Volume 13 • Number 5 • 2010
VALUE IN HEALTH

Cost-Effectiveness of Duloxetine: The Stress Urinary Incontinence Treatment (SUIT) Study

Borislava Mihaylova, DPhil,[1] Richard Pitman, PhD,[2] Douglas Tincello, BSc, MD,[3] Huub van der Vaart, MD, PhD,[4] Ralf Tunn, MD,[5] Louise Timlin, BSc, MSc,[6] Deborah Quail, BSc, CStat,[6] Adam Johns, BSc, MSc,[6] Mark Sculpher, PhD[2,7]

[1]Health Economics Research Centre, University of Oxford, Oxford, UK; [2]Oxford Outcomes Ltd, Oxford, UK; [3]Reproductive Science Section, Cancer Studies & Molecular Medicine, University of Leicester, Leicester Royal Infirmary, Leicester, UK; [4]Department of Perinatology and Gynaecology, University Medical Centre Utrecht, Utrecht, the Netherlands; [5]Urogynaecology, German Pelvic Floor Centre, St Hedwig Hospitals, Berlin, Germany; [6]Eli Lilly and Company Limited, Lilly Research Centre, Windlesham, UK; [7]University of York, Heslington, York, UK

ABSTRACT

Objective: To assess the cost-effectiveness of duloxetine compared with conservative therapy in women with stress urinary incontinence (SUI).

Methods: Cost and outcome data were taken from the Stress Urinary Incontinence Treatment (SUIT) study, a 12-month, prospective, observational, naturalistic, multicenter, multicountry study. Costs were assessed in UK £ and outcomes in quality adjusted life years using responses to the EuroQol (EQ-5D); numbers of urine leaks were also estimated. Potential selection bias was countered using multivariate regression and propensity score analysis.

Results: Duloxetine alone, duloxetine in combination with conservative treatment, and conservative treatment alone were associated with roughly two fewer leaks per week compared with no treatment. Duloxetine alone and with conservative treatment for SUI were associated with incremental quality-adjusted life-years (QALYs) of about 0.03 over a year compared with no treatment or with conservative treatment alone. Conservative treatment alone did not show an effect on QALYs. None of the interventions appeared to have marked impacts on costs over a year. Depending on the form of matching, duloxetine either dominated or had an incremental cost-effectiveness ratio (ICER) below £900 per QALY gained compared with no treatment and with conservative treatment alone. Duloxetine plus conservative therapy had an ICER below £5500 compared with no treatment or conservative treatment alone. Duloxetine compared with duloxetine plus conservative therapy showed similar outcomes but an additional cost for the combined intervention.

Conclusions: Although the limitations of the use of SUIT's observational data for this purpose need to be acknowledged, the study suggests that initiating duloxetine therapy in SUI is a cost-effective treatment alternative.

Keywords: cost-effectiveness analysis, EQ-5D, observational study, outcomes research, women's health.

Introduction

Urinary incontinence (UI), or the "involuntary leakage of urine" [1], affects between 10% and 60% of women worldwide [2–6]. The condition can be categorized into stress urinary incontinence (SUI), which is involuntary leakage of urine on exertion, sneezing, or coughing; or urge urinary incontinence (UUI), which is involuntary leakage of urine accompanied by, or immediately preceded by, urgency. Both symptoms are present in mixed urinary incontinence (MUI) [1]. SUI accounts for approximately half of all urinary incontinence in adult women [2,3,5].

Initial management of SUI is recommended in general practice using conservative treatments, including lifestyle interventions and pelvic floor muscle training (PFMT) [7]. The Prospective Urinary Incontinence Research study (PURE) was a European observational study to investigate the economic and social impact of UI [8]. PURE baseline data showed that 25–40% of women with symptoms of SUI were receiving medications which in Europe were unlicensed for the treatment of SUI, including alpha-adrenergic agonists, tricyclic antidepressants (TCA), and estrogens [9]. These drugs show limited evidence of efficacy in SUI, and some have side effects (e.g., hypertension, arrhythmias, abnormal vaginal bleeding, and breast cancer) [10]. Duloxetine has been evaluated as a treatment for SUI in a number of large, multinational randomized clinical trials, and is the first drug approved in Europe to treat moderate to severe SUI. A recent systematic review of nine randomized trials of duloxetine in SUI concluded that compared with placebo or no treatment, duloxetine improved health-related quality of life but commented on the common side effect of nausea [11]. Decision analytic modeling studies relating to the UK and Dutch health-care systems have used a trial and other evidence to assess the cost-effectiveness of duloxetine, but have commented on the considerable uncertainty in existing evidence [12,13].

The Stress Urinary Incontinence Treatment (SUIT) study was a 12-month, prospective, observational, naturalistic, multicenter, multicountry study designed to evaluate the cost-effectiveness of duloxetine compared with other forms of nonsurgical intervention in the treatment of women with symptoms of SUI [14]. The study included patients in Germany, United Kingdom, Sweden, and Ireland, with enrolment at the time of initiation of or change in SUI treatment. The participating physician, at his/her own discretion, and in consultation with the patient, determined the treatment initiation or change. Participating investigators observed and recorded data on a regular basis. Data were collected during visits at baseline (observation 1) and at 3 (observation 2), 6 (observation 3), and 12 (observation 4) months after baseline, plus or minus 6 weeks. The article reports on the primary objective of the SUIT study, which was to assess the cost-effectiveness of duloxetine using data from SUIT.

