

DEPARTMENT OF ECONOMICS

**Comparing the cost-effectiveness of Haloperidol,  
Risperidone and Olanzapine in the treatment of  
schizophrenia using the net-benefit regression approach**

**Annemieke De Ridder & Diana De Graeve**

**UNIVERSITY OF ANTWERP**  
**Faculty of Applied Economics**



Stadscampus  
Prinsstraat 13, B.213  
BE-2000 Antwerpen  
Tel. +32 (0)3 220 40 32  
Fax +32 (0)3 220 47 99  
<http://www.ua.ac.be/tew>

# FACULTY OF APPLIED ECONOMICS

DEPARTMENT OF ECONOMICS

Comparing the cost-effectiveness of Haloperidol, Risperidone and Olanzapine in the treatment of schizophrenia using the net-benefit regression approach

**Annemieke De Ridder & Diana De Graeve**

RESEARCH PAPER 2007-012  
JUNE 2007

University of Antwerp, City Campus, Prinsstraat 13, B-2000 Antwerp,  
Belgium  
Research Administration – room B.213  
phone: (32) 3 220 40 32  
fax: (32) 3 220 47 99  
e-mail: joeri.nys@ua.ac.be

**The papers can be also found at our website:**

**[www.ua.ac.be/tew](http://www.ua.ac.be/tew)  
(research > working papers)**

**D/2007/1169/012**

## **Comparing the cost-effectiveness of Haloperidol, Risperidone and Olanzapine in the treatment of schizophrenia using the net-benefit regression approach**

**Annemieke De Ridder and Diana De Graeve**

**OBJECTIVES:** This study determines the cost-effectiveness of 3 antipsychotics for the treatment of schizophrenia in Belgium.

**METHODS:** Data were retrieved from a prospective observational non randomized follow-up survey. Clinical investigators included 293 schizophrenic patients; 136 of those patients were assigned to Olanzapine, 129 to Risperidone and 28 to Haloperidol. Patients were followed for 2 years. Total health care costs were determined from the perspective of the public payer and calculated by multiplying resource use with official tariffs; effectiveness of the drugs was measured with EQ-5D. Several studies have already compared the cost-effectiveness of different antipsychotics for the treatment of schizophrenia, most of them are however flawed by methodological issues. This study therefore uses a new method that was developed to address these limitations but is not widely used yet: the net-benefit regression approach (NBRA). We show its merits by performing a cost-effectiveness analysis of Olanzapine, Risperidone and Haloperidol.

**RESULTS:** Models were checked for selection bias but drug choice was not endogenous; we therefore proceeded with simple OLS regressions. The results indicate that the drugs provide similar net monetary benefits to the patient (H vs O -4452.53 ( $p=0.645$ ), R vs O 4439.54 ( $p=0.425$ ), R vs H 8892.07 ( $p=0.366$ )). When we control for several patient characteristics R moves away further from H and O but the difference does not reach statistical significance (R vs O 5857.73 ( $p=0.332$ ), R vs H 15233.53 ( $p=0.178$ )). Several important patient subgroups were also identified; they indicate that a drug performs better in a specific patient group. Numerous sensitivity analyses confirm the robustness of the results.

**CONCLUSION:** We conclude by confirming that the NBRA is an important enrichment to the CEA methodology. As was demonstrated in this paper, it is often important to correct cost-effectiveness results for patient characteristics and to identify significant patient subgroups.

## **Introduction**

Schizophrenia is a serious mental disease with an early onset mostly between the ages 15-25. It is characterised by distortions of thinking and perception with inappropriate or blunted personal and social behaviour as a result. Schizophrenia is usually a chronic disease with acute relapses. It is therefore expensive in terms of direct treatment costs due to re-hospitalisations and lifelong maintenance treatment [1,2]. The disease is burdensome for the family who spends much time care-giving [3]. It is hard for the individual who experiences a low quality of life and who has disadvantaged employment experiences resulting in low income and marginalisation [4].

Various anti-psychotics, differing in costs and effects, can be used for treatment. An optimal treatment choice is important for the patient's health and for the costs generated by the disease. Such an optimal choice should achieve a better balance between the resources expended and the resulting outcomes. Economic evaluation studies typically compare costs and effects between different treatment alternatives, and can assist in choosing. Many evaluation studies have already been performed [5,6], but most of them suffer from methodological problems. Part of the studies are attached to RCTs. These studies have the advantage that they are comparing the effects and costs of treatments for similar groups of patients (because of randomization). At the same time trials do not reflect everyday clinical practice and often suffer from other shortcomings, such as a relatively limited follow-up time (up to maximum 1 year), a small number of patients (mostly less than 100 patients), inadequate consideration of uncertainty, etc.. In everyday practice treatments are not random but chosen by the practitioner on the basis of patient characteristics and / or experience of the practitioner. A simple comparison of treatment costs and effects can then be misleading, since the results are distorted by patient and physician characteristics. Retrospective cohort-based evaluation studies often fail to account for baseline differences. Despite the relatively large number of evaluation studies already performed, Basu [5] therefore concludes further studies of cost-effectiveness need to be carried out with careful consideration of the limitations of published analyses.

## **Research question**

In the present study we perform a cost-effectiveness analysis (CEA) of Olanzapine (O) as compared to Risperidone (R) and Haloperidol (H) on the basis of a Belgian prospective observational non randomised 2-year follow-up survey. Many CEA's have been performed already but, as stated, they all suffer from methodological problems. In our analysis we will try to accommodate to the cited methodological shortcomings by using the net-benefit regression approach to take account of baseline patient differences and uncertainty.

## **Methods**

### **Data**

The Schizophrenia Outcomes Survey (SOS) is a prospective observational non randomised follow-up survey for assessing costs and effectiveness of 3 drugs for the treatment of patients with schizophrenia.

All patients entering the study were initiated either on H, O or R in a natural clinical setting. They were then followed every 3 to 6 months for a 24-month period after inclusion (recruitment started in 2000). There were 6 visits for each patient, unless the patient withdrew from the survey or was lost to follow-up.

Patients were included in the SOS study if they were treated as outpatients or hospitalised part time after a hospitalisation because of a first episode or a relapse of schizophrenia; if they were older than 18; if they were not participating in a clinical trial; if they had been assigned therapy with H, O or R (through a physician's decision); and if they had signed an informed consent authorizing electronic collection and analysis of anonymous data.

A total of 53 clinical investigators took part in the survey. They included 294 patients suffering from schizophrenia or schizophreniform disorders in the study. 137 of those patients were assigned to O, 129 were assigned to R and 28 patients to H. Patients were allowed to switch between the treatments. We therefore performed an intent-to-treat analysis. The last-observation-carried-forward technique was used to deal with missing values.

