14)	158.6 (54.13)
Range	51 – 336	30 – 397	37 – 367	19 – 402
Mean number of urgency episodes (Grade 3 or 4) per 24 hours				
	n=480	*n=473*	*n=477*	*n=474*
Mean (SD)	5.76 (3.994)	5.69 (3.653)	5.94 (3.705)	5.77 (3.446)
Range	0 – 24.3	0 – 20.7	0 – 22.3	0 – 22.7
Mean level of urgency				
	n=480	*n=473*	*n=477*	*n=474*
Mean (SD)	2.37 (0.562)	2.40 (0.543)	2.45 (0.520)	2.41 (0.556)
Range	0 – 4.0	0.5 – 4.0	0.6 – 4.0	0.5 – 4.0
Mean number of nocturia episodes per 24 hours				
Mean (SD)	1.98 (1.412)	1.87 (1.293)	1.90 (1.356)	1.95 (1.412)
Range	0 – 9.7	0 – 6.3	0 – 8.0	0 – 8.3
178-CL-047 (ARIES)	**Placebo N=433**	**Mirabegron**		
		50 mg N=425	**100 mg N=412**	
Mean number of micturitions per 24 hours				
Mean (SD)	11.51 (3.269)	11.80 (3.458)	11.66 (3.389)	
Range	3.7 – 40.3	5.7 – 33.3	7.3 – 35.3	
Mean volume voided per micturition (mL)				
Mean (SD)	157.5 (58.68)	156.0 (58.69)	157.6 (60.19)	
Range	40 – 358	28 – 335	38 – 363	
Mean number of urgency episodes (Grade 3 or 4) per 24 hours				
Mean (SD)	5.61 (3.236)	5.88 (3.844)	5.95 (3.608)	
Range	0.7 – 16.5	0.0 – 33.3	0.6 – 20.7	
Mean level of urgency				
Mean (SD)	2.45 (0.537)	2.45 (0.534)	2.46 (0.544)	
Range	0.7 – 4.0	0.3 – 4.0	0.9 – 4.0	
Mean number of nocturia episodes per 24 hours				
Mean (SD)	1.93 (1.633)	1.90 (1.613)	2.04 (1.689)	
Range	0.0 – 13.0	0.0 – 12.3	0.0 – 11.3	

178-CL-074 (CAPRICORN)	Placebo N=415	Mirabegron	
		25 mg N=410	50 mg N=426
Mean number of micturitions per 24 hours			
Mean (SD)	11.48 (2.896)	11.68 (3.099)	11.66 (3.221)
Range	7.3 – 26.3	6.3 – 23.3	7.7 – 37.3
Mean volume voided per micturition (mL)			
Mean (SD)	164.0 (56.87)	165.2 (57.59)	159.3 (52.25)
Range	48 – 356	33 – 349	27 – 357
Mean number of urgency episodes (Grade 3 or 4) per 24 hours			
Mean (SD)	5.40 (3.310)	5.57 (3.617)	5.80 (3.567)
Range	0.3 – 26.0	1.0 – 21.7	1.0 – 18.7
Mean level of urgency			
Mean (SD)	2.36 (0.551)	2.37 (0.563)	2.41 (0.561)
Range	0.8 – 4.0	0.4 – 4.0	0.7 – 4.0
Mean number of nocturia episodes per 24 hours			
Mean (SD)	1.78 (1.274)	1.96 (1.516)	2.03 (1.537)
Range	0.0 – 6.7	0.0 – 9.0	0.0 – 12.0
Mean number of pads used per 24 hours			
Mean (SD)	0.92 (1.804)	0.77 (1.486)	0.83 (1.706)
Range	0.0 – 12.3	0.0 – 11.0	0.0 – 12.0
Abbreviations used in table: mg, milligram; mL, millilitre; OAB, overactive bladder; SD, standard deviation; SR, slow-release.			

OAB-related baseline characteristics in participants across randomised groups in TAURUS (reproduced from MS; Table 61, pg 157)

Baseline characteristics	Randomised groups		
	Mirabegron 50 mg N=812	Mirabegron 100 mg N=820	Tolterodine ER 4 mg N=812
Mean number of micturitions per 24 hours			
Mean (SD)	11.12 (2.809)	11.16 (2.917)	10.94 (2.668)
Range	6.3 – 31.7	5.7 – 29.3	4.3 – 26.3
Mean volume voided per micturition (mL)			
Mean (SD)	160.4 (58.80)	164.6 (58.62)	160.8 (56.98)
Range	28 – 346	17 – 350	36 – 354
Mean number of urgency episodes (Grade 3 or 4) per 24 hours			
Mean (SD)	5.66 (3.601)	5.61 (3.722)	5.44 (3.453)
Range	0.0 – 22.7	0.0 – 26.7	0.7 – 20.7
Mean level of urgency			
Mean (SD)	2.45 (0.544)	2.44 (0.525)	2.43 (0.519)
Range	0.3 – 4.0	0.4 – 4.0	0.5 – 4.0
Mean number of nocturia episodes per 24 hours			
Mean (SD)	1.83 (1.361)	1.85 (1.404)	1.77 (1.388)
Range	0.0 – 8.7	0.0 – 9.7	0.0 – 11.3
Mean number of pads used per 24 hours			
Mean (SD)	1.06 (1.872)	0.98 (1.769)	0.98 (1.759)
Range	0.0 – 12.7	0.0 – 12.7	0.0 – 12.7
Abbreviations: ER, extended-release; mg, milligram; mL, millilitre; SD, standard deviation.			

OAB-related baseline characteristics in participants across randomised groups in DRAGON (reproduced from MS; Table 134, pg 321)

178-CL-044 (DRAGON), mean (SD)	Placebo N=166	Mirabegron				Tolterodine SR 4mg N=85
		25 mg N=167	50 mg N=167	100 mg N=168	200 mg N=166	
Mean number of micturitions per 24 hours	11.67 (3.39)	11.87 (2.88)	11.85 (3.30)	11.81 (3.51)	11.34 (2.41)	12.31 (3.68)
Mean volume voided per micturition (mL)	161.38 (53.87)	160.83 (55.04)	153.62 (49.39)	152.67 (55.26)	156.10 (50.17)	157.00 (64.40)
Mean number of urgency episodes (Grade 3 or 4) per 24 hours	5.75 (3.95)	5.77 (4.12)	5.94 (3.87)	5.92 (3.89)	5.75 (3.57)	5.83 (3.72)
Mean level of urgency	2.36 (0.58)	2.32 (0.59)	2.39 (0.55)	2.38 (0.55)	2.34 (0.54)	2.34 (0.56)
Mean number of nocturia episodes per 24 hours	1.77 (1.12)	1.76 (1.17)	1.70 (1.02)	1.82 (1.08)	1.78 (1.17)	1.78 (0.98)
Mean number of incontinence episodes per 24 hours	2.45 (2.35)	2.92 (3.23)	2.41 (2.30)	2.49 (2.48)	2.47 (2.23)	2.85 (2.76)
Mean number of urgency incontinence episodes per 24 hours	2.21 (2.00)	2.88 (3.09)	2.21 (2.17)	2.39 (2.46)	2.36 (2.02)	2.63 (2.53)
Abbreviations: SR, slow-release; mg, milligram; mL, millilitre; SD, standard deviation.						

Appendix 3 Description of scales used to assess QoL and treatment satisfaction

European quality of life – five dimensions (EQ-5D)

The EQ-5D is a two-page questionnaire divided into the EQ-5D descriptive system and the EQ-5D VAS. The descriptive system consists of five dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression). Each dimension has three levels. The EQ-5D VAS is a 10 cm vertical visual analogue scale with the endpoints labelled as 'best imaginable health state' corresponding to a score of 100 and 'worst imaginable health state' corresponding to a score of 0.

Patient perception of bladder condition (PPBC)

PPBC is a six-point Likert scale on which a score of 1 indicates "no problems at all" and a score of 6 indicates "many severe problems". Negative change indicates improvement.

Patient perception of treatment benefit

Patient perception of treatment benefit is a three-point response to treatment scale.

Overactive bladder questionnaire (OABq)

OAB-q consists of 33 items that include coping, concern, sleep, social interaction and a symptom bother scale with eight symptoms. Higher scores on the HRQoL subscales and total score indicate a better QoL, and a positive change in the HRQoL scores indicates improvement. Scores for the symptom bother scale range from 0 to 100, with a score of 100 indicating worst severity. A negative change in symptom bother indicates improvement.

Treatment satisfaction – visual analogue scale (TS-VAS)

In the TS-VAS, patients are asked to put a vertical mark on a line that runs from 0 (No, not at all) to 10 (Yes, completely).

Work productivity and activity impairment: specific health problem (WPAI: SHP)

WPAI: SHP consists of six questions covering employment status, hours absent from work due to a specific health problem, hours absent from work due to other reasons, hours actually worked, impact of the health problem on productivity while working, impact of the health problem on productivity while doing regular daily activities other than work. A negative change from baseline indicates improvement.

Appendix 4. Summary of methodology of non-RCT 178-CL-051 (reproduced from MS; Table 54, pg 145)

Study no. (acronym)	**178-CL-051**
Study objective	Evaluation of safety and efficacy of long-term (52 weeks) treatment with 50 mg mirabegron in patients with OAB
Location	26 sites in Japan
Design	Phase 3 open-label, uncontrolled study of 204 enrolled patients

Inclusion criteria	**At enrolment**
	• Male or female outpatient aged ≥20 years at time of informed consent
	• Symptoms of OAB for ≥24 weeks prior to run-in period
	• Capable of walking to the lavatory unaided and measuring urine volume by him/herself
	At randomisation
	• ≥8 micturitions per 24 hours and:
	  ○ ≥1 urgency episode per 24 hours, and/or
	  ○ ≥1 urge incontinence episode per 24 hours

Exclusion criteria	**At enrolment** • No experience of urge incontinence before informed consent • Definite diagnosis of stress incontinence • Symptoms suggesting OAB is transient (e.g. drug-induced or psychogenic) • Complications of UTI • Complications/history of bladder or prostatic tumours • Clinically significant lower urinary tract obstructive disease • Treatment with medication for lower urinary tract obstructive disease within 4 weeks of run-in period • Indwelling catheter or practicing intermittent self-catheterisation • Radiotherapy affecting urinary tract function or thermotherapy for BPH • Surgical therapy potentially affecting urinary tract function within 24 weeks of run-in period • History of acute cerebrovascular disorder, serious cardiovascular disorder or clinically significant orthostatic hypotension within 24 weeks of run-in period • Uncontrolled hypertension (sitting SBP ≥180 mmHg or DBP ≥110 mmHg at Visit 1) • Pulse rate ≥110 or <50 bpm measured at Visit 1 • Clinically significant serious cardiac, hepatic, renal, immunological, pulmonary disease or malignant tumours • Hypersensitivity to β-receptor agonists • Treatment with other investigational medicines within 12 weeks prior to informed consent. • Previous treatment with mirabegron • Pregnancy/ breast feeding **At randomisation** • Polyuria >3000 mL/day • Confirmed PVR ≥100 mL or clinically significant lower urinary tract obstructive disease • Uncontrolled hypertension (sitting SBP ≥180 mmHg or DBP ≥110 mmHg at Visit 2) • Pulse rate ≥110 or <50 measured at Visit 2 • Abnormal electrocardiogram • AST/ALT 2.5xULN

	• Blood creatinine level ≥2.0 mg/dL
Duration of study	• 1-week run-in • 52 weeks on treatment
Method of randomisation	N/A: all patients took 50 mg mirabegron
Method of blinding (care provider, patient and outcome assessor)	N/A: open-label study
Interventions, N randomised	• 50 mg mirabegron, N=204 (Dose escalation to 100 mg from Week 8 where necessary.)
Comparators, N randomised	None
Permitted concomitant medications	**Antidepressants** • imipramine (Imidol®, Tofranil®) • amitriptyline (Tryptanol®, etc.) • nortriptyline (Noritren®) • clomipramine (Anafranil®) • dosulepin (Prothiaden®) • maprotiline (Ludiomil®, etc.) • milnacipran (Toledomin®) **Class I antiarrhythmic agents** • pirmenol (Pimenol®) • cibenzoline (Cibenol®) • disopyramide (Rythmodan®, Norpace®, etc.) **Antihistamines** • diphenylpyraline (Hy-stamin®, etc.) • cyproheptadine (Periactin®, etc.) • triprolidine (Venen®, etc.) • promethazine (Hiberna®, Pyrethia®, etc.) • homochlorcyclizine (Homoclomin®, etc.) • alimemazine (Alimezine®) • diphenhydramine (Restamin®, etc.) • clemastine (Tavegyl®, etc.) • chlorpheniramine (Polaramine®, etc.) • mequitazine (Zesulan®, etc.) **Anti-Parkinson drugs** • piroheptine (Trimol®) • mazaticol (Pentona®) • metixene (Methixart®, etc.) • profenamine (Parkin®) **Parasympathetic inhibitors/blockers (including drugs containing narcotics)** • tiquizium (Thiaton®, etc.) • piperidolate (Crapinon®, etc.) • propantheline (Pro-Banthine®, etc.) • timepidium (Sesden®, etc.) • methylscopolammonium (Daipin®, etc.)

