Economic Evaluations of Pharmacogenetic and Pharmacogenomic Screening Tests: A Systematic Review. Second Update of the Literature

Elizabeth J. J. Berm1*, Margot de Looff2*, Bob Wilffert1,4, Cornelis Boersma2,3, Lieven Annemans5, Stefan Vegter2,3, Job F. M. van Boven2, Maarten J. Postma2,6

1 Unit of Pharmacotherapy & Pharmaceutical Care, Department of Pharmacy, University of Groningen, Groningen, The Netherlands; 2 Unit of PharmacoEpidemiology & PharmacoEconomics, Department of Pharmacy, University of Groningen, Groningen, The Netherlands; 3 GSK, Zeist, The Netherlands; 4 Department of Clinical Pharmacy and Pharmacology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; 5 Faculty of Medicine & Health Science, Department of Public Health, Ghent University, Ghent, Belgium; 6 Institute of Science in Healthy Ageing & health care (SHARE), University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

* These authors contributed equally to this work.
* e.j.j.berm@rug.nl

Abstract

Objective
Due to extended application of pharmacogenetic and pharmacogenomic screening (PGx) tests it is important to assess whether they provide good value for money. This review provides an update of the literature.

Methods
A literature search was performed in PubMed and papers published between August 2010 and September 2014, investigating the cost-effectiveness of PGx screening tests, were included. Papers from 2000 until July 2010 were included via two previous systematic reviews. Studies’ overall quality was assessed with the Quality of Health Economic Studies (QHES) instrument.

Results
We found 38 studies, which combined with the previous 42 studies resulted in a total of 80 included studies. An average QHES score of 76 was found. Since 2010, more studies were funded by pharmaceutical companies. Most recent studies performed cost-utility analysis, univariate and probabilistic sensitivity analyses, and discussed limitations of their economic evaluations. Most studies indicated favorable cost-effectiveness. Majority of evaluations did not provide information regarding the intrinsic value of the PGx test. There were considerable differences in the costs for PGx testing. Reporting of the direction and magnitude of
bias on the cost-effectiveness estimates as well as motivation for the chosen economic model and perspective were frequently missing.

Conclusions
Application of PGx tests was mostly found to be a cost-effective or cost-saving strategy. We found that only the minority of recent pharmacoeconomic evaluations assessed the intrinsic value of the PGx tests. There was an increase in the number of studies and in the reporting of quality associated characteristics. To improve future evaluations, scenario analysis including a broad range of PGx tests costs and equal costs of comparator drugs to assess the intrinsic value of the PGx tests, are recommended. In addition, robust clinical evidence regarding PGx tests’ efficacy remains of utmost importance.

Introduction
Pharmacogenetics and pharmacogenomics investigate the influence of genetic and genomic variations on drug response in individuals [1]. The term pharmacogenetics covers the study of single genes, whereas pharmacogenomics is used to describe the study of several genes [1]. The abbreviation PGx is used to cover both pharmacogenetics and pharmacogenomics. PGx tests offer great opportunities for personalised medicine, by combining genetic information and corresponding phenotypes [2]. Ideally, by applying PGx, the most optimal, tailored pharmacotherapy can be determined, thereby increasing the treatment’s overall efficacy and decreasing the incidence of adverse events [1]. In the field of oncology it has been shown that for certain therapies the specific genetic variations in cancer cells can affect the drug efficacy and/or adverse events [2]. Hence, patients may benefit from the PGx tests by utilising an alternative therapy, or changing the drug dosage [3,4]. Therefore, PGx is nowadays often used as a synonym for personalised medicine, although personalised medicine is a much broader concept [2].

It is likely that an increasing amount of patient-specific genomic information will be available in the near future and this may result in an increased usage of PGx tests which needs evaluation of effects, but also cost effectiveness [5]. PGx has the potential to reduce the costs associated with inappropriate, expensive drug treatments and/or serious adverse drug reactions, in particular those that require hospitalisation [3]. Therefore, next to optimising health outcomes, PGx tests might be cost-saving [1,5]. However, in order to obtain valid, accurate, and relevant estimates of cost-effectiveness, reliable economic studies are required.

Economic evaluations of PGx tests entail some specific difficulties. Often there is no hard clinical evidence regarding the effects of the PGx test on the clinical utility and it is unlikely that such evidence will be available for the use of every genetic variant [6]. Furthermore, for PGx tests, compliance and adherence of clinicians to the test results might have an effect on the effectiveness of PGx tests which is hard to incorporate in a cost-effectiveness analysis [7]. Differences in costs for the PGx test can be substantial between countries, or even laboratories, and therefore it is advised to include different costs in scenario analysis [7]. In addition, the sensitivity and specificity of a PGx test can vary due to different ethnicities studied, or genetic variations analysed.

In the last decade, several reviews investigated economic evaluations of genetic tests [1,8–14]. These reviews showed that the level of consistency and quality could be improved. Many original studies lacked a thorough sensitivity analysis and moreover, in general a poor quality
of methodology was noticed [9,11]. Inconsistencies mainly resulted from e.g. lack of clinical evidence, different methodologies as well as statistics and modest heterogeneity among study designs and patient populations [3,9]. However, these different methodologies have not been in detail dealt with in previous systematic reviews. Recent developments have led to a bifurcation in the nature of the economic evaluations of PGx testing in, on the one hand, studies assessing the intrinsic value of a test and, on the other hand, studies assessing the value of the test in combination with an active compound. For example in colorectal cancer, the economic value of testing for KRAS as compared to no testing could be considered the “intrinsic value” of the PGx test. KRAS testing before treatment with cetuximab, is found to be dominant (i.e. cost-saving and better) as compared to no prior testing and therefore it is recommended before administration of cetuximab [15]. By its uptake in clinical guidelines, a shift in the comparator for future economic evaluations took place, as the combination of cetuximab and KRAS testing became usual care. In future evaluations, the intrinsic value of the KRAS test itself will no longer be assessed, but rather the combination of a drug and its test as compared with a new treatment option. Following this development, a distinction between the nature of economic evaluations of PGx tests is important for a fair comparison of studies.