Costs were calculated from the perspective of the public payer. For each SOS patient, the costs of hospitalization (full-time and part-time), emergency visits (general and psychiatric hospitals), visits to a psychiatrist or a GP, and of psychotherapeutical sessions were computed over the two-year follow-up period. Costs were also computed for all A/P drugs (H, O and R), concomitant (antidepressant), corrective (anticholinergic) and other A/P drugs taken by the patient. All costs were added to yield a final total publicly paid cost. It should be noted that for hospitalisations a distinction was made between hospitalization in a general hospital and hospitalization in a psychiatric hospital. Costs were calculated, taking insurance status into account (not VIPO, VIPO, unknown). After 12 months of follow-up, the costs were discounted at a rate of 3%. Costs were calculated by multiplying resource use with official tariffs.

Several scales measured the effectiveness of the three medicines; the Clinical Global Impression scale on severity, improvement and therapeutical index (CGI, CGIS and CGIT scale respectively), the Global Assessment of Functioning scale (GAF scale), the Brief Psychiatric Rating Scale (BPRS), the EQ-5D scale and the Subjective Well-being under Neuroleptics scale (SWN scale). We decided to work with EQ-5D because it is a well-known scale which is used in many CE analyses. Moreover, the EQ-5D index generates preference-based scores for Health Related Quality of Life (HRQoL) that can be used as weights for calculating QALY's [7]. A literature review revealed that EQ-5D generates reasonable construct and discriminative ability and is sensitive in detecting changes in QoL when schizophrenic patients are considered [7,8,9]. Therefore EQ-5D is believed to be a leading choice when determining the cost-effectiveness of interventions in schizophrenia [9]. We will consider the evolution of EQ-5D during the 2 years of follow up. Sensitivity analyses were performed using other scales.

### **Analysis**

Knowledge of the evolution of costs and effects is interesting but usually not sufficient. What can be concluded when one treatment is more expensive than another but also more effective? Cost-effectiveness of the 3 treatment drugs should be determined. Traditionally the ICER (incremental cost-effectiveness ratio) is used in CEA. The ICER reveals which

of the treatments is most cost-effective and consequently policy decisions can be made. However, researchers have to be careful in interpreting ICER's because many problems with the interpretation of this ratio have become evident: The sign of the ICER has to be interpreted carefully, the value of the ratio is not sufficient to give explicit treatment recommendations and statistical problems can arise when interpreting sampling uncertainty. Moreover, the ICER does not take into account patient characteristics. Since our data were collected in a non-randomized naturalistic survey controlling for baseline characteristics is very important.

Because of these problems with the ICER new methods for performing CEA's have been developed. Stinnet and Mullahy [10] developed the net-benefit framework. This new framework reformulates the CE problem to generate a linear net-benefit statistic. The linearity of the statistic offers important advantages because the interpretation is straightforward and it simplifies the handling of uncertainty in CEA. However, the net-benefit framework does not take into account patient characteristics. Therefore Hoch et al. [11] take the new approach a step further; they employ the linearity of the net-benefit statistic to apply regression methods on economic evaluation. This new approach has many advantages: researchers will be able to adjust for imperfect randomisation in trials and to use conventional econometric techniques. The key advantage of the regression approach is the ability to use standard regression techniques to examine the marginal impact of covariates on incremental CE instead of the usual approach of aggregating cost and effect differences across the arms of the trial. Therefore, when using this new technique, researchers can control for differences in patient groups at baseline and they can identify important subgroups [11].

The further course of this paper will be as follows: In order to compare the traditional CEA with the new methods we will first perform a traditional CEA by calculating the ICER's. Secondly we will calculate the net monetary benefits as suggested by Stinnet and Mullahy [10]. Finally, we will apply the new net-benefit regression technique to the SOS data. We will include important covariates and interaction effects in the model.

### **Incremental Cost Effectiveness Ratio (ICER)**

Denoting the expected values of cost and effect of H, R and O  $\mu_{CH}$ ,  $\mu_{CR}$ , and  $\mu_{CO}$  and  $\mu_{EH}$ ,  $\mu_{ER}$  and  $\mu_{EO}$  respectively; the ICER's are defined as follows (E.g. for the comparison of R and O):

$$ICER_{RO} = \frac{\mu_{CR} - \mu_{CO}}{\mu_{ER} - \mu_{EO}} = \frac{\mu_{\Delta C}}{\mu_{\Delta E}} = \pi$$

If  $\lambda$  represents the maximum WTP per unit of effect and  $\pi < \lambda$ , we can conclude that the health improvement generated by R is worth its cost.

When we want to estimate the ICER using trial data we use sample means for  $\Delta C$  and  $\Delta E$ :

$$\square ICER_{RO} = \frac{\Delta \hat{C}}{\Delta \hat{E}}$$

### **Net-benefit framework**

In the net-benefit framework  $\Delta E$  is given a monetary value, using  $\lambda$ . Through a simple reformulation of the ICER we find:

$$NMB_{RO} = \lambda \mu_{\Delta E} - \mu_{\Delta C}$$

NMB = net monetary benefit

If  $NMB_{RO} > 0$ , it can be concluded that the additional benefits generated by R are worth its costs, given the value of  $\lambda$ .

We will estimate NMB using the trial data:

$$\square NMB = \lambda \cdot \Delta \hat{E} - \Delta \hat{C}$$

Because of the linearity of the NMB it is straightforward to calculate the variance and the (1- $\alpha$ ) % CI:

$$\text{var}(\overline{NMB}) = \lambda^2 \text{var}(\Delta\hat{E}) + \text{var}(\Delta\hat{C}) - 2\lambda \text{cov}(\Delta\hat{E}, \Delta\hat{C})$$

$$CI = \overline{NMB} \pm z_{\frac{\alpha}{2}} \sqrt{\sigma_{\overline{NMB}}^2}$$

A possible disadvantage of the net-benefit framework is that it is a function of  $\lambda$ , the maximum WTP per unit of health gain, a value that is unknown in most cases. Stinnet and Mullahy however consider this as a strength because it forces explicit consideration of the value of  $\lambda$ . They emphasise the importance of sensitivity analyses and the estimation of cost-effectiveness acceptability curves (CEAC's) to examine different values of  $\lambda$  [10].

Because we work with the EQ5D-scale as the effectiveness measure in this paper, we can use a ratio threshold per QALY for  $\lambda$ . Based on the research of Owens [12] and Salomon [13], we decided to work with  $\lambda = \text{€}40\,000$ .

