IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

IQWiG-Berichte – Nr. 571

Analyse der Unsicherheit und Auswahl der Segmente zur Ermittlung von Erstattungspreisen mit dem Effizienzgrenzenansatz

Arbeitspapier

Auftrag: GA12-03
Version: 1.0
Stand: 20.12.2017

Impressum

Herausgeber:
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:
Analyse der Unsicherheit und Auswahl der Segmente zur Ermittlung von Erstattungspreisen mit dem Effizienzgrenzenansatz

Auftraggeber:
Bearbeitung im Rahmen des Generalauftrags

Interne Auftragsnummer:
GA12-03

Anschrift des Herausgebers:
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln

Tel.: +49 221 35685-0
Fax: +49 221 35685-1
E-Mail: berichte@iqwig.de
Internet: www.iqwig.de

ISSN: 1864-2500

Dieses Arbeitspapier wurde unter Beteiligung externer Sachverständiger erstellt.

Externe Sachverständige, die wissenschaftliche Forschungsaufträge für das Institut bearbeiten, haben gemäß § 139b Abs. 3 Satz 2 Sozialgesetzbuch – Fünftes Buch – Gesetzliche Krankenversicherung „alle Beziehungen zu Interessenverbänden, Auftragsinstituten, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie, einschließlich Art und Höhe von Zuwendungen“ offenzulegen. Das Institut hat von jedem der Sachverständigen ein ausgefülltes Formular „Offenlegung potenzieller Interessenkonflikte“ erhalten. Die Angaben wurden durch das speziell für die Beurteilung der Interessenkonflikte eingerichtete Gremium des Instituts bewertet. Die Selbstangaben der externen Sachverständigen zu potenziellen Interessenkonflikten sind in Anhang B dargestellt. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden.

Externe Sachverständige

- Maiwenn Al, Erasmus University Rotterdam, Rotterdam, Niederlande
- Isaac Corro Ramos, Erasmus University Rotterdam, Rotterdam, Niederlande

Das IQWiG dankt den extern Beteiligten für ihre Mitarbeit am Projekt.

Ehemalige Mitarbeiter des IQWiG englischsprachiger Abschlussbericht

- Andreas Gerber-Grote
- Stefan K. Lhachimi

Mitarbeiter des IQWiG deutscher Vorspanntext

- Miriam Luhnen
- Sarah Mostardt
- Anja Schwalm

Schlagwörter: Effizienzgrenze, Erstattungsbetrag, Unsicherheit

Keywords: Efficiency Frontier, Reimbursement Mechanisms, Uncertainty

Inhaltsverzeichnis

Seite

Abkürzungsverzeichnis

Abkürzung	Bedeutung
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KNB	Kosten-Nutzen-Bewertung
NHB	Net Health Benefit
PSA	probabilistische Sensitivitätsanalyse

1 Hintergrund

Im Rahmen des Generalauftrags wurde das Thema Analyse der Unsicherheit und Auswahl der Segmente zur Ermittlung von Erstattungspreisen mit dem Effizienzgrenzenansatz bearbeitet. Das Institut beauftragte eine Gruppe von externen Sachverständigen mit der Bearbeitung der Fragestellung.

Der Bericht der externen Sachverständigen ist in Anhang A dokumentiert. Nachfolgend wird der Inhalt des Berichtes der externen Sachverständigen in Anlehnung an dessen Kurzfassung wiedergegeben.

Die Ergebnisse dieses Projekts wurden zudem im Rahmen einer Publikation veröffentlicht [1].

2 Fragestellung

Mit dem Effizienzgrenzenansatz können Informationen zur Ermittlung des Erstattungsbetrags eines neuen Arzneimittels bereitgestellt werden. Nach § 130b Abs. 8 Satz 3 SGB V können die Ergebnisse einer solchen Kosten-Nutzen-Bewertung (KNB) helfen, einen angemessenen und zumutbaren Erstattungsbetrag zu verhandeln. Damit muss auch die Analyse der Unsicherheit, die in jeder KNB gegeben ist, verlässliche Aussagen zu einem Preiskorridor liefern. Sollte es auf der Basis eines festgelegten Schwellenwertes um eine Ja/Nein-Entscheidung bezüglich der Erstattung einer Intervention gehen, werden probabilistische Sensitivitätsanalysen (PSA) angewandt. Bei der Methodik der Effizienzgrenze entsteht eine weitere Ebene der Komplexität, da sich die Form der Effizienzgrenze mit jeder Wiederholung der PSA verändern kann. Hierdurch kann die Zahlungsbereitschaft, die durch die Segmente der Effizienzgrenze dargestellt wird, ebenfalls unterschiedlich ausfallen.

Da das Ergebnis einer KNB in Deutschland im Rahmen des AMNOG als Handlungsempfehlung für (weitere) Preisverhandlungen dient, sollten in diesem Arbeitspapier die Konsequenzen für den zu ermittelnden Preiskorridor für verschiedene Szenarien untersucht werden.

3 Methoden

Auf Grundlage des für die KNB der Antidepressiva entwickelten Modells [2] wurden unterschiedliche Simulationen durchgeführt.

3.1 Feste Effizienzgrenze

Als 1. Schritt wurde die Situation untersucht, in welcher die Effizienzgrenze als feststehend betrachtet wird (feste Effizienzgrenze). Für diese Analysen wurde aus dem Vorbericht des IQWiG zur KNB der Antidepressiva [2] die Effizienzgrenze zum Endpunkt Remission in Szenario 1 (studienbelegter Zeithorizont, GKV-Versichertengemeinschaft) ausgewählt, da sie von 3 Interventionen bestimmt wird und daher Vergleiche gegen unterschiedliche Segmente oder Schwellenwerte ermöglicht.

Es wurden mehrere Szenarien berücksichtigt, in denen eine Stichprobe mit 1000 Durchläufen für die Kosten und den Nutzen einer hypothetischen Intervention generiert werden konnte. Für die Simulation wurde angenommen, dass keine Korrelation zwischen den Kosten und dem Nutzen besteht. In jedem Szenario wurde eine neue Intervention simuliert, die eine idealtypische Position gegenüber der festen Effizienzgrenze darstellen sollte und nicht mit den 4 zu prüfenden Interventionen im Abschlussbericht (Duloxetin, Venlafaxin, Mirtazapin und Bupropion) zusammenhängt. Die Ergebnisse dieser Szenarien bildeten die gesamte Fläche des Kosten-Nutzen-Diagramms ab und stellten somit alle möglichen Positionen zur Effizienzgrenze dar. Um zu entscheiden, wann eine neue Intervention akzeptiert werden könnte und um die Kosten und damit die Preise für diese neue Intervention ausreichend bewerten zu können, wurden für jedes Szenario verschiedene Lage- und Streumaße (arithmetisches Mittel der Punktschätzer aus allen Durchläufen sowie arithmetisches Mittel und Median aller einzeln erhobenen horizontalen Abstände der Punkte zur Effizienzgrenze) berechnet.

Außerdem wurde der Anteil der Ergebnisse der Durchläufe ermittelt, der jeweils in verschiedenen zuvor definierten Flächen in Relation zur Effizienzgrenze liegt, z. B. oberhalb des letzten Segments.

3.2 Probabilistische Effizienzgrenze

Anschließend wurde die Situation untersucht, in welcher sich die Effizienzgrenze selbst verändert (probabilistische Effizienzgrenze). Für diesen Zweck wurde der PSA-Datensatz (welcher das Ergebnis von 100 000 PSA-Durchläufen ist) benutzt, der in der KNB des IQWiG zu den Antidepressiva beschrieben wurde (dort Tabelle 81 [2]). In jedem Durchlauf wurden eine neue Effizienzgrenze und ein exakter neuer Punktschätzer (Kosten und Nutzen) für die 4 zu prüfenden Interventionen generiert. Die Empfehlung basierte auf der Position, welche der einzelne (exakte) Punktschätzer relativ zur einzelnen Effizienzgrenze hatte.

Neben den für die feste Effizienzgrenze definierten Maßen wurden hier weitere Eigenschaften der Effizienzgrenze wie Anzahl und Anstieg der Segmente bei den Simulationen berechnet.

4 Ergebnisse

4.1 Feste Effizienzgrenze

In den 16 untersuchten Szenarien wurden die Lage der Punktwolke der PSA-Ergebnisse im Kosten-Nutzen-Diagramm und die Höhe der Unsicherheit der Kosten und des Nutzens verändert. In diesen Szenarien wurde beobachtet, dass die Wahrscheinlichkeit, die Kosten beziehungsweise den Preis einer neuen Intervention zu akzeptieren, von der Lage der PSA-Ergebnisse zur Effizienzgrenze im Kosten-Nutzen-Diagramm abhängt.

PSA-Ergebnisse, die links oberhalb der Effizienzgrenze liegen, werden am häufigsten als angemessen akzeptiert, während PSA-Ergebnisse, die unten rechts sowie unterhalb der Effizienzgrenze liegen, am häufigsten als unangemessen anzusehen sind.

Wenn der Grundsatz berücksichtigt würde, nur die PSA-Ergebnisse oberhalb des letzten rückwärts extrapolierten Segments der Effizienzgrenze zu akzeptieren, wäre die Wahrscheinlichkeit, die Kosten beziehungsweise den Preis einer neuen Intervention als angemessen zu akzeptieren, kleiner oder gleich jener Wahrscheinlichkeit, die im Vergleich zur gesamten Effizienzgrenze erreicht worden wäre. Einzelne Szenarien zeigten, dass die Auswirkungen dieses Grundsatzes sehr weitreichend sein können. Die Akzeptanzwahrscheinlichkeit wurde geringfügig von einer Erhöhung der Unsicherheit der Kosten und Nutzen beeinflusst und nahm je nach Position der PSA-Ergebnisse im Kosten-Nutzen-Diagramm zu oder ab.

Die Untersuchung der 3 Lagemaße bei der festen Effizienzgrenze ergab, dass die Wahl des richtigen Lagemaßes als Basis für die Kosten- beziehungsweise Preisempfehlung abhängig ist von der Verteilung der Abstände zur Effizienzgrenze.

Empfehlungen sollten zunächst auf dem arithmetischen Mittel der Punktschätzer aus allen Durchläufen basieren. Eine Erhöhung oder Reduktion der als angemessen zu akzeptierenden Kosten sollte aus dem horizontalen Abstand zwischen diesen arithmetisch gemittelten Punktschätzern und dem nächsten Segment der Effizienzgrenze kalkuliert werden. Die Empfehlung war abhängig von der Position des arithmetischen Mittels der Punktschätzer im Kosten-Nutzen-Diagramm: Liegt dieses oberhalb der Effizienzgrenze, wird eine Erhöhung der Kosten beziehungsweise des Preises als angemessen angesehen, anderenfalls wird eine Reduzierung vorgeschlagen.

Im hypothetischen Falle einer Empfehlung basierend auf dem arithmetischen Mittel oder dem Median der Abstände aller einzeln erhobenen horizontalen Abstände der Punkte zur Effizienzgrenze sollte das Ausmaß des Unterschiedes zur „korrekten“ Empfehlung basierend auf dem arithmetischen Mittel der Punktschätzer aller Durchläufe quantifiziert werden.

Wenn sich die Abstände aller Punktschätzer zur Effizienzgrenze symmetrisch verteilen, sind die 3 Maße ähnlich. Falls die Verteilung der Abstände zur Effizienzgrenze jedoch asymmetrisch ist, wird das arithmetische Mittel der Empfehlungen basierend auf den einzeln

erhobenen horizontalen Abständen der Punkte zur Effizienzgrenze von den hohen Werten am rechten Rand der Verteilung beeinflusst. In diesem Fall ermöglicht der Median der Empfehlungen basierend auf den einzelnen erhobenen horizontalen Abständen der Punkte zur Effizienzgrenze eine bessere Annäherung.

4.2 Probabilistische Effizienzgrenze

Die Anzahl der Interventionen (und damit die Anzahl der Segmente) auf der Effizienzgrenze kann je Simulation variieren. In ungefähr 75 % der 100 000 durchgeführten Simulationen hatte die Effizienzgrenze nur 1 Segment, 21 % hatten 2 Segmente, 3 % hatten (genau) 3 Segmente und 0,003 % der Simulationen hatten (genau) 4 Segmente. Das arithmetische Mittel der Steigung des letzten Segments der Effizienzgrenze betrug 0,00352. Der Mittelwert des Schwellenwertes für die Zahlungsbereitschaft (λ), welcher sich jeweils aus 1 dividiert durch die Steigung des letzten Segments errechnet, war 4,10 €*10^{13}. Dieses arithmetische Mittel ist extrem hoch, da es von Extremwerten beeinflusst wird, bei denen die Steigung nahezu 0 betrug, was (numerisch) einer unendlichen Zahlungsbereitschaft entspricht. In dieser Situation ermöglicht der Median nach 100 000 simulierten λ in Höhe von 237 € eine bessere Empfehlung.

Für die in den Simulationen zu bewertenden Interventionen wurde beobachtet, dass sich in einigen PSA-Durchläufen ein (numerisch) unendlicher Abstand zur Effizienzgrenze ergab (äquivalent einer unendlichen Kostenerhöhung). Dies zeigt, dass die Verteilung der Empfehlungen basierend auf den einzeln erhobenen Punktschätzern von Ausreißern beeinflusst wird. Auch in dieser Situation sollte der Median der einzeln erhobenen horizontalen Abständen der Punkte zur Effizienzgrenze anstelle ihres arithmetischen Mittels verwendet werden, um Empfehlungen abzuleiten.

Die Ergebnisse der Simulationsanalyse unter Verwendung des Medians legen nahe, dass die Kosten für Duloxetin, Venlafaxin, Mirtazapin und Bupropion bezogen auf den Endpunkt Remission im studienbelegten Zeithorizont um 131 €, 29 €, 12 € und 99 € reduziert werden sollten. Diese Empfehlungen wurden zur Überprüfung der Ergebnisse als feststehende Werte in eine erneute Simulation eingesetzt, in der für alle zu bewertenden Interventionen gezeigt wurde, dass sich durch die Reduktion der Unsicherheit der Prozentsatz der akzeptablen PSA-Ergebnisse deutlich erhöhte (auf 49 % für Duloxetin). Die „aktualisierte“ Empfehlung zur Absenkung oder Erhöhung der akzeptablen Kosten lag bei ungefähr 0 € für alle Interventionen. Darüber hinaus konnte festgestellt werden, dass sich die mit der Verteilung des Net Health Benefit (NHB) assoziierte Unsicherheit reduzierte. Nach Umsetzung der Empfehlungen lagen daher deutlich mehr PSA-Ergebnisse oberhalb der Effizienzgrenze. Die PSA-Ergebnisse lagen auch näher zum letzten Segment der Effizienzgrenze (aus diesem Grund war auch zu beobachten, dass sich die Verteilung des NHB nach Umsetzung der Kostenempfehlungen bei 0 zentrierte). Dass die Effizienzgrenze in den meisten Fällen (ca. 75 %) genau 1 Segment hatte, bedeutet, dass die PSA-Ergebnisse in der Tat näher zur Effizienzgrenze lagen.

5 Fazit

Bei einer festen Effizienzgrenze, die durch mehr als 1 Segment (also > 2 Interventionen) bestimmt wird, sollten die Empfehlungen auf dem horizontalen Abstand zwischen dem arithmetischen Mittel der Punktschätzer aus allen Durchläufen und dem am nächstgelegenen Segment der Effizienzgrenze basieren. Grundsätzlich ist dies nicht identisch mit dem arithmetischen Mittel aller einzeln erhobenen horizontalen Abstände der Punkte zur Effizienzgrenze. Letzteres war immer größer als oder gleich dem arithmetischen Mittel der Punktschätzer aus allen Durchläufen, da die Verteilung der Abstände zwischen jedem PSA-Ergebnis und der Effizienzgrenze asymmetrisch war und von Ausreißern beeinflusst wurde. Unter diesen Bedingungen ist der Median aller einzeln erhobenen horizontalen Abstände der Punkte zur Effizienzgrenze gegenüber dem arithmetischen Mittel dieser Abstände zu bevorzugen.

Bei einer probabilistischen Effizienzgrenze, die durch mehr als 1 Segment bestimmt wird, kann das arithmetische Mittel der Punktschätzer aller Durchläufe der Simulationen nicht gebildet werden, da in jedem Durchlauf auch eine andere Effizienzgrenze erzeugt wird. Die Verteilung der Abstände zwischen jedem PSA-Ergebnis und der Effizienzgrenze ist aller Wahrscheinlichkeit nach verzerrt und zeigt Ausreißer. Das arithmetische Mittel würde dann zu einer Überschätzung der Empfehlung führen. Wenn die Effizienzgrenze als probabilistisch angenommen wird, sollten Empfehlungen daher auf dem Median aller einzeln erhobenen horizontalen Abstände der Punkte zur Effizienzgrenze basieren.

6 Literatur

1. Corro Ramos I, Lhachimi SK, Gerber-Grote A, Al MJ. Cost recommendation under uncertainty in IQWiG's efficiency frontier framework. Med Decis Making 2017; 37(2): 162-172.

2. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin im Vergleich zu weiteren verordnungsfähigen medikamentösen Behandlungen: Vorbericht (vorläufige Kosten-Nutzen-Bewertung); Auftrag G09-01 [online]. 09.11.2012 [Zugriff: 11.03.2013]. URL: https://www.iqwig.de/download/G09-01_Vorbericht_Kosten-Nutzen-Bewertung-von-Venlafaxin-Duloxetin-Bupropion-und-Mirtazapin.pdf.

Anhang A – Abschlussbericht der externen Sachverständigen

Simulation of Uncertainty in the Efficiency Frontier and Exploration of the Selection of Segments and its Effect on the Maximum Reimbursable Price

Working Paper

Isaac Corro Ramos PhD
Stefan Lhachimi PhD
Andreas Gerber PhD
Maiwenn Al PhD

February 2015

Table of Contents

List of Tables

List of Figures

Abbreviations

Abbreviation	Definition
AMNOG	Arzneimittelmarkt-Neuordnungsgesetz (Act on the Reform of the Market for Medical Products)
CEAC	Cost-effectiveness Acceptability Curve
ICER	Incremental Cost-Effectiveness Ratio
INHB	Incremental Net Health Benefit
INMB	Incremental Net Monetary Benefit
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
NHB	Net Health Benefit
NICE	National Institute for Health and Care Excellence
NMB	Net Monetary Benefit
PSA	Probabilistic Sensitivity Analysis
QALY	Quality Adjusted Life Year

Summary

Background

While assessing the cost-effectiveness of antidepressants published by IQWiG [13], a potential methodological problem arose, relating to the analyses of uncertainty through a probabilistic sensitivity analysis (PSA) [3]. A PSA leads to a set of potential combinations of costs and effects on the cost-effectiveness plane, which could fall above two different segments of the efficiency frontier and not just to the right of the extension as assumed in the original concept [5, 12]. Because of that, the maximum reimbursable price of different interventions (or the same intervention) in a PSA would be measured against different thresholds representing a relation of benefits and cost as depicted in the cost-effectiveness plane. This could result in the problem that, depending on the place of a PSA outcome on the cost-effectiveness plane, the price of a given intervention would not have to be reduced or held/increased in a symmetric way like in a model of a fixed threshold.

Methods

We set up a simulation study based on the model developed by IQWiG to assess the cost-effectiveness of antidepressants [13].

We first studied the situation where the efficiency frontier was considered to be fixed. The efficiency frontier chosen for these analyses was the one obtained in the cost-effectiveness evaluation of antidepressants by IQWiG – Scenario 1 (Remission) on pages 198-199 [13]. We have chosen this efficiency frontier because it is determined by 3 interventions which allowed comparisons against different segments or thresholds. Furthermore, we have considered several scenarios where a random sample of size of 1,000 from costs and benefits of a hypothetical new intervention was generated. We assumed no correlation between costs and benefits. In each of these scenarios the 1,000 outcomes were located in a different area of the cost-effectiveness plane so that with all the scenarios we covered the entire plane. For each scenario we calculated several metrics to decide when a new intervention should be deemed acceptable and to properly assess cost recommendations for the new intervention.

We studied then the situation where the efficiency frontier was also considered to be probabilistic. For this scenario we used the real PSA dataset described in the cost-effectiveness evaluation of antidepressants by IQWiG – Table 81 [13]. In each iteration a new efficiency frontier and a point estimate (costs and benefits) of 4 new interventions were generated and the decision was based on the position of those single point estimates relative to that single efficiency frontier. Besides the metrics defined for the deterministic setting, we also computed here metrics regarding the statistical properties of the efficiency frontier.

Results

Fixed efficiency frontier

We have studied 16 different scenarios where the location of the PSA outcomes in the cost-effectiveness plane and the uncertainty in costs and benefits were changed.

Based on these scenarios we observed that the probability of accepting the new intervention depended on the position of the PSA outcomes in the cost-effectiveness plane. If the erroneous rule of accepting the PSA outcomes above the last segment on the efficiency frontier was considered, the probability of acceptance would be smaller than or equal to the one obtained with the whole efficiency frontier. The impact of using the wrong decision rule

can be very strong as shown in scenarios 3 and 4. The probability of acceptance was slightly affected by increasing the uncertainty around costs and benefits. It decreased or increased depending on the position of the PSA outcomes on the cost-effectiveness plane.

Cost recommendations should be based on the average point estimate of the PSA outcomes. An increase or decrease in cost recommendation should be then calculated as the horizontal distance between that average point estimate and its closest segment on the efficiency frontier. The cost recommendation depended on the position of the average point estimate of the PSA outcomes on the cost-effectiveness plane: when it lies above the efficiency frontier a cost increase is recommended and a cost decrease is recommended otherwise.