	• methyloctatropine (Valpin®) • scopolamine butylbromide (Buscopan®, etc.) • pipethanate ethobromide (Panpurol®) • prifinium (Padrin®) • butropium (Butropan®, etc.) • tiemonium (Visceralgine®) • oxapium (Esperan®, etc.) • valethamate (Shinmetane®) • trospium (Spasmex®) • dicyclomine (Resporimin®, etc.) • scopolia extract (Scopolia Extract®, etc.) • atropine (Atropine Sulfate®, etc.) • ipratropium (Atrovent®) • oxitropium (Tersigan®) • tiotropium (Spiriva®) • pridinol (Konlax®, etc.) • mepenzolate (Trancolon®, etc.) **sympathomimetic agents** • amezinium (Amegyl®, etc.) **α- and β-stimulants** • etilefrine (Effortil®, etc.) • methylephedrine (Methy-F®, etc.) • epinephrine (Bosmin®, etc.) • ephedrine (Ephedrine Hydrochloride®, etc.) • norepinephrine (Nor-Adrenalin) **β-stimulants** • isoproterenol (Proternol®, etc.) • methoxyphenamine (Asthma®, etc.) • trimetoquinol (Inolin®, etc.) • salbutamol (Venetlin®, etc.) • terbutaline (Bricanyl®) • tulobuterol (Hokunalin®, etc.) • procaterol (Meptin®, etc.) • fenoterol (Berotec®, etc.) • formoterol (Atock Dry®) • mabuterol (Broncholin®) • salmeterol (Serevent®) • dobutamine (Dobutrex®, etc.) • docarpamine (Tanadopa®) • denopamine (Kalgut®, etc.) • ritodrine (Utemec®, etc.) • isoxsuprine (Duvadilan®, etc.)
Disallowed concomitant medications	**Anticholinergics** • oxybutynin (Pollakisu®, etc.) • flavoxate (Bladderon®, etc.) • propiverine (BUP-4®, etc.)

	• solifenacin (Vesicare®) • tolterodine (Detrusitol®) • imidafenacin (Uritos®, Staybla®, etc.) **β-2 stimulant** • clenbuterol (Spiropent®) **Therapeutics for prostatic hypertrophy** • allylestrenol (Perselin®, etc.) • oxendolone (Prostetin®) • gestonorone (Depostat®) • chlormadinone (Prostal®, etc.) • tamsulosin (Harnal®, etc.) • terazosin (Hytracin®, Vasomet®, etc.) • prazosin (Minipress®, etc.) • silodosin (Urief®) • urapidil (Ebrantil®) • naftopidil (Flivas®, Avishot®) • mixtures (Eviprostat®, Paraprost®, etc.) • pollen extract containing drug (Cernilton®, etc.) **Substrates of CYP2D6 with a narrow therapeutic index** • flecainide (Tambocor®) • propafenone (Pronon®)
Discontinuation of study therapy	• Patient request/withdrawn consent • SAE/AE requiring a change in the protocol • Decision by investigator that termination was necessary • Insufficient efficacy • Patient lost to follow-up
Assessments	Visits at Weeks 8, 16, 28, 40, 52

Primary outcomes (including scoring methods and timings of assessments) Secondary outcomes (including scoring methods and timings of assessments)	**Efficacy endpoints** CFB, based on a 3-day micturition diary, to endpoint in: • micturitions per 24 hours • urgency episodes per 24 hours • incontinence episodes per 24 hours • urge incontinence episodes per 24 hours • nocturia episodes • QoL domain scores on the King's Health questionnaire **Safety endpoints** • Adverse events
Duration of follow-up	No follow up (with the exception of safety reporting) after the week52 final visit
Analysis populations	• FAS • QOL analysis set (patients in the FAS for whom ≥1 domain score could be calculated and who had taken the study drug for ≥14 days) • SAS
Statistical methods	• Minimum target sample size of 150 patients selected to allow for drop-outs and to ensure ≥100 patients received treatment for ≥1 year • Handling of missing data: If multiple observations were obtained within the same visit window for a patient, the value obtained closest to the target date was used. If deviations from the scheduled date were the same, the value obtained on the later date was used
Abbreviations in table: AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BPH, benign prostatic hyperplasia; bpm, beats per minute; CFB, change from baseline; DBP, diastolic blood pressure; FAS, full analysis set; mg, milligram; mm Hg; millimetres of mercury; N/A, not applicable; OAB, overactive bladder; PVR, post-void residual volume; QoL, quality of life; QTc, corrected QT interval; SAE, serious adverse event; SAS, safety analysis set; SBP, systolic blood pressure; ULN, upper limit of normal; UTI, urinary tract infection.	

Appendix 5 Quality assessment of included RCTs and non-RCTs

DRAGON, 178-CL-044			
Question	**How was the question addressed in the study? (description in MS[a])**	**Manufacturer's assessment**	**ERG's comment**
Was randomisation carried out appropriately?	Randomisation sequences prepared by the Contract Research Organisation IFE Europe GmbH, Essen, Germany, under responsibility of Biometrics Department of APEB.	Yes	Not clear
Was the concealment of treatment allocation adequate?	IVRS used to control the randomisation and clinical supply distribution. IVRS assigned medication numbers to patients fulfilling all selection criteria at Visit 2. Study medication packed in blister cards. Each card contained medication (mirabegron, tolterodine or placebo) for 1 week treatment for 1 patient.	Yes	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	Treatment groups well balanced for all demographic characteristics. No relevant differences between treatment groups with respect to medication and alcohol history at study entry. Proportion of patients with urge incontinence only also comparable across treatment groups.	Yes	Agree
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Each patient randomised to any treatment group was administered (swallowed) 3 tablets and 1 capsule each morning after breakfast throughout study. All treatments taken orally with glass of water and swallowed intact. Mirabegron tablets, tolterodine SR capsules and corresponding placebo tablets and capsules were indistinguishable (double dummy technique).	Yes	Agree
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Proportion of patients discontinuing study ranged from 4–10% across treatment groups. 32 patients with AEs leading to discontinuations (placebo, 5; 25 mg mirabegron, 10; 50 mg mirabegron, 3; 100 mg mirabegron, 5; 200 mg mirabegron, 7; 4 mg tolterodine, 2.	No	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	All outcomes planned to be measured in the study protocol appear to be reported in CSR.	No	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	FAS (randomised patients who had taken ≥1 dose of double-blind study medication and provided primary efficacy data at baseline and endpoint visit) was primary population for efficacy analysis. Safety population (all patients who had taken ≥1 dose of double-blind study medication) used for safety summaries and analyses. Per- protocol set included all patients in the FAS with no major protocol violations. For efficacy and safety data there was no imputation for missing data. Only patients with symptoms at baseline other than 0 were included in analysis of corresponding symptom.	Yes	Agree

[a] Reproduced from Section 10.3 (pg 287) of the MS.

Abbreviations used in table: AEs, adverse events; CSR, clinical study report; ERG, Evidence Review Group; FAS, full analysis set; IVRS, interactive voice response system; MS, manufacturer's submission.

178-CL-045			
Question	**How was the question addressed in the study? (description in MS[a])**	**Manufacturer's assessment**	**ERG's comment**
Was randomisation carried out appropriately?	Randomisation manager allocated study drugs randomly (1 subject in each group for each set, 4 patients in total) and retained sealed key code until code breaking.	Yes	Not clear
Was the concealment of treatment allocation adequate?	The mirabegron placebo tablet and its package were indistinguishable from the respective study drug tablets and their package, and only the randomisation manager knew the key code.	Yes	Not described
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	Overall, subject backgrounds were similar in all groups. No statistically significant imbalance was found between the groups in any of the items (the criterion for the two-sided significance level was 0.05).	Yes	Agree
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Randomisation manager confirmed appearance and package of the study drugs were indistinguishable before randomisation and code breaking.	Yes	Agree
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Major reasons for discontinuation during treatment period were AEs (6, 6, 8, and 8), protocol deviations (5, 1, 2, and 4), and consent withdrawal (1, 3, 2, and 0 in placebo, 25 mg, 50 mg, and 100 mg groups, respectively); no substantial difference between groups.	No	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	All outcomes planned to be measured in the study protocol appear to be reported in the CSR.	No	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	The FAS included all patients who received at least one dose of the study drug for treatment period and provided at least one efficacy data before initiation of the treatment period and during the treatment period. The SAF included all patients who received at least one dose of the study drug for treatment period. For patients who discontinued the study during the treatment period, the data obtained on the last day of each visit window (at Visits	Yes	Agree

	2, 3, 4, 5, and 6) or the last day of the study medication + 7 days (whichever was the earliest) was adopted as the data of the visit. Those who did not discontinue during the treatment period were handled in the same manner.		
[a] Reproduced from Section 10.3 (pg 287) of the MS. Abbreviations used in table: AEs, adverse events; CSR, clinical study report; ERG, Evidence Review Group; FAS, full analysis set; MS, manufacturer's submission.			

SCORPIO, 178-CL-046			
Question	**How was the question addressed in the study? (description in MS[a])**	**Manufacturer's assessment**	**ERG's comment**
Was randomisation carried out appropriately?	Patients randomised using computer-generated randomisation scheme prepared by Pierrel Research Europe GmbH. Randomisation stratified by country.	Yes	Agree
Was the concealment of treatment allocation adequate?	Patient numbers and randomised treatment allocated by the CIRT system. Study drugs provided in wallet (folded blister card) containing sufficient supply of study medication for 1 week. Each time study drug was dispensed, number of weekly wallets packed in box of study drug was equal to number of weeks between clinic visits plus 1 additional spare wallet. At the end of the screening visit, 1 box of study drugs containing 3 weekly wallets was dispensed.	Yes	Not clear. No description or reference provided for CIRT system
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	Demographic and baseline characteristics consistent across treatment groups for patients in FAS and SAS populations. Observations for demographic and baseline characteristics for per protocol set were similar to those for the FAS.	Yes	Agree
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Throughout study (placebo run-in period and post-randomisation), 2 study drug tablets (mirabegron 50 mg or matching placebo, mirabegron 100 mg or matching placebo) and 1 study capsule (tolterodine SR 4 mg or matching placebo) were taken by mouth with glass of water with or without food in the morning. Mirabegron (OCAS formulation) and placebo tablets to match the OCAS formulation were manufactured by Astellas Pharma Technologies. Tolterodine SR 4 mg capsules were over-encapsulated to maintain blind in a hard gelatine capsule shell. Investigator, study site personnel, patients, sponsor and sponsor representative blinded to identity of randomised drug assignment.	Yes	Agree
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Discontinuation rates similar across treatment groups. Placebo, 8.9%; mirabegron 50 mg, 11.5%; mirabegron 100 mg, 9.0%; tolterodine SR 4 mg, 10.1%.	No	Agree
Is there any evidence to suggest that the	All outcomes planned to be measured in the study protocol appear to be reported in the	No	Agree

authors measured more outcomes than they reported?	CSR.		
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	ITT (all randomised patients who took ≥1 dose of double-blind study drug and who had a baseline diary with micturition measurements), FAS (all randomised patients who took ≥1 dose of double-blind study drug and who had a micturition measurement in the baseline diary and ≥1 post baseline visit diary with a micturition measurement), SAF (all randomised patients who took ≥1 dose of double-blind study drug). LOCF methodology used where no values present for final visit.	Yes	Agree
[a] Reproduced from Section 10.3 (pg 287) of the MS. Abbreviations used in table: CSR, clinical study report; ERG, Evidence Review Group; ; FAS, full analysis set; ITT, intention-to-treat; LOCF, last observation carried forward; MS, manufacturer's submission; OCAS, oral controlled absorption system; SAS, safety analysis set; SR, slow release.			