Address correspondence to: Borislava Mihaylova, Health Economics Research Centre, University of Oxford, Old Road Campus, Headington, Oxford OX3 7LF, United Kingdom. E-mail: boby.mihaylova@dphpc.ox.ac.uk

10.1111/j.1524-4733.2010.00729.x

Methods

Overview

The purpose of this analysis was to study the cost-effectiveness of initiating duloxetine for treatment of SUI against non-surgical therapies. Cost-effectiveness was assessed in terms of the differential costs and quality-adjusted life-years (QALYs) between the SUI interventions observed during the 12 months of the study. Costs were assessed from the perspective of the health-care system plus costs to patients of protective materials, based on UK unit costs. The average number of leaks during the last 7 days reported at observation 4 was used to validate treatment effect on health related quality of life (HRQoL). Given the observational nature of the SUIT study, statistical methods were used to address the potential for selection bias accompanying a nonrandomized design.

Defining the Cohort for Analysis

In order to avoid the effects of prior treatments on SUI outcomes and to identify as homogeneous a group as possible, a specific cohort was defined for the analysis. Only study participants that were 1) not undergoing any treatment for SUI at observation 1; and 2) at observation 2 were reported to be undergoing one of the following four treatments: 2.a) duloxetine alone; 2.b) duloxetine plus conservative treatment for SUI; 2.c) conservative treatment for SUI; or 2.d) no treatment for SUI were included. Patients on other pharmaceutical treatments for SUI (n = 122) were omitted from the analysis cohort given the variety of drugs taken (including in combination) and the small numbers involved. The latter also explains why patients undergoing surgical treatments were excluded (n = 31).

Switches of treatments at observation 3 and observation 4 were observed and recorded in SUIT. Significant switches to alternative treatments were observed in the treatment cohorts studied between observation 2 and observation 4, ranging from 24% in "conservative treatment alone" to 45% in the "no treatment for SUI" cohort at observation 2. Therefore, the cost-effectiveness analysis evaluated the cost-effectiveness of treatment strategies of *initiating* patients not treated at the previous visit for SUI to 1) duloxetine alone; 2) duloxetine plus conservative treatment for SUI; 3) conservative treatment for SUI; and 4) no treatment for SUI.

Missing Data

Some missing study visits as well as missing individual data entries were observed in the study. One hundred eighty patients (11.9% of the analysis cohort) missed their visit 3, and 207 (13.7%) visit 4. Although the missing data on individual parameters seem minimal, if all patients with any missing data were excluded from the estimation, nearly a third of the study cohort would have been lost in some of the analyses. Therefore, a multivariate multiple imputation framework was developed, using the *ICE* program in Stata [15], to impute missing baseline characteristics, as well as missing outcome data, in a way that allowed for the estimates to account for the fact that part of data was imputed. The framework was used to impute both health outcomes and costs for SUI, as information from all visits was used to impute the missing data points. Five datasets were generated by imputing the missing data points. All analyses were processed on each of these datasets, and the results combined using Rubin's rules for combining the within-dataset and between-datasets variances [15].

Health Outcomes

EQ-5D [16] data were collected from SUIT participants at each study visit. The instrument has five domains (mobility, self care, usual activities, pain, depression/anxiety), each of which has three response variables (no problems, moderate problems and severe problems). Patients' responses to these domains locate them into one of 243 theoretically possible health states (245 with "dead" and "unconscious"). These EQ-5D states, together with a set of values (or utilities) on a zero (death) to one (full health) scale for these states from a general UK population [17], were used to evaluate HRQoL of study participants at baseline, at 3 months (observation 2), at 6 months (observation 3), and at 1 year (observation 4). Area under the curve methods [18], together with linear interpolation of utilities between the study visits, were used to estimate QALYs over the 1-year study period. Where patients' responses to EQ-5D domains were missing and imputed, the QALYs were reestimated following the imputation.

Resource Use and Costs

Five elements of resource use were measured directly in the study using case record forms: 1) visits to health-care professionals; 2) diagnostic procedures for UI; 3) pharmacologic treatments for SUI; 4) conservative treatment for SUI; and 5) protective material for UI.

Although SUIT was a multinational study, decisions about the use of medical technologies are taken at national level. Therefore, the current analysis was developed from the perspective of the UK NHS. Unit costs were taken, where available, from routine UK sources inflated to 2007 prices, where needed, using the Hospital and Community Health Services (HCHS) pay and price inflation [19]. The key unit costs used in the analysis, together with their sources, are shown in Appendix 1 found at: http://www.ispor.org/Publications/value/ViHsupplementary/ViH13i5_Mihaylova.asp. Costs were evaluated over the duration of the study of about one year.

Statistical Analysis

Two approaches were employed to evaluate treatment effects on costs and outcomes in a way that would address the possibility of selection bias: multivariate regression and propensity score matching.