The objective of this study was to give recommendations for improvement and an update of the literature about PGx tests, taking into account the difference between the intrinsic value of tests themselves and tests embedded into economic evaluations as usual care or best current care. Our new findings were placed in perspective with respect to findings from our previous reviews [1,9]. As such, our study links together a period from 2000- September 2014 of PGx testing and pharmacoconomics.

Methods

A search in PubMed was performed using combinations of the following terms [PubMed search: all fields] and their thesauri variants: [‘cost-effectiveness [including MeSH]’ OR ‘cost-utility’ OR ‘cost-benefit’ OR ‘cost-minimization’ OR ‘pharmacoconomics [including MeSH]’] AND [‘pharmaco-genetics’ OR ‘pharmacogenomics [including MeSH]’ OR ‘genetic screening’ OR ‘genetic testing [including MeSH]’ OR ‘genotyped’ OR ‘polymorphism screening’]. These search terms were in line with the terms that were used in previous reviews, performed in 2008 and 2010 [1,9]. The search was last updated in October 2014 and studies were included if they were: published between August 2010 and September 2014, peer reviewed, performed on a genetic screening method of the human genome, evaluating economic outcomes, written in English, and the genetic or genomic variations were shown to influence the drug efficacy or drug safety. Articles were first screened on title. If the title was not informative enough to form a decision with respect to these criteria, abstracts were assessed. Additional articles were identified through reference tracking.

From the selected studies, the following data were extracted: (I) area of disease or patient population, (II) gene(s) analysed by the pharmacogenetic test, (III) the costs of the pharmacogenetic test, (IV) pharmaceutical compound influenced by the genetic variation, (V) type of economic analysis, (VI) type of sensitivity analysis, (VII) time horizon, (VIII) discounting, (IX) perspective, (X) the outcome measurements, and (XI) the funding body. For interpretation of the outcome measure (i.e. cost-effectiveness), the conclusions as reported by the authors were used. Furthermore, an assessment of the papers ‘discussion on the study limitations’ was made. In this assessment, all limitations mentioned by the authors were captured to look for common and uncommon themes. As stated in the previous review, assessment of these points is expected to provide good information for an adequate interpretation of the studies design, reporting, robustness, methodologies used, and statistical analyses performed [9]. In addition
to these points, which were assessed in our previous reviews, we added (XII) reporting of analyti-
cal validity of the PGx test, (XIII) the cost-effectiveness threshold, (XIV) the country which
was used for the perspective of the economic evaluation, and (XV) a weighted quality assess-
ment for the studies included by this update. To assess the quality, the Quality of Health Eco-
nomic Studies (QHES) instrument was used [16]. This instrument was used to improve
generalisability of the results with respect to other reviews performed in the same field [10,13].
According to the QHES checklist a score between 0 and 100 was generated. A score of ≥ 75
was considered as a high quality score [11]. Two reviewers assessed the quality of the included
studies, if results were different, consensus was reached through discussion.

Data from before 2008, and for the period 2008—July 2010 were retrieved from the two pre-
vious reviews by Vegter et al. (2008, 2010) [1,9].

Results

Included studies

Results of the search strategy are provided in a PRISMA flow chart (Fig 1) [17]. The PubMed
search yielded 4408 hits. Duplicates were removed and out of the remaining 733 articles, 160
were selected for full text assessment. Three articles were identified through reference tracking.
After inspection of full texts, 122 studies were excluded, resulting in 38 included studies
[15,18–54]. Main reasons for exclusion were the type of genetic tests studied (i.e. not related to
pharmacogenetics) and review papers. Subsequently, 42 studies published before July 2010
were added based on the two reviews by Vegter et al. [1,9,55–96], resulting in a final inclusion
of 80 studies. Fig 2A provides an overview of the total of 80 studies, published from 2000 until
September 2014, sorted by the type of pharmacoeconomic analysis performed. Most studies
were published in 2012 (16%), with a total of 13 publications.

Type of analysis: intrinsic value or combination with new drug

Cost utility analysis (CUA) was the technique mostly applied, namely in 54 studies (68%). Cost-
effectiveness analysis (CEA) was performed in 20 studies (25%), cost minimization analysis
(CMA) was used 5 times (6%). Note that for the sake of clarity we explicitly differentiate between
CUA (with results expressed in cost per QALY) and CEA (with other parameters for effective-
ness). Before 2008, CEA was the most frequently applied study type. Since 2008, CUA was per-
fomed in most of the publications. Some studies directly assessed the intrinsic value of the PGx
test [15,18–22,24–28], while other studies applied scenarios in which equal costs and efficacy
were assumed for the drug related to the PGx tests and the alternative treatment [23,29]. Both
were considered cost-effectiveness estimates that provide an indication of the intrinsic value of
the PGx test (Table 1). However, we found that the majority of the newly included (i.e. since
2010) CEAs incorporated a PGx test strategy in combination with a drug and compared this
combination to another drug (Table 2). As such, the intrinsic value of the PGx test itself was not
assessed. For example, Handorf et al. compared therapy for non-small cell lung cancer with a
platinum combination to PGx selected treatment with erlotinib [44]. In this analysis, the cost-
effectiveness of the PGx test itself was not assessed, but the combination of the PGx test and erlo-
itinib. This makes the outcome mainly dependent on the price of erlotinib. The incremental cost-
effectiveness ratios (ICER) of base case scenarios are shown in Tables 1 and 2.