ARIES, 178-CL-047			
Question	**How was the question addressed in the study? (description in MS[a])**	**Manufacturer's assessment**	**ERG's comment**
Was randomisation carried out appropriately?	Patients randomised to 1 of 3 treatment groups (mirabegron 50 mg, mirabegron 100 mg or placebo) in 1:1:1 ratio using computer-generated randomisation scheme prepared by Pierrel Research Europe GmbH. Randomisation stratified by centre.	Yes	Agree
Was the concealment of treatment allocation adequate?	Patient numbers and randomised treatment allocated by the CIRT system. Study drugs packaged using double-dummy blinded method. They were provided in wallet (folded blister card) containing a sufficient supply of tablets for 1 week.	Yes	Not clear. No description or reference provided for CIRT system
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	Demographic and baseline characteristics consistent across treatment groups for patients in SAF population. Generally, demographic and baseline characteristics were similar across treatment groups in the FAS. Observations for demographic and baseline characteristics for the per protocol set were similar to those for the FAS.	Yes	Agree
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each	Investigator, study site personnel, patients, sponsor and sponsor representatives blinded to identity of randomised drug assignment. 2 tablets taken each day (mirabegron 50 mg or matching placebo, mirabegron 100 mg or matching placebo).	Yes	Agree

outcome)?			
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Proportion of patients randomised into double-blind treatment period that discontinued study was comparable across treatment groups. In each treatment group, the 2 most frequently cited primary reasons for discontinuation were an AE and consent withdrawal. The incidence of discontinuation due to an AE (primary reason) was 3.7%, 4.1% and 4.4% in the placebo, mirabegron 50 mg and mirabegron 100 mg groups, respectively.	No	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	All outcomes planned to be measured in the study protocol appear to be reported in the CSR.	No	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	ITT (all randomised patients who took ≥1 dose of double-blind study drug and who had a baseline diary with micturition measurements), FAS (all randomised patients who took ≥1 dose of double-blind study drug and who had a micturition measurement in the baseline diary and ≥1 post baseline visit diary with a micturition measurement), SAF (all randomised patients who took ≥1 dose of double-blind study drug. The values for the final visit were handled using LOCF methodology.	Yes	Agree
[a] Reproduced from Section 10.3 (pg 287) of the MS. Abbreviations used in table: AE, adverse event; CSR, clinical study report; ERG, Evidence Review Group; FAS, full analysis set; ITT, intention-to-treat; LOCF, last observation carried forward; MS, manufacturer's submission.			

178-CL-048			
Question	**How was the question addressed in the study? (description in MS[a])**	**Manufacturer's assessment**	**ERG's comment**
Was randomisation carried out appropriately?	Randomisation manager randomised study drugs (2 patients from each group per set; 6 patients total) and retained sealed randomisation code until code was broken.	Yes	Not clear
Was the concealment of treatment allocation adequate?	Investigators or sub-investigators assigned study medication to patients confirmed eligible for the study. Drug dispensed sequentially by allocated drug number.	Yes	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	Patient background factors generally similar in all treatment groups; no statistically significant imbalances between groups (significance level: 0.05, two-sided).	Yes	Agree
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were	Dosage forms and packaging for mirabegron placebo and tolterodine placebo indistinguishable from those of the active mirabegron 50 mg tablets and tolterodine 4 mg capsules. Randomisation managers confirmed that study drugs and their packaging were	Yes	Agree

not blinded, what might be the likely impact on the risk of bias (for each outcome)?	indistinguishable in appearance before randomisation and before code breaking after the study drugs were retrieved.		
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	In the respective treatment groups, 31, 31 and 23 patients withdrew from the treatment period. The most common reasons for withdrawal were AEs (9, 15 and 13 patients, respectively) and withdrawal of consent (12, 8 and 1 patients, respectively). The highest number of patients withdrawing consent was in the placebo group.	Yes	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	All outcomes planned to be measured in the study appear to be reported in the publication.	No	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	FAS (patients that took the study medication ≥1 and provided evaluable efficacy data for ≥1 variable before and after initiation of the treatment period), SAF (patients who took the study medication ≥1). Final decisions on disposition of missing data and outliers were made before code breaking, taking into account opinions and advice of medical expert and medical statistical advisor. If multiple observations were obtained within the same visit window for a patient, a value obtained close to the target date was used. If deviations from the scheduled date were the same, the value obtained on the later date was used. The day that the study medication was dispensed was counted as day 0 and the next day as day 1. This was considered appropriate.	Yes	Agree
[a] Reproduced from Section 10.3 (pg 287) of the MS. Abbreviations used in table: ERG, Evidence Review Group; FAS, full analysis set; MS, manufacturer's submission.			

TAURUS, 178-CL-049			
Question	**How was the question addressed in the study? (description in MS[a])**	**Manufacturer's assessment**	**ERG's comment**
Was randomisation carried out appropriately?	Patients randomised to 1 of the 3 treatment groups using computer-generated randomisation scheme prepared by Pierrel Research Europe GmbH.	Yes	Agree
Was the concealment of treatment allocation adequate?	Patient numbers allocated by the CIRT system. Study drugs provided in a wallet (folded blister card) containing a sufficient supply of study medication for 1 week.	Yes	Not clear. No description or reference provided for

			CIRT system
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	Demographic and baseline characteristics (from baseline for this study and not previous data from patients who rolled over from 178-CL-046 or 178-CL-047) consistent across treatment groups for patients in SAF. Generally, demographic and baseline characteristics were similar across treatment groups in FAS.	Yes	Agree
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	During the double-blind treatment period, investigator, study site personnel, patients, sponsor and sponsor's representatives were blinded to the identity of randomised drug assignment. Study drugs packaged using double-dummy blinded method. 2 tablets (mirabegron 50 mg or matching placebo, mirabegron 100 mg or matching placebo) and 1 capsule (tolterodine ER 4 mg or matching placebo) taken each day.	Yes	Agree
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Incidence of patients who discontinued study drug due to a TEAE was 5.9% in the mirabegron 50 mg group, 6.1% in the mirabegron 100 mg group and 5.7% in the tolterodine ER 4 mg group.	No	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	All outcomes planned to be measured in the study appear to be reported in the publication.	No	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	RAS (all randomised patients), FAS (all randomised patients who took ≥1 dose of double-blind study drug and who had a micturition measurement in the baseline diary and ≥1 post-baseline visit diary with a micturition measurement), SAF (all randomised patients who took ≥1 dose of double-blind study drug). For the safety and efficacy data, analysis based on Final Visit took into account patients who withdrew before month 12 and therefore did not have safety or efficacy measurements available for that month. The Final Visit analysis used a LOCF approach. This was considered appropriate.	Yes	Agree

[a] Reproduced from Section 10.3 (pg 287) of the MS.

Abbreviations used in table: ER, extended release; ERG, Evidence Review Group; FAS, full analysis set; LOCF, last observation carried forward; MS, manufacturer's submission; RAS, randomised analysis set; TEAE, treatment-emergent adverse event.

CAPRICORN, 178-CL-074			
Question	**How was the question addressed in the study? (description in MS[a])**	**Manufacturer's assessment**	**ERG's comment**
Was randomisation carried out	Patients randomised using computer-generated randomisation scheme prepared by Pierrel	Yes	Agree

appropriately?	Research Europe GmbH. Randomisation stratified by centre.		
Was the concealment of treatment allocation adequate?	Patient numbers and randomised treatment allocated by the CIRT system. Study drugs provided in a wallet (folded blister card) containing a sufficient supply of tablets for 1 week.	Yes	Not clear. No description or reference provided for CIRT system
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	Demographic and baseline characteristics consistent across treatment groups for patients in SAF. Generally, demographic and baseline characteristics similar across treatment groups in FAS. Demographic and baseline characteristics for per protocol set population similar to those for FAS.	Yes	Agree
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	During double-blind treatment and follow-up periods, investigator, study site personnel, patients, sponsor and sponsors representatives were blinded to identity of randomised drug assignment. Throughout the study, 2 study drug tablets (mirabegron 25 mg or matching placebo, mirabegron 50 mg or matching placebo) taken by mouth with glass of water with or without food in the morning.	Yes	Agree
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Proportion of patients randomised into double-blind treatment period that discontinued the study was numerically higher in the placebo group compared with mirabegron 25 mg and mirabegron 50 mg (15.2%, 10.6% and 12.3%, respectively). In each treatment group, the 2 most frequently cited primary reasons for discontinuation were an AE and withdrawal of consent. The incidence of discontinuation due to an AE (primary reason) was 3.5%, 3.9% and 2.7% in the placebo, mirabegron 25 mg and mirabegron 50 mg groups, respectively.	No	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	All outcomes planned to be measured in the study protocol appear to be reported in the CSR.	No	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	ITT-I (all randomised patients who took ≥1 dose of double-blind study drug and who had micturition measurements and ≥1 incontinence episode in the baseline diary), FAS (all randomised patients who took ≥1 dose of double-blind study drug and who had a micturition measurement in the baseline diary and ≥1 post baseline visit diary with a micturition measurement), SAF (all randomised patients who took ≥1 dose of double-blind study drug). For patients without a value at week 12 for an efficacy or safety variable, LOCF methodology was utilised for deriving final visit value.	Yes	Agree

[a] Reproduced from Section 10.3 (pg 287) of the MS.

Abbreviations used in table: CSR, clinical study report; ERG, Evidence Review Group; FAS, full analysis set; ITT-I, intention-to-treat – incontinence set; LOCF, last observation

carried forward; MS, manufacturer's submission.

178-CL-051 **Study question**	**Grade (yes/no/not clear)**
Were selection/eligibility criteria adequately reported?	Yes
Was the selected population representative of that seen in normal practice?	Yes
Was an appropriate measure of variability reported?	Yes
Was loss to follow-up reported or explained?	Not clear
Were at least 90% of those included at baseline followed up?	Not clear
Were patients recruited prospectively?	Yes
Were patients recruited consecutively?	Yes
Did the study report relevant prognostic factors?	Yes

Manufacturer's quality assessment of additional randomised controlled trials used to populate the network for the mixed treatment comparison (adapted from MS; Section 10.5, pg 291)

Study[a]	BLOSSOM 178-CL-008 (44)	Abrams 2006 (45)	Appell 2001 (46)	Birns 2000 (47)	Cardozo 2004 (48)	Chapple 2004 (51)	Chapple 2004 (50)	Chapple 2007 (49)	Choo 2008 (52)	Chu 2009 (53)	Corcos 2006 (54)	Diokno 2003 (55)
(1)	Yes	Unclear	Unclear	Yes	Unclear	No	Unclear	Unclear	Unclear	Yes	Yes	No
(2)	Yes	Unclear	Unclear	Unclear	Unclear	No	Unclear	Unclear	Unclear	Yes	Unclear	No
(3)	Yes	Unclear	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
(4)	Yes	Unclear	Yes	Yes	Unclear	Unclear	Unclear	Not clear	Unclear	Yes	Yes	Yes
(5)	Yes	Not clear	No	No	No	Yes	No	No	No	No	No	Yes
(6)	No	No	No	No	No	No	Yes	No	No	No	No	No
(7)	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes
Study[a]	**Dmochowski 2003 (56)**	**Herschorn 2008 (77)**	**Herschorn 2010 (58)**	**Herschorn 2010 (57)**	**Ho 2010 (59)**	**Homma 2003 (60)**	**Jacquetin 2001 (61)**	**Kaplan 2011 (62)**	**Khullar 2004 (63)**	**Lackner 2008 (64)**	**Lee 2002 (65)**	**Malone-Lee 2001 (66)**
(1)	No	Unclear	Yes	Yes	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes	Unclear
(2)	No	Unclear	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Yes	Unclear	Unclear	Unclear
(3)	Yes	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	No
(4)	No	Unclear	Unclear	Yes	No	Yes	Yes	Yes	Unclear	Yes	Unclear	Unclear
(5)	No	No	No	No	No	No	No	No	No	No	No	No
(6)	No	No	No	No	No	No	No	No	No	No	No	No
(7)	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes

Study[a]	Nitti 2007 [67]	Nitti 2010 [68]	Rackley 2006 [69]	Rogers 2008 [70]	Rudy 2006 [71]	Staskin 2007 [72]	van Kerrebroeck 2001 [73]	Yamaguchi 2007 [74]	Yamaguchi 2011 [75]	Zinner 2002 [76]		
(1)	Yes	Unclear	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Unclear	Yes		
(2)	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear		
(3)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
(4)	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear		
(5)	No	No	Unclear	No	No	No	No	No	No	No		
(6)	No	No	No	No	No	No	No	No	No	No		
(7)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		

[a] Key to questions:

(1) Was randomisation carried out appropriately?
(2) Was the concealment of treatment allocation adequate?
(3) Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?
(4) Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?
(5) Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?
(6) Is there any evidence to suggest that the authors measured more outcomes than they reported?
(7) Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?