### **Net-benefit regression approach (NBRA)**

In this paper we evaluate the CE of H, O and R. For ICER's and NMB's only 2 treatments can be compared simultaneously. For the NBRA however, we can compare several pairs of treatment at the same time.

We will first illustrate the NBRA as developed by Hoch et al.[11]: comparing only 2 treatments, using a regression equation with one treatment dummy. We chose to compare R and O in this first analysis.

Secondly, we elaborate the model in order to compare H with O and H with R. This is done by adding a treatment dummy to the model.

In our first analysis, the comparison of 2 drugs, it makes sense to compare R and O. H is a typical drug in the treatment of schizophrenia; R and O are more recently developed, atypical drugs. Many researchers have investigated and compared the cost-effectiveness of atypical drugs but the results are contradictory. In using the data from the SOS study we can compare the CE of R and O for the Belgian context.

To perform the NBRA we have to define a net-benefit value for each subject:

$$NMB_i = \lambda E_i - C_i$$

That way we can calculate the average individual net-benefit of each treatment drug. E.g.

$$\overline{NMB}_R \text{ and } \overline{NMB}_O$$

A simple linear model for  $NMB_i$  can then be constructed in the following way:

$$NMB_i = \alpha + \delta t_i + \varepsilon_i \quad \text{MODEL 1}$$

$t_i$  = treatment dummy;  $t_i = 0$  if patient takes O

$t_i = 1$  if patient takes R

$\delta$  provides an estimate of the average incremental net-benefit of R over O:  $\overline{NMB}_{RO}$

because:

$$\begin{aligned} \overline{NMB}_{RO} &= \lambda \Delta \bar{E} - \Delta \bar{C} \\ &= \lambda (\bar{E}_R - \bar{E}_O) - (\bar{C}_R - \bar{C}_O) \\ &= (\lambda \bar{E}_R - \bar{C}_R) - (\lambda \bar{E}_O - \bar{C}_O) \\ &= \overline{NMB}_R - \overline{NMB}_O \end{aligned}$$

The power of this approach is that it is straightforward to add additional explanatory variables in order to examine their impact on CE directly.

$$NMB_i = \alpha + \delta t_i + \sum_{j=1}^p \beta_j X_{ij} + \varepsilon_i \quad \text{MODEL 2}$$

We have included  $p$  covariates in the model<sup>1</sup>. These covariates contain important patient and investigator characteristics.

$\delta$  now provides an estimate of the average incremental net-benefit of R over O controlling for confounding variables.

The coefficients that are generated for the covariates in this model describe the impact on average net-benefits. Although average net-benefits are useful as a basis from which to obtain incremental benefits, they are not helpful for decision making on their own,

---

<sup>1</sup> In appendix a list of covariates that were included in the regression models can be found.

because they suffer from the same limitations as those associated with average CE ratios. The impact of covariates on the incremental net-benefit of a treatment is of fundamental interest. This impact can be examined by adding interaction effects to the model [11].

$$NMB_i = \alpha + \delta t_i + \sum_{j=1}^p \beta_j X_{ij} + t_i \cdot \sum_{j=1}^p \gamma_j X_{ij} + \varepsilon_i \quad \text{MODEL 3}$$

In this model significant  $\gamma_j$ 's reveal important patient subgroups; the magnitude of  $\gamma_j$  indicates how the CE of R is expected to vary at the margin.

After this first analysis we elaborate the model to compare H with O and H with R. This is done by adding a second treatment dummy to the model.

$$NMB_i = \alpha + \delta_1 tR_i + \delta_2 tO_i + \varepsilon_i \quad \text{MODEL 4}$$

$tR_i$  = treatment dummy;  $tR_i = 0$  if patient takes O or H

$tR_i = 1$  if patient takes R

$tO_i$  = treatment dummy;  $tO_i = 0$  if patient takes R or H

$tO_i = 1$  if patient takes O

$\delta_1$  = average incremental net-benefit of R over H

$\delta_2$  = average incremental net-benefit of O over H

As in our first analysis we can add covariates and interaction effects.

$$NMB_i = \alpha + \delta_1 tR_i + \delta_2 tO_i + \sum_{j=1}^p \beta_j X_{ij} + \varepsilon_i \quad \text{MODEL 5}$$

$$NMB_i = \alpha + \delta_1 tR_i + \delta_2 tO_i + \sum_{j=1}^p \beta_j X_{ij} + tR_i \cdot \sum_{j=1}^p \gamma_{1j} X_{ij} + tO_i \cdot \sum_{j=1}^p \gamma_{2j} X_{ij} + \varepsilon_i \quad \text{MODEL 6}$$

In this study patients were not randomly assigned to a specific drug, the drug choice is a decision of the investigators. Therefore our estimates need to account for possible endogeneity of the variable drug choice. We can correct for this by estimating a treatment effects model rather than using OLS.

## **Uncertainty**

Another major advantage of the NBRA is that it can easily handle uncertainty, which is not the case with ICER. The usual way to present uncertainty is to calculate confidence intervals (CI's). For ICER's this can however be problematic since ratios with similar signs but different interpretations should not be grouped together. The best way to present uncertainty of ICER's is to draw the CEAC. The CEAC is a method for summarizing the uncertainty in estimates of CE by presenting the probability that an intervention is CE compared with an alternative for a range of  $\lambda$  values [14].

When we apply the NBRA it becomes straightforward to illustrate uncertainty. Since the NMB is a linear statistic it is straightforward to calculate CI's and to draw CEAC's. In the results section we will illustrate the construction and interpretation of CEAC's.

## **Results**

In the SOS study 294 schizophrenic patients were followed. During the two years of the survey 75 patients withdrew from the study or were lost to follow up. Analyses were performed for all patients (293 because cost data are missing for one patient). In the regression analyses we can correct for the dropouts by adding the variable *follow*<sup>2</sup> to the model. The standard CEA, using ICER's, was also performed for all 293 patients to be able to compare the results with the NBRA. In table 1 some important baseline characteristics of the patients are summarized. Important descriptive statistics concerning the investigators are presented in table 2. It is clear from these tables that the 3 patient groups are quite divergent concerning several baseline characteristics, e.g. sex, social status, employment, ambulatory practice of the psychiatrist, etc.. It is therefore very important to control for these characteristics when performing a CEA.

---

<sup>2</sup> This variable represents the total number of days the patient was followed.