In the hypothetical case that the cost recommendation should be based on the average or the median of the cost recommendations based on the PSA outcomes separately, we were interested in quantifying to what extent these would differ from the correct recommendation based on average point estimate of the PSA outcomes. We observed that this depended on the distribution of the cost recommendations based on the PSA outcomes separately. When there is symmetry as in scenarios 1, 2, 5 and 6 the three metrics are similar, and so are the cost recommendations. However, when the distribution is skewed (as in scenarios 3, 4 and 7 to 16), the average of the cost recommendations based on the PSA outcomes separately is affected by the large values at the right tail of the distribution and the median provided a better approximation to the correct cost recommendation.

Probabilistic efficiency frontier

We first report the results regarding the distribution of the efficiency frontier. The number of interventions (and therefore the number of segments) on the efficiency frontier may vary per simulation. In particular, in approximately 75% of the simulations the efficiency frontier had only one segment, 21% of the simulations with exactly two segments, 3% of the simulations with exactly 3 segments and 0.003% of the simulations with exactly 4 segments. We obtained an *average* slope of the last segment on the efficiency frontier equal to 0.00352. The willingness-to-pay threshold (λ), which is computed as 1 divided by the slope of the last segment on the efficiency frontier, was €$4.1*10^{13}$. This average is extremely high due to the simulations where the slope was nearly zero, which corresponds to a (numerically) infinite willingness-to-pay. In this situation, the median of the 100,000 simulated λ's (€237) provided a better approximation to the standard ICER.

Regarding acceptance and cost changes of the new interventions considered in the simulation, we observed that when calculating the distances between every PSA outcome and every efficiency frontier, we found that for all the new interventions some simulations provided a (numerically) infinite distance (or equivalently an infinite cost increase). This shows that, the distribution of the cost recommendations based on the PSA outcomes separately is affected by the presence of outliers. In this situation, the median instead of the average should be used to determine the price recommendation. The results of the simulation analysis suggested that the costs of Duloxetin, Venlafaxin, Mirtazapin and Bupropion should be decreased by €131, €29, €12 and €99, respectively. We implemented the cost recommendations and repeated the analysis. The simulation results after the cost recommendations have been implemented showed that for all the new interventions, the percentage of acceptable PSA outcomes increased significantly (up to 49% for Duloxetin). The "updated" cost reduction/increase recommendation was (approximately) €0 for all the interventions. Moreover, we also observed that the uncertainty associated to the distribution

of the NHB was reduced. Thus, after implementing cost recommendations there were many more PSA outcomes above the efficiency frontier, but also the PSA outcomes were closer to the last segment on the efficiency frontier (for that reason we observed that the distribution of the NHB after implementing the cost recommendations is centered at zero). Since in most of the cases the efficiency frontier had exactly one segment (around 75%) it also means that the PSA outcomes were in fact closer to efficiency frontier.

Conclusions

With a fixed efficiency frontier cost recommendations should be based on the horizontal distance between the average point estimate of the PSA outcomes and the closest segment of the efficiency frontier (metric M4). In general, this is not the same as the average of the distances between every PSA outcome and the efficiency frontier (metric M5). We have observed that M5 was always larger than M4. This was because the distribution of the distances between every PSA outcome and the efficiency frontier was skewed and was affected by outliers. Under these conditions, the median (metric M6) is preferred to M5 as proxy for M4.

In the probabilistic efficiency frontier approach the cost recommendation cannot be based on the distance between the average point estimate of the PSA outcomes and the efficiency frontier (metric M4) simply because this cannot be computed. The distribution of the distances between every PSA outcome and the efficiency frontier is likely to be skewed and to have outliers. Thus, its average (metric M5) would overestimate the cost recommendation, which can be very wrong. The median (metric M6) is less sensitive to skewed distributions and outliers and when the efficiency frontier is probabilistic, cost recommendations should be based on it.

Keywords: Efficiency frontier, probabilistic sensitivity analysis (PSA), maximum reimbursable price, decision uncertainty

1. Background

After the “Act to promote competition among the statutory health insurance funds” (GKV-Wettbewerbsstärkungsgesetz) came into force on 1 April 2007, the assessment of the benefits and costs of drugs was introduced as a task for IQWiG (§ 139a (3) clause 5 SGB V). In response, IQWiG published its report on the general methods for the assessment of the relation between benefits and costs after extensive consultations in 2009 [12]. These methods were developed to facilitate the German approach to set a maximum reimbursable price according to § 35b SGB V, fixing the limit up to which health insurance funds can reimburse costs. With the Act on the Reform of the Market for Medical Products (Arzneimittelmarkt-Neuordnungsgesetz – AMNOG) of 22 December 2010, the § 35b SGB V was altered, but the scope of health economic evaluation to deliver information to arrive at an appropriate and affordable price for a new drug was retained.

Recently, the first assessment of costs and effects of new drugs within one therapeutic area was published by IQWiG. This was the assessment of the costs and effects of four new drugs in the intervention of depression [13]. While assessing these costs and effects using the methods guide [12], a potential methodological problem arose, relating to the analyses of uncertainty through a probabilistic sensitivity analysis (PSA) [3]. A PSA leads to a set of potential combinations of costs and effects on the cost-effectiveness plane. Such cloud of PSA outcomes could fall above two different segments of the efficiency frontier and not just to the right of the extension as assumed in the original concept [5, 12]. Subsequently, the maximum reimbursable price of different interventions (or the same intervention) in a PSA would be measured against different thresholds representing a relation of benefits and costs as depicted in the cost-effectiveness plane. This could result in the problem that, depending on the place of a PSA outcome on the cost-effectiveness plane, the price of a given intervention would not have to be reduced or held/increased in a symmetric way like in a model of a fixed threshold.

The current study aims to explore this potential problem and present possible solutions. The study was not primarily intended to define one optimal solution, but to assess the strengths and weaknesses of different approaches to deal with this problem.

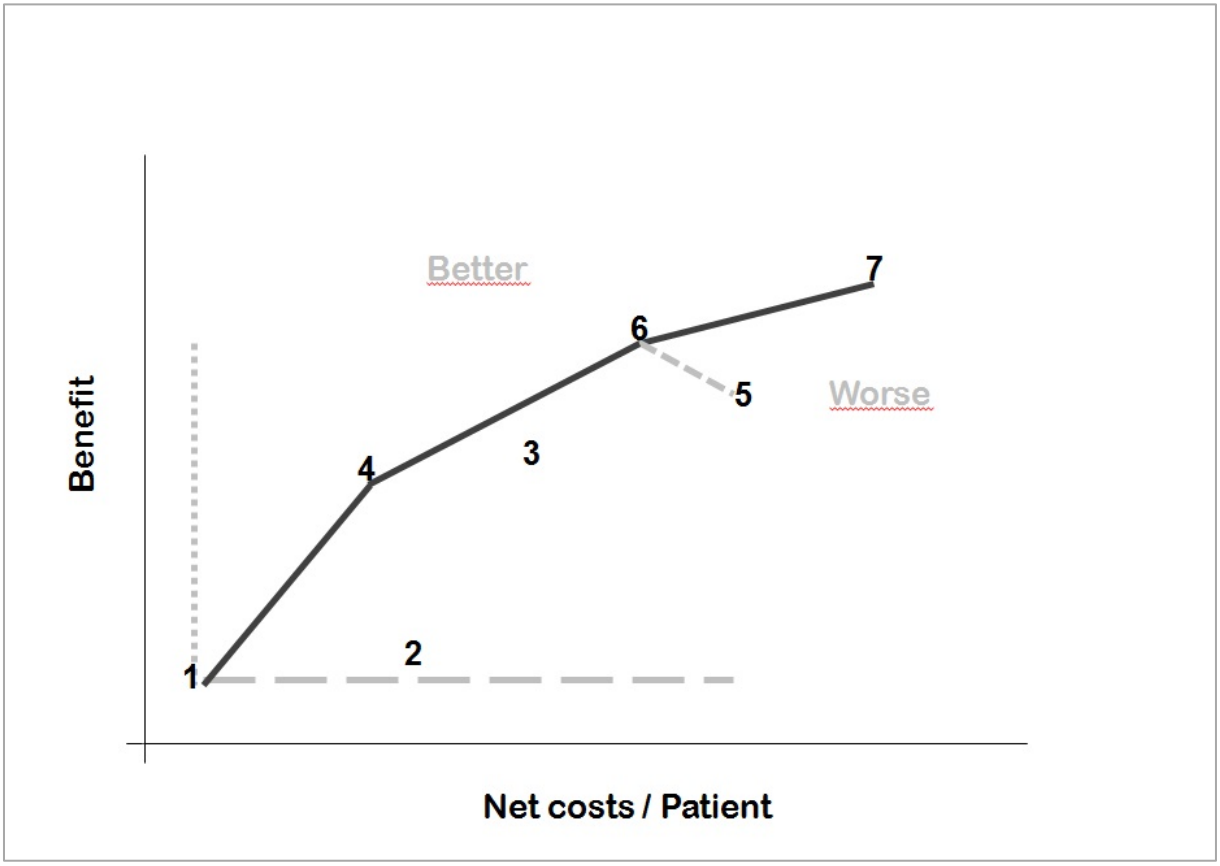
This report starts with a description of the theoretical framework used for the assessment of costs and effects within the IQWiG context. Then we will describe the simulation study that we set up and its results. Finally, we will present our conclusions and discuss these.

1.1 Efficiency frontier

The efficiency frontier is the method used by IQWiG to compare different health care interventions in health economic evaluations [5]. For the indication area under assessment, the efficiency frontier depicts graphically (in the cost-effectiveness plane) the net cost per patient (x-axis) and the health benefit (y-axis) of the health care interventions that are currently in use. The slope of the segment connecting any two points in the efficiency frontier represents the incremental benefit per incremental net costs of two different interventions (i.e. the inverse of the incremental cost-effectiveness ratio). Thus, a horizontal slope means no efficiency, a positive slope indicates additional benefit for increased costs (with a vertical slope indicating infinite efficiency) and a negative slope represents less benefit yet more costs. The point representing no intervention is usually taken as the origin of the cost-effectiveness plane (i.e. zero benefit and zero costs). However, this rarely represents reality

since the absence of intervention is normally associated with costs and health effects. This problem can be easily solved by shifting the no intervention point to the cost-effectiveness origin (i.e. subtracting the no-intervention benefits and costs to the benefits and cost of the other health interventions). An example of an efficiency frontier can be seen in Figure 1.

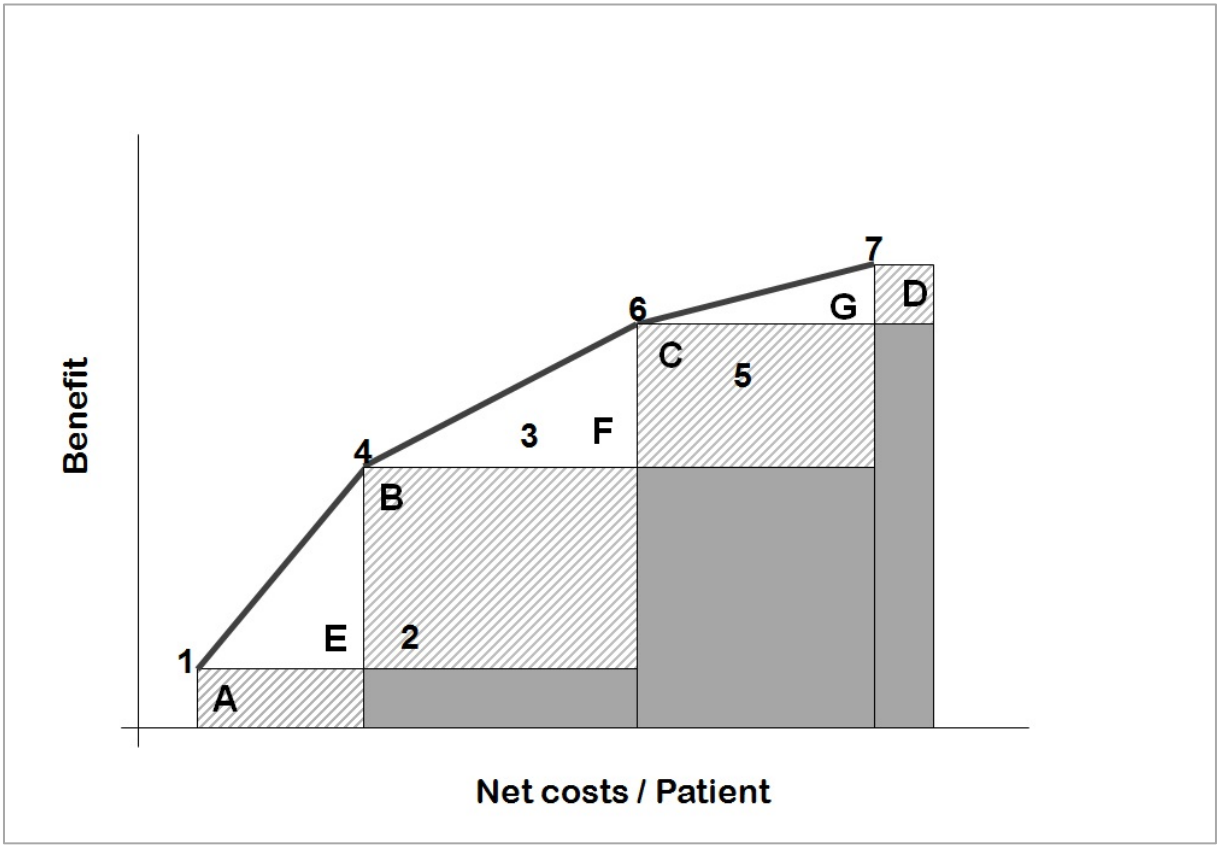
Figure 1*: Theoretical efficiency frontier



The efficiency frontier divides the cost-effectiveness plane into two main areas: an area above/left to the efficiency frontier representing better efficiency and an area below/right to the efficiency frontier representing worse efficiency. The area below the efficiency frontier can be further sub-divided by a series of rectangles (A-D) and triangles (E-G), as shown in Figure 2. Rectangles (A-D) represent the areas where interventions are inefficient (i.e. higher costs and less benefit) with respect to at least one intervention on the efficiency frontier (e.g. options 2 or 5 in Figure 2). The triangles (E-G) denote the areas where interventions are extendedly dominated (i.e. a combination of the two options forming the hypotenuse of the triangle will provide more benefit for lower costs). However, such a combination is not always feasible in reality. For example ethical considerations may arise in that equal care is not provided to all of the population [4]. Therefore, the options in the triangles might be part of the practical efficiency frontier.

* The figure was taken from the IQWiG general methods for the assessment of the relation of benefits to costs [12].

Figure 2[†]: Absolute versus extended dominance



Note that this method is an alternative to the standard approach based on incremental cost-effectiveness ratios [6]. Nevertheless, the concept of efficiency frontier is also well known in the literature but it is usually defined with reversed axes. Karlsson and Johannesson [14] for example presented this curve when describing optimal budget allocation across health care interventions. Barton et al 2008 [1], Goeree et al 2002 [9], Hallinen et al 2010 [10], and Rojnik et al 2008 [17] are examples of this approach. The same concept is referred as *cost-effectiveness frontier* in Briggs 2000 [2], as *expansion path* in Murray et al 2003 [15] and e*xpansion frontier* in Niessen et al 2003 [16]. A representation of the efficiency frontier on the cost-disutility plane (with costs on the y-axis) can be found in Eckermann et al 2008 [7].

[†] The figure was taken from the IQWiG general methods for the assessment of the relation of benefits to costs [12].

2. Methods

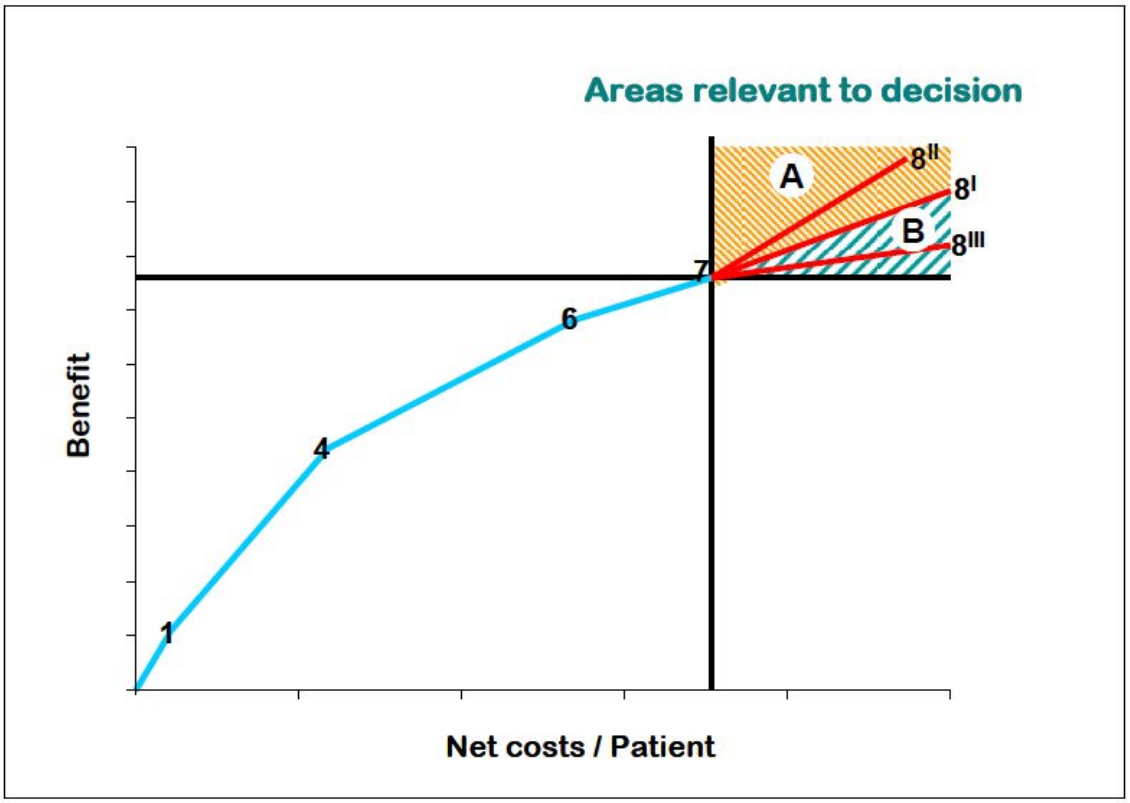
2.1 Current IQWiG's recommendations

The information provided by the efficiency frontier can be used as guidance for decision makers seeking a maximum reimbursable price for the health care intervention evaluated. The position on the cost-effectiveness plane of health care interventions can be driven by prices. For example, for those new interventions on the efficiency frontier, their price can be considered consistent with the efficiency that can be achieved with the interventions on the current market. For new interventions below or to the right to the efficiency frontier (i.e. showing lower efficiency), this price is considered too high and needs to be adjusted or justified. Finally, when a intervention lies above or to the left to the efficiency frontier (i.e. improved efficiency) its price is accepted.

Prior to AMNOG, the process of setting a maximum reimbursable price for a new health care intervention was subjected to an additional constraint: a recommended maximum reimbursable price was considered appropriate only when it did not lower the efficiency of the relevant therapeutic area. This corresponds to the shaded section A (8") in Figure 3. Note that implementing an intervention in this area leads to a re-defined frontier, where intervention 7 is now extendedly dominated. Interventions on the extension of the last segment on the efficiency frontier (8') were also considered to meet the criterion of appropriate costs. The shaded section B (8"') indicates superior benefits but a lower efficiency than the last intervention on the frontier. Thus, the price should not be considered appropriate and should be decreased to bring the intervention to the frontier. However, in general, decision makers could also consider a small decrement in efficiency reasonable. Based on the theory by Karlsson and Johannesson [14] it is clear that the budget defines the last implemented intervention (number 7 in Figure 3). Once the budget increases, new, less efficient, interventions may become acceptable (i.e. the budget determines what a reasonable decrement in efficiency is or equivalently how much of the shaded section B will be used in practice). This then increases the threshold ratio, also called the critical ratio, which is the ICER of the last intervention implemented on the efficiency frontier.

However, in the current situation superiority in terms of health benefits or efficiency with respect to the last intervention on the efficiency frontier is not required. Therefore, the new interventions are no longer limited to the areas A and B in Figure 3. They can be anywhere in the cost-effectiveness plane. This is the situation considered in the remainder of this study.

Figure 3[‡]: Areas relevant to decision making



2.2 Adding uncertainty

Several papers discuss a full incremental cost-effectiveness analysis approach, although most of them from a deterministic point of view. The paper by Fenwick et al. [8] describes how uncertainty may be dealt with when comparing more than 2 interventions. They present an extension of the cost-effectiveness acceptability curve (CEAC) introduced by van Hout et al. [11] that was called a family of acceptability curves. However, this methodology was developed in the NICE framework where a yes/no reimbursement decision is made based on a fixed threshold. However, in the IQWiG framework, the goal is to determine a reimbursable price. Thus, while in acceptability curves the price of the intervention is the input, in the IQWiG approach it is the output. Also, the threshold in the CEACs is varied deterministically, whereas it is uncertain here. Therefore, as an alternative to the standard CEAC, a price reimbursement acceptability curve was proposed [20], which shows for each suggested price the probability that the new intervention is efficient. Alternatively, one may ask for each new intervention considered, as in the cost-effectiveness evaluation of antidepressants performed by IQWiG [13], how likely it is to be acceptable. It was suggested that, as the last segment of the frontier determines the maximum threshold ratio, only the PSA outcomes above that line (including its backward and forward extension) would be acceptable.