Costs of PGx tests

When comparing the costs of the PGx tests, considerable differences between the costs of the
tests were observed. This is not surprising, since technology of genetic testing is developing and
costs are likely to be further reduced in the future. This was demonstrated by two studies of which one was performed in 2012 [41] and one in 2013 [40]. The costs for this particular screening test, US$72 and £20 respectively, were considerably lower compared to screening costs in studies performed in 2009 and 2010, which ranged from US$175 until US$575, respectively [63,78,96]. Another substantial difference was seen in the costs for the CYP2C9 and VKORC1 tests, which ranged from £20 to US$575 [40,78].
Fig 2. Type of outcome analysis (A) and sensitivity analysis (B) of PGx studies from 2000-September 2014.

doi:10.1371/journal.pone.0146262.g002

* two studies performed both a CEA and a CUACBE: Cost-benefit analysis; CEA: Cost-effectiveness analysis; CMA: Cost-minimization analysis; CUA: Cost-utility analysis
Table 1. Outcomes and funding sources of pharmaco-economic PGx studies on the intrinsic value of PGx test. Published between August 2010 and September 2014. Note that numbers were rounded towards hundreds.

<table>
<thead>
<tr>
<th>1st Author (reference)</th>
<th>PGx test</th>
<th>Analytical validity PGx test reported</th>
<th>Outcome Measure</th>
<th>Quantitative outcome or ICER (US$)</th>
<th>Cost effectiveness threshold ($/QALY)</th>
<th>Conclusion based on outcome</th>
<th>Funding</th>
<th>Country</th>
<th>QHES score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Klang 2010 [18]</td>
<td>21 gene assay</td>
<td>No</td>
<td>QALY</td>
<td>$10,800</td>
<td>No number</td>
<td>Cost-effective</td>
<td>Teva Pharmaceutical Industries Ltd.</td>
<td>Israel</td>
<td>83</td>
</tr>
<tr>
<td>Hall 2012 [20]</td>
<td>21 gene assay</td>
<td>No</td>
<td>QALY</td>
<td>$8,900</td>
<td>£20,00–30,00 ($31,200–46,700)</td>
<td>Cost-effective, though substantial uncertainty</td>
<td>Academic resources</td>
<td>UK</td>
<td>90</td>
</tr>
<tr>
<td>Vanderlaan2011 [21]</td>
<td>21 gene assay</td>
<td>No</td>
<td>QALY</td>
<td>$4,409 saved per patient per year</td>
<td>n.a.</td>
<td>Dominant</td>
<td>Genomic Health, Inc.</td>
<td>USA</td>
<td>31</td>
</tr>
<tr>
<td>Verhoef 2013 [22]</td>
<td>CYP2C9 and VKORC1</td>
<td>No</td>
<td>QALY</td>
<td>$2,700 ($3200)</td>
<td>$20,00 ($24,200)</td>
<td>Cost-effective</td>
<td>European grant</td>
<td>Netherlands</td>
<td>87</td>
</tr>
<tr>
<td>Dong 2012 [23]</td>
<td>HLA-B *1502</td>
<td>Yes</td>
<td>QALY</td>
<td>$29,800*</td>
<td>$50,000</td>
<td>Cost-effective for Singaporean Chinese and Malays, but not for Singaporean Indians</td>
<td>Academic resources</td>
<td>Singapore</td>
<td>63</td>
</tr>
<tr>
<td>Tiamkao 2013 [24]</td>
<td>HLA-B *1502</td>
<td>Yes</td>
<td>Costs</td>
<td>98,600 baht ($3,000) saved per 100 cases</td>
<td>n.a.</td>
<td>Cost-effective</td>
<td>None</td>
<td>Thailand</td>
<td>48</td>
</tr>
<tr>
<td>Blank 2011 [25]</td>
<td>KRAS and BRAF</td>
<td>Yes</td>
<td>QALY</td>
<td>KRAS and BRAF saves $3,300 ($2,500) per patient</td>
<td>n.a.</td>
<td>Cost saving</td>
<td>Academic resources</td>
<td>Switzerland</td>
<td>83</td>
</tr>
<tr>
<td>Behl 2012 [26]</td>
<td>KRAS and/or BRAF</td>
<td>No</td>
<td>LYS</td>
<td>KRAS alone saves $7,500 per patient</td>
<td>n.a.</td>
<td>Cost saving</td>
<td>Academic resources</td>
<td>USA</td>
<td>84</td>
</tr>
<tr>
<td>Shiffman 2012 [27]</td>
<td>LPA</td>
<td>No</td>
<td>QALY</td>
<td>$25,000</td>
<td>No number</td>
<td>Could be cost-effective</td>
<td>Berkeley HeartLab, Inc.</td>
<td>USA</td>
<td>34</td>
</tr>
<tr>
<td>Donnan 2011 [28]</td>
<td>TPMT</td>
<td>Yes</td>
<td>Life months</td>
<td>Cost no test CAN $700 ($600) per patient- Cost genetic test CAN $1,100 ($900) per patient</td>
<td>n.a.</td>
<td>Not cost-effective</td>
<td>Academic resources</td>
<td>Canada</td>
<td>77</td>
</tr>
<tr>
<td>Schackman 2013 [29]</td>
<td>UGT1A1</td>
<td>Yes</td>
<td>QALY</td>
<td>$2,000,000*</td>
<td>$100,000</td>
<td>Not cost-effective unless assay cost are low</td>
<td>Academic resources</td>
<td>USA</td>
<td>83</td>
</tr>
</tbody>
</table>