Table 1: Baseline characteristics of the patients

		<b>H (n = 28)</b>	<b>O (n=136)</b>	<b>R (n=129)</b>	<b>P value* (O vs R)</b>
Sex (% of male patients)		78.57%	63.97%	59.69%	.168 (.473)
Age (years)		42	36	35	.092 (.536)
Age at first psychiatric treatment		26	26	25	.362 (.194)
# of days patient was followed		597	629	640	.573 (.658)
Age dif		1.36	1.32	1.79	.598 (.331)
GAF score		46.93	52.26	52.73	.123 (.787)
% of stable patients		60.71%	76.47%	71.32%	.210 (.339)
% of VIPO patients		57.14%	36.03%	25.58%	.003 (.078)
Social status	Single	39.29%	26.47%	22.48%	.006 (.793)
	Couple	3.57%	11.03%	11.63%	
	Family	14.29%	43.38%	48.84%	
	Community	42.86%	19.12%	17.05%	
Living environment	At home alone	25.00%	19.85%	19.38%	.003 (.693)
	At home family	17.86%	52.94%	60.47%	
	Sheltered living	28.57%	18.38%	13.18%	
	Institutionalized	17.86%	5.15%	4.65%	
	Other	10.71%	3.68%	2.33%	
Employed		3.57%	14.71%	17.05%	.188 (.601)
Education	None	3.57%	2.21%	0.78%	.860 (.931)
	Primary school	14.29%	14.71%	23.26%	
	Secondary school	50.00%	60.29%	44.19%	
	Higher education	25.00%	15.44%	29.46%	
	other	7.14%	7.35%	2.33%	
Type of psychosis	Catatonic	0%	0.74%	3.10%	.128 (.722)
	Disorganized	10.71%	17.65%	14.73%	
	Paranoid	60.71%	59.56%	55.04%	
	Indifferent	3.57%	4.41%	5.43%	
	Résiduelle	17.86%	2.94%	3.88%	
	Not specified	0%	2.21%	1.55%	
	Schizophreniform troubles	7.14%	12.50%	16.28%	
First episode (first treatment)		10.71%	12.50%	27.91%	.003 (.002)

\*p values correspond to the comparison of the three groups, between brackets O versus R

Table 2: Baseline characteristics of the investigators

		<b>H (n = 28)</b>	<b>O (n = 136)</b>	<b>R (n = 129)</b>	<b>P value* (O vs R)</b>
Age		45	48	46	.055 (.050)
Experience (years)		14	18	17	.018 (.140)
Male		96.43%	71.32%	78.29%	.014 (.192)
Hospital practice	Psychiatric	82.14%	69.85%	85.27%	.009 (.003)
	General	17.86%	30.15%	14.73%	
Ambulatory practice	General hospital	28.57%	47.06%	34.11%	.045 (.032)
	Psychiatric hospital	71.43%	57.35%	71.32%	.044 (.018)
	Mental health unit	53.57%	25.00%	26.36%	.008 (.800)
	Private practice	82.14%	86.03%	82.95%	.747 (.488)

\*p values correspond to the comparison of the three groups, between brackets O versus R

### **Cost and effectiveness**

Based on the SOS dataset, total 2-year costs were calculated with STATA 9. Results can be found in table 3.

According to our calculations the patient group taking H generates more direct medical costs than the patient groups taking O and R. The difference in cost between O and R is small. O is the cheapest group.

The effectiveness of the 3 drugs was measured by the EQ-5D index. We will consider the evolution in EQ-5D score of the patients over the 2 years of follow up, which means that we consider the area under the curves in figure 1. The average areas under the curves for patients on H, O and R can be found in table 3. The health condition of the patients taking R and H (as measured with EQ-5D) are very similar, that of patients taking O is considerably lower. This is counterintuitive, in the literature it is usually found that R and O are the most effective drugs, while H is less effective. Here we have a reverse ordering. This can probably be explained by the small sample size of the H group. Only 28 patients in our sample take H, of whom only 18 patients are stable. We believe that this small sample is not representative for the general population group that takes H; these patients are usually older and sicker patients. We will therefore focus mainly on the comparison of patients taking R and O which have reasonable sample sizes. We also performed a number of sensitivity analyses in which we use different scales of effectiveness. Results of these analyses will be presented further in the paper.

Figure 1: Evolution of the EQ5D-index of the patients over the 24 months of follow up

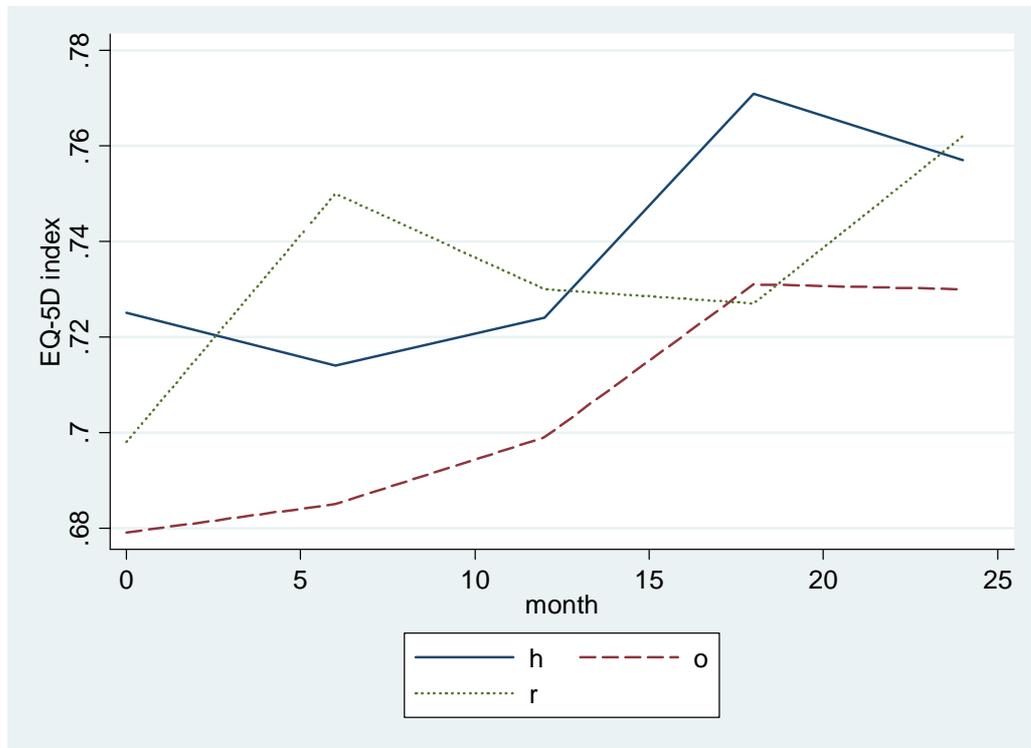


Table 3: average costs and effectiveness of the patients after 2 years of follow-up.