Appendix 6 Efficacy and safety results of non-randomised trial

Summary of non-RCT 178-CL-051 efficacy results (reproduced from MS; Table 55, pg 149)

Mean±SD (n)	All patients N=202	Patients maintained at 50 mg N=152	Patients increased to 100 mg N=50
Mean number of micturitions			
Week 8	–1.52 ± 2.201 *(196)*	–1.88 ± 2.256 *(146)*	–0.45 ± 1.634 *(50)*
Week 16	–2.08 ± 2.288 *(190)*	–2.16 ± 2.235 *(141)*	–1.86 ± 2.445 *(49)*
Week 28	–2.35 ± 2.460 *(185)*	–2.50 ± 2.454 *(137)*	–1.92 ± 2.454 *(48)*
Week 40	–2.13 ± 2.518 *(170)*	–2.22 ± 2.516 *(126)*	–1.87 ± 2.535 *(44)*
Week 52	–2.04 ± 2.595 *(165)*	–2.19 ± 2.708 *(123)*	–1.60 ± 2.201 *(42)*
Final assessment	–2.01 ± 2.599 *(196)*	–2.16 ± 2.673 *(146)*	–1.57 ± 2.341 *(50)*
Mean number of urgency episodes			
Week 8	–2.28 ± 2.549 (*196*)	–2.66 ± 2.541 *(146)*	–1.18 ± 2.253 *(50)*
Week 16	–2.84 ± 2.619 (*190*)	–3.08 ± 2.565 *(141)*	–2.16 ± 2.682 *(49)*
Week 28	–3.32 ± 2.866 (*185*)	–3.48 ± 2.946 *(137)*	–2.87 ± 2.599 *(48)*
Week 40	–3.25 ± 3.006 (*170*)	–3.28 ± 3.102 *(126)*	–3.14 ± 2.746 *(44)*
Week 52	–3.29 ± 3.030 (*165*)	–3.38 ± 3.092 *(123)*	–3.03 ± 2.863 *(42)*
Final assessment	–3.16 ± 2.935 (*196*)	–3.31 ± 2.948 *(146)*	v2.72 ± 2.884 *(50)*
Mean number of incontinence episodes			
Week 8	–1.02 ± 1.211 *(149)*	–1.18 ± 1.136 *(104)*	–0.66 ± 1.311 *(45)*
Week 16	–1.30 ± 1.454 *(145)*	–1.30 ± 1.150 *(101)*	–1.32 ± 2.001 *(44)*
Week 28	–1.54 ± 1.735 *(141)*	–1.55 ± 1.384 *(97)*	–1.52 ± 2.351 *(44)*
Week 40	–1.53 ± 1.634 *(128)*	–1.41 ± 1.568 *(88)*	–1.77 ± 1.766 *(40)*
Week 52	–1.45 ± 1.594 *(124)*	–1.34 ± 1.428 *(86)*	–-1.69 ± 1.918 *(38)*
Final assessment	–1.38 ± 1.656 *(149)*	–1.30 ± 1.400 *(104)*	–1.56 ± 2.143 *(45)*
Mean number of incontinence episodes			
Week 8	–1.04 ± 1.212 *(147)*	–1.23 ± 1.185 *(103)*	–0.58 ± 1.164 *(44)*
Week 16	–1.24 ± 1.317 *(143)*	–1.28 ± 1.132 *(100)*	–1.12 ± 1.681 *(43)*
Week 28	–1.39 ± 1.644 *(139)*	–1.43 ± 1.349 *(96)*	–1.30 ± 2.181 *(43)*
Week 40	–1.40 ± 1.579 *(126)*	–1.33 ± 1.609 *(87)*	–1.56 ± 1.517 (39)
Week 52	–1.37 ± 1.450 *(123)*	–1.31 ± 1.377 *(85)*	–1.48 ± 1.616 *(38)*
Final assessment	–1.33 ± 1.563 *(147)*	–1.32 ± 1.401 *(103)*	–1.33 ± 1.909 *(44)*
Mean number of nocturia episodes			
Week 8	–0.44 ± 0.821 *(165)*	–0.50 ± 0.794 *(122)*	–0.27 ± 0.882 *(43)*
Week 16	–0.53 ± 0.800 *(160)*	–0.50 ± 0.781 *(117)*	–0.60 ± 0.856 *(43)*
Week 28	–0.44 ± 0.844 *(156)*	–0.46 ± 0.816 *(113)*	–0.37 ± 0.920 *(43)*
Week 40	–0.51 ± 0.967 *(144)*	–0.51 ± 0.913 *(105)*	–0.50 ± 1.112 *(39)*
Week 52	–0.54 ± 0.916 *(139)*	–0.52 ± 0.854 *(102)*	–0.59 ± 1.079 *(37)*
Final assessment	–0.48 ± 0.899 *(165)*	–0.49 ± 0.832 *(122)*	–0.47 ± 1.077 *(43)*
Abbreviations in table: mg, milligram; SD, standard deviation.			

Summary of non-RCT 178-CL-051 QoL results, QoL (reproduced from MS; Table 56, pg 150)

Mean±SD *(n)*	**All patients N=202**	**Patients maintained at 50 mg N=152**	**Patients increased to 100 mg N=50**
General health perception (Domain 1)			
Week 28	–4.5 ± 22.99 *(182)*	–5.7 ± 23.65 *(135)*	–1.1 ± 20.82 *(47)*
Week 52	–7.8 ± 21.37 *(164)*	–8.4 ± 22.87 *(122)*	–6.0 ± 16.39 *(42)*
Final assessment	–6.3 ± 21.86 *(192)*	–6.9 ± 23.36 *(144)*	–4.2 ± 16.58 *(48)*
Incontinence impact (Domain 2)			
Week 28	–27.1 ± 29.49 *(182)*	–29.4 ± 28.23 *(135)*	–20.6 ± 32.27 *(47)*
Week 52	–22.8 ± 27.82 *(164)*	–23.8 ± 27.27 *(122)*	–19.8 ± 29.50 *(42)*
Final assessment	–22.7 ± 28.50 *(192)*	–24.3 ± 27.66 *(144)*	–18.1 ± 30.72 *(48)*
Role limitations (Domain 3)			
Week 28	–23.2 ± 25.06 *(182)*	–25.3 ± 24.84 *(135)*	–17.0 ± 24.94 *(47)*
Week 52	–19.5 ± 27.79 *(164)*	–21.6 ± 26.56 *(122)*	–13.5 ± 30.63 *(42)*
Final assessment	–19.3 ± 27.38 *(192)*	–21.5 ± 26.14 *(144)*	–12.5 ± 30.07 *(48)*
Physical limitations (Domain 4)			
Week 28	–22.1 ± 26.45 *(182)*	–23.7 ± 26.66 *(135)*	–17.4 ± 25.53 *(47)*
Week 52	–17.5 ± 28.29 *(164)*	–18.6 ± 29.24 *(122)*	–14.3 ± 25.39 *(42)*
Final assessment	–17.9 ± 28.00 *(192)*	–19.3 ± 29.08 *(144)*	–13.5 ± 24.23 *(48)*
Social limitations (Domain 5)			
Week 28	–11.6 ± 21.21 *(182)*	–13.1 ± 20.98 *(135)*	–7.3 ± 21.52 *(47)*
Week 52	–9.7 ± 22.82 *(164)*	–10.5 ± 23.94 *(122)*	–7.3 ± 19.26 *(42)*
Final assessment	–9.9 ± 22.90 *(192)*	–10.9 ± 24.01 *(144)*	–7.1 ± 19.10 *(48)*
Personal relationships (Domain 6)			
Week 28	–4.9 ± 13.12 *(128)*	–4.7 ± 12.55 *(95)*	–5.6 ± 14.83 *(33)*
Week 52	–4.7 ± 16.52 *(114)*	–3.6 ± 15.96 *(84)*	–7.8 ± 17.90 *(30)*
Final assessment	–4.7 ± 15.99 *(135)*	–4.2 ± 15.32 *(102)*	–6.1 ± 18.07 *(33)*
Emotions (Domain 7)			
Week 28	–19.2 ± 22.68 *(182)*	–19.3 ± 24.24 *(135)*	–18.9 ± 17.63 *(47)*
Week 52	–17.6 ± 24.33 *(164)*	–18.1 ± 26.00 *(122)*	–16.1 ± 18.88 *(42)*
Final assessment	–17.5 ± 24.66 *(192)*	–18.2 ± 26.40 *(144)*	–15.5 ± 18.58 *(48)*
Sleep/energy (Domain 8)			
Week 28	–13.4 ± 18.58 *(182)*	–13.0 ± 18.85 *(135)*	–14.5 ± 17.93 *(47)*
Week 52	–12.3 ± 20.46 *(164)*	–13.3 ± 18.17 *(122)*	–9.5 ± 26.07 *(42)*
Final assessment	–13.0 ± 20.21 *(192)*	–14.1 ± 18.31 *(144)*	–9.7 ± 24.99 *(48)*
Severity measures (Domain 9)			
Week 28	–14.4 ± 16.91 *(182)*	–15.8 ± 16.76 *(135)*	–10.6 ± 16.95 (47)
Week 52	–14.4 ± 16.11 *(164)*	–15.5 ± 16.04 *(122)*	–11.3 ± 16.10 *(42)*
Final assessment	–14.1 ± 16.57 *(192)*	–15.6 ± 16.32 *(144)*	–9.4 ± 16.61 *(48)*
Abbreviations used in table: mg, milligram; SD, standard deviation.			

Summary of non-RCT 178-CL-051 safety results (reproduced from MS; Table 57, pg 151)

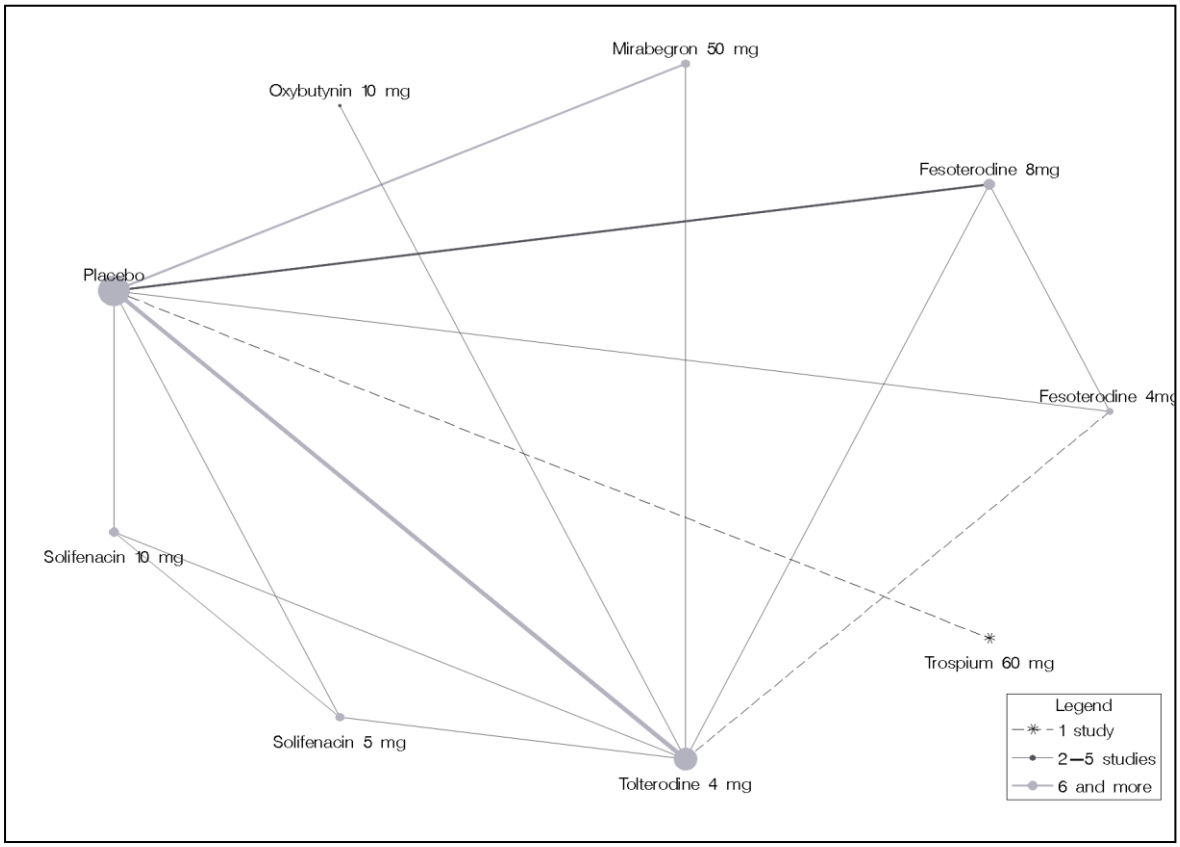
n (%)	All patients N=202	Patients maintained at 50 mg N=152	Patients increased to 100 mg N=50
TEAEs	189 (93.6)	139 (91.4)	50 (100.0)
Mild	175 (86.6)	129 (84.9)	46 (92.0)
Moderate	11 (5.4)	8 (5.3)	3 (6.0)
Severe	2 (1.0)	1 (0.7)	1 (2.0)
Treatment-related TEAEs†	66 (32.7)	51 (33.6)	15 (30.0)
Mild†	59 (29.2)	45 (29.6)	14 (28.0)
Moderate†	0	0	0
Severe†	0	0	0
SAEs	7 (3.5)	4 (2.6)	3 (6.0)
Treatment-related SAEs	0	0	0
TEAEs resulting in permanent discontinuation	15 (7.4)	10 (6.6)	5 (10.0)
Treatment-related TEAEs resulting in permanent discontinuation	5 (2.5)	4 (2.6)	1 (2.0)

† Mild/moderate/severe categories do not include AEs related to ECGs (where severity was not graded).