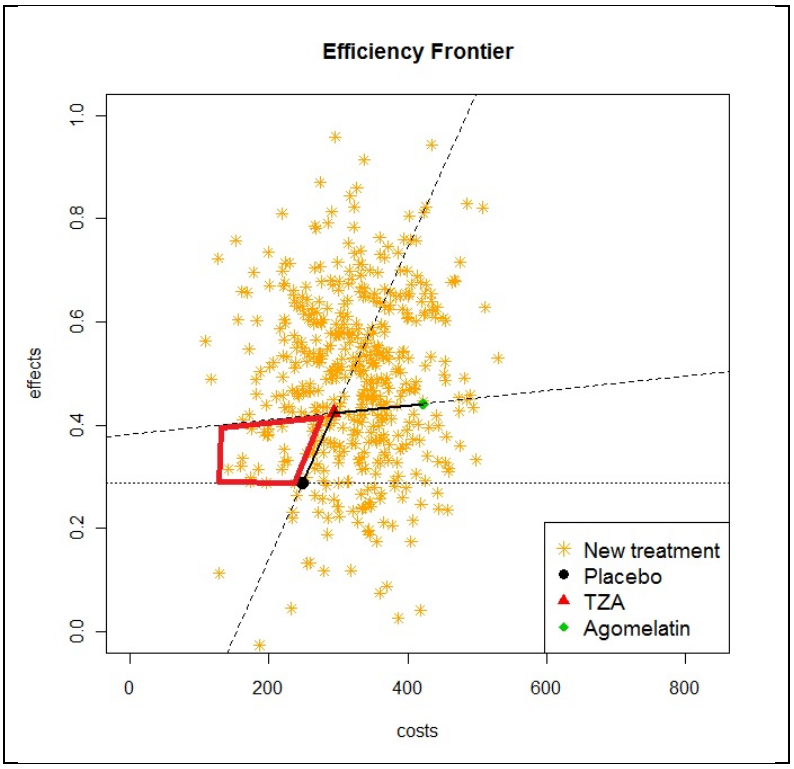
When should a new intervention be accepted?

In line with Karlsson and Johannesson [14] (and followers) the only theoretically correct way of determining whether a new intervention is acceptable is by judging it against the whole efficiency frontier. Therefore, the decision rule should be as follows: when a new intervention lies on the efficiency frontier or above, the intervention is accepted and in the latter case the

[‡] The figure was taken from the IQWiG general methods for the assessment of the relation of benefits to costs [12].

efficiency frontier has to be redefined. Below the efficiency frontier the cost of the new intervention is considered too high and the new intervention is rejected. Other options are not grounded in Health Economics theory. As a consequence, it is *incorrect* in general to use the backward extension of the last segment on the efficiency frontier to decide whether a new intervention is accepted or not. This is only true when the efficiency frontier is determined by two interventions so that the efficiency frontier consists of only one segment. Otherwise, all the interventions located below the backward extension of the last segment but above the efficiency frontier would be considered as unacceptable and this is wrong since compared to the next most effective alternative in the efficiency frontier the new intervention is in position of dominance or of greater efficiency. Hence, the new intervention should be accepted. This situation is depicted in Figure 4. If we (incorrectly) use the backward extension of the last segment on the efficiency frontier (instead of the whole efficiency frontier) to decide which PSA outcomes would be considered as acceptable, then all the PSA outcomes inside the polygon would not be accepted. For further details we refer to Appendix C.

Figure 4: PSA outcomes and acceptance



Cost reimbursement acceptability curves

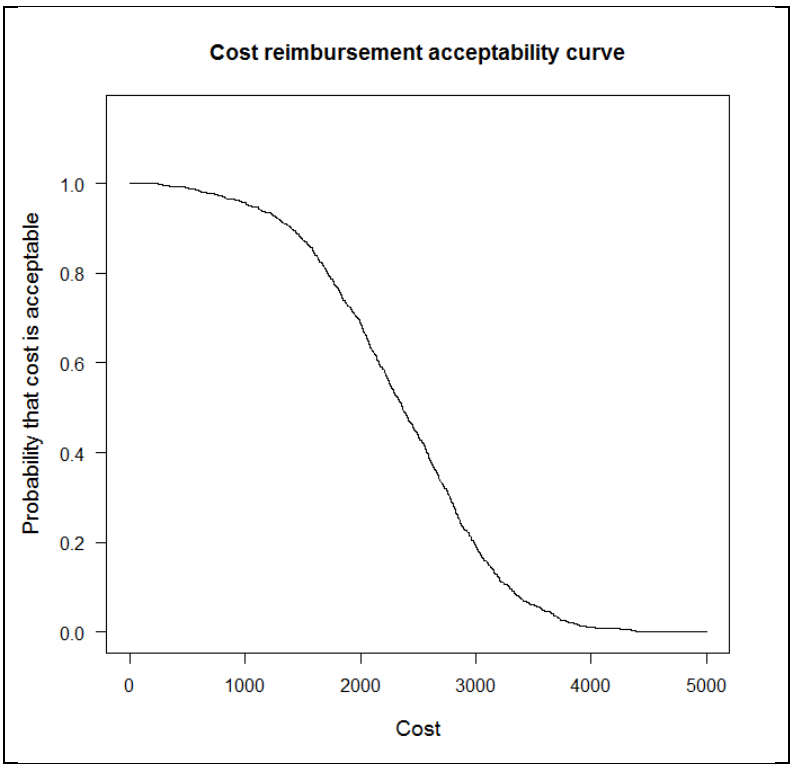
Price reimbursement acceptability curves were introduced as a method to estimate the probability that a suggested maximum price for a new intervention (as determined by the inverse of the slope of the last segment on the efficiency frontier) is going to be acceptable [20].

We propose here an alternative definition. For each run of the PSA we calculate the maximum cost that would be reimbursed as we will describe in detail below (see e.g. Figure 7). Thus, when a PSA outcome is located above the efficiency frontier, its cost is considered appropriate. However, given its level of benefits, a higher cost may also be acceptable. The

maximum acceptable cost for that level of benefits is defined as the cost that would bring the PSA outcome exactly to the efficiency frontier, thus the PSA cost estimate plus the horizontal distance to the efficiency frontier. Similarly, when a PSA outcome is below the efficiency frontier, its cost is considered too high and must be reduced. The maximum acceptable cost is defined then as the cost that would bring the PSA outcome exactly to the efficiency frontier, thus the PSA cost estimate minus the horizontal distance to the efficiency frontier. This procedure yields a set of maximum reimbursable costs. Our cost reimbursement acceptability curve shows for every possible cost (x-axis) the proportion of simulated reimbursable costs which are larger than or equal to this possible cost.

An example of a cost reimbursement acceptability curve based on a simulated PSA can be seen in Figure 5. In that example, if the suggested cost was €0, then this would have been accepted with probability one since all the simulated reimbursable costs were larger than €0. When the suggested cost was €2,284, half of the simulated reimbursable costs were larger than this suggested cost. Thus, €2,284 would have been accepted with probability 0.5 (median).

Figure 5: Cost reimbursement acceptability curve

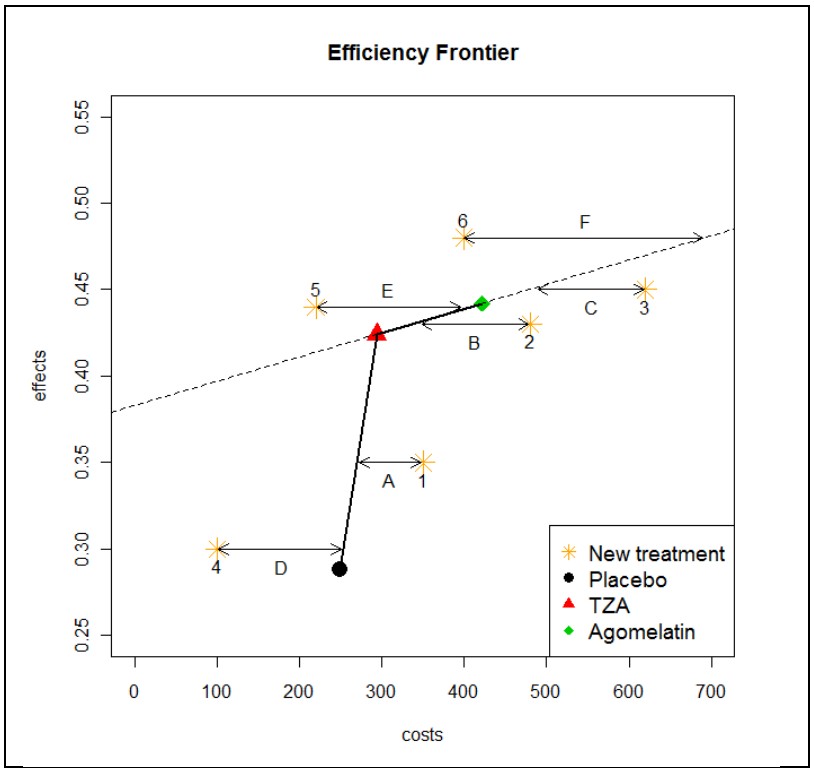


Cost reduction or increase

As mentioned above, when a new intervention is located below the efficiency frontier, its cost is considered too high and the new intervention is rejected. In this case, the question to be answered is: what the cost of the new intervention should be so that it becomes acceptable (i.e. cost reduction to bring it exactly to the efficiency frontier). Note that this implies that the location of the new intervention in the cost-effectiveness plane might only be modified horizontally, i.e. the benefits are considered to be given but the cost could be adjusted (in this case reduced), for example by the intervention's manufacturer. This idea is illustrated in

Figure 6, where, the outcomes labelled as 1, 2 and 3 would not be accepted since they fall below the efficiency frontier (note that the forward extension of the last efficiency frontier segment is considered part of the efficiency frontier but the backward extension is not). Thus, the cost reductions needed to bring them to the efficiency frontier are the horizontal distances A (€77), B (€131) and C (€130), respectively. Note that, the segment used to determine the cost reduction depends on the position of the outcomes on the cost-effectiveness plane. In this case, we used the Placebo-TZA segment for outcome 1 and the TZA-Agomelatin segment (including forward extension) for outcomes 2 and 3. Similarly, when a new intervention is located above the efficiency frontier, its cost is considered appropriate and should be reimbursed. In fact, above the efficiency frontier, for a given level of benefits, a higher cost may also be considered as appropriate. Thus, the question to be answered now is: what the cost of the new intervention could be so that it remains acceptable (i.e. cost increase to bring it exactly to the efficiency frontier). This is also illustrated in Figure 6. The outcomes labelled as 4, 5 and 6 are accepted since they fall above the efficiency frontier (note that the forward extension of the last efficiency frontier segment is considered part of the efficiency frontier but the backward extension is not). Thus, the cost increases that would bring them to the efficiency frontier are the horizontal distances D (€150), E (€175) and F (€290), respectively.

Figure 6: PSA outcomes and cost adjustment



Finally, note that the *cost* reduction (or increase) mentioned above is not equal to the *price* reduction (increase) for a given intervention as reported in the cost-effectiveness evaluation of antidepressants by IQWiG [13]. However, there is a (linear) relationship between the price of the new intervention and total cost associated to it so that when the cost reduction (increase) is known the price reduction should be easily calculated. Since this relationship was unknown to us we present all the results in terms of cost instead of price.

Cost reduction or increase under uncertainty

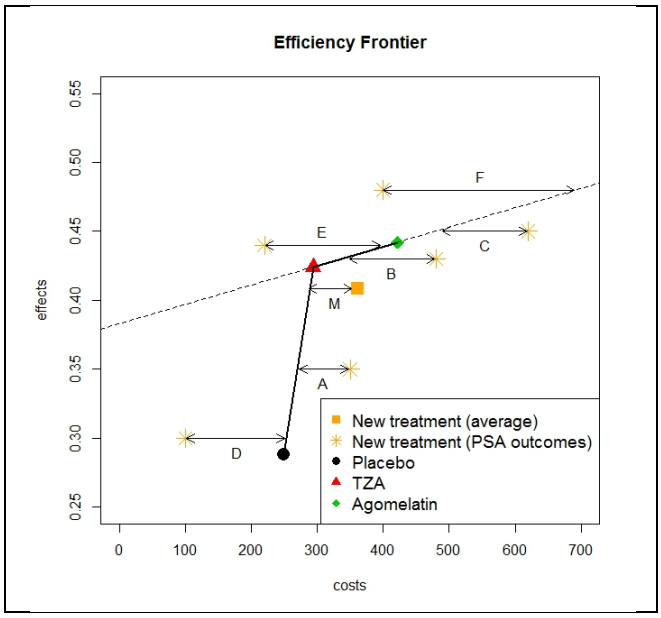
The decision whether to accept a new intervention or not is often based on a single (deterministic) point estimate of costs and benefits. When this point estimate lies on the efficiency frontier or above, the new intervention is accepted, and when it lies below the efficiency frontier the new intervention is rejected. Afterwards, a probabilistic sensitivity analysis (PSA) is often performed in order to study the uncertainty around that point estimate, which gives an indication of the uncertainty about the decision. The basic idea behind a PSA is to randomly obtain a large sample of point estimates of costs and benefits. When the joint distribution of costs and benefits is symmetric, the average costs and benefits converge to the deterministic point estimate as the PSA sample increases. Therefore, in presence of symmetry, if the PSA sample is large enough, then we may assume that the average (probabilistic) point estimate is closed to the deterministic point estimate. Thus, if based on the PSA a decision whether the new intervention is going to be accepted or not has to be made, then the average point estimate should be compared against the whole efficiency frontier and again if it lies below the efficiency frontier then the new intervention is rejected and otherwise it is accepted. However, symmetry does not always occur and there might be discrepancy between the decision based on the deterministic point estimate and the decision based on the average (probabilistic) point estimate[§]. Nevertheless, for the purposes of this section let us assume that the decision is solely based on the average point estimate. Depending on the position of that average point estimate relative to the efficiency frontier we may consider a cost reduction or increase (i.e. cost reduction or increase to bring the average point estimate exactly to the efficiency frontier). Note that, as explained above, this reduction or increase should be computed as the horizontal distance between the average point estimate and its closest segment on the efficiency frontier.

It is important to emphasize that the distance between the average point estimate and the efficiency frontier is not necessarily the same as the average of the distances between every single PSA outcome and the efficiency frontier. Furthermore, the latter may lead to erroneous decisions. We will explain this with the help of Figure 7. In that scenario, the average point estimate (solid square) lies below the efficiency frontier, thus the new intervention would be rejected and the horizontal distance M suggests a €75 cost reduction. However, when computing all the individual distances we got cost reductions A (€77), B (€131) and C (€130), and cost increases D (€150), E (€175) and F (€290). On average, this would suggest a €46 cost increase, which contradicts the €75 cost reduction suggested by M. Note also that a cost increase would keep the average point estimate below the efficiency frontier (thus it would not be accepted) and may decrease the probability of accepting the new intervention since some of the outcomes located to the left of the efficiency frontier may move to the right of the efficiency frontier (and thus become rejected). Therefore, it does not seem appropriate in general to base cost recommendations on the average of the distances between every single PSA outcome and the efficiency frontier.

This result is especially important in case of a probabilistic efficiency frontier since in that case the distance between the average point estimate and the efficiency frontier cannot be computed. We will come back to this issue in Section 2.4 and Section 3.2.

[§] This issue is illustrated with scenarios 9 and 10 in Section 3.1.

Figure 7: PSA outcomes and average cost adjustment



2.3 Simulation analysis

We set up a simulation study in order to assess the impact of using various definitions of when a new intervention is deemed acceptable. The assumptions we made in our simulations analyses are summarized in Table 1.

Table 1. Simulation assumptions.

Assumptions
1. The new health care intervention does not need to lower the efficiency of the relevant therapeutic area. Thus, any possible outcomes for the new intervention were allowed in the cost-effectiveness plane so that all the areas below and above the efficiency frontier were used.
2. We have considered a general situation where the first intervention included in the efficiency frontier (the one with the lowest costs and benefits) was not placed at (0, 0) in the cost-effectiveness plane.
3. We have assumed that the costs of the new intervention were always positive and the benefits were constrained between zero and one.
4. We considered Placebo as 'doing nothing'. Thus, when the first intervention included in the efficiency frontier was Placebo then every value of a new intervention providing less benefit than Placebo was considered unacceptable, i.e. it would be unethical to 'do something' as opposed to 'do nothing' so that the patient loses health, even if it leads to cost.
5. In our analyses we assumed that there are at least two interventions determining the efficiency frontier. Thus, we did not consider the situation where only one intervention constituted the efficiency frontier.

Data and scenarios

The simulation analyses were based on the model developed by IQWiG to assess the cost-effectiveness of antidepressants [13]. We considered two main setups depending on whether the efficiency frontier was chosen to be fixed (deterministic) or probabilistic.

Deterministic efficiency frontier

We first studied the situation where the efficiency frontier was considered to be fixed. The efficiency frontier chosen for these analyses was the one obtained in the cost-effectiveness evaluation of antidepressants by IQWiG – Scenario 1 (Remission) on pages 198-199 [13]. We have chosen this efficiency frontier because it is determined by 3 interventions (i.e. it has 2 segments) which allowed comparisons against different segments or thresholds. The other efficiency frontiers presented in Section 6.7 of the IQWiG report were determined by 2 or 1 intervention only (thus, one segment or no segment).

Furthermore, we have considered several scenarios where a random sample of size of 1,000 from costs and benefits of a hypothetical new intervention was generated. We assumed independency between costs and benefits, i.e. there was no correlation assumed between costs and benefits. In each of these scenarios the 1,000 outcomes (referred further as PSA outcomes) were located in a different area of the cost-effectiveness plane so that with all the scenarios we covered the entire plane. We also studied the effect of increasing the uncertainty in costs and benefits. Thus, each scenario was performed twice: one time with "base case" uncertainty and another time with increased uncertainty. For each scenario we calculated the metrics defined in Table 2.

Table 2. Metrics computed in the scenarios where the efficiency frontier was assumed to be deterministic.

Metrics – deterministic efficiency frontier scenarios
M1. The percentage of PSA outcomes above the efficiency frontier and above the minimum (i.e. Placebo) benefits on the efficiency frontier: this is the percentage of PSA that we considered to be

accepted given the current efficiency frontier.
M2. The percentage of PSA outcomes above the last segment of the efficiency frontier: although we have made clear that this is an erroneous decision rule, we have decided to compute this metric to study the differences with the correct decision rule M1.
M3. The percentage of PSA outcomes below the efficiency frontier and above the maximum benefits on the efficiency frontier: to illustrate what occurs in area B in Figure 3.
M4. For the average of the PSA outcomes, we computed the cost reduction or increase (i.e. horizontal distance) needed to bring it to the efficiency frontier.
M5. For all (accepted and rejected) PSA outcomes separately, we computed the cost reduction or increase needed to bring it to the efficiency frontier and then we took the average of all these.
M6. For all (accepted and rejected) PSA outcomes separately, we computed the cost reduction or increase needed to bring it to the efficiency frontier and then we took the median of all these.

Although we have explained in Section 2.3 that the correct way of calculating the cost recommendation should be based on the horizontal distance needed to bring the average of the PSA outcomes to the efficiency frontier (metric M4), we have also computed metrics M5 and M6 and compared them with M4. As we will explain in the next section, metrics M5 and M6 become relevant when the efficiency frontier is probabilistic since in that case metric M4 cannot be computed.

Probabilistic efficiency frontier

For the probabilistic efficiency frontier scenario we used a real PSA dataset provided by IQWiG. This dataset was the result of the PSA described in the cost-effectiveness evaluation of antidepressants by IQWiG – Table 81 [13]. In each iteration a new efficiency frontier and a point estimate (costs and benefits) of 4 new interventions (i.e. Venlafaxin, Duloxetin, Bupropion and Mirtazapin) were generated and the decision was based on these. Besides the metrics defined Table 2, in this case we also computed the metrics defined in Table 3.

Table 3. Metrics computed in the scenarios where the efficiency frontier was assumed to be probabilistic.

Metrics – probabilistic efficiency frontier scenarios.
M7. The number of segments on the efficiency frontier in each simulation.
M8. The position of the comparators (i.e. placebo, TZA, SSRI minimum, SSRI maximum, Agomelatin and Trazodon) on the efficiency frontier in each simulation.
M9. The mean of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier.
M10. The median of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier.
M11. The 2.5% percentile of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier.
M12. The 97.5% percentile of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier.
M13. The minimum of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier.
M14. The maximum of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier.

It is important to emphasize that in this situation, where the efficiency frontier is probabilistic, the horizontal distance needed to bring the average of the PSA outcomes to the efficiency frontier (metric M4 in Table 2) cannot be calculated. Whereas it is possible to compute the average of the PSA outcomes for every new intervention, the distance from this average

point estimate to *the* efficiency frontier cannot be measured since a single reference efficiency frontier does not exist, i.e. the efficiency frontier is different in each PSA simulation.

3. Results

3.1 Fixed (deterministic efficiency frontier)

We have studied 16 different scenarios where the location of the PSA outcomes in the cost-effectiveness plane and the uncertainty in costs and benefits were varied. A detailed description of the 16 scenarios can be found in Appendix A. The efficiency frontier was considered to be fixed and equal to the one obtained in Scenario 1 (Remission) of the cost-effectiveness evaluation of antidepressants performed by IQWiG [13]. Thus, the efficiency frontier is determined by three interventions (two segments), namely placebo, TZA and Agomelatin. Table 4 summarizes the results obtained.

Table 4. PSA metrics for scenarios 1 to 16. A negative monetary value means cost reduction whereas a positive one means cost increase.