	<b>H</b> <b>(n = 28)</b>	<b>O</b> <b>(n = 136)</b>	<b>R</b> <b>(n = 129)</b>
<b>Total costs (€ 2002)</b>	30483.77	20569.69	20915.33
<b>Qaly (EQ-5D)</b>	2.95	2.81	2.93

### ICER

When we compare R and O we find the following ICER:

$$ICER_{RO} = \frac{20915.33 - 20569.65}{2.934019 - 2.81439} = 2889.60$$

The result shows that R improves patients' health over O at a price of €2889.60 per QALY. This is worthwhile paying for, if  $\lambda$ , the maximum willingness to pay per unit of health gain is higher. We have fixed the value of  $\lambda$  at €40000;  $2889.60 < 40000$ , consequently, R should be chosen.

To be able to compare the ICER results with the NBRA we have also calculated the ICERs for the comparison of O and H and R and H.

$$\text{ICER}_{\text{HO}} = \frac{30483.77-20569.65}{2.950929-2.81439} = 72610.20$$

$$\text{ICER}_{\text{HR}} = \frac{30483.77-20915.33}{2.950929-2.934019} = 565845.00$$

In both cases we find an ICER that exceeds  $\lambda$  (=€40 000). This means that both O and R are more CE than H.

We conclude that R is the most CE drug, followed by O and H respectively.

### **Net-benefit framework**

To illustrate the NMB approach we have calculated the respective NMB's for the 3 drugs.

$$\begin{aligned} \text{NMB}_{\text{RO}} &= \lambda*(2.934019 - 2.81439) - (20\ 915.33 - 20569.65) \\ &= 40000*(0.119629) - (345.68) \\ &= 4439.48 \end{aligned}$$

The net-benefit of R over O is positive, which means that R is more CE than O.

$$\begin{aligned} \text{NMB}_{\text{RH}} &= \lambda*(2.934019 - 2.950929) - (20\ 915.33 - 30483.77) \\ &= 40000*(-0.01691) - (-9568.77) \\ &= 8892.37 \end{aligned}$$

The net-benefit of R over H is also positive, so R is more CE than H.

$$\begin{aligned} \text{NMB}_{\text{OH}} &= \lambda*(2.81439 - 2.950929) - (20569.65 - 30483.77) \\ &= 40000*(-0.136539) - (-9914.12) \\ &= 4452.56 \end{aligned}$$

The net-benefit of O over H is €4314.12; this means that O is more CE than H.

Because we do not want our results to depend too strongly on the choice of  $\lambda$  we performed a sensitivity analysis by drawing the CEAC in which we vary the value of  $\lambda$ . Results can be found in the section on sensitivity analyses.

### **Net-benefit regression approach**

Before estimating our different models we need to check the possibility of selection bias (endogeneity of the variables  $t$ ,  $tR$  and  $tO$ ). When selection bias is found we need to estimate treatment effect models instead of OLS models. We instrument drug choice by the sex, language and age of the investigator and by the sex and the age of the patient. We first verified whether these instruments are valid by calculating Hansen J statistics; validity of the instruments was confirmed. We then tested whether  $t$ ,  $tO$  and  $tR$  are endogenous in our models by estimating treatment effect models and calculating the rho statistic. For none of our models the variable of drug choice was endogenous ( $\rho=0$  for all models), we can therefore proceed by performing simple OLS regressions<sup>3</sup>.

Table 4 summarizes the regressions results comparing O and R for models 1, 2 and 3. For model 1 we find that the coefficient of  $t$  amounts to 4439.54. This means that the incremental net benefit of R over O is €4439.54. Note that this corresponds with the NMB found in the previous section. Consequently it is confirmed that R is more CE than O. The difference in NMB is however not significantly different from zero.

In model 2 some baseline characteristics of the patients were added. Again there is no significant effect of the variable drug choice. After correction for the covariates it does become more apparent that R is the most CE drug since the incremental net benefit of R over O rises to 5143.19 ( $p=0.391$ ). A number of covariates have a significant influence on the NMB of the patient. The variable *stable* has a significantly positive influence on NMB, which means that stable patients (who do not switch between the 3 drugs) have higher net benefits than the switchers. The sex of the patient also has a significant influence on the average net benefits: male patients have higher net benefits than female

---

<sup>3</sup> More detailed results can be retrieved from the authors

patients. Patients that are entitled to higher reimbursements (*high\_reimb*) and patients that have been partially hospitalized in the year before the survey (*prev\_hosp*) have significantly lower net benefits. The income variable also has a significant influence: patients with a higher monthly income at baseline have higher average net benefits. The characteristics of the investigator were also included in the model. We found that *priv\_prac* has a significant positive influence on the net benefits of the patients. Patients that consult a therapist with a private practice have higher average net benefits than patients who visit a therapist in a hospital or in a mental health unit.

In model 2 we find the covariates that significantly influence the average net benefits of the patients. It is however more interesting to investigate whether one drug performs better than another for a certain patient group. That is why we added interaction effects in model 3. To avoid problems of multicollinearity, only interaction effects of the significant covariates ( $p < 0.05$ ) of model 2 were included. In the third model the coefficient  $\delta$  further rises to 7343.58 ( $p = 0.725$ ). None of the interaction effects has a significant influence, there are therefore no significant subgroups found.

Table 4: Regression results of the comparison of R and O

	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>
<b>Constant term</b>	92005.9	55504.30	55526.25
<b>Treatment dummy</b>	4439.54	5143.19	7343.58
<b>Covariates</b>			
follow		-7.68	-8.78
stable		17325.70**	17263.14**
age_dif		135.36	-49.13
male		24653.25***	30117.61***
high_reimb		-12745.59**	-12766.06*
prev_hosp		-14396.40*	-11079.24
income		5988.67***	5645.98***
age_psy		30.08	60.19
Psy_hosp		-1211.98	-1567.62
Priv_prac		20387.34*	16486.50
<b>Interaction effects</b>			
t_male			-11325.32
t_prev_hosp			-7076.42
t_priv_prac			8180.91
<b>R-squared</b>	0.0024	0.2361	0.2414
<b>F</b>	0.64	5.78	4.62
<b>Prob &gt;F</b>	0.424	0.000	0.000

\* significant at the 10% level

\*\* significant at the 5% level

\*\*\* significant at the 1% level

### H versus O versus R

Table 5 summarizes the results of models 4, 5 and 6. From the results of model 4 we derive that the average incremental net benefit of R over H is 8892.07; the average incremental net benefit of O over H is 4452.53. The incremental net benefit of R over O is then the difference between 8892.07 and 4452.53, which is 4439.54 (also found in model 1). These amounts correspond with the NMB's found in the previous section. None of the coefficients is significantly different from zero; there are therefore no significant differences between the 3 drugs according to the net monetary benefit they provide to the patient.