*Authors assumed equal costs and efficacy for different pharmaceuticals

BRAF, v-Raf murine sarcoma viral oncogene homolog B1; HLA, human leukocyte antigen; HSR, hyper sensitivity reaction; ICER, incremental cost-effectiveness ratio; KRAS, Kirsten rat sarcoma viral oncogene homolog; LPA, lipoprotein-a; LYG, life-years-gained; LYS, life-years-saved; PGx, pharmacogenetic; QALY, quality-adjusted-life-year; TPMT, thiopurine S-methyltransferase; UGT, UDP-glucuronosyltransferase.

doi:10.1371/journal.pone.0146262.t001
<table>
<thead>
<tr>
<th>1st Author (reference)</th>
<th>PGx test</th>
<th>Analytical validity PGx test reported (y/n)</th>
<th>Outcome Measure</th>
<th>Quantitative outcome or ICER (US$)</th>
<th>Cost-effectiveness threshold ($/QALY)</th>
<th>Conclusion based on outcome</th>
<th>Funding</th>
<th>Country</th>
<th>QHES score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olgiati 2012 [30]</td>
<td>5-HTTLPR</td>
<td>No</td>
<td>QALW</td>
<td>Euro A $1,100</td>
<td>&lt; 3 times the GDP per capita</td>
<td>Cost-effective in high-income countries</td>
<td>Unknown</td>
<td>Europe</td>
<td>61</td>
</tr>
<tr>
<td>Serretti 2011 [31]</td>
<td>5-HTTLPR</td>
<td>No</td>
<td>QALW</td>
<td>$2,900</td>
<td>$50,000</td>
<td>Not cost-effective</td>
<td>Unknown</td>
<td>Italy</td>
<td>87</td>
</tr>
<tr>
<td>Reed 2011 [32]</td>
<td>8–14 risk alleles</td>
<td>No</td>
<td>QALY</td>
<td>$98,100</td>
<td>$100,000</td>
<td>Cost-effective</td>
<td>National Cancer Institute</td>
<td>USA</td>
<td>83</td>
</tr>
<tr>
<td>Djalalov 2012 [33]</td>
<td>APOE ε4</td>
<td>yes</td>
<td>QALY</td>
<td>CAN$38,000 ($32,700)</td>
<td>No number</td>
<td>May be economically attractive</td>
<td>Academic resources</td>
<td>Canada</td>
<td>90</td>
</tr>
<tr>
<td>Kazi 2014 [34]</td>
<td>CYP2C19</td>
<td>No</td>
<td>QALY</td>
<td>Extendedly dominated</td>
<td>$50,000</td>
<td>may improve cost effectiveness</td>
<td>American Heart Association and academic resources</td>
<td>USA</td>
<td>90</td>
</tr>
<tr>
<td>Reese 2012 [35]</td>
<td>CYP2C19</td>
<td>No</td>
<td>CVE avoided</td>
<td>Cost saving</td>
<td>No number</td>
<td>Dominant</td>
<td>Unknown</td>
<td>USA</td>
<td>64</td>
</tr>
<tr>
<td>Sorich 2013 [36]</td>
<td>CYP2C19</td>
<td>No</td>
<td>QALY</td>
<td>AUS$30,00–50,000 ($24,500–40,800)</td>
<td>Cost-effective</td>
<td>Heart Foundation of Australia</td>
<td>Australia</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Crespin 2011 [37]</td>
<td>CYP2C19 *2</td>
<td>Yes</td>
<td>QALY</td>
<td>$10,100</td>
<td>$50,000</td>
<td>Cost-effective</td>
<td>Academic resources</td>
<td>USA</td>
<td>93</td>
</tr>
<tr>
<td>Lala 2013 [38]</td>
<td>CYP2C19 *2</td>
<td>Yes</td>
<td>QALY</td>
<td>Dominant</td>
<td>n.a.</td>
<td>Dominant</td>
<td>Academic resources</td>
<td>USA</td>
<td>83</td>
</tr>
<tr>
<td>Panattoni 2012 [39]</td>
<td>CYP2C19 *2</td>
<td>No</td>
<td>QALY</td>
<td>NZ$24,600 ($19,200)</td>
<td>NZ$50,000 ($39,000)</td>
<td>Cost-effective</td>
<td>Academic resources</td>
<td>New Zealand</td>
<td>75</td>
</tr>
<tr>
<td>Pink 2013 [40]</td>
<td>CYP2C9 and VKORC1</td>
<td>No</td>
<td>QALY</td>
<td>£13,200 ($20,600)</td>
<td>£20,000–30,000 ($31,200–46,700)</td>
<td>Cost-effective</td>
<td>Academic resources</td>
<td>UK</td>
<td>93</td>
</tr>
<tr>
<td>You 2012 [41]</td>
<td>CYP2C9 and VKORC1</td>
<td>No</td>
<td>QALY</td>
<td>Dominated by genotype-guided approach</td>
<td>$13,800 per QALY</td>
<td>Dominated by dabigatran 150 mg</td>
<td>Cost-effective</td>
<td>Research Grants Council of the Hong Kong special administrative Region, China</td>
<td>USA</td>
</tr>
<tr>
<td>You 2014 [42]</td>
<td>CYP2C9 and VKORC1</td>
<td>No</td>
<td>QALY</td>
<td>$2,800 per QALY</td>
<td>$50,000</td>
<td>Cost-effective</td>
<td>Research Grants Council of the Hong Kong special administrative Region, China</td>
<td>USA</td>
<td>70</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>1&lt;sup&gt;st&lt;/sup&gt; Author (reference)</th>
<th>PGx test</th>
<th>Analytical validity PGx test reported (y/n)</th>
<th>Outcome Measure</th>
<th>Quantitative outcome or ICER (US$)</th>
<th>Cost-effectiveness threshold ($/QALY)</th>
<th>Conclusion based on outcome</th>
<th>Funding</th>
<th>Country</th>
<th>QHES score</th>
</tr>
</thead>
<tbody>
<tr>
<td>de Lima Lopes 2012 [43]</td>
<td>EGFR</td>
<td>No QALY</td>
<td>Dominant</td>
<td>n.a.</td>
<td>Dominant</td>
<td>AstraZeneca Pte Ltd.</td>
<td>Singapore</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>Handorf 2012 [44]</td>
<td>EGFR</td>
<td>No QALY</td>
<td>$110,600</td>
<td>$100,000</td>
<td>Cost-effective</td>
<td>OSI Pharmaceuticals/ Genentech</td>
<td>USA</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Zhu 2013 [45]</td>
<td>EGFR</td>
<td>No QALY and LYG</td>
<td>$57,000 per QALY</td>
<td>&lt; 3 times the GDP per capita of China ($16,300)</td>
<td>Not cost-effective*</td>
<td>Shanghai Health Bureau</td>
<td>China</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Kauf 2010 [46]</td>
<td>HLA-B*5701</td>
<td>No HSR avoided</td>
<td>$300 (60 days' time horizon)</td>
<td>No number.</td>
<td>Cost-effective</td>
<td>Glaxo Smith Kline, Inc.</td>
<td>USA</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>Rattanavipaphong 2013 [47]</td>
<td>HLA-B*1502</td>
<td>Yes QALY</td>
<td>Epilepsy patients</td>
<td>THB$220,000 ($6,700)</td>
<td>THB$32,502,000 ($984,800)</td>
<td>Not cost-effective</td>
<td>Academic resources</td>
<td>Thailand</td>
<td>70</td>
</tr>
<tr>
<td>Liu 2012 [48]</td>
<td>IL-28B</td>
<td>No QALY</td>
<td>$50,400</td>
<td>No number</td>
<td>Not clear</td>
<td>Academic and governmental</td>
<td>USA</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>Greeley 2011 [49]</td>
<td>KCNJ11 and ABCC8</td>
<td>Yes QALY</td>
<td>Dominant</td>
<td>n.a.</td>
<td>Dominant</td>
<td>Academic resources</td>
<td>USA</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>Parthan 2013 [50]</td>
<td>KIF6</td>
<td>No QALY</td>
<td>$45,000</td>
<td>$100,000</td>
<td>May be cost-effective</td>
<td>Celera corporation</td>
<td>USA</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>Vijayaraghavan 2012 [51]</td>
<td>KRAS</td>
<td>Yes LYS</td>
<td>Cost saving</td>
<td>Cost saving</td>
<td>Cost-saving in both US and Germany</td>
<td>Roche Molecular Systems, Inc.</td>
<td>USA and Germany</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Hagaman 2010 [53]</td>
<td>TPMT</td>
<td>No QALY</td>
<td>$29,700</td>
<td>$50,000</td>
<td>Cost-effective</td>
<td>Unknown</td>
<td>USA</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Thompson 2014 [52]</td>
<td>TPMT</td>
<td>No QALY</td>
<td>Negative ICER</td>
<td>n.a.</td>
<td>Cost-saving but also health loss</td>
<td>Department of Health UK</td>
<td>UK</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>Pichereau 2010 [54]</td>
<td>UGT1A1*28</td>
<td>No neutropenia avoided</td>
<td>€900–1,100 ($1100–1300)</td>
<td>No number</td>
<td>Cost-effective</td>
<td>No specific financial support for this study</td>
<td>France</td>
<td>84</td>
<td></td>
</tr>
</tbody>
</table>