	Acceptance			Cost recommendation		
	M1	M2	M3	M4	M5	M6
Scenario 1	100%	100%	0.0%	€2176	€2176	€2198
Scenario 2	98.9%	96.9%	0.0%	€2170	€2191	€2214
Scenario 3	80.9%	42.4%	0.0%	€123	€285	€136
Scenario 4	70.3%	44.0%	0.0%	€123	€412	€157
Scenario 5	99.0%	99.0%	0.5%	€1731	€1733	€1753
Scenario 6	93.0%	93.0%	2.9%	€1638	€1654	€1645
Scenario 7	17.7%	17.7%	8.8%	-€361	-€209	-€348
Scenario 8	28.3%	28.3%	7.6%	-€359	-€52	-€328
Scenario 9	64.7%	61.5%	0.2%	€223	€394	€188
Scenario 10	64.9%	60.2%	0.1%	€255	€543	€305
Scenario 11	46.7%	43.3%	0.0%	-€61	€200	-€16
Scenario 12	47.8%	44.7%	0.4%	-€64	€307	-€7
Scenario 13	59.7%	48.0%	0.1%	€10	€425	€69
Scenario 14	34.3%	34.3%	8.7%	-€398	€7	-€330
Scenario 15	28.2%	23.2%	0.0%	-€35	€283	-€68
Scenario 16	28.3%	23.5%	0.1%	-€37	€292	-€71

M1=percentage of PSA outcomes above the efficiency frontier and above Placebo benefits; M2=percentage of PSA outcomes above the last segment of the efficiency frontier; M3=percentage of PSA outcomes below the efficiency frontier and above the maximum benefits (area B in Figure 3); M4=cost recommendation based on the average point estimate of the PSA outcomes; M5=average of the cost recommendations based on the PSA outcomes separately; M6=median of the cost recommendations based on the PSA outcomes separately.

First note that the probability of accepting the new intervention (given by metric M1) depended on the position of the PSA outcomes in the cost-effectiveness plane. If the erroneous rule of accepting the PSA outcomes above the last segment on the efficiency frontier (metric M2) was considered, the probability of acceptance would be smaller than or equal to the one obtained with the whole efficiency frontier (metric M1). The impact of using the wrong decision rule can be very strong as shown in scenarios 3 and 4, where the probability of acceptance decreased by approximately 40% and 26%, respectively. The percentage of PSA outcomes below the efficiency frontier but above maximum benefits (area B in Figure 3) was low in general, being 8.8% the highest observed in scenario 7. The

probability of acceptance was slightly affected by increasing the uncertainty around costs and benefits. It decreased or increased depending on the position of the PSA outcomes on the cost-effectiveness plane.

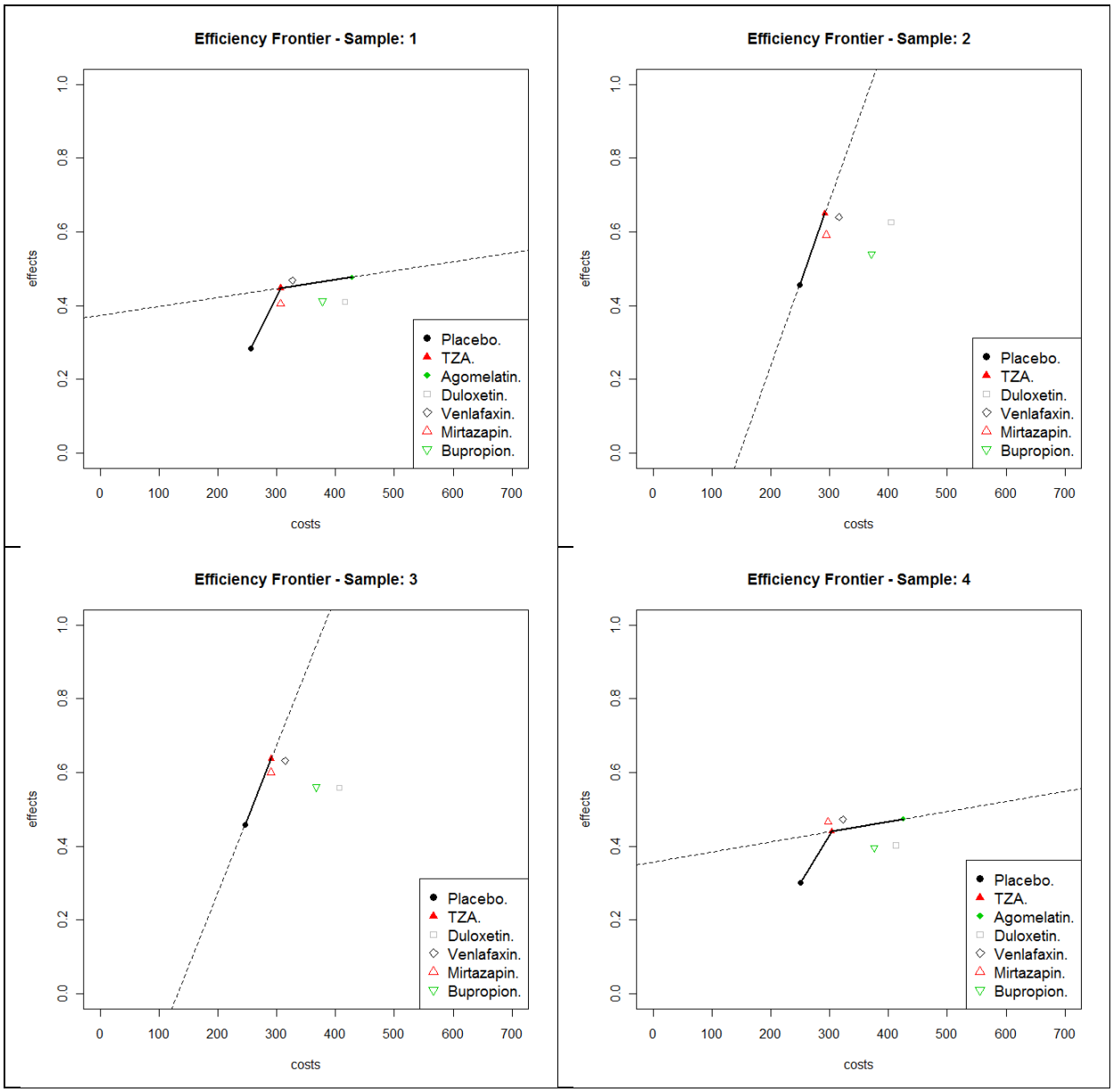
As mentioned in Section 2.3, cost recommendations should be based on the average point estimate of the PSA outcomes. An increase or decrease in cost recommendation should be then informed with metric M4 and it depended on the position of the average point estimate of the PSA outcomes on the cost-effectiveness plane: when it lies above the efficiency frontier a cost increase is recommended and a cost decrease is recommended otherwise.

In the hypothetical case that the cost recommendation should be based on the average (metric M5) or the median (metric M6) of the cost recommendations based on the PSA outcomes separately, we were interested in quantifying to what extent these would differ from the recommendation based on metric M4. We observed that this depended on the distribution of the cost recommendations based on the PSA outcomes separately. When there is symmetry as in scenarios 1, 2, 5 and 6 (see Figure 13 and Figure 17) the three metrics are similar, and so are the cost recommendations. However, when the distribution is skewed (as in scenarios 3, 4 and 7 to 16), metric M5 (the average) is affected by the large values at the right tail of the distribution and metric M6 (the median) provided a better approximation to metric M4.Thus, when a proxy is sought for metric M4, in this hypothetical situation, cost recommendations should be based on the median (metric M6) rather than the average (metric M5).

3.2 Probabilistic efficiency frontier

For this scenario, in which the uncertainty around the efficiency frontier is also taken into account, we considered a real PSA dataset (which contains 100,000 PSA outcomes) provided by IQWiG which corresponds to the cost-effectiveness evaluation of antidepressants performed by IQWiG in Table 81 [13]. Besides the point estimates for the new interventions (Venlafaxin, Duloxetin, Bupropion and Mirtazapin), the efficiency frontier is also randomly sampled in each iteration. An example of the first four configurations can be seen in Figure 8. We can observe that the efficiency frontier and the position of the new interventions with respect to the efficiency frontier change for each iteration.

Figure 8: PSA samples with probabilistic efficiency frontier.



Distribution of the probabilistic efficiency frontier

We first report the results regarding the distribution of the efficiency frontier. As can be observed in Figure 8, the number of interventions (and therefore the number of segments) on the efficiency frontier may vary per simulation. In particular, there were 75,040 simulations

where the efficiency frontier had only one segment (thus, determined by two interventions), 21,645 simulations with exactly two segments (three interventions), 3,276 simulations with exactly 3 segments (four interventions) and 39 simulations with exactly 4 segments (five interventions). In all simulations placebo was the first element on the efficiency frontier (cheapest and with lowest benefits). The second element on the efficiency frontier was TZA in 94,913 simulations, SSRI minimum in 5,086 simulations and Trazodon in only one. The third element on the efficiency frontier was Agomelatin in 20,498 simulations, TZA in 4,185 simulations, Trazodon in 228 simulations and SSRI maximum in 49 simulations. The fourth element on the efficiency frontier was Agomelatin in 3,252 simulations and Trazodon in 63. Finally, in the 39 simulations with five interventions, Agomelatin was in the fifth position.

As a result from the 100,000 simulations we obtained an *average* slope of the last segment on the efficiency frontier equal to 0.00352 with median 0.00421 and (0.00002, 0.00641) as 95% percentile interval. The willingness-to-pay threshold (λ) is computed as 1 divided by the slope of the last segment on the efficiency frontier. Thus, the average of these 100,000 estimates of the willingness-to-pay was €$4.1*10^{13}$ with median €237 and (€156, €40,022) as 95% percentile interval. Note that this average is extremely high due to the simulations where the slope was nearly zero (the minimum of the last slopes computed was of the order e^{-18}). Note however that this average does not correspond to the *standard* definition of the ICER as $\overline{\Delta C}/\overline{\Delta E}$. In fact, as explained in Section 2.4, the *median* of the 100,000 simulated λ's is a better approximation to the standard ICER than the average of the inverse slopes of the last segment (see e.g. Ross handbook [18] – Section 3.3). All the metrics regarding the slope of the last segment of the efficiency frontier and the willingness-to-pay are summarized in Table 5.

Table 5. Descriptive statistics for the slope of the last segment of the efficiency frontier and the maximum willingness-to-pay.

	M9	M10	M11	M12	M13	M14
Slope last segment	0.0035	0.0042	0.00002	0.0064	$3*10^{-18}$	0.0091
Willingness-to-pay (λ)	€$4.1*10^{13}$	€237	€156	€40,022	€109	€$3*10^{17}$

M9=mean of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier; M10=median of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier; M11=2.5% percentile of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier; M12=97.5% percentile of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier; M13=minimum of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier; M14=maximum of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier.

We would like to emphasize that the results shown in Table 5 were obtained using the (100,000) last segments obtained in our simulation, thus, as mentioned above, in 75,040 of these simulations the efficiency frontier had only one segment (first segment usually means large slope – more vertical than horizontal – or equivalently low willingness-to-pay). Therefore, the results in Table 5 are mainly driven by the situation where only one segment forms the efficiency frontier. For that reason, we show in Table 6, the same descriptive statistics as in Table 5, but for each segment separately. That way, we can observe what the contribution of each segment to the willingness-to-pay might be. However, the descriptive statistics in Table 6 must be interpreted with caution, especially when the number of observations used to compute them is small. Thus, the descriptive statistics for the 4th segment are based on 39 observations only, whereas for the 1st, 2nd and 3rd segment we

used 100,000 (the whole sample), 24,960 and 3,315 observations, respectively. This gives an indication about the reliability of those descriptive statistics.

Table 6. Descriptive statistics for the slope of each segment of the efficiency frontier and their associated willingness-to-pay.

	n**	M9	M10	M11	M12	M13	M14
Slope segment 1	100,000	0.0043	0.0043	0.0025	0.0064	0.0007	0.0091
Willingness-to-pay 1	100,000	€242	€231	€155	€395	€109	€1,314
Slope segment 2	24,960	0.0004	0.0001	0.000006	0.0028	0	0.0042
Willingness-to-pay 2	24,960	€35,527	€5,713	€344	€146,634	€237	∞
Slope segment 3	3,315	0.0002	0.0002	0.00001	0.0006	0.0000003	0.0016
Willingness-to-pay 3	3,315	€13,817	€4,136	€1,440	€76,630	€607	€2,692,600
Slope segment 4	39	0.0001	0.0001	0.00002	0.0005	0.00001	0.0005
Willingness-to-pay 4	39	€13,700	€6,795	€1,950	€47,567	€1,813	€74,956

M9=mean of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier; M10=median of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier; M11=2.5% percentile of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier; M12=97.5% percentile of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier; M13=minimum of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier; M14=maximum of the slopes of the last segment (maximum willingness-to-pay) on the efficiency frontier.

First note that the median willingness-to-pay threshold associated to the last segment on the efficiency frontier in Table 5 (resulting from the 100,000 simulations) was €237 which is very similar to the mean (€242) and median (€231) willingness-to-pay threshold associated to the 1st segment on the efficiency frontier in Table 6. This is because, as mentioned above, about 75% of the 100,000 segments used to compute the median willingness-to-pay in Table 5 were actually the 1st segment on the efficiency frontier.

We also found that 49 of the 100,000 simulations produced a 2nd segment on the efficiency frontier with a slope equal to zero (horizontal segment) or equivalently an infinite willingness-to-pay. For that reason, the metrics regarding the 2nd segment shown in Table 6 have been computed ignoring these simulations (thus, especially the mean, may be underestimated). The mean and median willingness-to-pay thresholds associated to the 2nd segment on the efficiency frontier in Table 6 were €35,527 and €5,713, respectively. In this case the mean is larger due to the observations where the 2nd segment had a slope very close to zero. Conclusions about the 3rd segment are similar to those about the 2nd one, but no infinite willingness-to-pay was observed. Finally, the descriptive statistics for the 4th segment were based on 39 only. Therefore, it is hard to extract any reliable conclusion from these statistics.

Acceptance and cost change with the probabilistic efficiency frontier

Next we discuss the results regarding acceptance and cost changes of the new interventions considered in the simulation. First note that in each PSA iteration a *single* point estimate for the new interventions (Venlafaxin, Duloxetin, Bupropion and Mirtazapin) and a new efficiency frontier were randomly sampled. Thus, the decision whether to accept or not the new intervention per iteration is based on the position of that single point estimate relative to that single efficiency frontier. Therefore, the *average* point estimate for costs and benefits and the

** Number of segments observed in the 100,000 simulations.

cost increase or reduction needed to bring that average to the efficiency frontier (i.e. metric M4) cannot be computed here. Table 7 summarizes the results obtained in this scenario.

Table 7. Metrics for the probabilistic efficiency frontier scenario for Duloxetin, Venlafaxin, Mirtazapin and Bupropion. A negative monetary value means cost reduction whereas a positive one means cost increase.

	Acceptance			Cost recommendation	
	M1	M2	M3	M5 [††]	M6
Duloxetin	0.28%	0.28%	0.27%	-€128	-€131
Venlafaxin	13.92%	13.91%	10.71%	€33	-€29
Mirtazapin	12.22%	10.35%	0.00%	€1	-€12
Bupropion	0.10%	0.10%	0.06%	-€99	-€99

M1=percentage of PSA outcomes above the efficiency frontier and above Placebo benefits; M2=percentage of PSA outcomes above the last segment of the efficiency frontier; M3=percentage of PSA outcomes below the efficiency frontier and above the maximum benefits (area B in Figure 3); M5=average of the cost recommendations based on the PSA outcomes separately; M6=median of the cost recommendations based on the PSA outcomes separately.

First note that when calculating the distances between every PSA outcome and every efficiency frontier, we found that for all the new interventions some simulations provided a (numerically) infinite distance (or equivalently an infinite cost increase). This shows that, as we expected, the distribution of the cost recommendations based on the PSA outcomes separately is affected by the presence of outliers. In this situation, the median (metric M6) instead of the average (metric M5) should be used to determine the price recommendation. In any case, the average reported in Table 7 has been computed without considering the simulations with an infinite price recommendation.

For all the new interventions the percentage of acceptable PSA outcomes (above efficiency frontier and above minimum benefits) was low, with 13.92% for Venlafaxin being the highest. Venlafaxin and Mirtazapin had similar percentage of acceptance. The main difference between these two was that for Venlafaxin 10.71% of the PSA outcomes were located above the maximum benefits but below the efficiency frontier (area B in Figure 3). Therefore, (some of) these PSA outcomes could become acceptable if λ was increased. For Duloxetin and Bupropion the probability of being acceptable was almost zero. Note finally that the erroneous rule of accepting only the PSA outcomes above the last segment on the efficiency frontier (metric M2) affects mostly Mirtazapin since it reduces its acceptance probability in almost 2%.

The average cost of Duloxetin was €408, with a 95% percentile interval (€392, €425). When calculating the average of the distances between every PSA outcome and every efficiency frontier, we found that 16 simulations provided a (numerically) infinite distance (or equivalently an infinite cost increase). Thus, the average shown in Table 7 has been computed without these simulations. After that, the average of the distances between every PSA outcome and every efficiency frontier implied a -€128 cost reduction. The median was -€131 and the 95% percentile interval was equal to (-€144, -€110). Thus, excluding possible

[††] Simulations providing an infinite cost increase were excluded from the computation of the average. Therefore, it may be expected that the true average is larger than the one reported here.

'outliers', the results for Duloxetin seem to be consistent with a cost reduction recommendation.

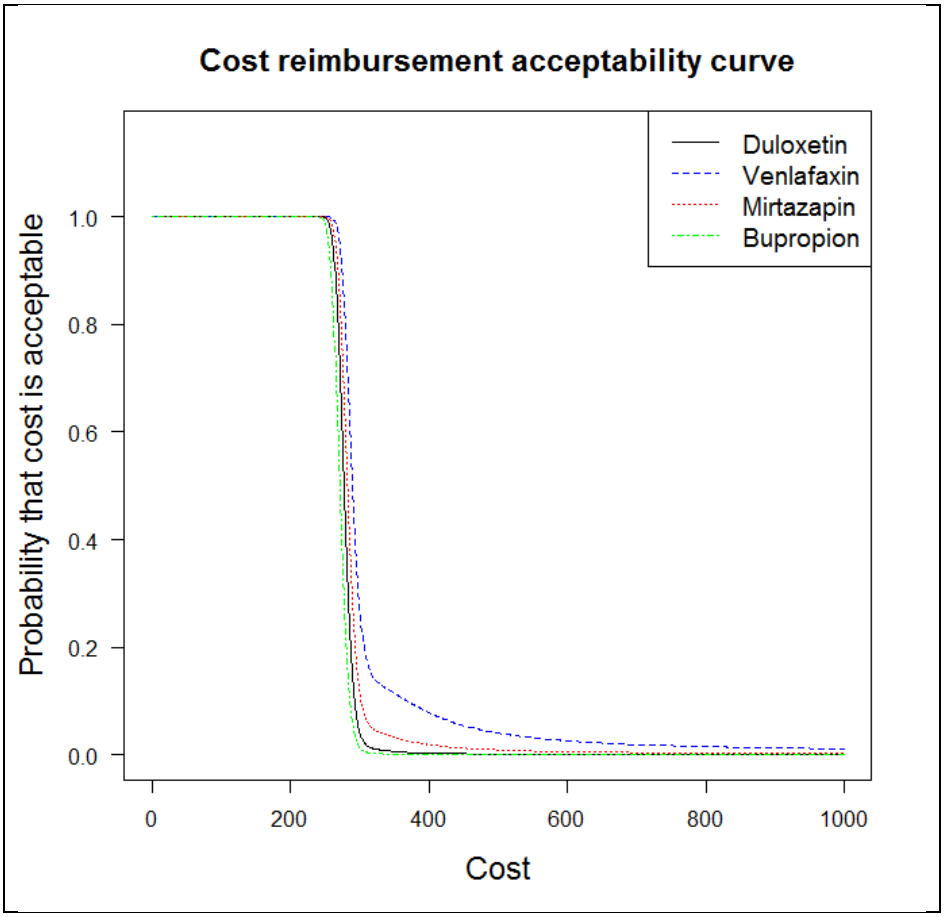
In case of Venlafaxin, the average cost was €319, with a 95% percentile interval (€301, €338). When calculating the average of the distances between every PSA outcome and every efficiency frontier, we also found that 49 simulations provided an infinite cost increase, which were excluded for the average calculation. In this case, the average of the distances between every PSA outcome and every efficiency frontier would imply a €33 cost increase. However, the median was -€29, implying a cost reduction. The 95% percentile interval was equal to (-€44, €287), thus including both positive and negative values. Therefore, the results for Venlafaxin seem to be more uncertain than those for Duloxetin.

The average cost was €295 for Mirtazapin, with a 95% percentile interval (€276, €315). In this case 30 simulations provided an infinite distance (or equivalently an infinite cost increase recommendation), which were excluded from the calculation of the average. After that, the average of the distances between every PSA outcome and every efficiency frontier would imply a €1 cost increase (thus approximately exactly on the efficiency frontier). However, the median was -€12, implying a cost reduction. The 95% percentile interval was equal to (-€29, €79). Therefore, as occurred with Venlafaxin, the results for Mirtazapin seem to be somewhat uncertain.

Finally, for Bupropion we observed that the average cost was €372, with (€354, €390) as 95% percentile interval. When calculating the average of the distances between every PSA outcome and every efficiency frontier, we found that only 2 simulations provided an infinite cost increase. Thus, the average shown in Table 7 has been computed without these simulations. After that, the average of the distances between every PSA outcome and every efficiency frontier implied a -€99 cost reduction. The median was also -€99 and the 95% percentile interval was equal to (-€116, -€80). Thus, the results for Bupropion seem to strongly suggest a -€99 cost reduction recommendation.

As mentioned in Section 2.3, cost reimbursement acceptability curves can be used to estimate the probability that a suggested cost for a new intervention is going to be acceptable. We have calculated the cost reimbursement acceptability curves for the four interventions considered here. These can be seen in Figure 9. We observed that Venlafaxin was always the intervention with the highest probability of being reimbursed, followed by Mirtazapin, Duloxetin, and Bupropion. The largest differences were found in the range between (approximately) €300 and €400, where the reimbursement probabilities for Mirtazapin, and especially for Venlafaxin, were relatively large compared to the reimbursement probabilities for the other two interventions. Recommended costs smaller than or equal to €239, €251, €241 and €229 for Duloxetin, Venlafaxin, Mirtazapin and Bupropion, respectively, were always reimbursed (i.e. accepted with probability 1). As the recommended costs increase, the reimbursement probability decreases for all interventions. In particular, we found a median maximum reimbursable cost (i.e. accepted with probability 0.5) equal to €277, €290, €282 and €271 for Duloxetin, Venlafaxin, Mirtazapin and Bupropion, respectively. Note that this is in line with the results described above, i.e. the median maximum reimbursable cost is approximately the average cost of the intervention minus the median cost reduction. For example, for Duloxetin its average cost is €408, minus its median cost reduction €131 is exactly equal to €277. Due to the simulations providing an infinite cost increase, the acceptability curves do not converge to zero in any case (although the probability of accepting high costs is very small).