In the multivariate adjustment approach, outcomes and costs were modeled as functions of treatment allocation and observed baseline characteristics in a linear regression modeling framework. The main assumption was that outcomes are independent of treatment selection after conditioning on the observed factors within the modeling framework selected. A stepwise backward elimination algorithm (at 0.05 level of statistical significance) with addition to model (at 0.049 level of statistical significance if subsequently proven important) was used to identify covariates significant at 5% for each of the five imputed datasets. This was followed by using all identified important covariates within a framework combining regression results over the five datasets, and subsequently excluding covariates that were not statistically significant based on a backward elimination algorithm. Treatment cohort indicators (as per observation 2) were kept in the model even if not statistically significant at 5%. The full list of factors used in the initial models was based on discussion with clinical coauthors and included baseline EQ-5D, average number of leaks at baseline, a range of comorbidities, previous surgery for UI, age, and body mass index. Country of treatment was also introduced into the regression, and interactions between country and treatment cohorts were studied. A seemingly unrelated

regression [20] was undertaken to account for the correlation between the QALYs and costs at 1 years.

In the propensity score analysis, for each treatment comparison, the probability of treatments (propensity scores) were evaluated as a function of observed covariates [21,22]. For each comparison, patients in the two treatment cohorts were then matched based on these propensity scores in order to adjust for observed imbalances and to enable comparison across the matched patient strata. Comparison of outcomes between treatments was performed after employing three alternative propensity scores matching algorithms: nearest neighbor, Kernel matching, and stratification [23]. The three methods of matching apply different trade-offs between the quality and the quantity of the matches, and none of them is a priori superior to the others. Therefore, a degree of robustness can be achieved by considering them jointly for the purpose of interpreting the results.

Cost-Effectiveness

Cost-effectiveness was evaluated using the incremental cost-effectiveness ratio (ICER), which shows the incremental cost per additional QALY gained from one treatment compared with another. The constraints of the propensity scores implementation means that it was not possible to simultaneously compare all available treatments in the same matched cohort, so a series of pairwise comparisons was undertaken. Reflecting the joint uncertainty in costs and outcomes, the probability that a given intervention is more cost-effective than its comparator is presented based on the propensity score analysis. This uses a cost-effectiveness threshold of £20,000 per QALY gained, based on the lower range of the National Institute of Health and Clinical Excellence's stated threshold [24].

Results

Analysis Cohort

A total of 1973 (53%) out of 3739 women participating in SUIT were not on treatment for their SUI at observation 1. Of these, 1541 were in one of the alternative treatment cohorts of interest. After excluding participants that received a surgical procedure for their UI during the study (n = 31), a cohort of 1510 SUIT participants remained, on whom cost-effectiveness analyses were based. All of these patients attended their 3-month study visit and 1330 of them attended their 6-month study visit, while 1303 attended their 1-year visit.

Table 1 outlines the main baseline characteristics by cohort and overall. The treatment cohorts at observation 2 were used to compare outcomes of different treatments for SUI. It should be noted that although all women in SUIT had symptoms of SUI, only about 40% had pure SUI symptoms, with the rest reporting symptoms of both stress and urge urinary incontinence.

The demographic characteristics of this cohort were largely similar to those of the full study population [25], with several notable differences. Specifically, the analysis cohort was more likely to be diagnosed with pure SUI (43% vs. 37.2%).

Resource Use and Costs

Details of the resource use measured in SUIT are provided in Appendix 2 at: http://www.ispor.org/Publications/value/ViHsupplementary/ViH13i5_ Mihaylova.asp. To some extent, the pattern of resource use varied between countries. For example, in terms of visits to clinical professionals, in the United Kingdom and Ireland, most women visited their primary care doctor and relatively few visited a hospital-based physician. This

Table 1 Baseline characteristics of Stress Urinary Incontinence Treatment cost-effectiveness cohort

	Treatment cohort at observation 2				
Baseline characteristic	Duloxetine alone	Duloxetine plus conservative Tx for SUI	Conservative Tx for SUI	No treatment for SUI	Total
Age, mean (SD)	59 (13)	55 (12)	55 (14)	57 (13)	56 (13)
BMI, mean (SD)	28 (5)	28 (5)	28 (6)	27 (5)	28 (5)
EQ index, mean (SD)	0.82 (0.20)	0.82 (0.22)	0.79 (0.25)	0.79 (0.23)	0.80 (0.23)
Mean number of leaks in last 7 days at observation 1 (SD)	16 (15)	16 (14)	11 (12)	14 (15)	13 (14)
Comorbidity(ies) affecting QoL,* % of cohort affected	55%	50%	52%	50%	53%
Comorbidity(ies) affecting incontinence,† % of cohort affected	52%	47%	36%	41%	43%
Higher education, %	33%	44%	49%	41%	43%
Baseline Severity (% severe SUI)	11%	13%	8%	11%	10%
Specialty of study doctor‡ (%)	PCP (17%), GYN (74%), URO (8%), Other (1%)	PCP (19%), GYN (69%), URO (12%)	PCP (38%), GYN (52%), URO (10%), Other (1%)	PCP (25%), GYN (65%), URO (9%), Other (1%)	PCP (28%), GYN (62%), URO (9%), Other (1%)
Pure SUI (%)	42%	37%	47%	37%	43%
Any employment (%)	35%	48%	42%	41%	41%
Previous surgery for SUI (%)	12%	6%	4%	11%	8%
Total (% of cost-effectiveness cohort)	394 (26%)	212 (14%)	623 (41%)	281 (19%)	1510

*"Yes" if the answer was "yes" for the presence of the following conditions: urinary tract infection, obstructive lung disease or persistent cough, neurological disorder that affects the lower urinary tract, diabetes, depression, myocardial infarction, hypertension, cancer (excluding skin cancer), and chronic heart failure. "No" if the answer was "no" for all of the listed conditions. Otherwise, the answer was considered missing.
†"Yes" if any of the medical conditions hysterectomy, fecal incontinence, constipation, pelvic organ prolapse was present; otherwise "No."
‡Primary care physician (PCP), gynecologist (GYN), urologist (URO).
BMI, body mass index, QoL quality of life; SUI, stress urinary incontinence.