ABCC, ATP-binding cassette transporter sub family C; APOE, apolipoprotein-E; CYP, cytochrome P-450; EGFR, epidermal growth factor receptor; GPD: gross domestic product; HLA, human leukocyte antigen; HTTLPR, serotonin-transporter-linked polymorphic region; ICER, incremental cost-effectiveness ratio; KCNJ, Potassium inwardly-rectifying channel, subfamily J; KRAS, Kirsten rat sarcoma viral oncogene homolog; LYG, life-years-gained; LYS, life-years-saved; PGx, pharmacogenetic; QALW, quality-adjusted-life-week; QALY, quality-adjusted-life-year; THB, Thai Baht; TPMT, thiopurine S-methyltransferase; UGT, UDP-glucuronosyltransferase; UK, United Kingdom; VKORC, Vitamin K epoxide reductase complex

* With the gefitinib patient assistance program (sponsored therapy after first six months) it might be a cost-effective treatment option.

doi:10.1371/journal.pone.0146262.t002
Sensitivity analysis

Until 2008, only 5 out of the 31 (16%) studies performed both a univariate and probabilistic sensitivity analysis, whereas since 2008 this were 35 out of the 49 (71%) studies (Fig 2B). One of the major uncertainty factors was the lack of robust clinical evidence for clinical utility of the PGx test itself. Hence, almost all authors had to define assumptions which were sometimes solely based on expert opinion. Other frequently mentioned uncertainty factors were the costs of the PGx tests and their real-world utility and performance. As a consequence of the uncertainty around the included genotyping costs, several studies included a costs range in their analysis. For example, Schackman et al. (2013) demonstrated that at a test cost of US$107, genetic testing was not cost-effective. However, at a price of US$10 per test, the PGx test was cost-effective [29]. Although variance in genotyping costs was frequently found to have a major impact on the ICER, 11 out of the 38 studies (30%) assessed in this update did not include a range of genotype costs in their sensitivity analysis (S1 Table). With respect to drug costs, some studies showed that drug driven costs did not influence the study’s outcome. For example, Crespin et al. (2011) showed in their sensitivity analysis that even if the costs of the drugs guided by PGx testing dropped substantially, the non-PGx-test-guided drug remained cost-effective [37].

