ADVISHE: A new tool to report validation of health-economic decision models
Vemer, P.; Corro Ramos, I.; Van Voorn, G.; Al, M.J.; Feenstra, T.L.

Published in:
Value in Health

DOI:
10.1016/j.jval.2014.08.1831

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Early version, also known as pre-print

Publication date:
2014

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
population are 18.5% for males and 9.8% for females. CONCLUSIONS: Random sampling from patient level data provided the best approximation of actual NHANES population predicted CVD rates. The cholesery deposition approach was slightly limited since only continuous variables could be utilized which could explain the deviation from the population predicted CVD rates. Independent sampling under-estimation of variance by 15% – 20%, an interesting finding as many individual simulation models created patients with this approach. Researchers should be cautious in their use of summary statistics when populating individual simulation models.

**PRM74 VALIDATION OF THE SPHR DIABETES PREVENTION MODEL**

**Thomas C., Watson P., Squires H., Chilcott J., Brennan A.**

University of Sheffield, Sheffield, UK

OBJECTIVES: We have developed a model to evaluate type-2 diabetes prevention interventions. We aimed to validate this model against external data to test the accuracy of the model. **Methods**: An individual patient simulation was developed to predict longitudinal trajectories of HbA1c, 2-hr glucose, FPG, BMI, systolic blood pressure, total cholesterol and HDL cholesterol based on statistical analyses of the Whitehall II longitudinal cohort. Criteria for diabetes diagnosis were flexibly specified. Data from an automated control protocol were utilized to ensure that the number of complications of diabetes were estimated from the UKDF5 outcomes model. Several validations were performed to compare model outcomes with reported data from external sources. We assessed the predicted diabetes incidence using data from the EPIC Norfolk cohort. Data from the Health Survey for England (HSE) 2003 cohort was simulated for eight years to compare predicted disease incidence and metabolic outcomes in a diabetic population with those observed in the UKDF5. We assessed the performance of the model in predicting the results of the ADDITION trial for diabetes screening. **Results**: We found that the model overestimated three-year incidence of diabetes and was therefore in high risk (HbA1c >6.0) individuals, underesti-
mated diabetes incidence in medium risk individuals (HbA1c 5.5-5.9) compared with the EPIC-Norfolk data. Predictions from HSE 2003 were fairly accurate. Predictions for mortality were similar to those obtained from the Whitehall II follow-up. Mortality was slightly under-predicted. The model replicated the non-significant difference seen between control and intervention arms of the ADDITION trial, but overestimated total mortality and cardiovascular disease. CONCLUSIONS: The SPHR Diabetes model appears to be fairly accurate at predicting complications of diabetes, was able to overestimate mortality rates in a newly diagnosed diabetic cohort, and underestimate cardiovascular disease and mortality compared with the UKDF5.

**PRM75 USE OF MODEL AVERAGING TECHNIQUES IN COST-EFFECTIVENESS ANALYSIS IN ONCOLOGY**

Le HH, Ozur Shillman I

Evidera, Boston, MA, USA

OBJECTIVES: Often in cost-effectiveness analysis (CEA) of oncologic drugs, survival data are available from a randomized controlled trial extrapolated to a lifetime horizon using parametric regression techniques. To capture parameter uncertainty in the analysis, regression parameters along with other model parameters are varied in probabilistic sensitivity analyses. However, structural uncertainty in the choice of regression models is rarely investigated. This study developed a probabilistic model of CEA framework and provides an example to address structural uncertainty in CEA. **METHODS**: Using a cohort partition model, the numbers of patients in "progression-free", "progressed", "에너지", "age of the target population", "quality of life", and "drug innovation" that were added to the aggregate index scores. In the drugs studied, the $R^2$ was 0.808 and the mean deviation from the population predicted CVD rates. Independent sampling under-
estimation of variance by 15% – 20%, an interesting finding as many individual simulation models created patients with this approach. Researchers should be cautious in their use of summary statistics when populating individual simulation models.

**PRM76 COMPARING THREE DIFFERENT METHODS OF HALF-CYCLE CORRECTION**

Nicolai E, Steker V

National Institute of Quality- and Organizational Development in Healthcare and Medicines, Budapest, Hungary

OBJECTIVES: To compare three different half-cycle correction methods and their effect on the final results of Markov models. **METHODS**: To assess the relative performance of the alternatives to the standard half-cycle correction we constructed a 5-state Markov model where the courses of the number of patients in health states follow different shapes to represent the most likely cases in modelling practise. We applied three different correction methods (standard half-cycle correction, Simpson’s method and using the mid-cycle values) and we also looked at the results without any correction and with the Markov cycle, the pattern members could score items in three e-mail rounds. Based on the Delphi method, the panel members could score items in three e-mail rounds.