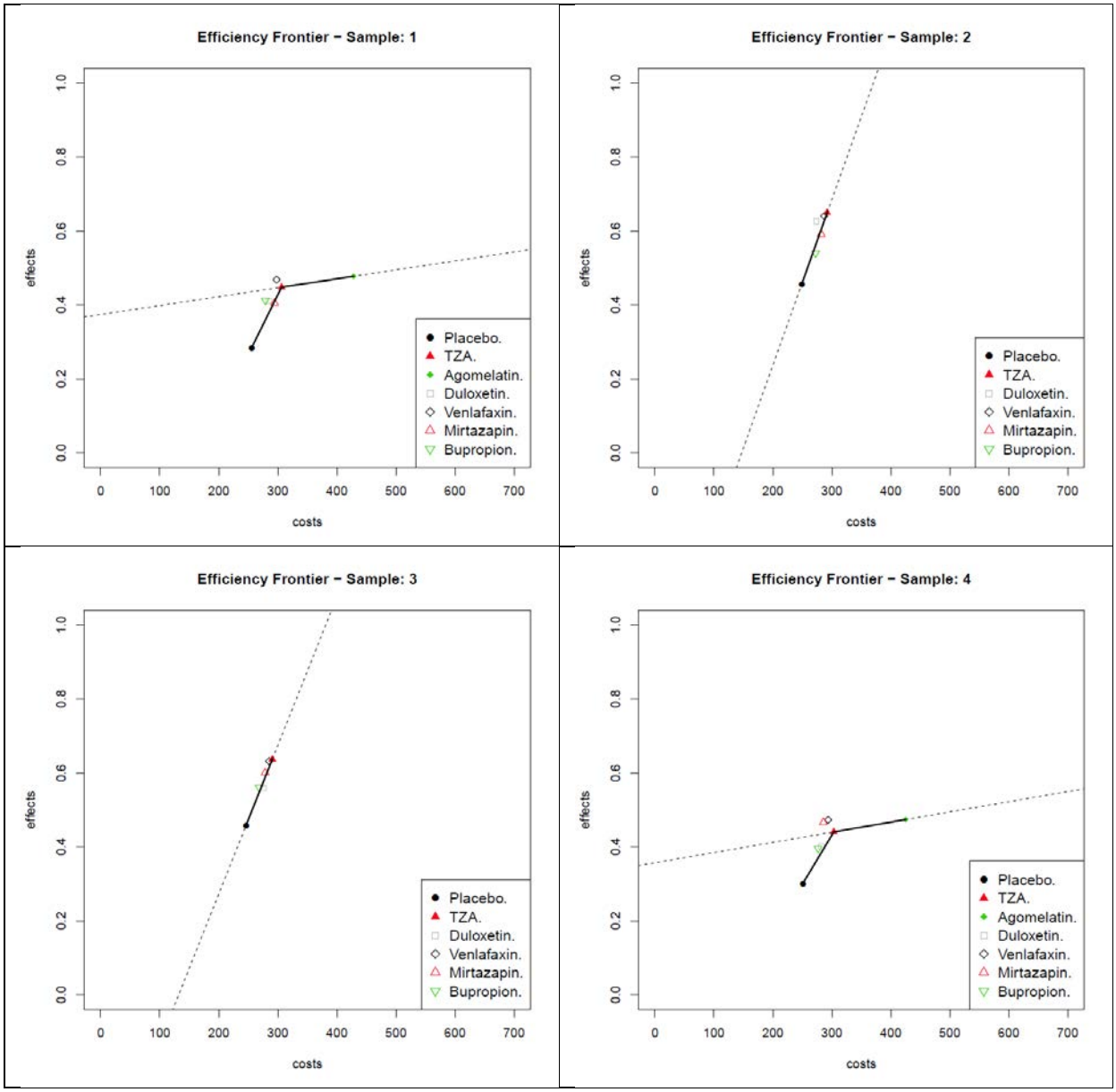
Figure 9: Cost reimbursement acceptability curves for Duloxetin, Venlafaxin, Mirtazapin and Bupropion.



Implementing cost recommendations

The results of the previous simulation analysis suggest a cost reduction for all the new interventions as determined by the median of the cost recommendations based on the PSA outcomes separately (metric M6 in Table 7). More specifically, we found that the costs of Duloxetin, Venlafaxin, Mirtazapin and Bupropion should be decreased by €131, €29, €12 and €99, respectively. Thus, for this scenario we implemented the cost recommendations and repeated the analysis. An example of the first four configurations after implementing the cost recommendations can be seen in Figure 10. We can observe that compared to Figure 8, the position of the new interventions has been simply shifted to the left.

Figure 10: PSA samples with probabilistic efficiency frontier after implementing cost recommendations.



The simulation results after the cost recommendations have been implemented are shown in Table 8.

Table 8. Metrics for the probabilistic efficiency frontier scenario with cost recommendations implemented for Duloxetin, Venlafaxin, Mirtazapin and Bupropion. A negative monetary value means cost reduction whereas a positive one means cost increase.

	Acceptance			Cost recommendation	
	M1	M2	M3	M5	M6
Duloxetin	49.65%	33.09%	0.00%	€3	€0
Venlafaxin	48.67%	36.07%	0.00%	€62	€0
Mirtazapin	47.96%	36.07%	0.00%	€13	€0

Bupropion	43.81%	28.98%	0.00%	€0	-€1

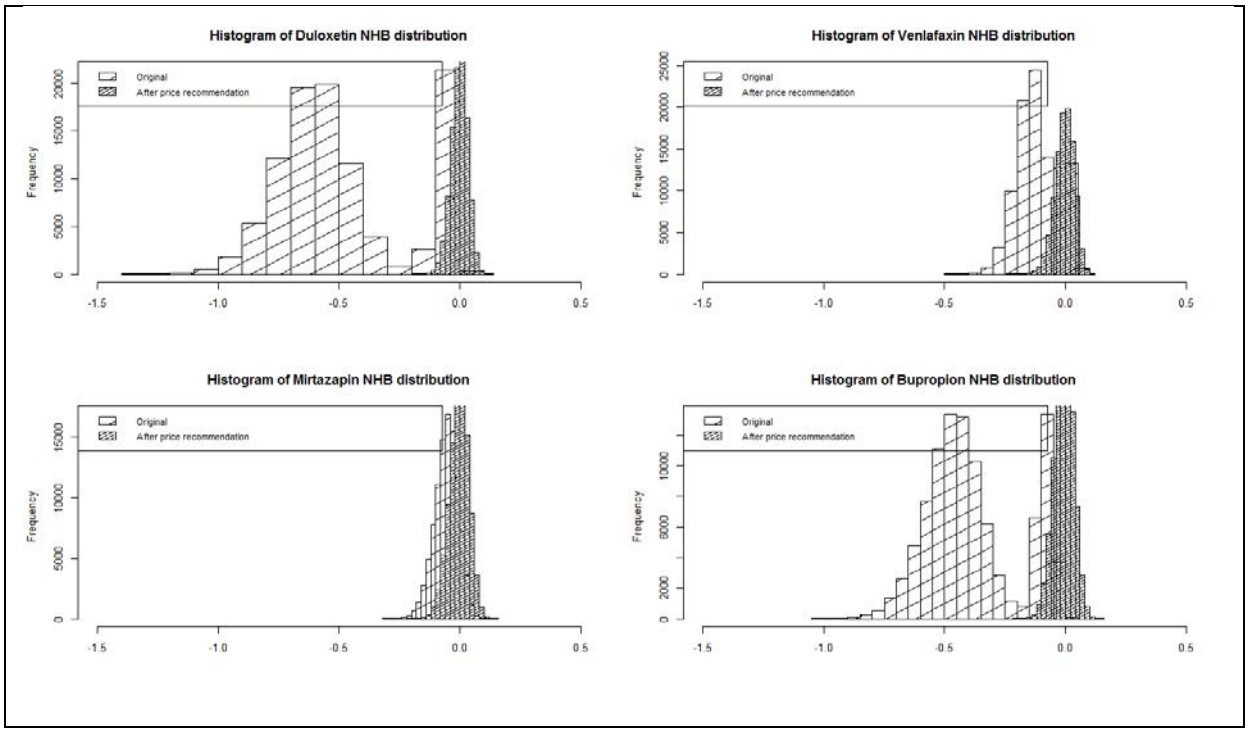
M1=percentage of PSA outcomes above the efficiency frontier and above Placebo benefits; M2=percentage of PSA outcomes above the last segment of the efficiency frontier; M3=percentage of PSA outcomes below the efficiency frontier and above the maximum benefits (area B in Figure 3); M5=average of the cost recommendations based on the PSA outcomes separately; M6=median of the cost recommendations based on the PSA outcomes separately.

For all the new interventions, the percentage of acceptable PSA outcomes increased significantly, with 43.81% for Bupropion being the lowest, and 49.65% for Duloxetin being the highest. The erroneous rule of accepting only the PSA outcomes above the last segment on the efficiency frontier (metric M2) would affect all the interventions as well since it would reduce the acceptance probability in at least 12% (Mirtazapin). Note also that no PSA outcomes were located above the maximum benefits but below the efficiency frontier (i.e. area B in Figure 3).

The average cost of Duloxetin was now €277, i.e. the cost before recommendation (€408) minus the recommended reduction (€131), with a 95% percentile interval (€261, €294). The average of the distances between every PSA outcome and every efficiency frontier implied a €3 cost increase. The median, as expected, was €0. For the other three interventions the results can be explained in a similar way as for Duloxetin, i.e. the average costs, average of the distances between every PSA outcome and every efficiency frontier, and the median resulted from subtracting the recommended cost reduction to the “original” costs and distances in Table 7. The only (really minor) exception is the median for Mirtazapin of -€1 where €0 should be expected (which may be caused by rounding up the costs).

Moreover, we have also observed that the uncertainty associated to the distribution of the NHB has been reduced after implementing the cost recommendation as can be seen in Figure 11. Note that the NHB was graphically defined as the vertical distance between a point estimate and the last segment (or its extension – backward or forward) on the efficiency frontier [20]. Thus, it is not only that there are many more PSA outcomes above the efficiency frontier, but also that the PSA outcomes are closer to the last segment on the efficiency frontier (for that reason we observed that the distribution of the NHB after implementing the cost recommendations is centered at zero). Since in most of the cases the efficiency frontier had exactly one segment (around 75%) it also means that the PSA outcomes are in fact closer to efficiency frontier.

Figure 11: NHB distribution before and after implementing cost recommendations.



4. Discussion

In this paper we have explored the potential problems (and solutions) arising when, within the IQWiG efficiency frontier approach, a maximum reimbursable price for a new intervention is sought through probabilistic sensitivity analyses.

We have first explained that the only theoretically correct way of determining whether a new intervention is acceptable is by using the whole efficiency frontier as decision rule [14]. Thus, only when a new intervention lies on the efficiency frontier or above, the intervention should be deemed acceptable. Other options are not grounded in Health Economics theory. As a consequence of this, it is incorrect in general to use the last segment on the efficiency frontier and its backward extension to decide whether a new intervention is accepted or not. This relies on the erroneous decision rule of accepting the interventions whose NHB or NMB is positive.

The conclusions extracted from the simulation analysis with a fixed efficiency frontier are very logical and intuitive. Thus, the percentage of acceptable PSA outcomes depends on their position on the cost-effectiveness plane and the uncertainty surrounding the PSA outcomes. Increasing the uncertainty around costs and benefits had in general little effect on the probability of acceptance. However, the direction of the uncertainty may have an impact on this. Since we assumed independency between costs and benefits, i.e. there was no correlation assumed between costs and benefits, the cloud of PSA outcomes was somewhat circular. Under correlation, the cloud of PSA outcomes would have the form of an ellipse, leaning to the left or to the right depending on whether the correlation is negative or positive. Therefore, the independency assumption has only influence on the shape of the cloud of the PSA outcomes but not on the methodology or the conclusions drawn from this study.

When the efficiency frontier is fixed, cost recommendations should be based on the average point estimate of the PSA outcomes. An increase or decrease in cost recommendation is then calculated as the horizontal distance between the average point estimate and the efficiency frontier (metric M4). In our scenarios the efficiency frontier had two segments. Because of this, the distance between the average probabilistic point estimate and the efficiency frontier (metric M4) and the average of the distances between every PSA outcome and the efficiency frontier (metric M5) were not the same. In fact, we have observed that the average of the distances between every PSA outcome and the efficiency frontier (metric M5) was always larger than the distance between the average point estimate and the efficiency frontier (metric M4). This is because the distribution of metric M5 was affected by very large values (outliers). In this case, the median of the distances between every PSA outcome and the efficiency frontier (metric M6) provided a better approximation to metric M4 and it is preferred to metric M5. This is clearly illustrated in scenario 12 (shown in Figure 22). The average point estimate was below the efficiency frontier (thus the new intervention was rejected) and the reduction needed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 22 – right – and the segment Placebo-TZA) was -€64 (metric M4). However, in this case the average of the distances between every PSA outcome and the efficiency frontier implied a cost increase of €307 (metric M5). Hence, in this scenario these two recommendations would be contradictory. The consequences of adopting either this cost reduction or increase are properly illustrated in scenarios 13 and 14. In particular, the effect of wrongly increasing the cost would produce a

less favorable mean of the PSA outcomes (resulting in a subsequent cost decrease recommendation) and a decreased overall probability of acceptance. It is expected that as the number of segments on the efficiency frontier and their steepness increase, the difference between metric M4 and metric M5 becomes larger, thus indicating that M5 cannot be used as a proxy for M4. Note finally that if the maximum reimbursable cost is going to be determined according to a certain probability of being reimbursed, then cost reimbursement acceptability curves should be used.

In the probabilistic efficiency frontier approach the cost recommendation cannot be based on the distance between the average point estimate of the PSA outcomes and the efficiency frontier (metric M4) since it cannot be computed. What we can compute is the average of the distances between every single PSA outcome and the efficiency frontier, which is defined in Table 2 as M5 and that it is not the same as M4 (as shown in Section 2.3). What happens is that the distribution of the metric M5 can be highly skewed and it may have unusually large values (outliers). As already observed for the fixed frontier, the median (metric M6 in Table 2) is more representative of the typical observation than the average (metric M5 in Table 2) and it is the preferred measure of central tendency of the population under study. This is a general result in statistics (see e.g. Ross handbook [18] – Section 3.3). Thus, the average (metric M5) would overestimate the cost recommendation, which can be very wrong. The median (metric M6) is less sensitive to skewed distributions and outliers and when the efficiency frontier is probabilistic, cost recommendations should be based on it. After implementing the costs recommendations based on the median, we have seen that the probability of acceptance increased dramatically, the uncertainty around the NHB distribution decreased and the “updated” cost recommendation was €0. Finally, note that we had always at least one segment in the efficiency frontier in order to determine the needed or allowed costs. This will not generally be the case, as can be observed in the cost-effectiveness evaluation of antidepressants performed by IQWiG [13]. Here, various situations (depending on scenario and outcome used) occur where placebo dominates all current intervention options. In such situation, there is no clear decision rule to state which outcomes are acceptable, as there is no revealed willingness to pay. How to deal with those occurrences should be subject of further discussion.

5. Conclusions

With a fixed efficiency frontier cost recommendations should be based on the horizontal distance between the average point estimate of the PSA outcomes and the closest segment of the efficiency frontier (metric M4). In general, this is not the same as the average of the distances between every PSA outcome and the efficiency frontier (metric M5). We have observed that M5 was always larger than M4. This was because the distribution of the distances between every PSA outcome and the efficiency frontier was skewed and was affected by outliers. Under these conditions, the median (metric M6) is preferred to M5 as proxy for M4.

In the probabilistic efficiency frontier approach the cost recommendation cannot be based on the distance between the average point estimate of the PSA outcomes and the efficiency frontier (metric M4) simply because this cannot be computed. The distribution of the

distances between every PSA outcome and the efficiency frontier is likely to be skewed and to have outliers. Thus, its average (metric M5) would overestimate the cost recommendation, which can be very wrong. The median (metric M6) is less sensitive to skewed distributions and outliers and when the efficiency frontier is probabilistic, cost recommendations should be based on it.

Literature

1. Barton GR, Briggs AH, Fenwick EA. Optimal cost-effectiveness decisions: the role of the cost-effectiveness acceptability curve (CEAC), the cost-effectiveness acceptability frontier (CEAF), and the expected value of perfection information (EVPI). Value in Health. 2008 Sep-Oct;11(5):886-97

2. Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics. 2000 May;17(5):479-500.

3. Briggs A. Probabilistic analysis of cost-effectiveness models: statistical representation of parameter uncertainty. Value Health. 2005;8(1):1–2.

4. Cantor, S.B. Cost-effectiveness analysis, extended dominance, and ethics: a quantitative assessment. Medical Decision Making 1994; 14: 259 – 265.

5. Caro JJ, Nord E, Siebert U, McGuire A, McGregor M, Henry D, Kolominsky-Rabas P. The efficiency frontier approach to economic evaluation of health-care interventions. Health economics 2010, 19(10), 1117-1127.

6. Drummond MF, Sculpher MJ, Torrance G, O'Brien J, Stoddart GL. Methods for the economic evaluation of health care programmes. 3rd edition. New York: Oxford University Press; 2005.

7. Eckermann S, Briggs A, Willan AR. Health intervention assessment in the cost-disutility plane. Med Decis Making. 2008 Mar-Apr;28(2):172-81.

8. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ. 10: 779–787 (2001)

9. Goeree R, O'Brien BJ, Blackhouse G, Marshall J, Briggs A, Lad R. Cost-effectiveness and cost-utility of long-term management strategies for heartburn. Value Health. 2002 Jul-Aug;5(4):312-28.

10. Hallinen TA, Soini EJ, Eklund K, Puolakka K. Cost-utility of different intervention strategies after the failure of tumour necrosis factor inhibitor in rheumatoid arthritis in the Finnish setting. Rheumatology (Oxford). 2010 Apr;49(4):767-77.

11. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ. 1994 Sep-Oct;3(5):309-19.

12. IQWiG (2009). General methods for the assessment of the relation of benefits to costs Version 1.0. Available at: https://www.iqwig.de/download/General_Methods_for_the_Assessment_of_the_Relation_of_Benefits_to_Costs.pdf

13. IQWiG (2012). Vorbericht G09-01. Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin im Vergleich zu weiteren verordnungsfähigen medika-mentösen Behandlungen. Available at: https://www.iqwig.de/download/G09-01_Vorbericht_Kosten-Nutzen-Bewertung-von-Venlafaxin-Duloxetin-Bupropion-und-Mirtazapin.pdf

14. Karlsson, G. Johannesson , M (1996). The decision rules of cost-effectiveness analysis. Pharmacoeconomics 9 (2), 113-120

15. Murray CJ, Lauer JA, Hutubessy RC, Niessen L, Tomijima N, Rodgers A, Lawes CM, Evans DB. Effectiveness and costs of interventions to lower systolic blood pressure and cholesterol: a global and regional analysis on reduction of cardiovascular-disease risk. Lancet. 2003 Mar 1;361(9359):717-25.

16. Niessen LW, Dijkstra R, Hutubessy R, Rutten GE, Casparie AF. Lifetime health effects and costs of diabetes intervention. Neth J Med. 2003 Nov;61(11):355-64.

17. Rojnik K, Naversnik K, Mateović-Rojnik T, Primiczakelj M. Probabilistic cost-effectiveness modeling of different breast cancer screening policies in Slovenia. Value Health. 2008 Mar-Apr;11(2):139-48.

18. Ross, S.M. Introductory Statistics. Elsevier Science. 2010.

19. Sheskin, DJ. Handbook of Parametric and Nonparametric Statistical Procedures: Third Edition. Taylor & Francis. 2003.

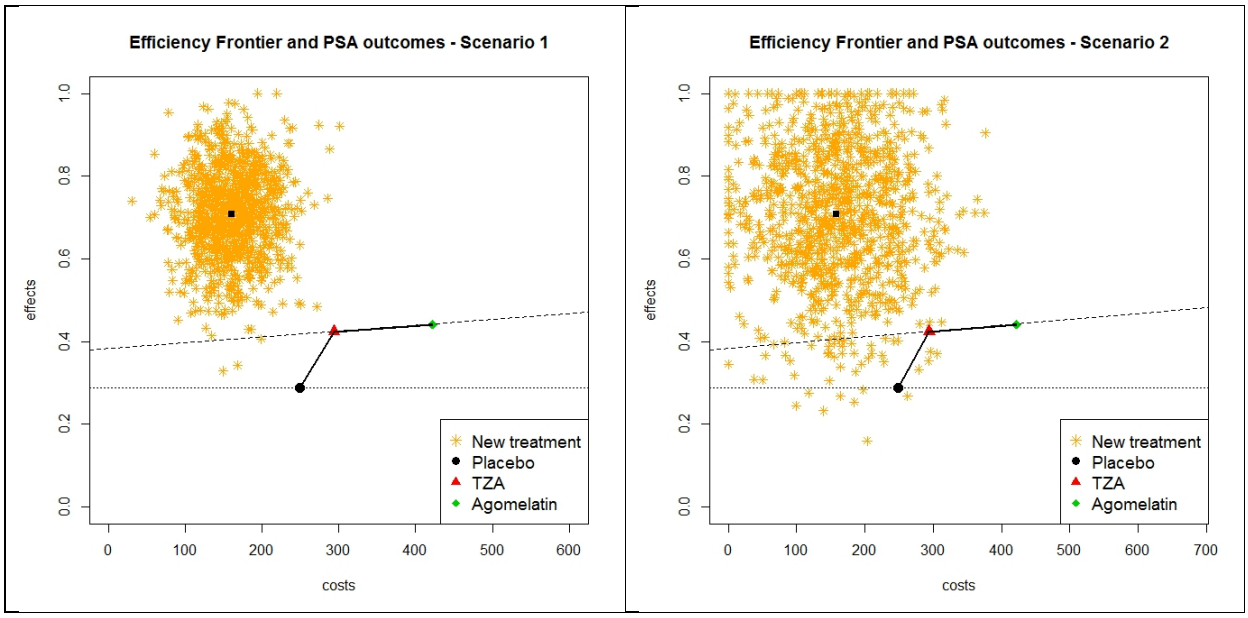
20. Stollenwerk B, Lhachimi SK, Briggs A, Fenwick E, Caro JJ, Siebert U, Danner M, Gerber-Grote A. Communicating the parameter uncertainty in the IQWiG efficiency frontier to decision-makers. Health Economics. Article first published online 2014.