Genes investigated and analytical validity of PGx test

In the period from 2000 until 2014, the majority of studies investigated the TPMT gene (17 studies). Among the studies presented in this update (i.e. from August 2010 onwards), most investigations concerned the CYP2C19 screening tests (Fig 3A). We noticed that the CYP2C19 gene was studied in many different scenarios. Some studies assessed if the use of a CYP2C19 independent drug like prasugrel or ticagrelor would be a cost-effective treatment option [35,36,38]. Similar results were found for CYP2C9 and VKORC1 testing in combination with warfarin treatment, which was compared with CYP2C9 and VKORC1 independent novel oral anticoagulants, like dabigatran.[41].

In general the analytical validity of PGx tests is high (>95%), nevertheless variation in the analytical validity will in combination with the prevalence of a certain genetic trait determine the predictive value of a test [97]. This can be of influence on the cost-effectiveness and was nicely demonstrated by Kauf et al., who identified the negative predictive value as an important input parameter for the costs-effectiveness of HLA-B*5701 genotyping in their sensitivity analysis [46]. Nevertheless, underlying assumptions about the analytical validity of the test were not reported. It was found that only 11 out of the 38 (30%) studies included since 2010, did report underlying analytical validity of the PGx test (Table 1).

Outcome of the studies

There were 21 out of 80 (26%) studies which concluded that PGx testing was dominant (i.e. resulting in clinical benefits as well as cost-savings). From 2010 onwards, most authors concluded that PGx testing was cost-effective, while only four studies concluded that it was not cost-effective (Fig 3B). In the period from 2010 until 2014, several studies provided the specific conditions at which genetic testing might become cost-effective (Tables 1and 2). For example, Dong et al. (2012) and Rattanavipapong et al. (2013) showed that the PGx tests could be cost-effective depending on either the patient population or the disease [23,47]. Due to an imperfect capability of PGx tests to differentiate between carriers of a genetic variant, some patients might be misclassified and receive suboptimal treatment. As a result, two studies assessing KRAS and BRAF testing and one study assessing TMPT testing found the PGx testing strategy was cost-saving, but with a small health loss compared to the non-testing strategy [25,26,52].
Time horizon and discounting

To capture all costs, savings and effects of an intervention, a lifelong time horizon often seems the best time horizon, although in some scenarios it may be argued that a shorter period is acceptable. Before 2008, 12 out of 26 studies (46%) applied a time horizon of 12 months, and only five (19%) studies applied a lifelong time window. Among the studies published since 2010, there was a broad range in the applied time horizons, from two weeks until lifetime (S1 Table). In addition, out of the 38 studies 17 (45%) applied a lifelong time horizon, six (16%)
studies applied a 30 year time horizon of which two studies also used a lifetime horizon. To deal with uncertainty around the appropriated time horizon, different time horizons can be used. We found that six out of 38 studies (16%) varied their time horizons [18,20,26,38,46,49]. In this way, insight into short- and long-term outcomes was given.

Discounting was applied for almost all studies published after 2006 that applied a time horizon longer than one year. The majority of the studies used discounting at 3% annually for costs and effects similarly (S1 Table).

Discussion of limitations
The limitations and uncertainties of an economic analysis have to be acknowledged in order to judge the study on its merits. Nearly all of the newly included papers (i.e. published between August 2010 and October 2014) discussed their limitations, uncertainties and possible shortcomings of their economic analysis (S2 Table). The topic most commonly discussed was the lack of solid clinical evidence. As a result, many studies had to make assumptions or relied on experts’ opinions. However, as mentioned before, most authors did include different efficacy scenarios in their sensitivity analysis. Not all papers gave clear information about assumptions concerning sensitivity and specificity of the PGx tests, and only three out of 38 (8%) mentioned these assumptions as a limitation [45,49,51]. Another typical limitation for the effects of PGx tests is the time in which test results will become available for clinical decision making. In current analyses, test results were often assumed to be immediately available. However, this might be an unrealistic assumption and only one out of 38 studies explicitly mentioned this assumption as a limitation [23]. More general limitations were the lack of data with respect to heterogeneity in patient populations, hampering extrapolation of results to patients of different ethnicities, subpopulations and/or country specific populations. Moreover, several papers mentioned the difficulties in extrapolating long-term clinical utility results from the short-term clinical trials and lab studies.

Role of funding
Out of the 78 selected studies, 11 (14%) were funded by pharmaceutical companies. Before 2008, none of the studies was (directly) funded by pharmaceutical companies. In 2008 and 2009 only two (12.5%) out of 16 studies published in that period were funded by a pharmaceutical company [70,71]. Between August 2010 and September 2014, 9 (24%) out of the 38 selected studies were funded by pharmaceutical companies (Tables 1 and 2). Regarding outcomes and conclusions all of these studies concluded that PGx tests were dominant, cost-saving, or cost-effective. Among the remaining studies which were funded by other resources, 14% concluded PGx tests were not cost-effective.

Quality assessment
The studies included through this update received a quality score according to the QHES. There was 2% disagreement between the reviewers for which consensus was reached through discussion. The average quality score was 76. The score which was given to each study is shown in Tables 1 and 2. The majority of the studies (71%) were of high quality. On average, studies concerning testing for EGFR received the highest rating and studies about the 21-gene assay received the lowest rating. Some items were scored negative for the majority of the studies. One of them was the item concerning the perspective of the analysis (i.e. societal, health care payers, etcetera). Most studies did not explain why the perspective of the analysis was chosen. In addition, low scoring was received for separate reporting of the short-and long-term outcomes.
Most studies did not make such a distinction. Lastly, the direction and magnitude of potential biases were often not discussed.