Table 2 Mean (SD) health-care costs and protective materials for urinary incontinence by type of resource used and treatment cohort (£ 2007)*

Cohort	Duloxetine alone	Duloxetine plus conservative treatment for SUI	Conservative treatment for SUI	No treatment for SUI	All
Number of patients*	318	186	545	202	1251
Visits to health-care professional	357 (229)	392 (257)	328 (293)	386 (305)	354 (276)
Diagnostics for UI	199 (324)	149 (232)	126 (271)	202 (388)	160 (303)
Pharmacologic treatments for SUI†	176 (43)	170 (58)	18 (45)	44 (76)	85 (91)
Conservative treatment for SUI	17 (99)	172 (475)	133 (453)	96 (290)	103 (376)
Protective material	74 (120)	68 (74)	68 (75)	81 (99)	72 (92)
All costs	823 (502)	952 (624)	672 (770)	809 (708)	774 (686)

*Only for patients with all available data over the year.
†Duration limited to 365 days.
SUI, stress urinary incontinence; UI, urinary incontinence.

position was reversed in Germany and Sweden. Furthermore, fewer women in the UK received diagnostic tests than those in other countries. In terms of treatments used to manage SUI, the proportions of women that underwent pharmacologic treatments varied from 37% in the United Kingdom to 65% in Germany, with most of these patients having taken duloxetine, which had the longest mean duration of use of 258 days over the year. The proportions of women that received conservative treatments ranged from 66% in Germany to 94% in Ireland, with 88% in the United Kingdom. For most women, this took the form of self-administered pelvic floor exercises.

Table 2 summarizes the average costs (SD) by treatment cohort and has not been adjusted for covariates. Average costs were largest in the duloxetine plus conservative therapy group, followed by the duloxetine alone group, while the costs in the conservative treatment group were lower. This observation could be because of the fact that most patients in the duloxetine cohorts were located in Germany, where the costs of resource use of health care for SUI were higher, or that patients in the duloxetine cohorts seemed to have somewhat more severe symptoms (Table 1). Therefore, without statistically adjusting for factors determining costs independently of interventions, it is difficult to determine the effects of treatment strategies on costs.

Multivariate Regression

The results of the regression analysis are provided in Table 3. For QALYs at one year, significantly lower QALY scores were estimated for increased baseline age, baseline BMI, and being treated in the UK. Significantly higher QALY scores were estimated for a higher baseline EQ-5D score. Compared with no treatment, higher incremental QALY scores were predicted for duloxetine and for duloxetine plus conservative treatment, and lower (non-significant) QALY scores were predicted for conservative therapy alone.

For total cost at 1 year, higher costs were predicted for those women who had previous surgery and those managed by an urologist compared with a gynecologist. Lower costs were predicted for women of a higher age, employed, treated in the United Kingdom or Sweden (compared with Germany), and treated by a primary care doctor or other health-care provider (excluding urologist) compared with a gynecologist. Women suffering from stress urinary incontinence alone were also estimated to incur lower cost compared with those with mixed urinary incontinence symptoms. No statistically significant interactions between country of residence and treatment cohort for QALYs and costs were identified.

Table 4 summarizes the incremental costs and QALYs for pairwise treatment comparisons based on the multivariate regression model. It also provides details of the number of leaks avoided in the last week. Duloxetine alone, duloxetine in combination with conservative treatment, and conservative treatment alone were associated with roughly two fewer leaks per week compared with no treatment. Duloxetine alone and duloxetine in combination with conservative treatment for SUI were associated with incremental QALYs of about 0.03 over a year compared with no treatment or with conservative treatment alone. Conservative treatment alone did not show an effect on QALYs. None of the interventions considered appeared to have statistically significant impacts on costs over a year. No important differences were observed in any outcome between initiating duloxetine alone and duloxetine in combination with conservative treatment.

Propensity Scores

With propensity score analysis, no significant treatment effects were estimated in any treatment comparison when the nearest neighbor approach was employed. The estimates, though, were consistent with those based on kernel matching and stratification

Table 3 Multivariate regression model of costs (£ 2007) and QALYs at 1 year

	Estimate	(SE)
Part1: QALY at 1 years		
EQ 5D index at baseline	0.575	(0.014)†
Age (years)	−0.001	(0.0002)†
BMI (kg/m^2)	−0.002	(0.001)†
Country (United Kingdom)	−0.056	(0.009)†
Tx (Duloxetine alone)	0.027	(0.009)†
Tx (Duloxetine plus conservative treatment)	0.025	(0.011)*
Tx (Conservative treatment alone)	−0.005	(0.009)
Constant term	0.527	(0.026)†
Part 2: Total cost at 1 years		
Previous surgery for SUI	141	(63)*
Age (years)	−3	(1)*
Country (United Kingdom; comparator Germany)	−236	(74)†
Country (Sweden; comparator Germany)	−169	(72)*
Tx (Duloxetine alone)	−3	(53)
Tx (Duloxetine plus conservative treatment)	110	(59)
Tx (Conservative treatment alone)	−17	(49)
Other health-care provider (comparator: gynecologist)	−410	(179)*
Primary care doctor (comparator: gynecologist)	−278	(72)†
Urologist (comparator: gynecologist)	189	(60)†
Employed (yes)	−105	(40)†
SUI alone (comparator mixed UI)	−186	(37)†
Constant term	1166	(102)†

*$P < 0.05$; †$P < 0.01$.
BMI, body mass index; EQ-5D, EuroQol; QALY, quality-adjusted life-year; SUI, stress urinary incontinence; UI, urinary incontinence.