Appendix A – Deterministic efficiency frontier: description of scenarios

Scenarios 1 and 2

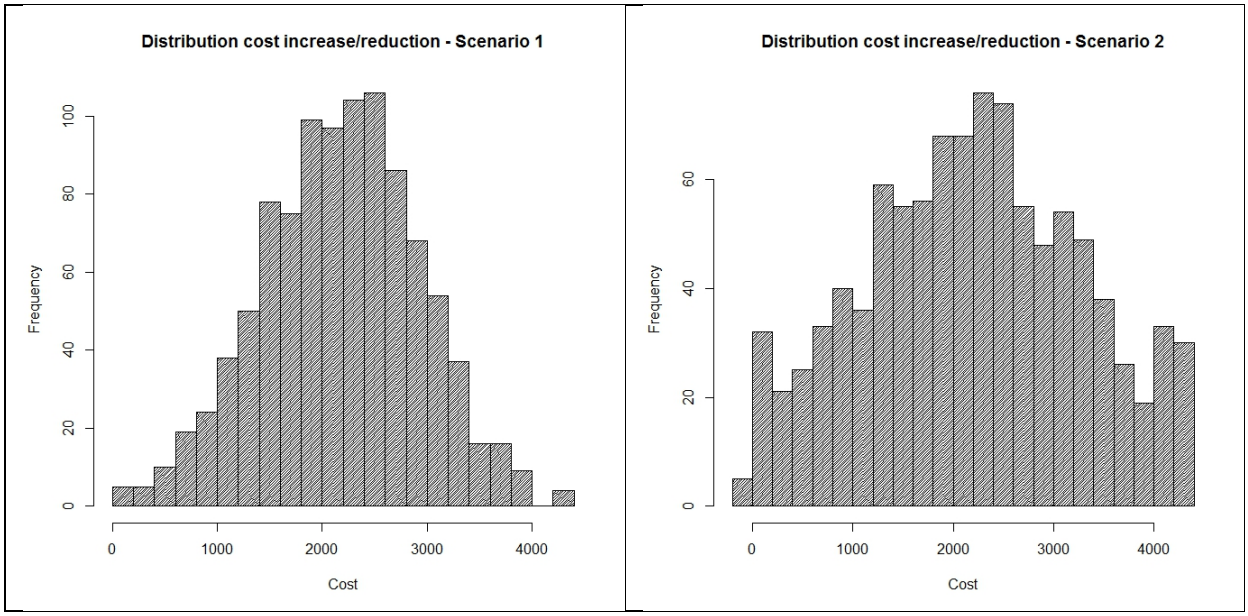
We first considered the two scenarios depicted in Figure 12, where the PSA outcomes of the new intervention were mostly located on the top-left corner of the cost-effectiveness plane and above the efficiency frontier.

Figure 12: PSA scenarios 1 (left) and 2 (right).



Note that in scenario 1 all the PSA outcomes can be found above the efficiency frontier and all provided more benefits than Placebo (which is the intervention in the efficiency frontier with the lowest benefits). Therefore, all the PSA outcomes would be accepted in this scenario (M1=100%). The average cost of the PSA outcomes was €158, with a 95% percentile interval (€83, €231). The average point estimate was above the efficiency frontier, thus the new intervention would be accepted. Moreover, the increase allowed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 12 –left– and the forward extension of the segment TZA-Agomelatin) was €2,176 (metric M4). The average and the median of the distances between every PSA outcome and the efficiency frontier were €2,176 (metric M5) and €2,198 (metric M6), respectively. Note that in this case these three metrics are very similar. This is due to the symmetry of the distribution of the distances between every PSA outcome and the efficiency frontier (almost all the distances are measured against the segment TZA-Agomelatin) as can be seen in Figure 13 (left).

Figure 13: Distribution of the distances between every PSA outcome and the efficiency frontier in scenarios 1 (left) and 2 (right).

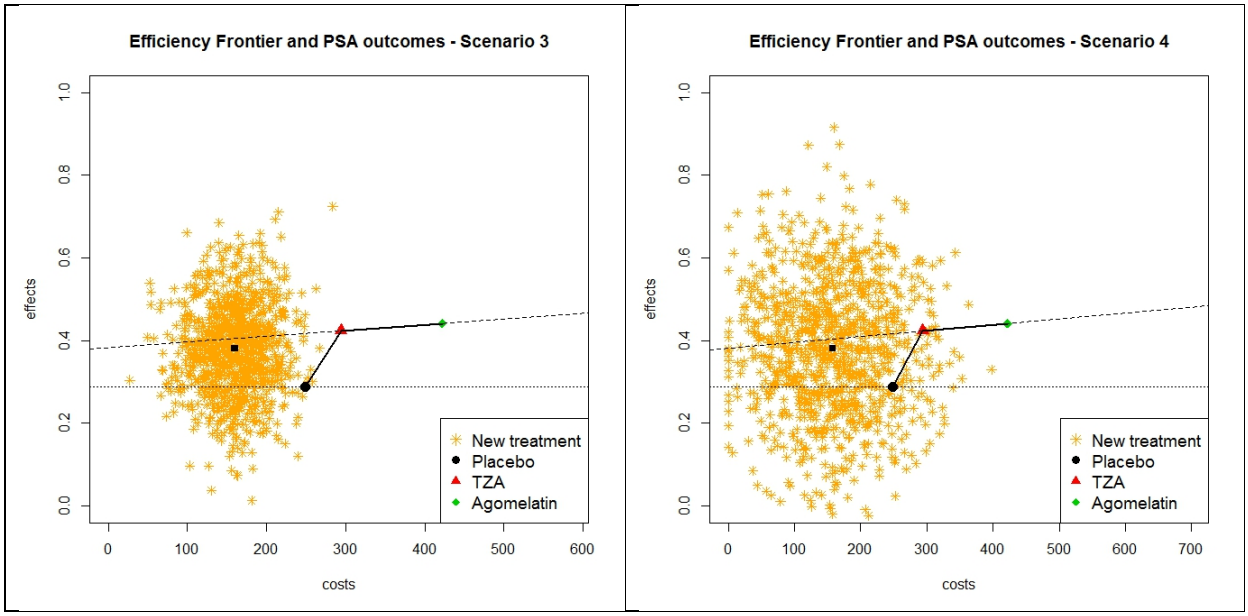


The effect of increasing the uncertainty in scenario 2 (see Figure 12 - right) was small with respect to the probability of accepting the new intervention since M1 was equal to 98.9%. We obtained (as expected) similar cost estimates (averages) but wider percentile intervals. In particular, the average cost was €160, with a 95% percentile interval (€3, €305). The average point estimate was above the efficiency frontier, thus the new intervention would be accepted. Furthermore, the increase costs allowed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 12 –right– and the forward extension of the segment TZA-Agomelatin) was €2,170 (metric M4). The average and the median of the distances between every PSA outcome and the efficiency frontier were €2,191 (metric M5) and €2,214 (metric M6), respectively (thus slightly higher than in scenario 1). In this case the differences between the three metrics are higher than in scenario 1 (due to the increase in uncertainty there are more PSA outcomes compared to the Placebo-TZA segment than in scenario 1).

Scenarios 3 and 4

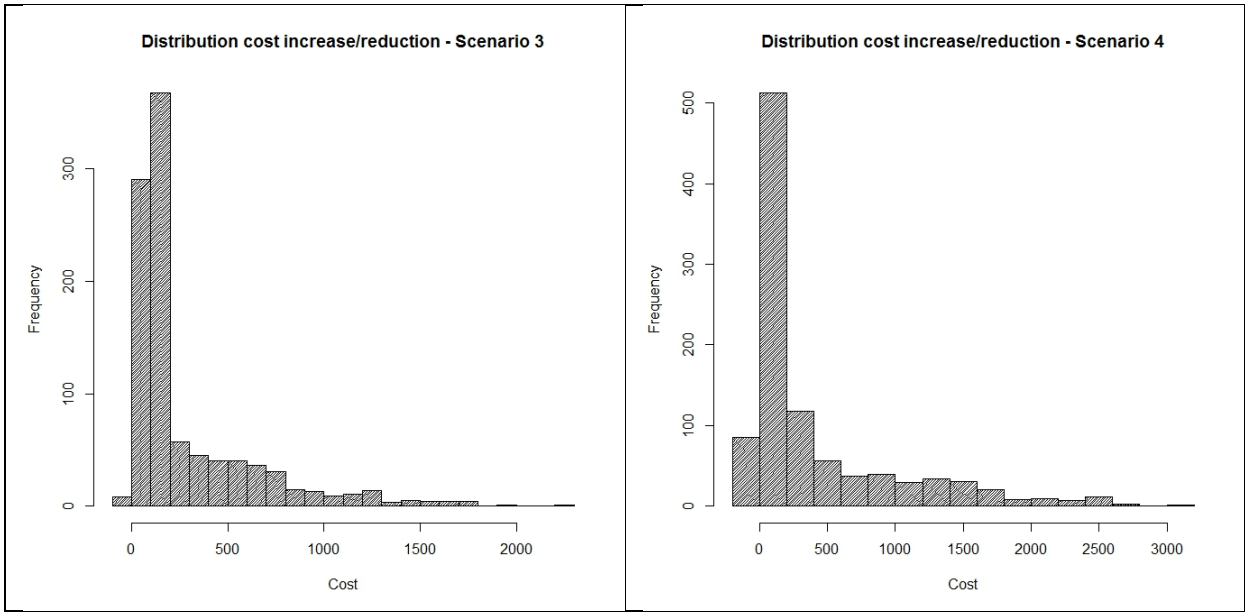
Next we considered the two scenarios depicted in Figure 14 where the PSA outcomes of the new intervention were mostly located at the bottom-left corner of the cost-effectiveness plane and to the left to the efficiency frontier (but with many of them below Placebo benefits, thus rejected).

Figure 14: PSA scenarios 3 (left) and 4 (right).



In scenario 3 about 20% of the PSA outcomes provided less benefit than Placebo. If we considered the (erroneous) rule of accepting the PSA outcomes above the last segment on the efficiency frontier, we would accept only 42.4% (metric M2) of the PSA outcomes, which shows the strong impact of using the wrong decision rule. In this scenario we obtained an average cost equal to €158, with a 95% percentile interval (€80, €234). The average point estimate was above the efficiency frontier, thus the new intervention would be accepted. Moreover, the increase costs allowed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 14 –left– and the segment Placebo-TZA) was €123 (metric M4). The average and median of the distances between every PSA outcome and the efficiency frontier were €285 (metric M5) and €136 (metric M6), respectively. Note that in this case the difference between metrics M4 and M5 is more than double. This is due to the skewness of the distribution of the distances between every PSA outcome and the efficiency frontier as can be seen in Figure 15 (left). This can be explained by the relatively large amount of PSA outcomes that were compared with the TZA-Agomelatin segment (in contrast with the average outcome which is measured against the segment Placebo-TZA). Thus, it seems to be clear that besides the symmetry of the cloud of the PSA outcomes, the number of segments of the efficiency frontier and their steepness play also an important role in explaining the differences between metrics M4 and M5.

Figure 15: Distribution of the distances between every PSA outcome and the efficiency frontier in scenarios 3 (left) and 4 (right).

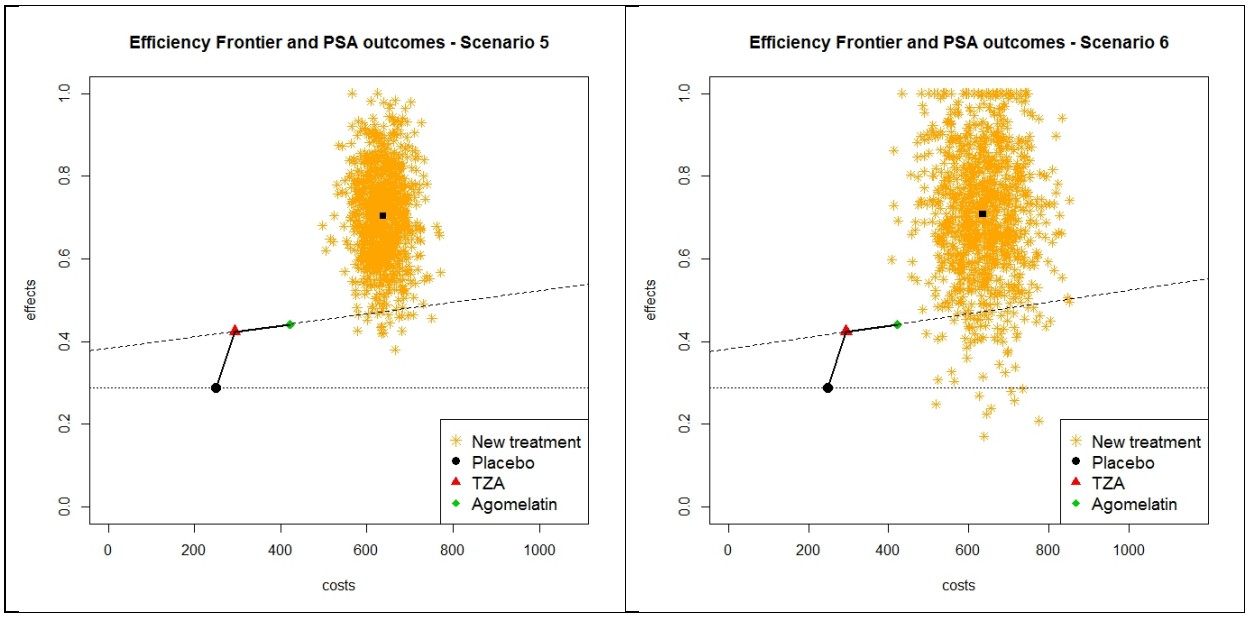


In scenario 4 the probability of accepting the new intervention decreased to 70.3% (metric M1) due to the large amount of PSA outcomes providing less benefit than Placebo. The average cost of the PSA outcomes was now €161, with a 95% percentile interval (€14, €304). The average point estimate was above the efficiency frontier, thus the new intervention would be accepted. Furthermore, the increase allowed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 14 –right– and the segment Placebo-TZA) was also €123 (metric M4). The average and median of the distances between every PSA outcome and the efficiency frontier were €412 (metric M5) and €157 (metric M6), respectively.

Scenarios 5 and 6

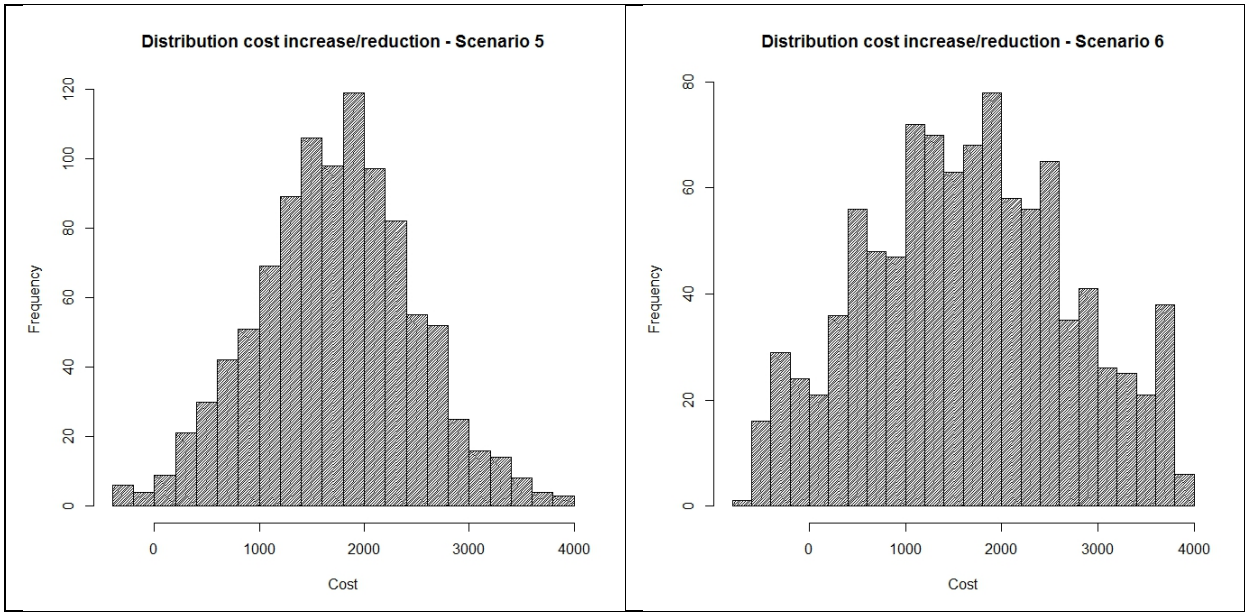
In scenarios 5 and 6 in Figure 16, the PSA outcomes of the new intervention were mostly located on the top-right corner of the cost-effectiveness plane and above the forward extension of the last segment of the efficiency frontier.

Figure 16: PSA scenarios 5 (left) and 6 (right).



In scenario 5 almost all the PSA outcomes (M1 = 99%) were found above the efficiency frontier and all of them provided more benefits than Placebo. Note that this is the first scenario where we found PSA outcomes below the efficiency frontier but above maximum benefits given in the efficiency frontier (Agomelatin) which corresponds to area B in Figure 3 (M3 = 0.5%). In our simulations we considered that the PSA outcomes located in this area would not be accepted. The average cost was €638, with a 95% percentile interval (€564, €709). The average point estimate was above the efficiency frontier, thus the new intervention would be accepted. Moreover, the increase allowed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 16 –left– and the forward extension of the segment TZA-Agomelatin) was €1,731 (metric M4). The average and median of the distances between every PSA outcome and the efficiency frontier were €1,733 (metric M5) and €1,753 (metric M6), respectively. Note that in this case the three metrics are very similar. This is due to the symmetry of the distribution of the distances between every PSA outcome and the efficiency frontier as can be seen in Figure 17 (left).

Figure 17: Distribution of the distances between every PSA outcome and the efficiency frontier in scenarios 5 (left) and 6 (right).

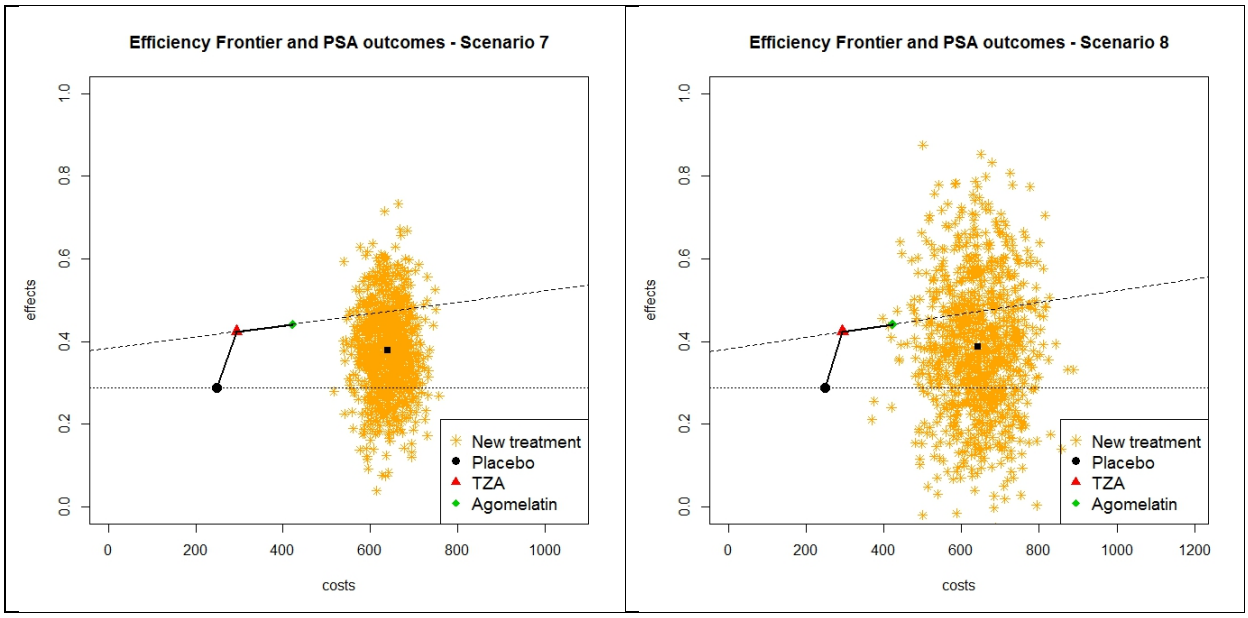


When the overall uncertainty was increased in scenario 6 (see Figure 16 - right) M1 decreased to 93% (6% reduction compared to scenario 5). The average cost of the PSA outcomes was €636, with a 95% percentile interval (€487, €791). The average point estimate was above the efficiency frontier, thus the new intervention would be accepted. The increase allowed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 16 - right – and the forward extension of the segment TZA-Agomelatin) was €1,638 (metric M4). The average and median of the distances between every PSA outcome and the efficiency frontier were €1,654 (metric M5) and €1,645 (metric M6), respectively.