In model 5 baseline characteristics were added. The incremental net benefit of R over H increases to 15233.53, the incremental net benefit of O over H increases to 9375.80, the net benefit of R over O rises to 5857.73. It is clear that R moves away from O and H and is evidently the preferred treatment option. The results on covariates correspond to those in model 2; *high\_reimb* and *prev\_hosp* have a significantly negative influence on the

average net benefits of the patients; *stable*, *male*, and *income* positively influence the net benefits.

To identify subgroups interaction effects are added in model 6. To avoid problems of multicollinearity, some covariates were dropped and only interaction effects of the significant covariates in model 5 were included. We again find that *stable* and *high\_reimb* significantly influence the net-benefits. Further 3 significant interaction effects are found. The coefficient of *tR\_prev\_hosp* is significantly negative which means that there is a significant interaction between the fact whether the patient has been partially hospitalised during the year before the survey and the use of R. It means that patients that were previously hospitalised incur a lower net benefit from R than patients that were not previously hospitalised. The other significant interaction effects are *tO\_male* and *tR\_male*, the coefficients are positive. This means that male patients incur more benefits from O and R than female patients. It is therefore important to treat these subgroups (prev\_hosp versus no prev\_hosp and male versus female patients) separately.

Table 5: regression results of the comparison of H, O and R

	<b>Model 4</b>	<b>Model 5</b>	<b>Model 6</b>
<b>Constant term</b>	87553.37	51320.49	75653.81
<b>Treatment dummies</b>			
tR	8892.07	15233.53	5303.86
tO	4452.53	9375.80	-7433.25
<b>Covariates</b>			
follow		-12.37	-8.97
stable		19168.97***	19693.69***
age_dif		68.75	-314.71
male		27411.43***	
high_reimb		-12400*	-14441.89**
prev_hosp		-16834.00**	
income		4908.63**	
age_psy		104.76	-55.73
Psy_hosp		-410.28	-586.73
Priv_prac		13251.05	11492.65
<b>Interaction effects</b>			
tO_male			30254.08***
tR_male			24715.26***
tO_prev_hosp			-8574.16
tR_prev_hosp			-24416.81**
<b>R-squared</b>	0.0040	0.1995	0.1995
<b>F</b>	0.56	5.89	5.89
<b>Prob &gt;F</b>	0.573	0.000	0.000

\* significant at the 10% level

\*\* significant at the 5% level

\*\*\* significant at the 1% level

## Sensitivity analyses

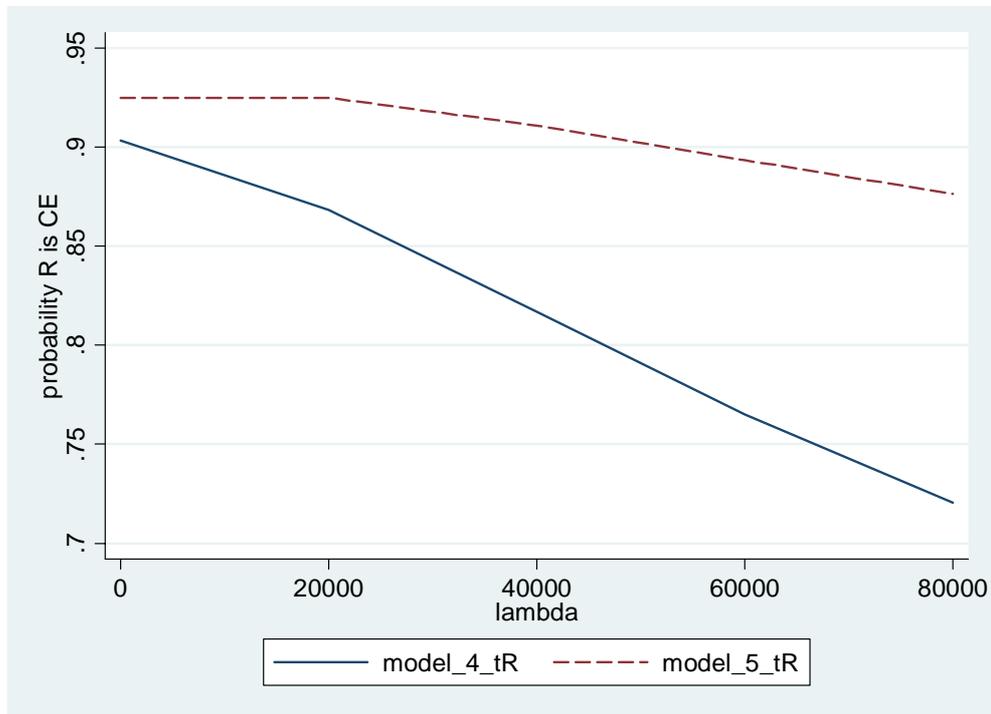
### *Uncertainty*

As stated before it is very important to present uncertainty results when performing a CEA. Uncertainty is usually illustrated by calculating CI's and drawing CEAC's. For the NBRA it is however unnecessary to calculate CI's because we have significance results (p-values) for all parameters in the model.

Because it is important to explore the sensitivity of our results to the chosen  $\lambda$  (€40 000), we constructed CEAC's in which the value of  $\lambda$  is varied from €0 to €80 000.

When we construct the CEAC for model 1 and 4 this is identical to constructing the CEAC for the ICER's, but far less time consuming.