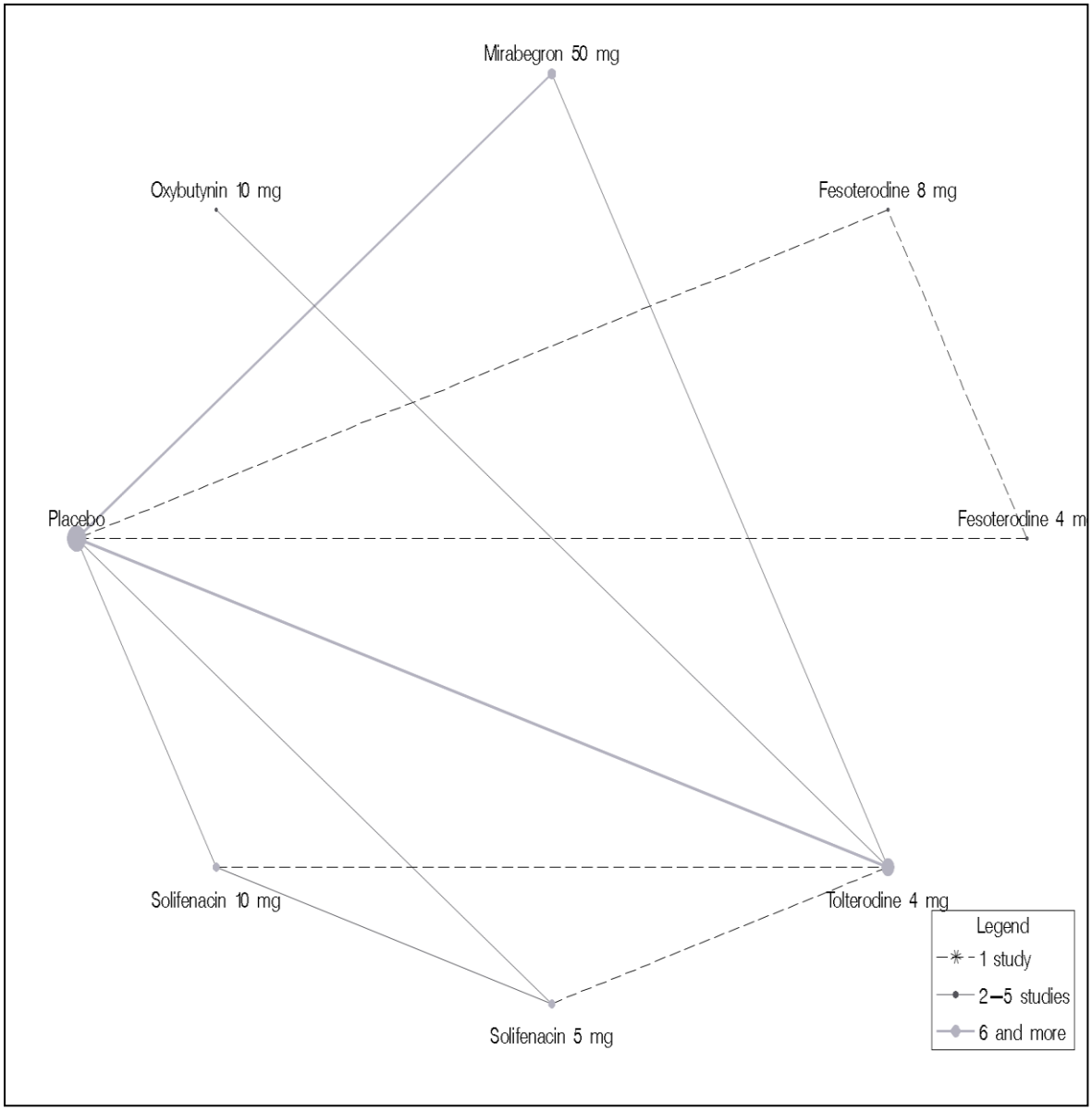
Abbreviations in table: mg, milligram; SAE, serious adverse event; TEAE, treatment-emergent adverse event.