Table 4 Summary of number of leaks avoided per week, incremental QALYs and incremental costs (£ 2007) at 1 year for individual treatment comparisons

	Number of leaks avoided per week (SE)		Incremental QALY (SE)		Incremental cost (SE)	
Duloxetine vs. No treatment for SUI	−2.1	(0.6)†	0.027	(0.009)†	−3	(53)
Conservative treatment for SUI vs. No treatment for SUI	−1.5	(0.5)†	−0.005	(0.009)	−17	(49)
Duloxetine plus Conservative treatment for SUI vs. No treatment for SUI	−2.3	(0.7)†	0.025	(0.011)*	110	(59)
Duloxetine vs. Duloxetine plus Conservative treatment for SUI	0.2	(0.9)	0.002	(0.015)	−113	(79)
Duloxetine vs. Conservative treatment for SUI	−0.5	(0.8)	0.031	(0.013)*	14	(72)
Duloxetine plus Conservative treatment for SUI vs. Conservative treatment for SUI	−0.7	(0.9)	0.030	(0.014)*	127	(77)

*$P < 0.05$; †$P < 0.01$.
QALY, quality-adjusted life-year; SUI, stress urinary incontinence.

methods. The results based on kernel matching and stratification methods were very similar. Based on their results, although duloxetine alone and duloxetine in combination with conservative treatment for SUI seemed to marginally reduce the weekly number of leaks by about two compared with no treatment, comparisons of duloxetine (with or without conservative treatment) with conservative treatment alone did not show significant reductions in average number of leaks. Duloxetine (with and without conservative treatment) consistently but marginally improved QALYs by about 0.02 to 0.04 (result not statistically significant for "Duloxetine plus conservative treatment" compared with "no treatment for SUI" and when "nearest neighbor" matching employed) over a year compared with no treatment or with conservative treatment alone. Conservative treatment alone did not show an effect on QALYs when compared with no treatment; neither did it show an effect in combination with duloxetine when compared with duloxetine alone. None of the interventions considered appeared to have statistically significant impact on costs over a year. No important differences were observed in any outcome between duloxetine alone and duloxetine in combination with conservative treatment.

Table 5 summarizes the cost-effectiveness results from the three propensity score matching methods in the form of a series of pairwise comparisons. The differential costs and QALYs are presented with the treatment strategy in the row being the reference (i.e., a positive differential means that treatment has higher costs or QALYs). The ICER is presented for this strategy in the first column together with the probability that the treatment is the more cost-effective, assuming a cost-effectiveness threshold of £20,000 per QALY gained. At NICE's cost-effectiveness threshold (£20,000 to £30,000 per QALY gained), duloxetine alone and in combination with conservative therapy were cost-effective strategies for initial management.

It can be seen that depending on the form of matching, duloxetine either dominated or had an ICER below £900 per QALY gained compared with no treatment or with conservative treatment alone. Duloxetine had a probability of being the more cost-effective than no treatment and than conservative treatment, at a £20,000 per QALY threshold, of between 0.86 and 1.00. Duloxetine plus conservative therapy had an ICER below £5500 compared with no treatment or conservative treatment alone, and probabilities of being the more cost-effective from 0.68 to 0.99. Duloxetine compared with duloxetine plus conservative therapy showed very similar outcomes but an additional cost for the combined intervention; the probability of either strategy being the most cost-effective was approximately 0.50.

Discussion

The SUIT study provides a rare example of a prospective naturalistic study in the context of stress urinary incontinence in women. Using a cohort of 1510 patients, the study provided estimates of 1-year costs and QALYs associated with the use of duloxetine (with or without conservative therapy), conservative therapy alone, and no treatment. The statistical modeling provides estimates of the differential costs and benefits of therapies, which are more reliable than a crude (unadjusted) comparison of the groups. The study suggests that in terms of average effects, duloxetine (with or without conservative therapy) dominated conservative therapy and no treatment, or had cost per QALY gained below standard thresholds such as those used by NICE. Thus, duloxetine, either alone or in combination with conservative treatment, can be considered a cost-effective treatment for SUI. Nevertheless, it should be emphasized that the differences between the therapeutic groups were small, both in terms of costs and QALYs, over the period of a year.