**Discussion**

**Principal findings**

Since 2004, there is an increase in the number of studies evaluating the economic value of PGx tests and this increase accelerated from 2008 onwards. There were not that many economic evaluations of PGx tests available as one might expect given the unravelling of the whole human genome in 2003, the clinical possibilities, and the fast development and decreasing costs of genetic tests [98]. This could be related to limited implementation of pharmacogenetic knowledge into daily clinical practice [5]. Reasons for this are the uncertainty about clinical relevance, concerns about the availability of genetic data and considerable differences in cost-effectiveness which are found, in particular between countries [5,99,100].

Many studies included through this update did not assess the intrinsic value of the PGx test itself, but compared a PGx test treatment combination with an alternative treatment. For example, the tests for *CYP2C19* or, both *CYP2C9* and *VKORC1* were incorporated in several models as the current treatment option in combination with clopidogrel or warfarin treatment, respectively. The alternative treatment options, which were assessed in the included studies, were independent of pharmacogenetics and were found to be cost-effective. Therefore, in cost-effectiveness analysis of antithrombotic therapy, there seems an ongoing movement away from pharmacogenetic testing, towards treatment options with compounds that are, so far, considered to be independent of pharmacogenetics.

Before 2008, most analyses were cost-effectiveness analyses, however since 2008 there has been a trend towards the use of cost-utility analyses. Cost-utility analysis is currently considered as the preferred type of analysis for health care choices as is advised in several national guidelines, although other types can be suitable depending on the specific study ([1,102]. Before 2008, most studies performed only a simple univariate sensitivity analysis, if any [9]. We found that more recent studies performed both univariate and probabilistic sensitivity analyses. This combination is also advised in several national guidelines for pharmacoeconomic evaluation and is a part of the QHES checklist [16,101,102]. As a result of these more comprehensive sensitivity analyses, the quality of economic evaluations appears to be improved over the last decade. However, we found that differences in genotyping costs were not always included in the sensitivity analysis which leaves room for improvement.

Among the new studies included in this update (i.e. since 2010), considerable differences in the length of the time horizon were noticed, varying from two weeks to lifelong. A time horizon shorter than a year was primarily related to the expected relevant clinical outcomes. Interestingly, there were differences in the time horizon between studies investigating the same genes. For example, Kazi *et al*. and Panattoni *et al*. applied a lifelong time horizon [34,39], whereas others applied a time horizon of 15 months [35,38]. All studies discussed at least some limitations and a lack of robust clinical evidence was most frequently mentioned as an important limitation contributing to uncertainty in the analysis. Although studies frequently made assumptions about analytical validity of PGx tests as well as about the rapid clinical availability, these were often not discussed as limitations or studied in sensitivity or scenario analyses. Besides the analytical validity, the clinical validity of PGx tests is important. In general this is the same as an effect estimate of a PGx test. As such, the clinical validity is correctly embedded in an economic evaluation. However, this does not apply for all PGx tests, because some PGx tests involve an analysis of multiple genetic variations. These variations can all contribute to a similar genotype prediction and therefore the clinical validity depends on the number and type
of variant alleles analysed. For example, CYP2C19 genotyping which depends on general molecular genetic analysis of single nuclear polymorphisms to detect variant alleles. Several studies included in this review studied only the CYP2C19*2 allele [37–39]. However, other allele like the *3 allele can also effect CYP2C19 activity and give a similar clinical effect as the *2 allele [103]. When more variant alleles are analysed the clinical validity of the PGx test will increase, although some variants are rare and will contribute little. With the ongoing and rapid increase of knowledge about variant alleles, it is important economic evaluations report the variant alleles on which their assumptions about effects of genotyping (i.e. clinical validity) were based. For the example of CYP2C19, two out of the six studies involving the CYP2C19 gene, did not specify which alleles were included [34,36]. This hampers the extrapolation of their findings towards populations of a different ethnicity and other laboratories.

Compared to studies published before 2008, studies were more commonly funded by pharmaceutical companies. Interestingly, all of these sponsored studies concluded that PGx tests were dominant, cost-saving, or cost-effective whereas the few studies which concluded otherwise were financed by academic, governmental, or unknown resources. It is known that studies funded by pharmaceutical companies publish more positive results when compared to studies funded by other resources which is in line with our results [104]. These positive biased results are not related towards the quality of the studies, but to the comparison which is made or publication bias [105]. For many of the economic evaluations in this review, assumptions about the effect of the genetic test were made. Furthermore, analytical validity was often not included in the model. This leaves room to bias result in favour of a preferred treatment strategy. Therefore, attention should be given toward assumptions about these aspects, especially when studies are funded by pharmaceutical companies or if the funding is not reported.

Most studies included in this review concluded that the application of PGx tests was cost-effective. Yet, the conclusions were not unambiguous, often due to the uncertainties in the economic models. Another reason for this was that most of the newly included studies did not assess the intrinsic value of the PGx test itself, but a scenario involving one or more PGx tests. Although such estimates are important for the economic impact of the application of personalized medicine, they do not provide information about the cost-effectiveness of the PGx test itself and therefore outcomes can be different. To improve generalizability between studies, an additional scenario analysis in which equal costs and efficacy of the compared treatment strategies are assumed, to assess the intrinsic value of the PGx tests, could be used. Such an approach would especially be interesting when the comparator drug is under patent and drug costs are likely to decrease in the future. Another aspect creating different conclusions was the genetic variety between study populations. This was clearly described by some papers which mentioned the specific conditions like a specific geographic region or a disease for which the genetic testing was cost-effective. For example, PGx testing for HLA-B*5702 was cost-effective in Singaporean Chinese and Malays, but not in Singaporean Indians [23]. In addition, Rattanavipapong et al. (2013) found that HLA-B*5702 testing was cost-effective in epileptic patients, but not in neuropathic pain patients [47].