Appendix 7. MTC network figures

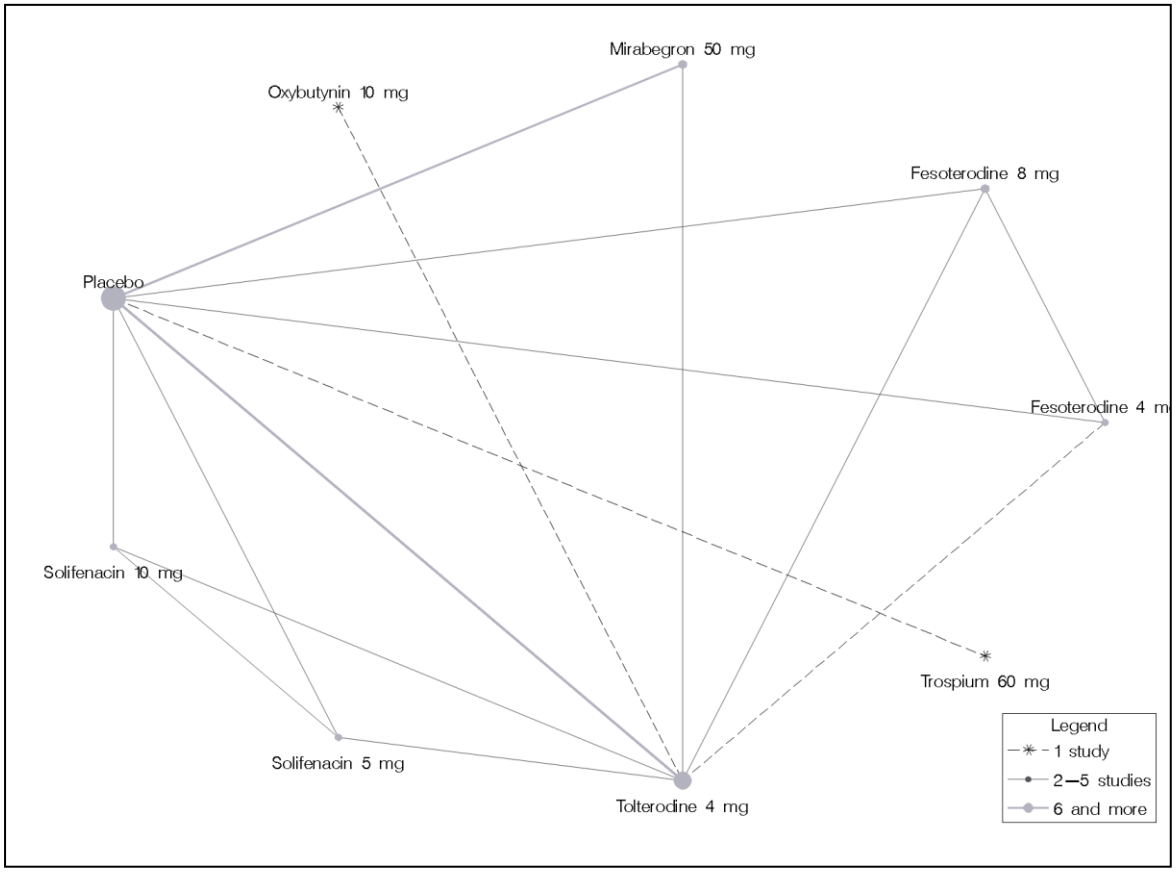
MTC network, micturitions



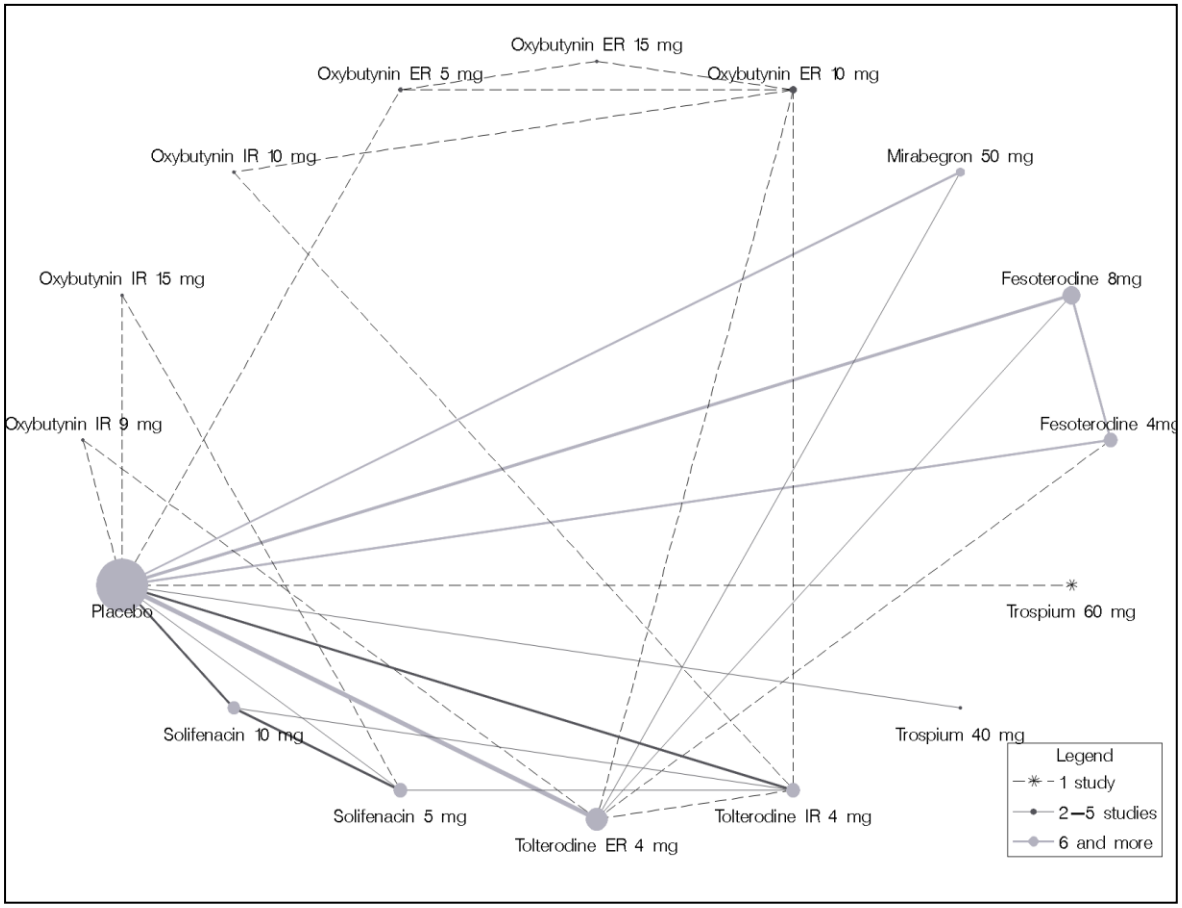
MTC network, incontinence episodes



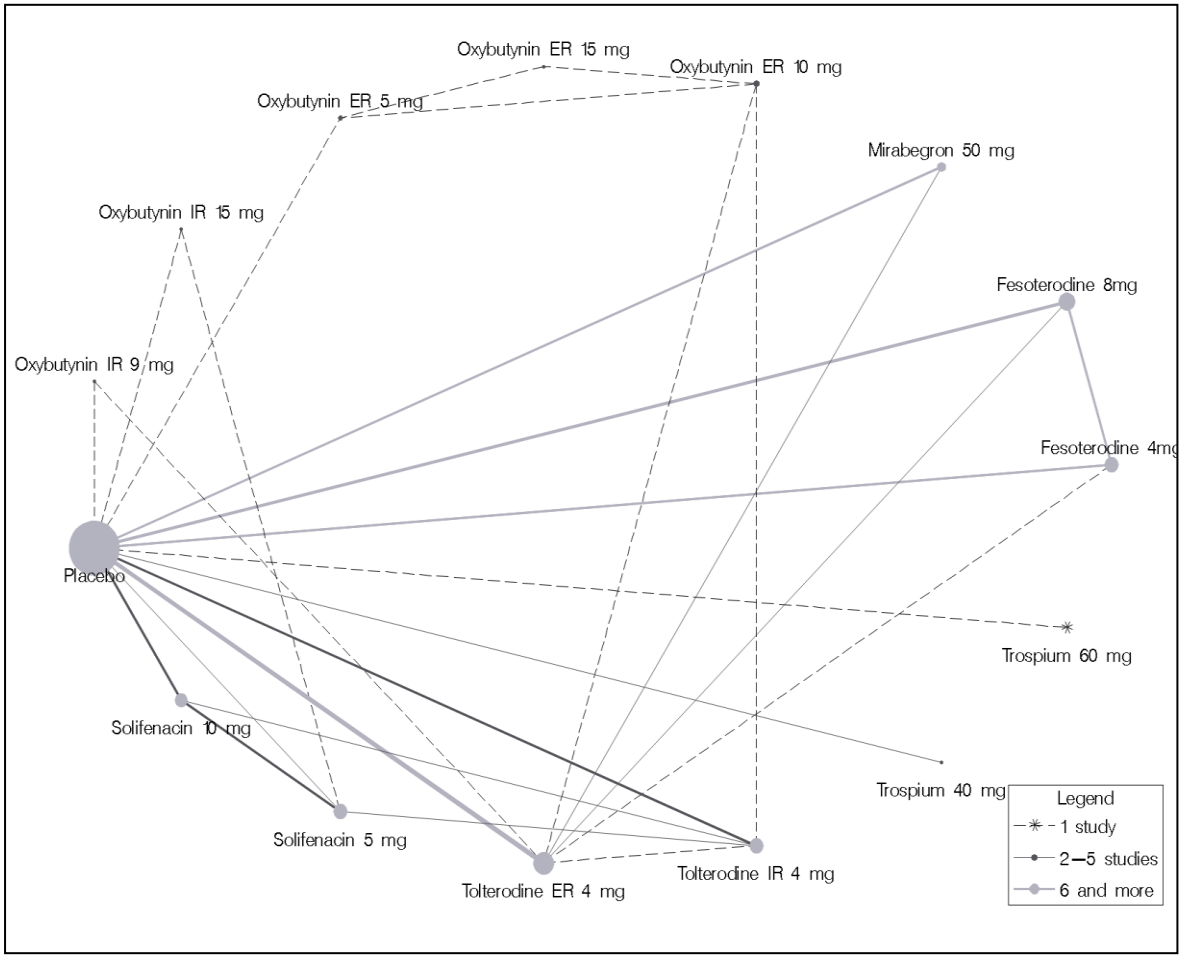
MTC network, urge incontinence episodes



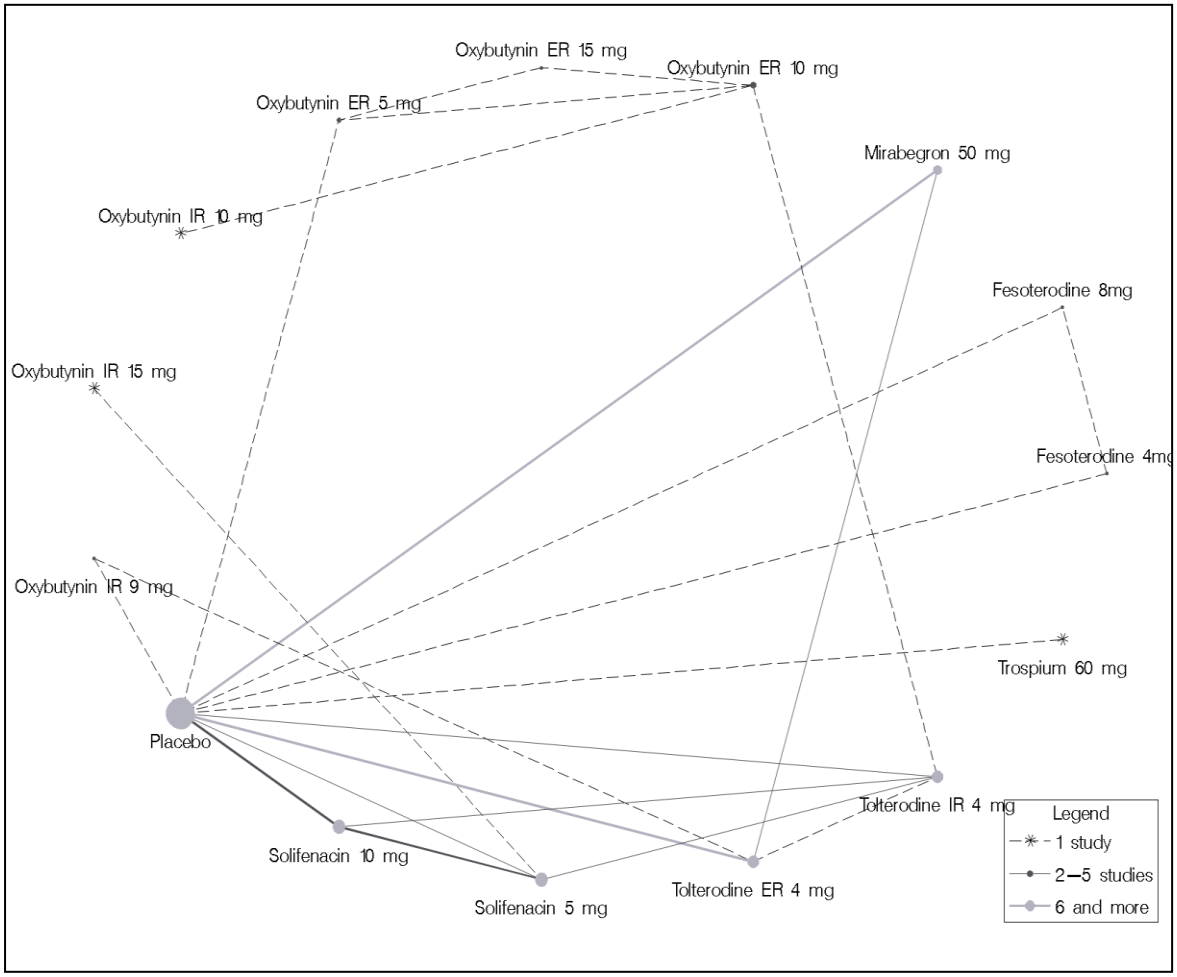
MTC network, dry mouth



MTC network, constipation



MTC network, blurred vision



Appendix 8. Beta-coefficients of multinomial regression model used to inform probability of transition between levels of symptom severity

Beta-coefficients for micturition derived from optimisation using the initial beta-coefficients for solifenacin 5 mg or tolterodine ER 4 mg or mirabegron 50 mg (reproduced from manufacturer's clarification response; pg 59)

Initial betas		**Mirabegron 50 mg**	**Tolterodine 4 mg**	**Fesoterodine 4 mg**	**Fesoterodine 8 mg**	**Oxybutynin 10 mg**	**Placebo**	**Solifenacin 10 mg**	**Solifenacin 5 mg**	**Trospium 60 mg**
Solifenacin 5 mg	1.0687	0.8705409	0.8381488	0.875946	0.9209121	0.8720473	0.1884837	1.46458	1.1288648	1.0057509
	0.6566	0.6488051	0.5738913	0.565284	0.6536128	0.5662379	0.5445493	0.7065623	0.6648672	0.6573182
	0.301	0.3112693	0.4828114	0.4972612	0.3038383	0.4956869	0.5346519	0.2044482	0.2858888	0.2992674
	0.2789	0.2926035	0.5181286	0.5373098	0.2826101	0.5352181	0.5858384	0.1472374	0.2588457	0.2765607
	0	0	0	0	0	0	0	0	0	0
Tolterodine 4 mg	0.2683	0.7804671	0.6036642	0.6248522	0.8785082	0.6227023	0.0668459	1.6448588	0.9902683	1.0604196
	0.183	0.2355082	0.3056946	0.3077109	0.4078512	0.3076435	0.167527	–2.9525588	0.3703752	0.5250727
	0.0484	–0.008958	–0.0254884	–0.0275863	–0.0873292	–0.0274639	0.0559453	0.0433259	–0.0842591	–0.1514099
	0.2193	0.0154003	0.0720633	0.0687468	–0.0384601	0.0691174	0.2360527	-0.355984	–0.0802148	–0.1215486
	0	0	0	0	0	0	0	0	0	0
Mirabegron 50 mg	0.6037	0.7712062	0.6192253	0.6386378	0.8179794	0.6367083	0.0298257	1.3642957	0.9977145	0.8892316
	0.3803	0.4192267	0.3838897	0.3883839	0.4296279	0.3879342	0.2832229	0.4148676	0.4933007	0.4508974
	0.1454	0.1042972	0.141603	0.1368493	0.0931167	0.1373248	0.2290496	0.0037059	0.0383556	0.0728419
	0.0665	0.0125956	0.0615523	0.0553335	–0.001556	0.0559679	0.181445	–0.1735636	–0.0728545	–0.0301982
	0.0000	0	0	0	0	0	0	0	0	0

Beta-coefficients for incontinence derived from optimisation using the initial beta-coefficients for solifenacin 5 mg or tolterodine ER 4 mg or mirabegron 50 mg

Initial betas		Mirabegron 50 mg	Tolterodine 4 mg	Fesoterodine 4 mg	Fesoterodine 8 mg	Oxybutynin 10 mg	Placebo	Solifenacin 10 mg	Solifenacin 5 mg	Trospium 60 mg
Solifenacin 5 mg	0.9604	0.9081774	0.7545774	0.8245374	–0.2320188	0.3316369	-0.0235821	1.3587263	1.3528821	1.120266
	0.4778	0.4846169	0.4728726	0.3729926	1.0348199	0.7727254	0.353877	0.5364531	0.5355928	0.5010967
	0.8625	0.8556963	0.8677258	0.9684563	0.3010612	0.5810428	0.9421343	0.8034044	0.8042684	0.838922
	0.9816	0.9661372	0.9931987	1.2213599	–0.2888879	0.338933	1.2119896	0.844323	0.8462698	0.9284365
	0	0	0	0	0	0	0	0	0	0
Tolterodine 4 mg	0.1431	0.681784	0.5335206	0.4891803	0.2834052	0.4365229	–0.6018398	1.148042	1.1416982	0.8936735
	0.1768	0.3691849	0.3161592	0.300623	0.2264321	0.2819257	0.5944108	0.5900898	0.5866602	0.4461532
	–0.3271	–0.2868434	–0.2978404	–0.3011007	–0.316662	–0.3050327	–0.2388623	–0.2662257	–0.2653566	–0.2715431
	–0.0298	–0.0785594	–0.0652071	–0.0613521	–0.0423402	–0.0566709	–0.1349218	–0.128158	–0.1277416	–0.0977835
	0	0	0	0	0	0	0	0	0	0
Mirabegron 50 mg	0.3617	0.6533808	0.4966448	0.4494091	0.1503967	0.4924673	–0.9893631	1.1469386	1.1403215	0.8761466
	0.4634	0.579962	0.5297018	0.5148741	0.5981099	0.2143695	0.8044326	0.7364411	0.734266	0.6492994
	-0.0251	0.0284532	0.0263288	0.0257067	0.0292271	0.0130657	0.043188	0.0347324	0.0346597	0.0313464
	0.2040	0.1859377	0.210901	0.2178889	0.1785082	0.3603739	0.0872105	0.1125987	0.1136287	0.1540873
	0.0000	0	0	0	0	0	0	0	0	0