Although randomized trials offer high internal validity in assessing the efficacy of interventions, they often fail to provide reliable evidence on the costs and health effects of particular interventions, as they are used in routine clinical practice. This is particularly the case with randomized trials of pharmaceutical products, which often have short periods of follow-up, focus on clinical outcomes rather than the impact on patients' perceived health, fail to compare with all relevant existing treatments and take place prior to the widespread use of the product [26]. There is therefore a potential role for more naturalistic studies, in addition to RCTs, where measurement of outcomes takes place in the context of "real-life" clinical management [27]. A systematic review of the randomized trials, which evaluated duloxetine in SUI, found a positive impact on HRQoL, subjective cure rates, and improvement in incontinence, but also a high rate of side effects [11]. The most troublesome side effect was nausea, which occurred in up to 25% of subjects, and might be considered to have potentially adverse effects on HRQoL. The characteristics of these trials suggests potential value in evidence from SUIT: relatively short follow-up (mostly 12 weeks, but as little as 3 weeks and no more than 36 weeks); a focus on disease specific HRQoL measures rather than generic measures preferred for cost-effectiveness, and little evidence on the use of duloxetine in combination with conservative therapy [11]. A Cochrane systematic review of the effectiveness of conservative treatment suggested that it leads to about one fewer incontinence episode per day on average (Hay–Smith, 2006 #8937) (compared with an estimate of about 1.5 fewer episodes weekly in the current analysis). Nevertheless, the trials in the review were small (25–43 participants), over shorter time periods (8 weeks to 6 months), and likely involved more systematic training and monitoring.

Nevertheless, naturalistic studies are typically observational in that they do not randomize patients to interventions. To estimate the differential effect on costs and benefits of interventions in such studies, it is necessary to use appropriate techniques to deal with potential selection bias because of confounding, namely factors

Table 5 Incremental QALYs (SE), incremental costs (SE) (£ 2007) and cost-effectiveness for the separate treatment comparisons based on the propensity scores approach

	Compared with:		
	No treatment for UI	Conservative treatment only	Duloxetine plus conservative treatment
(a) Nearest neighbor matching			
Conservative treatment only	ΔQALY = –0.012 (0.022) ΔC = –34 (80) ICER = 2970 $P(CE)_{20,000}$ = 0.35		
Duloxetine treatment only	ΔQALY = 0.026 (0.023) ΔC = 21 (85) ICER = 812 $P(CE)_{20,000}$ = 0.86	ΔQALY = 0.032 (0.021) ΔC = –99 (113) ICER = dominant $P(CE)_{20,000}$ = 0.91	ΔQALY = 0.0004 (0.029) ΔC = –120 (96) ICER = dominant $P(CE)_{20,000}$ = 0.59
Duloxetine plus conservative	ΔQALY = 0.021 (0.024) ΔC = 112 (119) ICER = 5375 $P(CE)_{20,000}$ = 0.74	ΔQALY = 0.015 (0.024) ΔC = 56 (170) ICER = 3852 $P(CE)_{20,000}$ = 0.68	
(b) Kernel matching			
Conservative treatment only	ΔQALY = –0.011 (0.013) ΔC = –68 (54) ICER = 6379 $P(CE)_{20,000}$ = 0.28		
Duloxetine treatment only	ΔQALY = 0.029 (0.009)** ΔC = –9 (54) ICER = dominant $P(CE)_{20,000}$ = 1.00	ΔQALY = 0.031 (0.010)** ΔC = –40 (61) ICER = dominant $P(CE)_{20,000}$ = 1.00	ΔQALY = –0.004 (0.016) ΔC = –148 (60)* ICER = dominant $P(CE)_{20,000}$ = 0.58
Duloxetine plus conservative	ΔQALY = 0.021 (0.011) ΔC = 95 (61) ICER = 4494 $P(CE)_{20,000}$ = 0.92	ΔQALY = 0.029 (0.012)* ΔC = 33 (89) ICER = 1124 $P(CE)_{20,000}$ = 0.99	
(c) Stratification method			
Conservative treatment only	ΔQALY = –0.012 (0.016) ΔC = –70 (60) ICER = 5775 $P(CE)_{20,000}$ = 0.29		
Duloxetine treatment only	ΔQALY = 0.031 (0.013)* ΔC = –7 (55) ICER = dominant $P(CE)_{20,000}$ = 0.99	ΔQALY = 0.031 (0.012)* ΔC = –38 (60) ICER = dominant $P(CE)_{20,000}$ = 0.99	ΔQALY = –0.005 (0.016) ΔC = –129 (60)* ICER = dominant $P(CE)_{20,000}$ = 0.54
Duloxetine plus conservative	ΔQALY = 0.027 (0.017) ΔC = 114 (61) ICER = 4282 $P(CE)_{20,000}$ = 0.90	ΔQALY = 0.029 (0.015)* ΔC = 40 (83) ICER = 1363 $P(CE)_{20,000}$ = 0.96	

*$P < 0.05$; **$P < 0.01$.

Dominated means the strategy in the first column has higher QALYs and lower costs than the comparator strategy. $P(CE)_{20,000}$ is probability of intervention being cost-effective at a cost-effectiveness threshold of £20,000/QALY.

ΔC, incremental cost; ICER, incremental cost-effectiveness ratio; ΔQALY, incremental quality-adjusted life-year.

related to treatment selection, which also determine treatment outcomes. Here, two general methods were employed for this purpose: multivariate regression and propensity scoring. Multivariate regression uses observed patient and health system characteristics to adjust estimates of cost and benefit to allow for differences in the characteristics of patients undergoing the different treatments. Propensity scores take this a step further by seeking to match patients undergoing different treatments according to the likelihood of their receiving a given intervention.