Scenarios 7 and 8

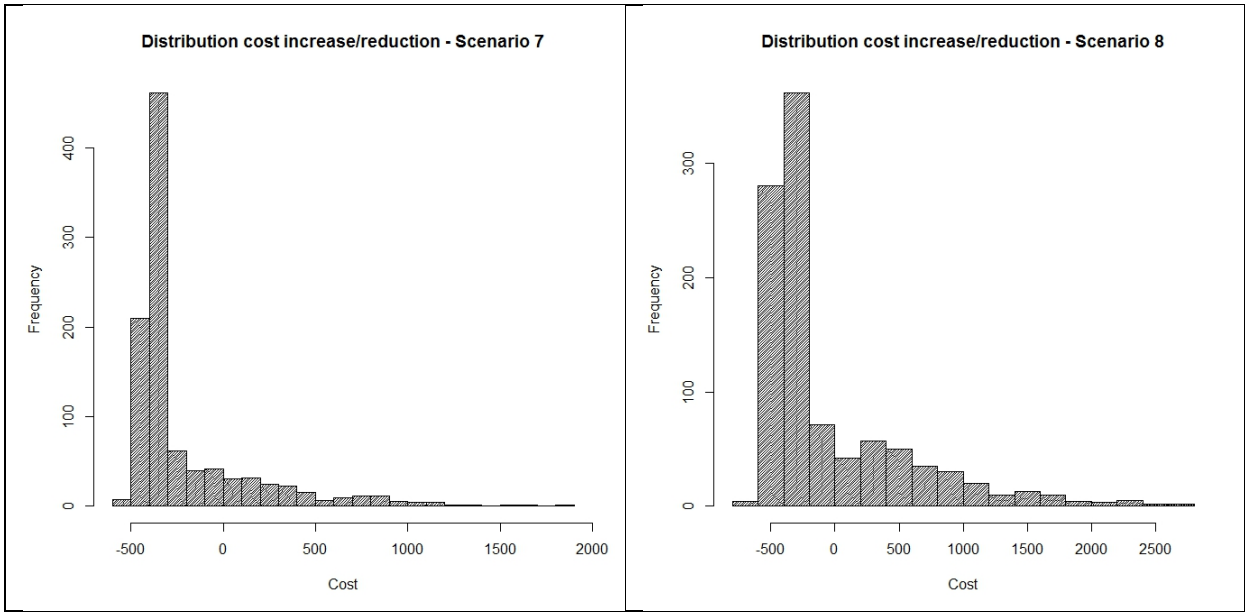
In scenarios 7 and 8 shown in Figure 18, the PSA outcomes of the new intervention were mostly located on the bottom-right corner of the cost-effectiveness plane and below the efficiency frontier (with many of them below the Placebo benefit).

Figure 18: PSA scenarios 7 (left) and 8 (right).



As can be observed in Figure 18 (left), most of the PSA outcomes in scenario 7 were below the efficiency frontier. In particular, only about 18% (metric M1) would be accepted. M3 was now almost 9%, the highest observed so far. The average cost was €637, with a 95% percentile interval (€565, €708). The average point estimate was below the efficiency frontier, thus the new intervention would be rejected. Moreover, the *reduction* needed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 18 –left– and the segment Placebo-TZA) was -€361 (metric M4). The average and median of the distances between every PSA outcome and the efficiency frontier were -€209 (metric M5) and -€348 (metric M6), respectively. Note that in this case the difference between the two averages was approximately €150. This can be explained by the skewness of the distribution of the distances between every PSA outcome and the efficiency frontier as can be seen in Figure 19 (left).

Figure 19: Distribution of the distances between every PSA outcome and the efficiency frontier in scenarios 7 (left) and 8 (right).



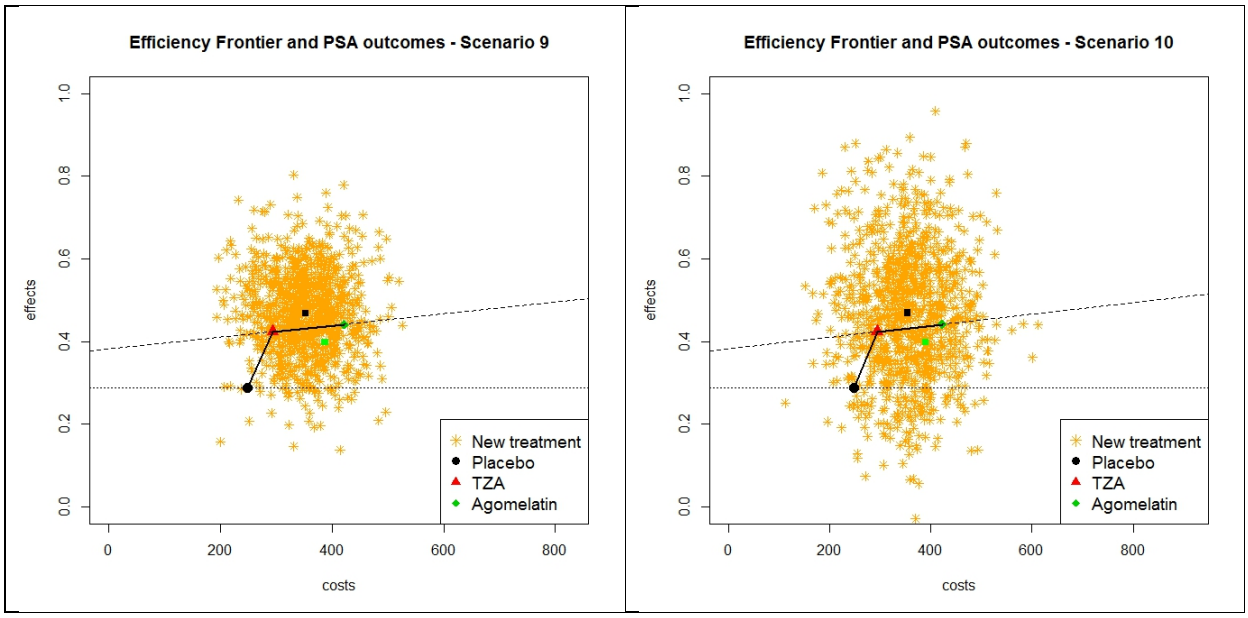
When the overall uncertainty was increased in scenario 8 (see Figure 18 - right) M1 increased (compared to scenario 7) to 28.3%. The average cost of the PSA outcomes was €636, with a 95% percentile interval (€499, €778). The average point estimate was below the efficiency frontier, thus the new intervention was rejected. Moreover, the reduction needed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 18 –right– and the segment Placebo-TZA) was -€359 (metric M4). The average and median of the distances between every PSA outcome and the efficiency frontier were -€52 (metric M5) and -€328 (metric M6), respectively.

Scenarios 9 and 10

In scenarios 9 and 10 (shown in Figure 20) we also studied the (possible) differences between the deterministic and the probabilistic cost and benefits point estimates.

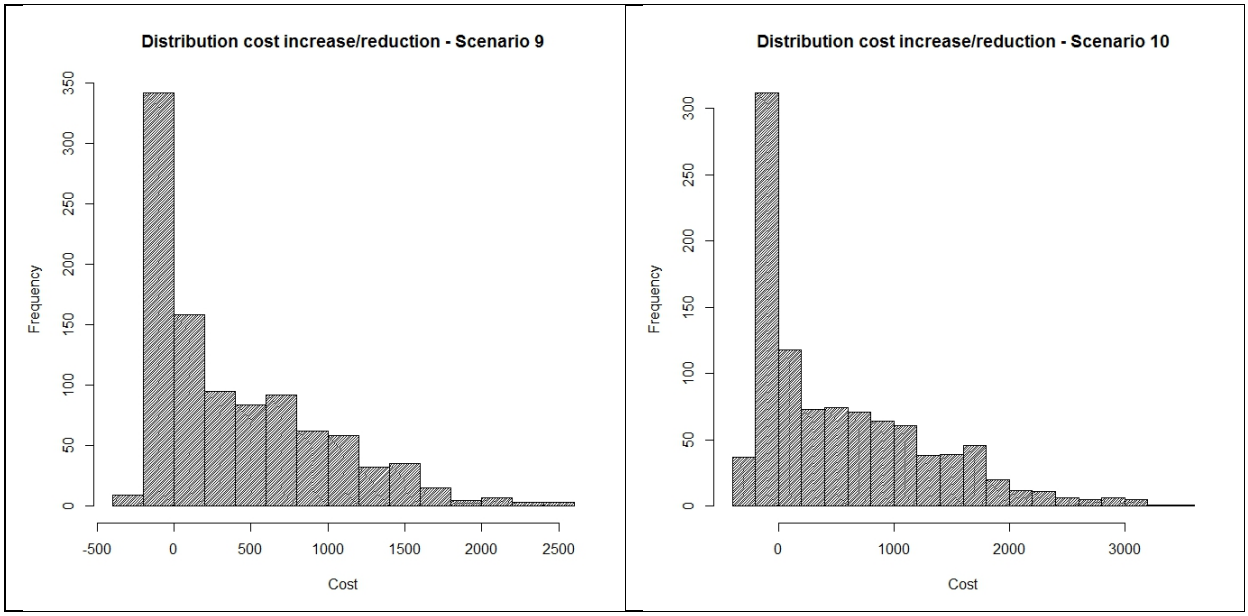
The PSA outcomes of the new intervention were located approximately in the middle of the cost-effectiveness plane with the average point estimate (black solid square) slightly above the efficiency frontier but the deterministic point estimate (green solid square) below the efficiency frontier. These scenarios show how the average (probabilistic) point estimate does not necessarily converge to the deterministic point estimate. Furthermore, in these scenarios the two point estimates lead to different decisions: based on the deterministic point estimate the new intervention would be rejected (and a cost reduction would be suggested) but based on the average (probabilistic) point estimate the new intervention would be accepted (and a cost increase would be suggested). As it occurs with all the other scenarios presented in this section, it is assumed that the decision whether the new intervention is going to be accepted or not is based on the PSA, thus the average point estimate (black solid square) is the one that should be compared against the whole efficiency frontier (if it lies below the efficiency frontier then the new intervention is rejected and otherwise it is accepted).

Figure 20: PSA scenarios 9 (left) and 10 (right).



In scenario 9 we observed that 64.7% of the PSA outcomes were above the efficiency frontier (metric M1). There were almost no PSA outcomes (M3 = 0.2%) in area B (below the efficiency frontier but above maximum benefits). The average cost of the PSA outcomes was €349, with a 95% percentile interval (€229, €455). The average point estimate was above the efficiency frontier, thus the new intervention would be accepted. Moreover, the *increase* allowed to bring it to the efficiency frontier (computed as the horizontal distance between the black solid square in Figure 20 –left– and the forward extension of the segment TZA-Agomelatin) was €223 (metric M4). The average and median of the distances between every PSA outcome and the efficiency frontier were €394 (metric M5) and €188 (metric M6), respectively. The difference between the three metrics was again due to the skewness of the distribution of the distances between every PSA outcome and the efficiency frontier as can be seen in Figure 21 (left).

Figure 21: Distribution of the distances between every PSA outcome and the efficiency frontier in scenarios 9 (left) and 10 (right).

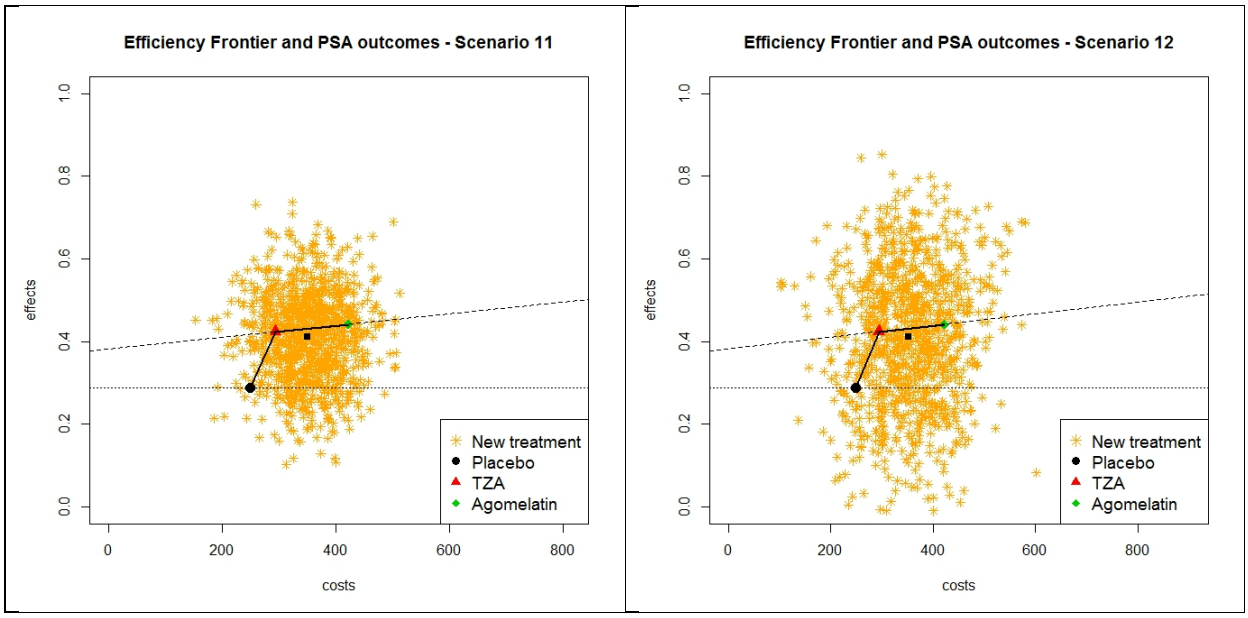


When the overall uncertainty was increased in scenario 10 (see Figure 20 - right) M1 was almost the same as in scenario 9. The average cost was €347, with a 95% percentile interval (€202, €500). The average point estimate was above the efficiency frontier, thus the new intervention would be accepted. Moreover, the increase allowed to bring it to the efficiency frontier (computed as the horizontal distance between the black solid square in Figure 20 – right– and the forward extension of the segment TZA-Agomelatin) was €255 (metric M4). The average and median of the distances between every PSA outcome and the efficiency frontier were €543 (metric M5) and €305 (metric M6), respectively. In this case, the difference between the metrics M4 and M5 increased in about €150 compared to scenario 9, reflecting the effect of increasing the overall uncertainty.

Scenarios 11 and 12

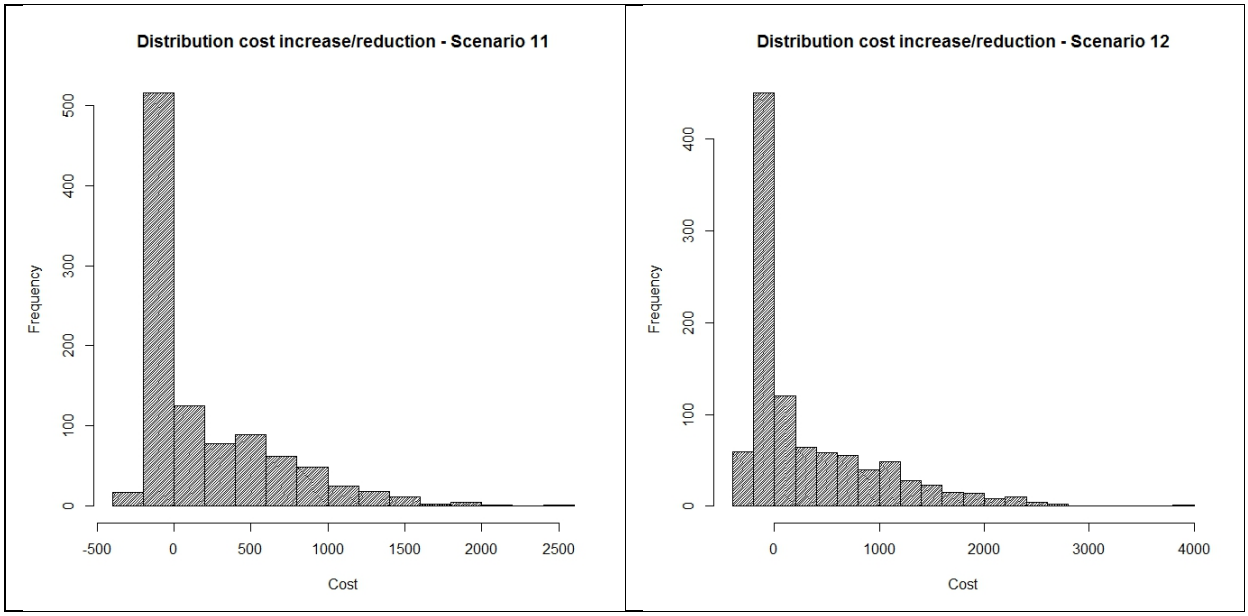
In scenarios 11 and 12 (shown in Figure 22) the PSA outcomes of the new intervention were located approximately in the middle of the cost-effectiveness plane with the average point estimate slightly below the efficiency frontier.

Figure 22: PSA scenarios 11 (left) and 12 (right).



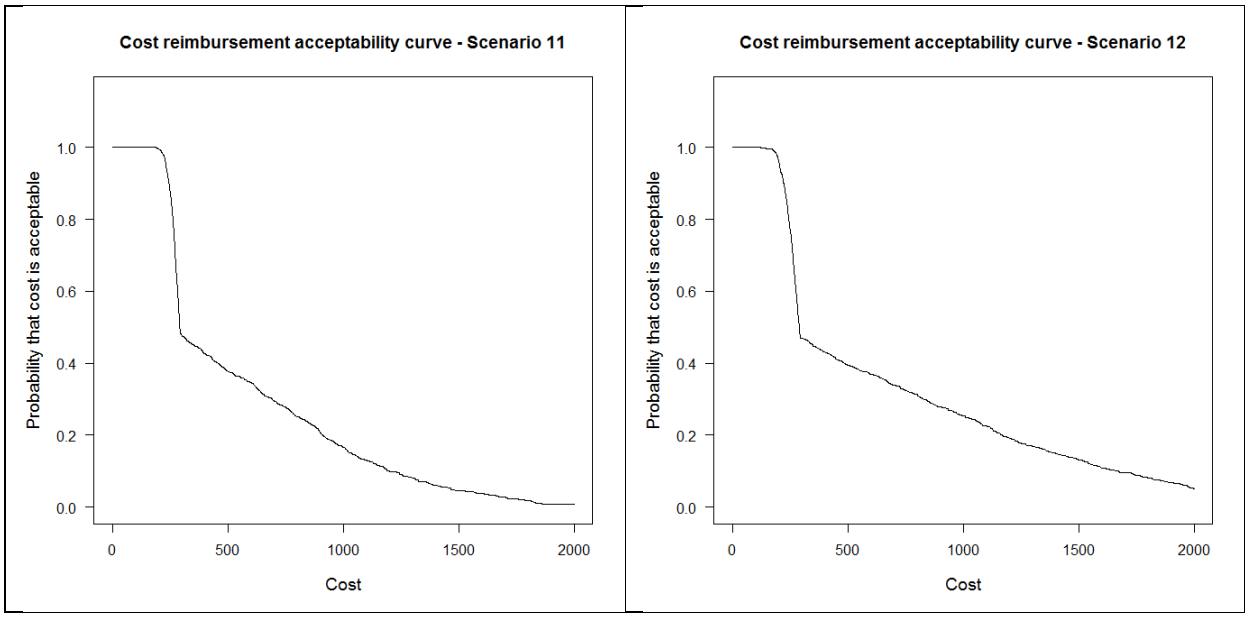
In scenario 11 only 46.7% of the PSA outcomes were above the efficiency frontier (metric M1). The average cost was €354, with a 95% percentile interval (€241, €467). The average point estimate was below the efficiency frontier, thus the new intervention would be rejected. Moreover, the reduction needed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 22 - left – and the segment Placebo-TZA) was -€61 (metric M4). However, the average of the distances between every PSA outcome and the efficiency frontier (metric M5) implied a €200 cost increase. Note that this is the first scenario where we found that these two metrics would provide contradictory recommendations. Finally, the median of the distances between every PSA outcome and the efficiency frontier (metric M6) suggested a -€16 cost reduction (thus, the same recommendation as with metric M4 but much smaller). Once more the differences between the three metrics were due to the skewness of the distribution of the distances between every PSA outcome and the efficiency frontier as can be seen in Figure 23 (left).

Figure 23: Distribution of the distances between every PSA outcome and the efficiency frontier in scenarios 11 (left) and 12 (right).



In this case we also computed the cost reimbursement acceptability curve for scenario 11 (Figure 24 - left). Note that this curve shows for every possible cost its estimated probability of being acceptable. The curve starts at 1 since all the simulated reimbursable costs were larger than €0. When the suggested cost was €293 (i.e. the average cost of the PSA outcomes, €354, minus the average reduction needed to bring it to the efficiency frontier, €61), approximately half of the simulated reimbursable costs were larger than this suggested cost. Thus, €293 would have been accepted with probability 0.5 (median). The last suggested cost shown in the cost reimbursement acceptability curve was €2,000. Even for this high cost there was a 0.013 probability of being acceptable (due to the few PSA outcomes with very high benefits which provided a large cost increase).

Figure 24: Cost reimbursement acceptability curves in scenarios 11 (left) and 12 (right).