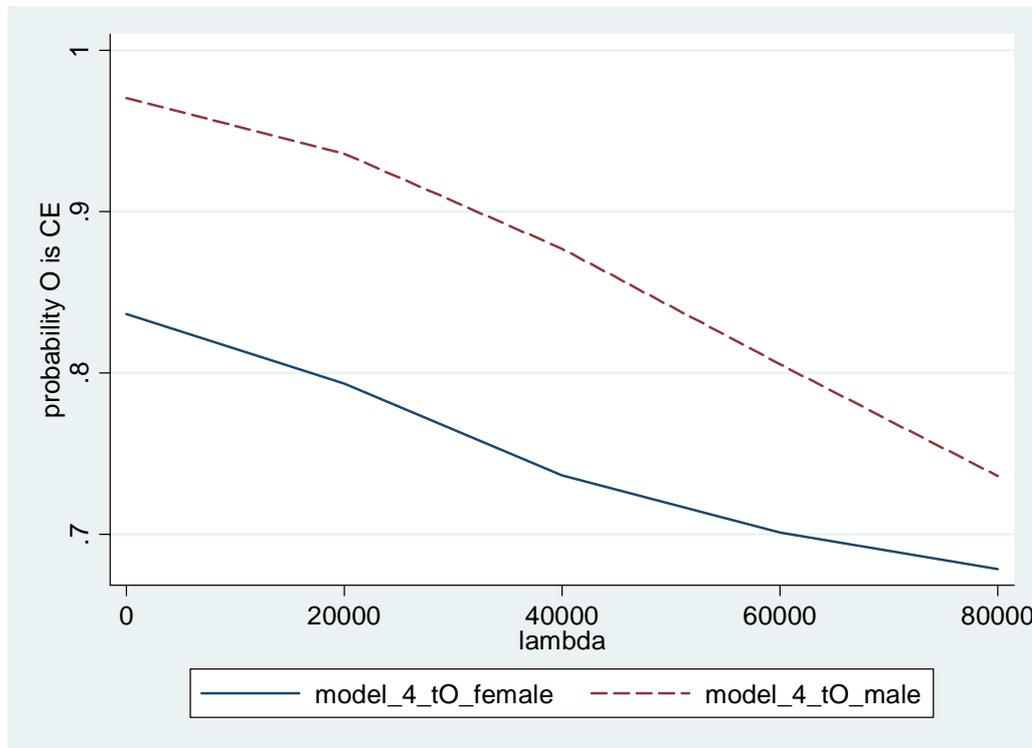
Figure 2: CEAC for R in comparison to H: models 4 and 5



In figure 2 the CEAC's of the comparison of R and H are presented for both model 4 (solid line) and 5 (dashed line). The probability that R is more CE than H declines in  $\lambda$ ; this is intuitive since H is the most effective drug. Although we can derive here that R is better value for money than H, the difference between R and H is never significantly different from zero, as the CEAC's never rise above 95%. In model 4 (solid line) no covariates are taken into account, when  $\lambda$  is 40 000, the probability that R is more CE is about 82%, when  $\lambda$  rises to 80 000, the probability decreases to 73%.

Using the NBRA we were able to introduce patient characteristics in the models. From figure 2 it is obvious that the introduction of these covariates has an important impact on the CE of R (dashed line). By correcting for the selected covariates the CE of R is higher, for example, when  $\lambda$  is 80 000, the probability is about 88%.

Figure 3: CEAC for O in comparison to H: stratified analysis for male and female (model 4)



A very important advantage of the NBRA is that subgroups can be identified for which the CE of a particular treatment is different. In our analysis we found that there is a significant interaction effect for the variable  $tO\_male$ . We can illustrate this effect with a CEAC. We performed stratified regression analyses for the variable *male*. Figure 3 illustrates the CEAC's of male (dashed line) and female patients (solid line).

It is clear from figure 3 that the sex of a patient is of great importance for the CE of O compared to H. For both male and female patients the CE of O declines as the WTP for an extra unit of health rises. However, for male patients the probability of O being more CE than H is much higher (varies between 97% and 74%) than for female patients (varies between 84 and 69%). Consequently it is important to perform separate analyses and to treat male and female patients differently.

### *Justification of the use of LOCF*

Because a number of patients drop out the study during the two years of follow up, we had missing values in our dataset. To account for these missing values we used the well known last-observation-carried-forward technique (LOCF). To justify the use of this technique we performed two sensitivity analyses; we performed the NBRA with the original dataset (with the missing values) and we performed the NBRA with standardized costs and effects (standardized for 1 year). Both analyses provide strongly resembling results compared to the original analysis described earlier. The sign, order of magnitude and significance of the coefficients is almost always alike, only a few times a non-significant variable has another sign and a few times a variable is significant while it was not in the original analysis or vice versa<sup>4</sup>. We can therefore be confident that the use of LOCF does not bias our results.

### *Justification of the use EQ-5D as effectiveness measure*

As described and argued in the methods section we chose to work with the evolution of the EQ-5D index values over the two years of follow up as the effectiveness measure in our study. To test the robustness of our results we carried out a number of sensitivity analyses in which we performed the NBRA for several other effectiveness measures, namely the gain in EQ-5D index between visit 1 and 6, the EQ-5D index at visit 6, the EQ-5D VAS score, BPRS and GAF score. The results of these sensitivity analyses are very similar to those of the original analysis, concerning the sign, order of magnitude and significance of the coefficients<sup>4</sup>. We found that some variables were significant while they were not in the original analysis, or vice versa. This however never happened systematically, except for the variable ‘follow’. For all but one of the sensitivity analyses the coefficient for the variable ‘follow’ was significantly different from zero; patients that were longer in the study have lower net benefits. This effect is not captured when we use the evolution in EQ-5D score.

It was also striking that for the EQ-5D VAS score the coefficient for the variable tR in model 6 was significantly different from zero, which means that R incurs significantly higher net-benefits than H when we control for covariates and interaction effects. This

---

<sup>4</sup> Detailed results of these sensitivity analyses can be retrieved with the authors

was not found in the original analysis. Further it sometimes happens that a coefficient has another sign, but this never happens for a significant variable.

We conclude that, using the evolution in the EQ-5D index as the effectiveness measure, our results are robust when we compare them with other measures.

## **Discussion**

In this paper we investigate the cost-effectiveness of Haloperidol, Olanzapine and Risperidone, 3 drugs for the treatment of schizophrenia, for Belgium. Because many problems have arisen with the interpretation of ICER's, the usual tool for CEA, we decided to use the new approach developed by Hoch et al. [11]. Main advantages of this new methodology are that patient characteristics are taken into account when calculating the incremental net benefits of a drug and that handling uncertainty becomes straightforward.

A literature review revealed a large ambiguity in the results of CEA's concerning H, O and R. None of the studies however takes patient characteristics into account when interpreting CEA results [5,6]. In our study great cost differences are found between the patient groups. The patient group taking H generates about €10 000 more direct costs than the patient groups taking O and R (over the 2 years of follow-up). The effectiveness results are less divergent. It was found that, according to the EQ5D scale, the effectiveness of R and H is similar, while O is obviously less effective. Neither cost differences, nor effectiveness differences are significant however.