It is important for economic analyses to relate to specific jurisdictions in terms of relevant evidence. Rather than present four separate analyses here (one for each country recruiting into SUIT), the analysis was undertaken from the viewpoint of the UK given the availability of suitable cost data and the widespread use of economic analyses to support decision-making. This analysis used UK-specific costs, but resource use and outcome data from all four countries. The statistical models have sought to adjust for differences between countries in these measures and no important interactions between country and treatment were observed. Nevertheless, the cost-effectiveness results could vary with unit costs from other countries.

The study's limitations should be acknowledged. For the cost-effectiveness comparison, it was necessary to identify a cohort to facilitate a meaningful comparison of alternative treatment strategies. In order to identify as homogeneous a group as possible, a specific cohort was defined for the analysis, which included women who were not undergoing any treatment for SUI at observation 1, and were reported to be undergoing one of the specific treatments of interest at observation 2. This process of cohort identification meant that only 1510 out of the 3739 women who participated in SUIT were included in the analysis. Although this loss of data reduces the precision with which costs and benefits could be estimated, it is unlikely to increase the potential for bias. In principle, other interventions could have been added to the comparison, such as surgical treatment and other pharmaceuticals. Nevertheless, the numbers of patients undergoing such treatments who were untreated at observation 1 were too small for reliable estimates of costs and benefits.

As with most observational studies, missing data was a feature of the SUIT dataset. This is a particular problem for economic analysis, because estimates of total costs and QALYs require the use of outcomes measured at different points of follow-up. If one of these data points is missing, the whole cost/QALY estimate is effectively missing. Therefore, it is necessary to impute the missing data. We used a multiple multivariate imputation technique that implicitly assumes that after conditioning on the covariates, the missing data is missing at random, and kept covariate selection in the separate imputation models at a minimum in an attempt to avoid introducing biases. The estimated uncertainty took into account the fact that part of the data was imputed, and the results following multiple imputation were in line with those based on available data only (results not shown), although they on occasions gained statistical significance only after imputation.

Both multivariate regression and propensity score matching approaches rely on the potential sources of bias (i.e., characteristics of patients that influence both treatment selection and outcomes) being measured in the study and appropriately used as covariates in the statistical models. While care has been taken in the development of the multivariate linear regression framework, this approach is more limited in its abilities to control for confounding in the estimation of treatment effects in observational studies (e.g., although no statistically significant interactions between treatment cohort and covariates were identified, imbalances might still remain). The propensity scores approach directly minimizes imbalances between the treatment groups with respect to the covariates in order to reduce the bias in the estimation of treatment effects, and therefore is likely more appropriate for the estimation of cost-effectiveness. Therefore, while in the present study the estimated treatment effects on health outcomes and costs from the two approaches are very close, we summarize cost-effectiveness for the propensity scores approach only. The use of another statistical technique, instrumental variables (IV), can adjust for both observed and unobserved confounding variables by breaking the link between treatment choice and outcomes reported [28,29]. The use of IV was explored in this study, but appropriate instruments (which predict treatment selection but not outcomes) were not identified.

A principle of economic evaluation to support decision-making is that all relevant evidence is included. Ideally, this would include factoring in the results of randomized trials, including duloxetine and conservative therapies. Given the heterogeneity in these trials, the different outcomes measured, including the absence of utility data in the trials, and, in the case of the drugs trials, the selective patient populations, the synthesis of the trial and SUIT data was considered infeasible. The SUIT dataset has the advantage of greater generalizability and the inclusion of the major comparators, but the absence of the trial data from the analysis should be borne in mind when interpreting the results.

In conclusion, the SUIT study provides a rare opportunity to assess the costs, effects, and cost-effectiveness of a pharmaceutical as used in routine clinical practice in a large number of patients and against appropriate comparators. Although the limitations of the use of observational data for this purpose need to be acknowledged, the study suggests that initiating duloxetine therapy in stress urinary incontinence is a cost-effective treatment alternative.

Acknowledgments

The authors are grateful to the patients and physicians who participated in the SUIT Study, and to Dr Christian Falconer, who served as a member of the SUIT advisory board and contributed to the study design and analysis and interpretation of results.

Source of financial support: This study was funded by Eli Lilly and Company Ltd. and Boehringer Ingelheim GmbH. This analysis was undertaken by Oxford Outcomes Ltd. under a contract with Lilly.

Mark Sculpher, Douglas Tincello, Huub Van der Vaart, and Ralf Tunn received consultancy fees for time spent attending Advisory Boards to advise on the design, conduct, and analysis for this study, Louise Timlin, Deborah Quail, and Adam Johns are employed by Eli Lilly.

Borislava Mihaylova occasionally acts as paid consultant to pharmaceutical industry. She has received consultancy payments from Oxford Outcomes Ltd. Mark Sculpher is a director and shareholder of Oxford Outcomes Ltd. He has also received consultancy payments from Eli Lilly. Douglas Tincello has been an investigator/advisory board member for clinical trials funded by Eli Lilly (this study), Pfizer, and Ethicon (Johnson & Johnson). He has received research grants for investigator initiated studies from Astellas, ucb Pharma, and Ethicon, and has received financial support to attend scientific conferences from Pfizer, American Medical Systems, and Astellas. All consultancy payments and research grants have been managed by the University of Leicester Research and Business Office.