Appendix 9. Results of ERG's sensitivity analysis of beta coefficients used to inform manufacturer's secondary base case model

Results of ERG's sensitivity analysis using beta-coefficients derived from optimisation on tolterodine data on cost-effectiveness results of comparisons made in the manufacturer's secondary base case model

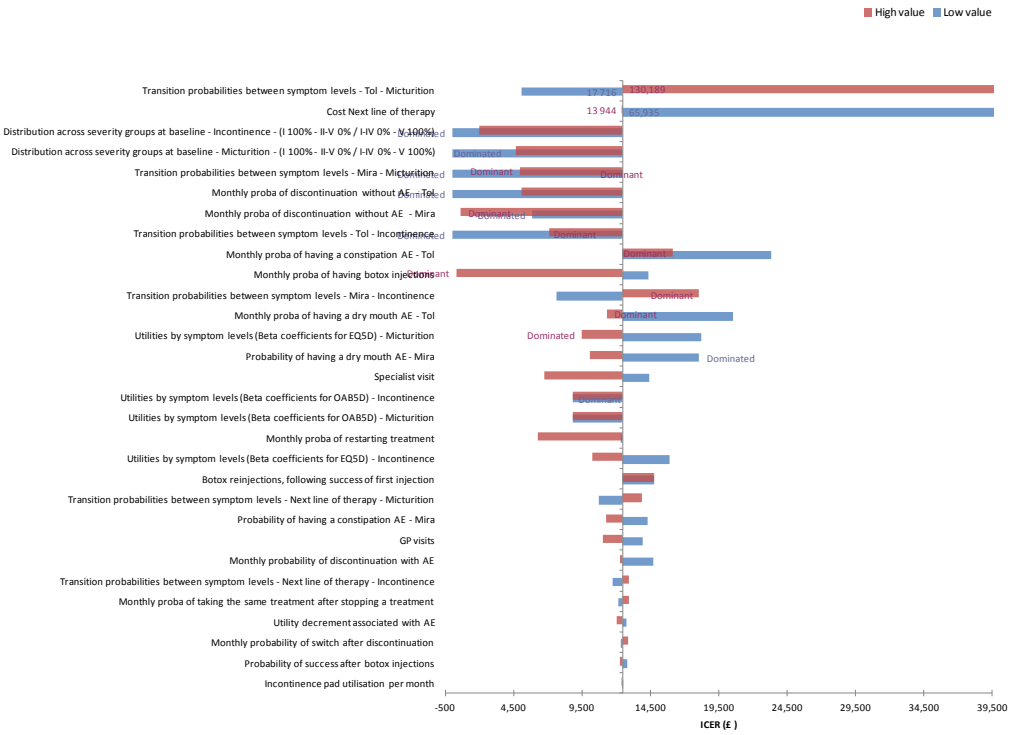
Treatment	Total		Incremental		ICER (£/QALY) versus mirabegron
	Costs (£)	QALYs	Costs (£)	QALYs	
Solifenacin 10 mg	1,647.01	3.764	4.12	0.008	503
Fesoterodine 4 mg	1,599.13	3.760	40.36	0.009	4,479
Tolterodine ER 4 mg	1,599.44	3.760	40.05	0.009	4,621
Oxybutynin ER 10mg	1,586.70	3.756	42.48	0.010	4,145
Trospium chloride MR 60 mg	1,550.83	3.761	84.93	0.007	12,208
Solifenacin 5 mg	1,592.04	3.768	59.09	0.004	15,787
Oxybutynin IR 10 mg	1,420.75	3.752	208.43	0.014	14,705
Abbreviations used in table: ER, extended release; ICER, incremental cost-effectiveness ratio; MR, modified release; QALY, quality adjusted life year.					

Results of ERG's sensitivity analysis using beta-coefficients derived from optimisation on solifenacin data on cost-effectiveness results of comparisons made in the manufacturer's secondary base case model

Treatment	Total		Incremental		ICER (£) versus mirabegron
	Costs (£)	QALYs	Costs (£)	QALYs	
Solifenacin 10 mg	1,647.61	3.761	3.53	0.0108	326
Fesoterodine 4 mg	1,602.12	3.757	37.37	0.0117	3,195
Tolterodine ER 4 mg	1,602.77	3.758	36.72	0.011	3,208
Oxybutynin ER 10mg	1,591.75	3.753	37.43	0.013	2,808
Trospium chloride MR 60 mg	1,552.05	3.758	83.71	0.010	8,345
Solifenacin 5 mg	1,592.93	3.767	58.20	0.005	11,181
Oxybutynin IR 10 mg	1,424.38	3.750	204.80	0.016	12,466
Abbreviations used in table: ER, extended release; ICER, incremental cost-effectiveness ratio; MR, modified release; QALY, quality adjusted life year.					

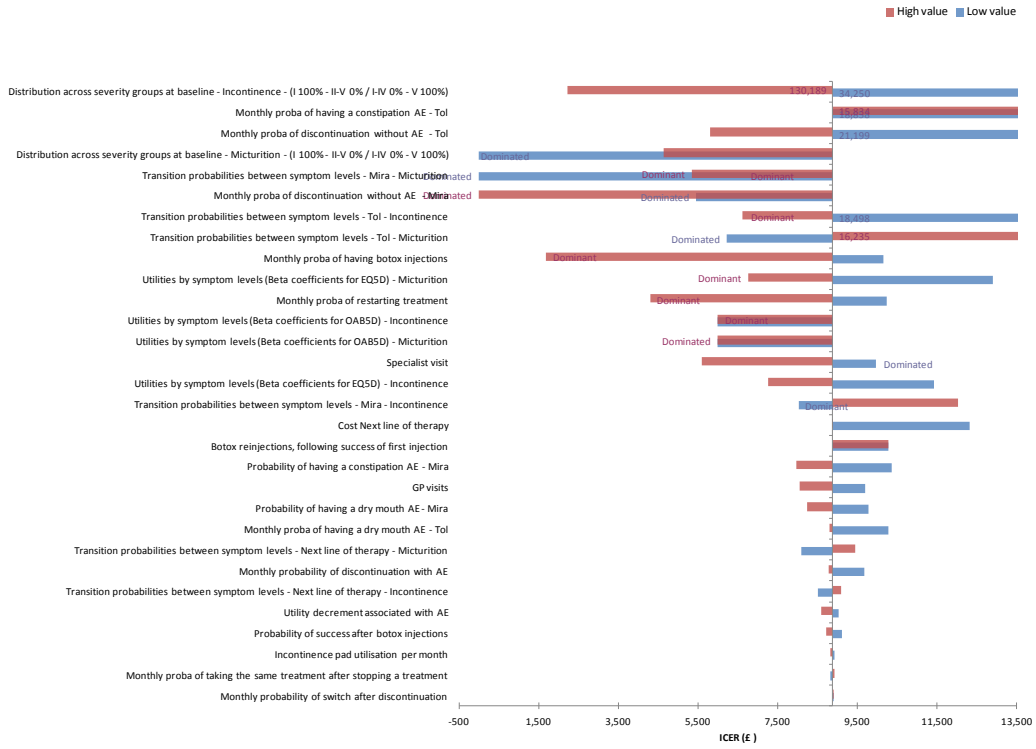
Appendix 10. Results of deterministic sensitivity analysis carried out on the manufacturer's secondary base case cost-effectiveness results

Results of manufacturer's one-way sensitivity analysis on mirabegron 50 mg versus solifenacin 5 mg cost-effectiveness results (reproduced from manufacturer's secondary base case model)

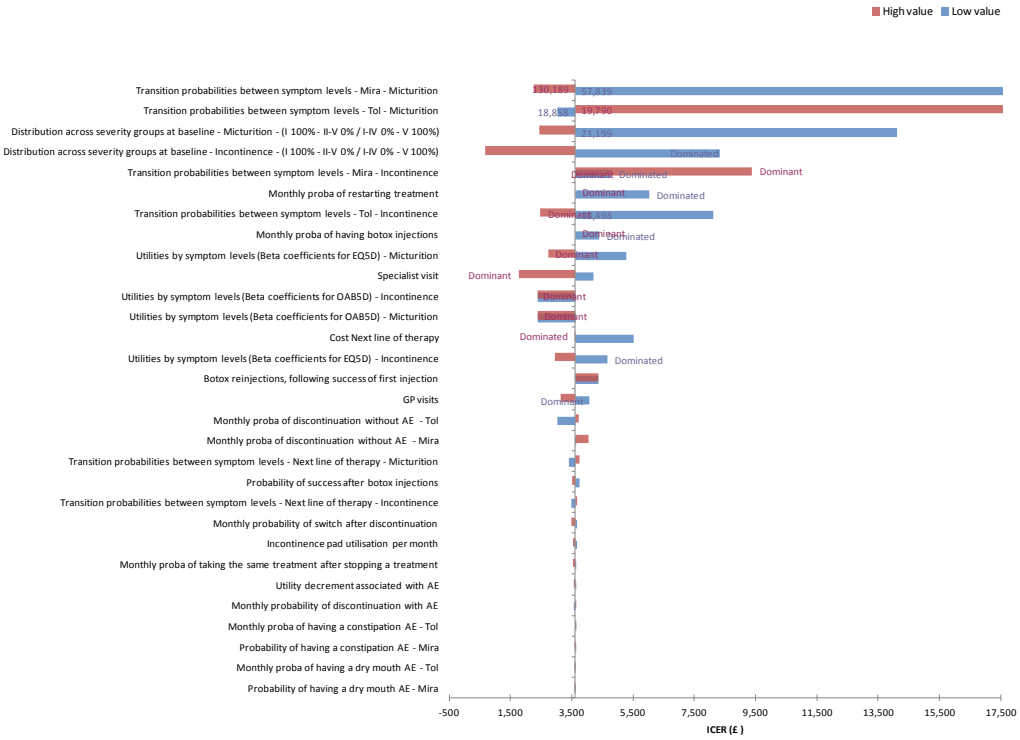


Results of manufacturer's one-way sensitivity analysis on mirabegron 50 mg versus tolterodine ER 4 mg cost-effectiveness results (reproduced from manufacturer's secondary base case model)

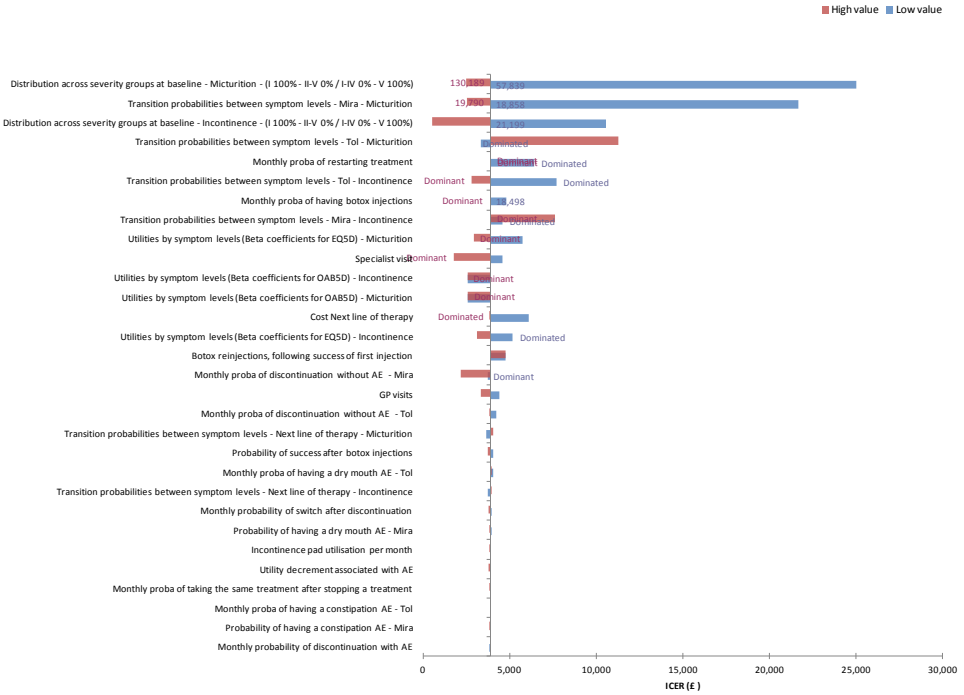
Results of manufacturer's one-way sensitivity analysis on mirabegron 50 mg versus trospium chloride 60 mg cost-effectiveness results (reproduced from manufacturer's secondary base case model)