The main result of this study is that furthermore, there is no significant difference between the 3 drugs concerning cost-effectiveness. This is counterintuitive as the literature indicates that O and R are more CE than H. As stated before, this unexpected result can probably be explained by the small sample of H-patients we have in our study. Only 28 patients take H; of them only 18 are stable patients. The fact that O and R perform similarly is in accordance with a number of other studies. We point out that all 3 analyses, ICER, NMB and NBRA suggest that R is more CE than O; the NBRA however indicates that the difference in CE between R and O is not significantly different from zero.

Several patient characteristics influence the incremental net benefits of the 3 drugs, it is therefore important to control for them. When patient characteristics are taken into account in the NBRA, R moves further away from the other 2 drugs. The difference however never reaches statistical significance at conventional levels: for a WTP of 40000 per QALY ( $\lambda = \text{€}40000$ ) the probability that R is more CE than O is about 80%, the probability that R is more CE than H is about 91%. We also identified several important patient subgroups. Patients that were partially hospitalized in the year before the survey obtain lower net-benefits from R than patients that were not; male patients perform better on O and R than female patients.

An important limitation to this study is that all regression analyses were performed using standard OLS. More efficient estimation results will be obtained by using robust regression techniques and by correcting for possible skewness in the data. In our future research we will analyse these data with more sophisticated techniques and verify whether this changes our results. The objective of this paper was to investigate the CE of the 3 drugs controlling for patient characteristics and not to solve methodological issues. A second limitation is that, like many CEA's for schizophrenia, we use an intent-to-treat approach to account for patients that switch between the drugs. Several sensitivity analyses however indicate that this will not bias our results. Recently a new approach has been developed to deal with switching patients, the epoch approach [15]. This approach could be used to adjust our results.

We can conclude by confirming that the net-benefit regression approach is an important enrichment to the CEA methodology. It is often very important to correct CE results for patient characteristics and to identify significant patient subgroups. In our analyses we found indications that patient and physician characteristics influence the cost-effectiveness of schizophrenia treatment. We did however not find a statistically significant difference in the CE between H, R and O for the treatment of schizophrenia, in general; significant differences were found for some specific subgroups. The enormous variability in costs, together with a relatively large number of patient characteristics and our small sample size preclude a more decisive conclusion.

## **Acknowledgements**

The study was funded and followed up by Eli Lilly Belgium and designed by the study committee (J. Peuskens, UPC KULeuven, Campus Sint-Jozef Kortenberg; D. De Graeve, Universiteit Antwerpen; M. Declercq; UCL Bruxelles; P. Gilis; Eli Lilly Belgium). Data have been made available by Eli Lilly for the statistical analyses in this paper.

We would like to thank Dr. Irina Cleemput for reading and commenting in this paper.

## Bibliography

1. Knapp M, Chisholm D, Leese M, Amaddeo F, Tansella M, Schene A, Thornicroft G, Vazquez-Barquero J-L, Knudsen H-C, Becker T, and the EPSILON Study Group. Comparing patterns and costs of schizophrenia care in five European countries: the Epsilon study. *Acta Psychiatrica Scandinavica* 2002; 105:42-54.
2. De Hert M, Thys E, Boydens J, et al. Health care expenditure on schizophrenia patients in Belgium. *Schizophr Bull* 1998;24:519–27.
3. Lowyck B, De Herdt M, Peeters E, Wampers M, Gilis P, Peuskens J. A study of the family burden of 150 family members of schizophrenic patients. *European Psychiatry* 2004; 19: 394-401.
4. Percudani M, Barbui C, Tansella M. Effect of Second-Generation Antipsychotics on Employment and Productivity in Individuals with Schizophrenia. An Economic Perspective. *Pharmacoeconomics* 2004; 22(11): 701-18.
5. Basu A. Cost-effectiveness analysis of pharmacological treatments in schizophrenia: critical review of results and methodological issues. *Schizophrenia Research* 2004; 71: 445– 462
6. Hargreaves WA, Gibson PJ. Effectiveness and cost of risperidone and olanzapine for schizophrenia: a systematic review. *CNS Drugs* 2005;19:393–410.
7. König H., Roick C, Angermeyer M.C. Validity of the EQ-5D in assessing and valuing health status in patients with schizophrenic, schizotypal or delusional disorders. *European psychiatry* 2007; 22(3): 177-187
8. van de Willige G. et al. Changes in quality of life in chronic psychiatric patients: a comparison between EuroQol (EQ-5D) and WHOQoL. *Qual Life Res* 2005; 14: 441-451
9. Prieto L. et al. Psychometric validation of a generic health-related quality of life measure (EQ-5D) in a sample of schizophrenic patients. *Current Medical Opinion* 2004; 20(06):827-835
10. Stinnet AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decision Making* 1998; 18(Special issue on Pharmacoeconomics): 68-80

11. Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: a framework for marriage of health econometrics and cost-effectiveness analysis. *Health Econ* 2002; 11: 415-430
12. Owens D. Interpretation of cost-effectiveness analyses. *Archives of Internal Medicine* 1998; 13:716-717
13. Salomon JA, Weinstein MC, Goldie SJ. Taking account of future technology in cost effectiveness analysis. *British Medical journal*, 2004; 329: 733-736
14. Löthgren M., Zethraeus, N. Definition, interpretation and calculation of cost-effectiveness acceptability curves. *Health Economics* 2000; 9:623-630
15. Windmeijer F. et al. Methodological approach for assessing the cost-effectiveness of treatments using longitudinal observational data: the SOHO study. *International Journal of Technology Assessment in Health Care* 2006; 22(4): 460-468

### **Appendix 1:** list of covariates

- **Follow** = # of days patients were followed in the survey
- **Stable** = dummy variable (=1 if patient does not switch between the 3 drugs during the survey; =0 if patient switches at least once between the 3 drugs)
- **Age\_dif** = difference between age at first psychiatric treatment and age at the start of the disease
- **Male** = dummy variable (=1 if patient is male, =0 if patient is female)
- **High\_reimb** = dummy variable (=1 if patient is entitled to higher reimbursements; =0 if patient is not entitled to higher reimbursements)
- **Prev\_hosp** = dummy variable (=1 if patient was partially hospitalised during the year before the survey; = 0 if patient was not hospitalised during the year before the survey)
- **Income** = net monthly income of the patient
- **Age\_psy** = age at first psychiatric visit
- **GAF** = GAF score
- **Psy\_hosp** = dummy variable (= 1 if investigator of the patient works in a psychiatric hospital; =0 if investigator of the patient works in another ambulatory practice)
- **Priv\_prac** = dummy variable (= 1 if investigator of the patient works has a private practice; =0 if investigator of the patient does not have a private practice)