The publication of study results is not contingent on the approval of the sponsors. Proprietary Statement: Data, models and methodology used in the manuscript are nonproprietary.

References

1 Abrams P, Cardozo L, Fall M, et al. The standardisation of terminology in lower urinary tract function: report from the standardisation sub-committee of the International Continence Society. Urology 2003;61:37–49.
2 Hunskaar S, Lose G, Sykes D, Voss S. The prevalence of urinary incontinence in women in four European countries. BJU Int 2004;93:324–30.
3 Hunskaar S, Burgio K, Diokno A, et al. Epidemiology and natural history of urinary incontinence in women. Urology 2003;62(Suppl. 1):S16–23.
4 Vandoninck V, Bemelmans BL, Mazzetta C, et al. The prevalence of urinary incontinence in community-dwelling married women: a matter of definition. BJU Int 2004;94:1291–5.
5 Hannestad YS, Rortveit G, Sandvik H, Hunskaar S, Norwegian EsEoIitCoN-T. A community-based epidemiological survey of female urinary incontinence: the Norwegian EPINCONT study. Epidemiology of Incontinence in the County of Nord-Trondelag. J Clin Epidemiol 2000;53:1150–7.
6 Thom D. Variation in estimates of urinary incontinence prevalence in the community: effects of differences in definition, population characteristics, and study type. J Am Geriatr Soc 1998;46:473–80.
7 National Collaborating Centre for Women's and Children's Health. Urinary Incontinence—The Management of Urinary Incontinence in Women. London: Royal College of Obstetricians and Gynaecologists (RCOG), 2006.
8 Sykes D, Castro R, Pons ME, et al. Characteristics of female outpatients with urinary incontinence participating in a 6-month observational study in 14 European countries. Maturitas 2005; 52(Suppl. 2):13–23.
9 Papanicolaou S, Pons ME, Hampel C, et al. Medical resource utilisation and cost of care for women seeking treatment for urinary incontinence in an outpatient setting. Examples from three countries participating in the PURE study. Maturitas 2005;52(Suppl. 2):35–47.
10 Viktrup L, Bump RC. Pharmacological agents used for the treatment of stress urinary incontinence in women. Curr Med Res Opin 2003;19:485–90.
11 Mariappan P, Alhasso A, Ballantyne Z, et al. Duloxetine, a serotonin and noradrenaline reuptake inhibitor (SNRI) for the treatment of stress urinary incontinence: a systematic review. Eur Urol 2007;51:67–74.

12 Das Gupta R, Caiado M, Bamber L. An evaluation of the cost-effectiveness of duloxetine as a treatment for women with moderate-to-severe stress urinary incontinence. J Med Econ 2006;9:27–44.
13 Brunenberg DEM, Joore MA, Veraart CPWM, et al. Economic evaluation of duloxetine for the treatment of women with stress urinary incontinence: a Markov model comparing pharmacotherapy with pelvic floor muscle training. Clin Ther 2006;28: 604–18.
14 Van der Vaart H, Falconer C, Quail D, et al. Patient reported outcomes tools in an observational study of female stress urinary incontinence. Neurourol Urodyn 2010;29:348–53.
15 Royston P. Multiple imputation of missing values. Stata J 2004;4:227–41.
16 Kind P. The EuroQoL instrument: an index of health-related quality of life. In: Spilker B, ed. Quality of Life and Pharmacoeconomics in Clinical Trials (2nd ed.). Philadelphia, PA: Lippincott-Raven, 1996.
17 Dolan P, Gudex C, Kind P, Williams A. A Social Tariff for EuroQol: Results from a UK General Population Survey. Centre for Health Economics Discussion Paper 138. York: Centre for Health Economics, University of York: CHE, 1995.
18 Matthews JNS, Altman D, Campbell MJ. Analysis of serial measurements in medical research. BMJ 1990;300:230–5.
19 Unit Costs of Health and Social Care 2007. Canterbury, Kent: Personal Social Services Research Unit, University of Kent; http://www.pssru.ac.uk/uc/uc2007contents.htm 2007; accessed on March 3, 2008.
20 Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. Health Econ 2004;13:461–75.
21 Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. J Am Stat Assoc 1984;79:516–24.
22 Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41–55.
23 Becker SO, Ichino A. Estimation of average treatment effects based on propensity scores. Stata J 2002;2:358–77.
24 National Institute for Health and Clinical Excellence (NICE). Guide to the Methods of Technology Appraisal. London: NICE, 2008.
25 Tincello D, Sculpher M, Tunn R, et al. Patient characteristics impacting health state index scores, measured by the EQ-5D of females with stress urinary incontinence symptoms. Value Health 2010;13:112–18.
26 Sculpher M, Claxton K, Akehurst R. It's just evaluation for decision making: recent developments in, and challenges for, cost-effectiveness research. In: Smith PC, Ginnelly L, Sculpher M, eds. Health Policy and Economics. Opportunities and Challenges. Maidenhead: Open University Press, 2005.
27 Drummond MF. Experimental versus observational data in the economic evaluation of pharmaceuticals. Med Decis Making 1998;18(Suppl.):S12–18.
28 Hadley J, Polsky D, Mandelblatt JS, et al. An exploratory instrumental variable analysis of the outcomes of localized breast cancer treatments in a medicare population. Health Econ 2003; 12:171–86.
29 McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables [See comment]. JAMA 1994;272:859–66.