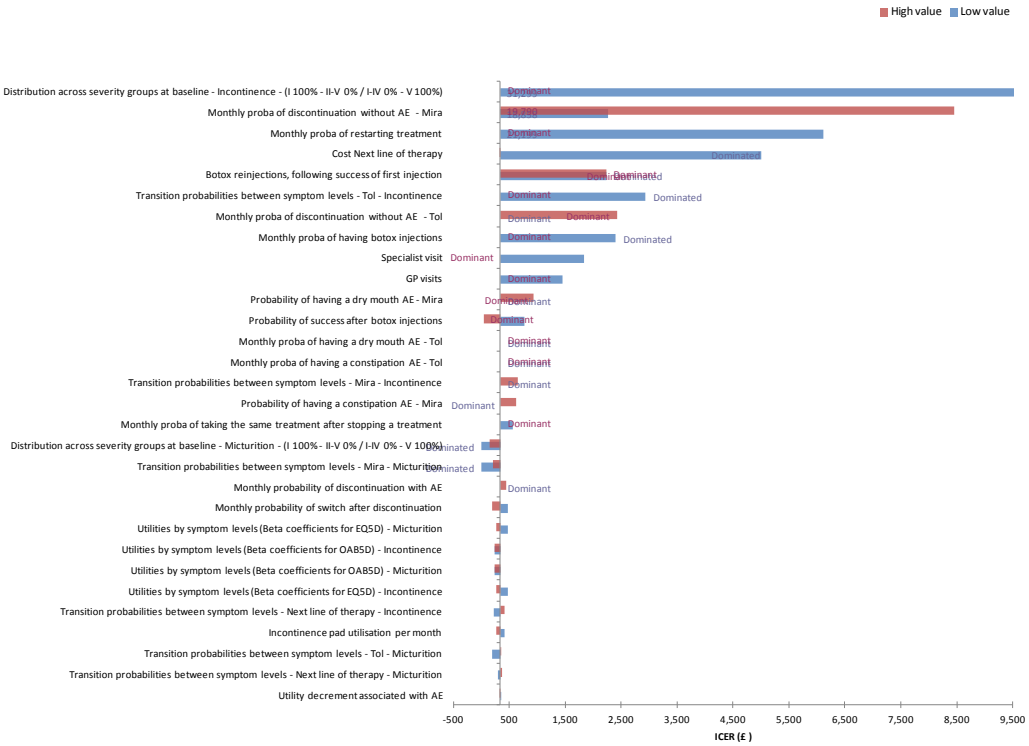
Results of manufacturer's one-way sensitivity analysis on mirabegron 50 mg versus fesoterodine 4 mg cost-effectiveness results (reproduced from manufacturer's secondary base case model)



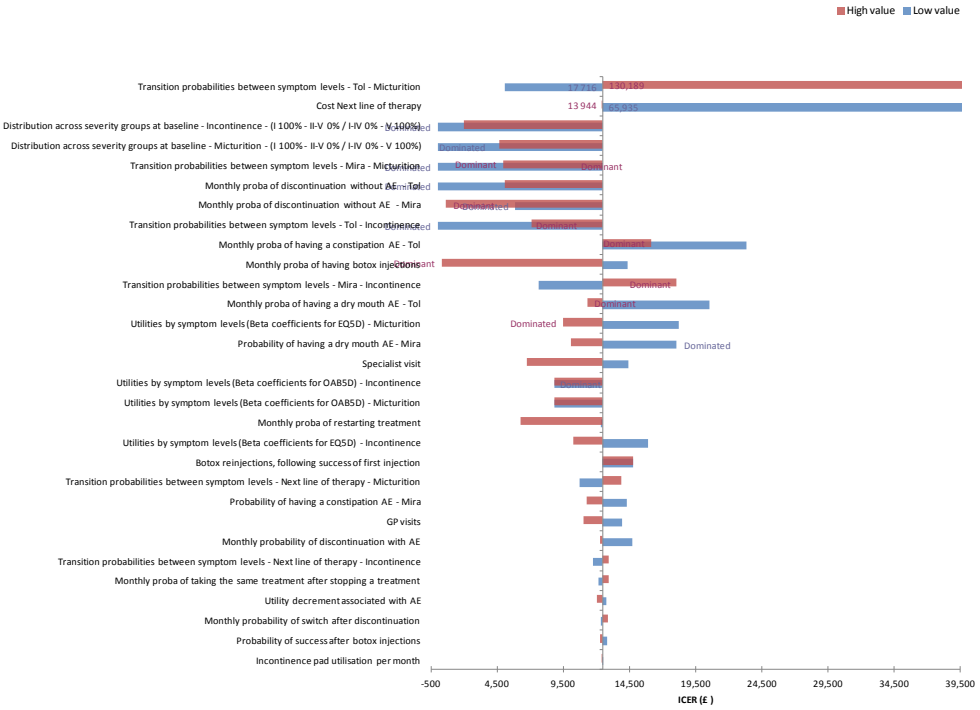
Results of manufacturer's one-way sensitivity analysis on mirabegron 50 mg versus oxybutynin ER 10 mg cost-effectiveness results (reproduced from manufacturer's secondary base case model)



Results of manufacturer's one-way sensitivity analysis on mirabegron 50 mg versus solifenacin 10 mg cost-effectiveness results (reproduced from manufacturer's secondary base case model)



Results of manufacturer's one-way sensitivity analysis on mirabegron 50 mg versus oxybutynin IR 10 mg cost-effectiveness results (reproduced from manufacturer's secondary base case model)



Appendix 11. Full results of cumulative impact of ERG's sensitivity analyses on manufacturer's incremental secondary base case results

Impact of assuming 28% persistence rate for mirabegron on manufacturer's incremental secondary base case results

Treatment	Total		Incremental (versus previous therapy)		ICER (£/QALY) versus oxybutynin IR 10 mg	Incremental ICER (£/QALY)
	Cost (£)	QALY	Cost (£)	QALY		
Oxybutynin IR 10 mg	1,421.00	3.752	–	–	–	–
Trospium chloride MR 60 mg	1,551.86	3.759	130.86	0.007	18,816.13	18,816.13[a]
Oxybutynin 10mg ER	1,587.06	3.755	35.20	–0.003	44,127.74	Strictly dominated[b]
Solifenacin 5 mg	1,592.94	3.768	5.89	0.012	10,812.95	4,591.75[c]
Fesoterodine 4 mg	1,601.40	3.758	8.46	–0.009	26,336.26	Strictly dominated[d]
Tolterodine ER 4 mg	1,601.64	3.759	0.23	0.000	25,017.43	Strictly dominated[e]
Mirabegron 50 mg	1,639.16	3.769	37.52	0.010	12,595.41	32,571.50[f]
Solifenacin 10 mg	1,647.60	3.762	8.44	–0.007	22,261.85	Strictly dominated[g]

[a] versus oxybutynin IR 10 mg.

[b] by trospium chloride MR 60 mg.

[c] versus trospium chloride MR 60 mg.

[d] by solifenacin 5 mg.

[e] by solifenacin 5 mg.

[f] versus solifenacin 5 mg.

[g] by mirabegron 50 mg.

Abbreviations used in table: ER, extended release; ICER, incremental cost-effectiveness ratio; IR, immediate release; MR, modified release; QALY, quality adjusted life year.

Impact of assuming 28% persistence rate for mirabegron and 0% of patients reinitiating original therapy on manufacturer's incremental secondary base case results

Treatment	Total		Incremental (versus previous therapy)		ICER (£/QALY) versus oxybutynin IR 10 mg	Incremental ICER (£/QALY)
	Cost (£)	QALY	Cost (£)	QALY		
Oxybutynin IR 10 mg	1,441.91	3.753	–	–	–	–
Trospium chloride MR 60 mg	1,546.04	3.759	104.13	0.006	18,689.91	18,689.9[a]
Oxybutynin 10mg ER	1,574.36	3.756	28.32	–0.003	44,010.19	Strictly dominated[b]
Solifenacin 5 mg	1,579.56	3.766	5.21	0.010	10,643.49	£4,553.91 [c]
Fesoterodine 4 mg	1,586.40	3.759	6.84	–0.007	26,091.89	Strictly dominated[d]
Tolterodine ER 4 mg	1,586.64	3.759	0.24	0.000	24,766.80	Strictly dominated [e]
Mirabegron 50 mg	1,618.90	3.768	32.26	0.009	12,332.34	£27,727.62[f]
Solifenacin 10 mg	1,622.54	3.761	3.64	–0.006	22,181.50	Strictly dominated [g]

[a] versus oxybutynin IR 10 mg.

[b] by trospium chloride MR 60 mg.

[c] versus trospium chloride MR 60 mg.

[d] by solifenacin 5 mg.

[e] by solifenacin 5 mg.

[f] versus solifenacin 5 mg.

[g] by mirabegron 50 mg.

Abbreviations used in table: ER, extended release; ICER, incremental cost-effectiveness ratio; IR, immediate release; MR, modified release; QALY, quality adjusted life year.

Impact of assuming 28% persistence rate for mirabegron, 0% of patients reinitiating original therapy and NHS reference costs for BoTox on manufacturer's incremental secondary base case results

Treatment	Total		Incremental (versus previous therapy)		ICER (£/QALY) versus oxybutynin IR 10 mg	Incremental ICER (£/QALY)
	Cost (£)	QALY	Cost (£)	QALY		
Oxybutynin IR 10 mg	1,352.14	3.753	–	–	–	–
Trospium chloride MR 60 mg	1,459.88	3.759	107.73	0.006	19,337.50	19,337.50[a]
Oxybutynin 10mg ER	1,489.38	3.756	29.50	–0.003	45,603.77	Strictly dominated[b]
Solifenacin 5 mg	1,499.64	3.766	10.26	0.010	11,404.93	5,401.51[c]
Fesoterodine 4 mg	1,505.83	3.759	6.19	–0.007	27,751.83	Strictly dominated[d]
Tolterodine ER 4 mg	1,506.43	3.759	0.60	0.000	26,402.42	Strictly dominated[e]
Solifenacin 10 mg	1,534.98	3.761	28.54	0.002	22,452.16	Strictly dominated[f]
Mirabegron 50 mg	1,545.30	3.768	10.32	0.006	13,458.81	32,181.63[g]

[a] versus oxybutynin IR 10 mg.

[b] by trospium chloride 60 mg MR.

[c] versus trospium chloride 60 mg MR.

[d] by solifenacin 5 mg.

[e] by solifenacin 5 mg.

[f] by solifenacin 5 mg.

[g] versus solifenacin 5 mg.

Abbreviations used in table: ER, extended release; ICER, incremental cost-effectiveness ratio; IR, immediate release; MR, modified release; QALY, quality adjusted life year.

Impact of assuming 28% persistence rate for mirabegron, 0% of patients reinitiating original therapy and NHS reference costs for BoTox and outpatient specialist visits on manufacturer's incremental secondary base case results

Treatment	Total		Incremental (versus previous therapy)		ICER (£/QALY) versus oxybutynin IR 10 mg	Incremental ICER (£/QALY)
	Cost (£)	QALY	Cost (£)	QALY		
Oxybutynin IR 10 mg	1,329.37	3.753	–	–	–	–
Trospium chloride MR 60 mg	1,437.46	3.759	108.09	0.006	19,337.50	19,401.56[a]
Oxybutynin 10mg ER	1,467.09	3.756	29.63	–0.003	45,603.77	Strictly dominated[b]
Solifenacin 5 mg	1,477.88	3.766	10.80	0.010	11,404.93	5,491.22[c]
Fesoterodine 4 mg	1,484.00	3.759	6.12	–0.007	27,751.83	Strictly dominated[d]
Tolterodine ER 4 mg	1,484.65	3.759	0.65	0.000	26,402.42	Strictly dominated[e]
Solifenacin 10 mg	1,512.41	3.761	27.77	0.002	22,452.16	Strictly dominated[f]
Mirabegron 50 mg	1,524.29	3.768	11.88	0.006	13,458.81	32,711.50[g]

[a] versus oxybutynin IR 10 mg.

[b] by trospium chloride 60 mg MR.

[c] versus trospium chloride 60 mg MR.

[d] by solifenacin 5 mg.

[e] by solifenacin 5 mg.

[f] by solifenacin 5 mg.

[g] versus solifenacin 5 mg.

Abbreviations used in table: ER, extended release; ICER, incremental cost-effectiveness ratio; IR, immediate release; MR, modified release; QALY, quality adjusted life year.