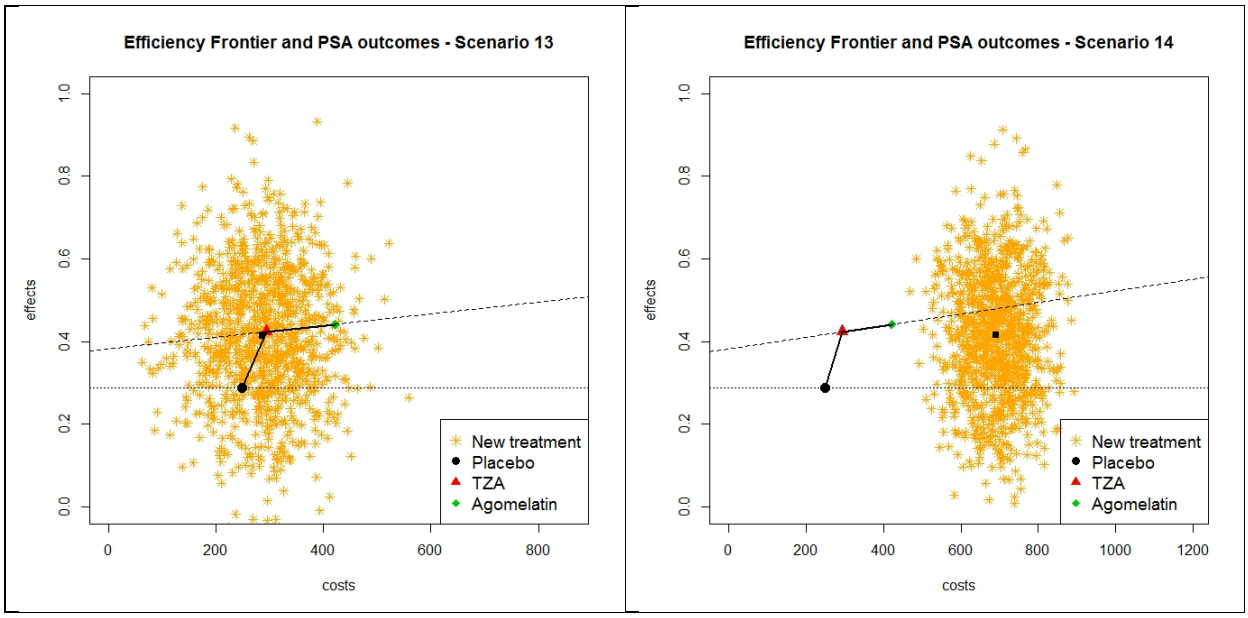


When the overall uncertainty was increased in scenario 12 (see Figure 22 - right) M1 increased by 1% with respect to scenario 11. The average cost of the PSA outcomes was €350, with a 95% percentile interval (€192, €503). The average point estimate was also below the efficiency frontier (thus the new intervention was rejected) and the reduction needed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 22 - right – and the segment Placebo-TZA) was -€64 (metric M4). However, also in this case the average of the distances between every PSA outcome and the efficiency frontier (metric M5) implied a cost increase of €307 (metric M5). Hence, also in this scenario the two recommendations were contradictory. The median of the distances between every PSA outcome and the efficiency frontier (metric M6) suggested as in scenario 11 the correct recommendation (-€7 cost reduction) but this was also much smaller than with metric M4. The cost reimbursement acceptability curve for scenario 12 (Figure 24 - right) was very similar to that one in scenario 11.

Scenarios 13 and 14

In case of contradictory recommendations, as in scenarios 11 and 12 above, we were interested in studying the consequences of adopting either cost reduction or increase. For that reason we built scenarios 13 and 14 based on the results obtained in scenario 12. Thus, we assumed in scenario 13 an average cost reduction of -€64 and in scenario 14 an average cost increase of €307 with respect to scenario 12. These scenarios are shown in Figure 25.

Figure 25: PSA scenarios 13 (left) and 14 (right).



In scenario 13, as expected, the average of the PSA outcomes was located (almost) on the efficiency frontier. The average point estimate was slightly above the efficiency frontier (thus the new intervention would be accepted) and the cost increase needed to bring it to the efficiency frontier (computed as the horizontal distance between the solid square in Figure 25 –right– and the segment Placebo-TZA) was €10 (metric M4), thus close to €0, which is what we would get when the average point estimate lies exactly on the efficiency frontier. Moreover, 59.7% of the PSA outcomes were above the efficiency frontier (metric M1). Thus, 11.9% increase with respect to scenario 12. However, the average and the median of the distances between every PSA outcome and the efficiency frontier still suggested a cost increase of €425 (metric M5) and €69 (metric M6), respectively.

In scenario 14, the average point estimate was below the efficiency frontier (thus the new intervention would be rejected), resulting on a cost decrease of -€398 (computed as the horizontal distance between the solid square in Figure 25 –right– and the segment Placebo-TZA) in order to bring it to the efficiency frontier (metric M4). Moreover, only 34.3% of the PSA outcomes were above the efficiency frontier (metric M1). Thus, almost 14% decrease with respect to scenario 12. This clearly illustrates the effect of wrongly increasing the cost: not only that the mean of the PSA outcomes became less favorable than in scenario 12 but also the overall probability of acceptance decreased.

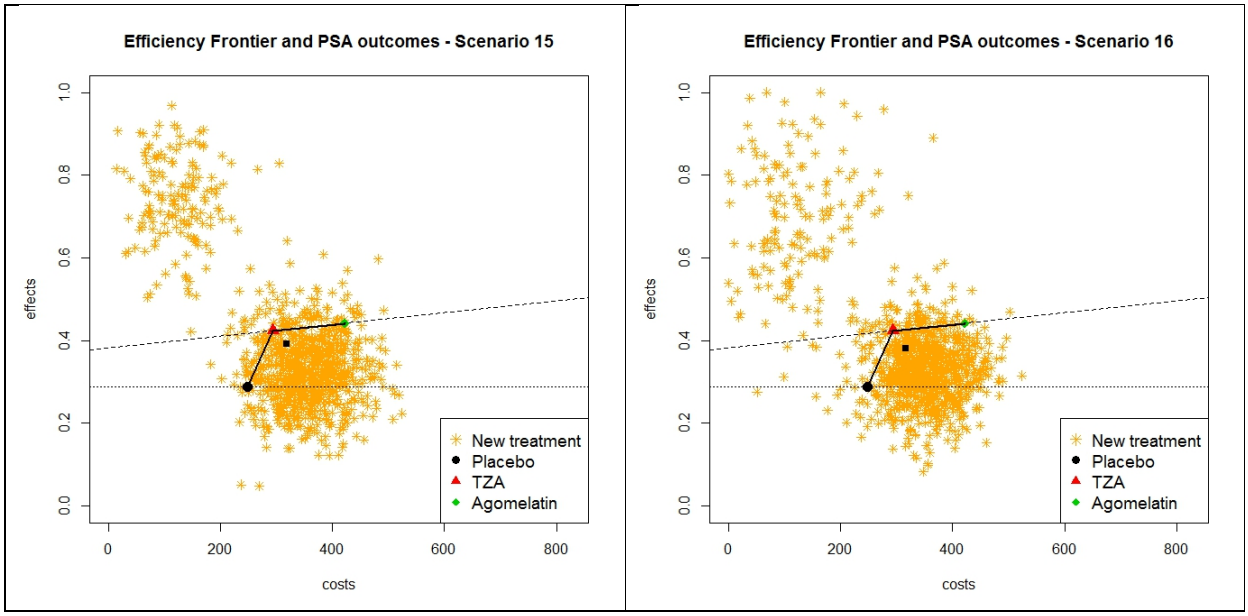
Scenarios 15 and 16

Finally, in scenarios 15 and 16 we explored the effect of the asymmetry of uncertainty in costs and benefits (see Figure 26).

In scenario 15 most of the PSA outcomes of the new intervention were located below the efficiency frontier but relatively close to it (relatively small distances between PSA outcomes and efficiency frontier which means relatively small cost increase or reduction). On top of that, 15% of the PSA outcomes we generated with relatively large uncertainty above the

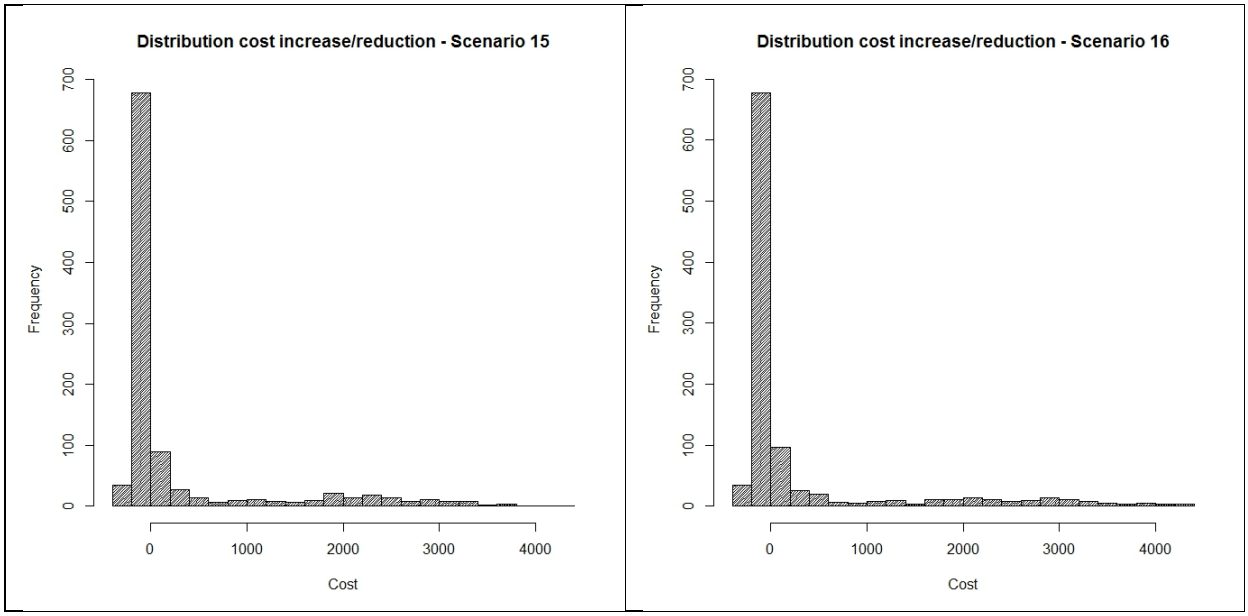
efficiency frontier (relatively large distances between some PSA outcomes – e.g. outliers – and the last segment of the efficiency frontier which means relatively large cost increase). Note that scenario 16 (Figure 26 - right) resulted from scenario 15 with increased uncertainty in both costs and benefits.

Figure 26: PSA scenarios 15 (left) and 16 (right).



In scenario 15, only 28.2% of the PSA outcomes were located above the efficiency frontier (metric M1). The average cost was €316, with a 95% percentile interval (€85, €456). The average point estimate was below the efficiency frontier (thus the new intervention would be rejected), resulting on a cost decrease of -€35 (computed as the horizontal distance between the solid square in Figure 26 –left– and the segment Placebo-TZA) in order to bring it to the efficiency frontier (metric M4). The average and median of the distances between every PSA outcome and the efficiency frontier were €254 (metric M5) and -€68 (metric M6). The difference between the three metrics is due to the large cost increases produced by the 'outliers' which are measured against the segment TZA-Agomelatin, which is also reflected in the distribution of the distances between every PSA outcome and the efficiency frontier as can be seen in Figure 27 (left).

Figure 27: Distribution of the distances between every PSA outcome and the efficiency frontier in scenarios 15 (left) and 16 (right).



In scenario 16, the results were similar to those in scenario 15 but with wider percentile intervals. The average cost of the PSA outcomes was €319, with a 95% percentile interval (€72, €457). The average point estimate was still below the efficiency frontier (thus the new intervention was rejected), resulting on a cost decrease of -€37 (computed as the horizontal distance between the solid square in Figure 26 –right– and the segment Placebo-TZA) in order to bring it to the efficiency frontier (metric M4). Similarly to scenario 15, only 28.3% of the PSA outcomes were above the efficiency frontier. The average and median of the distances between every PSA outcome and the efficiency frontier were €292 (metric M5) and -€71 (metric M6), respectively. Thus, the conclusions from this scenario are similar to those from scenario 15.

Appendix B – Assessment of outliers (probabilistic efficiency frontier scenario)

The method we have used to indentify outliers is a simple popular method based on interquartile ranges and known as Tukey fences [19]. According to this method, outliers are those values below Q1-1.5*IQR or above Q3+1.5 IQR, where Q1 denotes the first quartile, Q3 the third quartile and IQR the interquartile range, thus Q3 minus Q1. More sophisticated methods for the detection of outliers exist but that is not the main purpose of this appendix. We simply want to show that in the probabilistic efficiency frontier scenario there are extreme values associated to the distribution of the distances between every PSA outcome and the efficiency frontier. For that reason a cost increase or reduction recommendation based on the median (metric M6) is preferred over a recommendation based on the average (metric M5), as we explained in Section 2.4.

In fact, we have modified here the Tukey method and instead of using Q1 and Q3 we are considering the 2.5% and 97.5% percentiles to compute the fences. That way we are pushing the fences away and only values beyond the 95% percentile interval are considered outliers in this analysis. Besides the lower and upper fences provided by the Tukey method, we also computed the percentage (based on 100,000 PSA outcomes) of cost recommendations below the lower fence and above the upper fence in the PSA sample, and, if any, the minimum recommended cost below the lower fence and the maximum recommended cost above the upper fence. The results can be seen in Table 9.

Table 9. Assessment of outliers in cost recommendations for Duloxetin, Venlafaxin, Mirtazapin and Bupropion in the scenario with probabilistic efficiency frontier. A negative monetary value means cost reduction whereas a positive one means cost increase.

	LF	UF	% below LF	Min. below LF	% above UF	Max. above UF[‡‡]
Duloxetin	-€195	-€59	0.00%	NA	0.00625%	€31,083
Venlafaxin	-€551	€802	0.00%	NA	0.00966%	€375,272
Mirtazapin	-€191	€241	0.00%	NA	0.00747%	€86,683
Bupropion	-171€	-€25	0.00%	NA	0.00162%	€6,251

LF=modified lower fence; UF=modified upper fence.

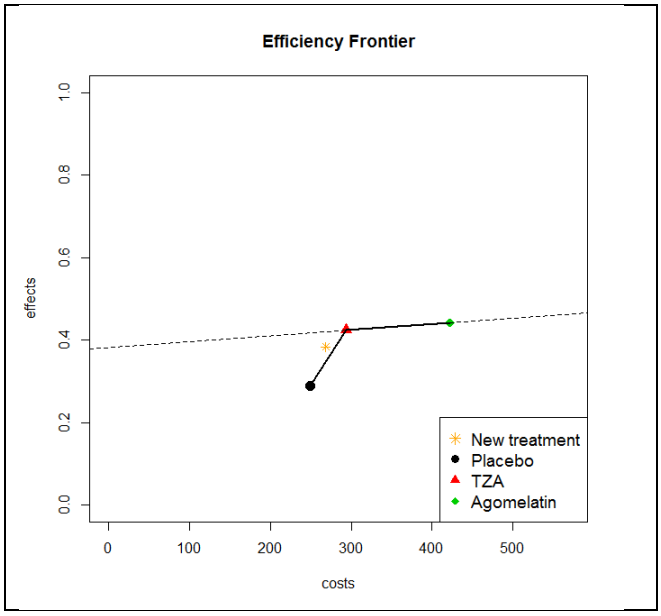
Venlafaxin and to a lower extent Mirtazapin are the interventions where there is more uncertainty about the cost recommendation since the difference between the upper and lower fences is high and allows both cost reduction (negative values) and increase (positive values). With our definition of fences, there are only outliers above the upper fence and although the percentages of outliers are very small, these are very large (some of them are numerically infinitely large – not shown in the table though). That would explain why the average of the distances between every PSA outcome and the efficiency frontier (metric M5) would overestimate the price recommendation and would justify the use of the median instead.

[‡‡] For all interventions the maximum was (numerically) infinite. The ones shown in the column are those after removing the infinite cases.

Appendix C – Net (health or monetary) benefit as decision rule

Confusion may arise when looking at the Net Health Benefit (NHB) or Net Monetary Benefit (NMB) of the new intervention solely and use it to decide whether to accept the new intervention or not. Suppose the situation depicted in Figure 28. The efficiency frontier is determined by three interventions: Placebo-TZA-Agomelatin.

Figure 28: The new intervention is accepted although the NHB is negative.



The corresponding incremental cost-effectiveness ratios (as determined by the slope of the segments of the efficiency frontier) are shown in Table 10.

Table 10. Full incremental analysis for the interventions on the efficiency frontier.

	Expected Benefit	Expected Cost	Incremental Benefit	Incremental Cost	ICER
Placebo	0.288	€249			
TZA	0.424	€294	0.136	€45	€330.88
Agomelatin	0.442	€422	0.018	€128	€7111.11

*ICER=Incremental cost-effectiveness ratio

In the situation illustrated in Figure 28 the new intervention will be regarded as acceptable since it lies above the efficiency frontier. In terms of incremental cost-effectiveness ratios, the new intervention is acceptable because the *incremental* cost-effectiveness ratio (compared

to placebo) is lower than the incremental cost-effectiveness ratio of the next more effective alternative (TZA), as shown in Table 11.

Table 11. Full incremental analysis for the interventions on the efficiency frontier and the new intervention.

	Expected Benefit	Expected Cost	Incremental Benefit	Incremental Cost	ICER
Placebo	0.288	€249			
New intervention	0.383	€271	0.095	€22	€231.57
TZA	0.424	€294	0.041	€23	€560.97
Agomelatin	0.442	€422	0.018	€128	€7111.11

*ICER=Incremental cost-effectiveness ratio

Note that the last program implemented (Agomelatin in this case) determines the critical threshold (λ=€7111.11) and that in the example above this has not changed when the new intervention became available. The threshold determined by the last program implemented (i.e. determined by the available budget) will be used to compute the NHB (or the NMB). That way, λ can be computed as 1 divided by the slope of the last segment on the efficiency frontier. With that value of λ, the NHB (NMB) can be graphically defined as the vertical (horizontal) distance between a intervention point estimate and the last segment (or its extension – backward or forward) on the efficiency frontier [20]. The problem arises when the NHB (NMB) is used as a decision rule, i.e. the rule "accept the intervention whose NHB (NMB) is positive" is wrong and may contradict the general principle of accepting any intervention above the whole efficiency frontier (dominance). As emphasized by Karlsson and Johannesson [14], nothing can be concluded based on NHB (or NMB), the incremental NHB (or NMB) have to be calculated instead. This is also illustrated with the example in Figure 28. If we assume that λ is determined by the last segment on the efficiency frontier, then the new intervention has negative NHB (and NMB) since it is below the backward extension of the last segment. Thus, the rule based on the NHB solely would reject the new intervention. However, that is wrong because the new intervention is in position of increased efficiency. The missing argument is that Placebo has also negative NHB (and NMB) but then the incremental NHB (INHB) of the new intervention compared to Placebo is positive. Or equivalently, the ICER determined by the new intervention and Placebo is smaller (and thus more efficient) than the ICER determined by TZA and Placebo as shown in Tables 10 and 11.

Note finally that to properly compute the NHB/NMB the first treatment on the last segment of the efficiency frontier (i.e. TZA in this case) has to be the origin of the cost-effectiveness plane. This is illustrated in Table 12.

Table 12. Full incremental analysis for the interventions on the efficiency frontier and the new intervention, when TZA is the origin of the cost-effectiveness plane.

	Expected Benefit	Expected Cost	Incremental Benefit	Incremental Cost	ICER	NMB	INMB	NHB	INHB
Placebo	-0.136	-€45				-€922		-0.129	
New intervention	-0.041	-€23	0.095	€22	€231	-€268	€653	-0.037	0.091
TZA	0	€0	0.041	€23	€560	€0	€268	0	0.037
Agomelatin	0.018	€128	0.018	€128	€7111.11	€0	€0	0	0

*ICER=Incremental cost-effectiveness ratio; NMB=Net monetary benefit; INMB=Incremental net monetary benefit; NHB=Net health benefit; INHB=Incremental net health benefit.

Anhang B – Darlegung potenzieller Interessenkonflikte (externe Sachverständige)

Im Folgenden sind die potenziellen Interessenkonflikte der externen Sachverständigen dargestellt. Alle Informationen beruhen auf Selbstangaben der einzelnen Personen anhand des „Formblatts zur Offenlegung potenzieller Interessenkonflikte“ mit Stand 12/2011. Das aktuelle Formblatt ist unter www.iqwig.de abrufbar. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung.

Externe Sachverständige

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
Al, Maiwenn[1]	nein	nein	nein	ja	nein	nein
Corro Ramos, Isaac[1]	nein	nein	nein	ja	nein	nein

[1] Formblatt zur Offenlegung potenzieller Interessenkonflikte Version 12/2011

Im „Formblatt zur Offenlegung potenzieller Interessenkonflikte“ (Version 12/2011) wurden folgende 6 Fragen gestellt:

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor angestellt bei einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere bei einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor ein Unternehmen, eine Institution oder einen Interessenverband im Gesundheitswesen, insbesondere ein pharmazeutisches Unternehmen, einen Hersteller von Medizinprodukten oder einen industriellen Interessenverband, direkt oder indirekt beraten?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband, Honorare erhalten für Vorträge, Stellungnahmen oder Artikel?

Frage 4: Haben Sie und / oder hat die Einrichtung[2], die Sie vertreten, abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband, finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

Frage 5: Haben Sie und / oder hat die Einrichtung[2], bei der Sie angestellt sind bzw. die Sie vertreten, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren ohne wissenschaftliche Gegenleistung) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile eines Unternehmens oder einer anderweitigen Institution, insbesondere von einem pharmazeutischen Unternehmen oder einem Hersteller von Medizinprodukten? Besitzen Sie Anteile eines „Branchenfonds“, der auf pharmazeutische Unternehmen oder Hersteller von Medizinprodukten ausgerichtet ist?

[2] Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.