Comparison with previous literature reviews

For the new studies included in this update (i.e. since 2010), an average QHES-quality score of 76 was found. This was in line with the results from the review from Wong et al. about pharmacoeconomics of PGx tests. They found an average quality score of 77 [13]. However, it was lower compared to a related review of Djalalov et al., who found an average quality score of 90 [10]. This is likely due to the subjective nature of some items in the quality assessment. For example, ‘item 3’ asks if the used estimates for the analysis were from the best available source,
and is therefore quite sensitive to the interpretation of the reviewer [16]. Previous review studies pointed out that the methodology of economic evaluations of PGx tests is often heterogeneous and of insufficient quality [9,11]. Assasi et al. found that economic PGx studies which were of low quality (i.e. QHES score of < 50), frequently failed to handle uncertainty, did not inform about the study’s limitations, and did not discuss direction and magnitude of potential uncertainty [11]. Among the studies included in this update (i.e. since 2010), most of the authors discussed uncertainties and except for the study by Tiamkao et al., all incorporated uncertainty in a sensitivity analysis [24]. However, direction and magnitude of potential bias was still not sufficiently discussed and remains a major point for improvement.

Limitations of our approach
This review has some limitations. Firstly, we did not include studies from other databases than PubMed or grey literature and we only included English written studies. Therefore some studies might have been missed. In general, studies that are not indexed in MEDLINE or written in English do not have a large impact on reviews’ outcomes [106]. Nevertheless, these studies are frequently of lower quality, and therefore the average quality of the studies included in our review might have been slightly overestimated. However, based on the comparison with other reviews we found a slightly lower average quality score. Lastly, publication bias can always influence the findings of a review. Therefore cost-effectiveness of PGx tests could be overestimated.

Recommendations
Based on the quality assessment, reporting of the reasons behind the chosen perspective and the type of economic model can improve the quality of economic evaluations of PGx tests. In addition, reporting of both short- and long-term outcomes and the influence of potential bias, in terms of direction and magnitude on the cost-effectiveness estimates could be improved. Although a substantial and persistent increase in the use of both univariate and probabilistic sensitivity analyses was observed since 2008, there is still room for improvement by using a combination of these techniques instead of one technique, among several studies published since 2008. Publication bias or biased comparators might favour cost-effectiveness of PGx tests. Among studies funded by companies with conflicting interests, the risk on this kind of bias should be critically assessed. In our previous reviews the main limitation that was identified was the unavailability of clinical evidence [1,9]. Although this remains an important issue, based on our new findings we can add some recommendations which are more applicable to implementation of PGx tests in clinical practice. First, the clinical validity of a test, i.e. the capability of the test to predict phenotypes with a clinical effect and the analytical validity should be reported as is recommended by the US Academy of Managed Care Pharmacy [107]. Both parameters should be included in a sensitivity analysis. In addition, the variant alleles on which these parameters were based should be reported. For some PGx tests, these estimates might be unknown. In this case, a better approach towards this problem would be inclusion of an univariate sensitivity analysis with different cut-off values for the analytical and clinical validity of the PGx test. In this way, a minimum for the analytical and clinical validity can be generated. Note that with an increase in analysed alleles and as such the clinical validity, the price for the PGx test usually increases as well. Secondly, different turnaround times in which PGx test results would become available for healthcare professionals after requesting the test would be very informative. If direct availability of the genetic test is assumed, for example in the case of pre-emptive genotyping, this should be clearly stated in the method section. This way, the generalizability of results to other countries where PGx tests are available would improve. In
addition, a range of costs for the genetic test should be evaluated in univariate and if applicable probabilistic sensitivity analyses. Furthermore, we recommend the addition of a scenario analysis in which drug costs between comparator groups are equalized to give information about the intrinsic value of the PGx test itself.

Conclusion
Application of PGx tests was mostly found to be a cost-effective or cost-saving strategy, although some studies concluded otherwise which underlines the importance of future studies assessing the cost-effectiveness of PGx tests. We found that only the minority of recent pharmacoeconomic evaluations assessed the intrinsic value of the PGx tests. New compounds that are not affected by genetics, are emerging as cost-effective alternatives for pharmacogenetic testing strategies. Over the last decade, there was an increase in the number of studies and in the reporting of quality associated characteristics. Due to rapid development in analytical techniques, reporting of analytical and clinical validity of the assessed PGx test is recommended for future evaluations. Furthermore robust clinical evidence regarding PGx tests’ efficacy is warranted.

Supporting Information
S1 PRISMA Checklist. (DOCX)

S1 Table. Overview of pharmacoeconomic PGx studies published between August 2010 and September 2014 analysing the intrinsic value of a PGx test (A), or comparing different treatment strategies involving PGx testing (B). (DOCX)

S2 Table. Main limitations that were discussed in the papers published between August 2010 and September 2014. (DOCX)

Author Contributions
Conceived and designed the experiments: BW SV MP. Analyzed the data: ML EB JB. Wrote the paper: EB ML JB. Commented on the first draft and provided input for the final manuscript: LA SV BW CB. Supervised the overall progress and provided final revision, comments and review: MP.

References


Neuropsychopharmacology 2009 Sep; 34(10):2227–2236. doi:10.1038/npp.2009.50 PMID: 19494805