**RESULTS**: Our Markov model Simpson’s method provided the most accurate results with the difference from results with the mid-values 0.1% in 67 of 80 cases. The second most accurate method was using the mid-values. The standard half-cycle correction method provided more accurate results than calculations with-
out any method of half-cycle correction with the exception of one set of input param-
eters. **CONCLUSIONS**: Based on our model the most accurate method for half-cycle correction is Simpson’s method as in most cases it was the closest to real data. It is important to note that with a few exceptions even the standard method’s results were more accurate than in cases where no half-cycle correction was applied.

**PRM77 APPLICATION OF A MODEL OF DECISION BASED ON FUZZY LOGIC TO PHARMACOECONOMICS: TREATMENT OF CROHN’S DISEASE WITH ANTITNF IN OUT OF LABEL USE**

Alonso Hernandez NA, Gonzalez Cuello A

Hospital Los Arcos Mar Menor, San Javier (MURCIA), Spain, 2Murcia University, MURCIA, Spain

OBJECTIVES: We present a model based on decision fuzzy logic, and apply to off labels used of antiTNF in Crohn’s disease (CD) (Infliximab (IFB) $30 mg/kg/week$, adalimumab (ADA) $80 mg/2weeks$, Certolizumab (CZB) $200mg/2weeks$). The term “fuzzy logic” (FL) was introduced in 1965 by LAZadie. Compared to traditional logic, FL variables may have a truth value in degree. FL has been applied to many fields, from electronics to medicine.

**METHODS**: According to a decision analysis model based on FL four fuzzy variables that affect the choice of treatment are defined: treatment success (expressed as a probability), cost, success of failure (expressed as inverses), and other conditions about the cost (negotiation, handling of drugs...). Based on the value of these fuzzy variables, three linguistic variables(high, Medium, Low) are defined to express convenience of choice. The combination of the three possible values for each of the variables gives us 81 possible decision rules, so that the (RHHI) would be the most favorable option and (LILLI) the most unfavorable. So a new fuzzy variable called “ranking” was established for classifying these options with 7 possible values (very-unfavorable, unfavorable, slightly-unfavorable, neutral, slightly-favorable, favorable, very-favorable) The value of the fuzzy variables for antiTNF at 52 weeks of treatment, were established based recent meta-analysis and reviews. **RESULTS**: The ranking of treatment was similar to the cardiovascuar outcomes from the UKPDS outcomes model. Cardiovascular events were estimated from the QRISK2 algorithm. Microvascular complications of diabetes were estimated from the UKPDS outcomes model. Several conclusions were obtained. The SPHR Diabetes model appears to be fairly accurate at predicting complications of diabetes, was able to overestimate mortality rates in a newly diagnosed diabetic cohort, and underestimate cardiovascular disease and mortality compared with the UKDF5.

**PRM78 MULTI-CRITERIA DECISION ANALYSIS (MCDA): TESTING A PROPOSED MCDA MODEL FOR ORPHAN DRUGS**

Schebi C., Connolly M*

Unit of Pharmacoeconomics & Pharmacometrics, Groningen, The Netherlands, *University of Groningen, Groningen, The Netherlands

OBJECTIVES: Since the introduction of the orphan drugs in Europe, it has been sug-
gested that the general method of appraising drugs' reimbursement is not necessarily suitable for orphan drugs. The National Institute for Health and Clinical Excellence indicated that several criteria other than cost and efficacy could be considered in reimbursement decisions for orphan drugs. The aim of this study was to assess the potential of a 3 criteria MCDA framework to add value to the decision making process for a range of orphan drugs in different diseases to test the correlation between drug price and aggregated MCDA scores for each product. **METHODS**: A MCDA framework was developed using the nine criteria suggested by Hughes-Wilson et al. A supplementary literature review was conducted to identify other attributes described in the application of MCDA in rare diseases. A numerical scoring matrix is used to scale the criteria to three levels (positive, neutral, negative) for each criterion. Conclusions between the average annual cost of the drugs and aggregate MCDA scores were tested and plotted graphically. Different weightings for each of the attributes were also tested. A further analysis was conducted to test the impact of including the drug cost as an attribute in the aggregate index scores. **RESULTS**: The literature review identified further commonly cited criteria: ‘convenience of administra-
tion’, ‘age of the target population’, ‘quality of life’, and ‘drug innovation’ that were added to the aggregate index scores. In the drugs studied, the Kappa was 0.808 and 0.704 when costs were included and not included, respectively. The standard error of the slope varied from 771.9 to 1143.3 when costs were included and not included, respectively. **CONCLUSIONS**: This quantitative study provided insight into using MCDA and its framework that was proposed by Hughes-Wilson et al. to a range of orphan drugs in different diseases to test the correlation between drug price and aggregated MCDA scores for each product. **METHODS**: A MCDA framework was developed using the nine criteria suggested by Hughes-Wilson et al. A supplementary literature review was conducted to identify other attributes described in the application of MCDA in rare diseases. A numerical scoring matrix is used to scale the criteria to three levels (positive, neutral, negative) for each criterion. Conclusions between the average annual cost of the drugs and aggregate MCDA scores were tested and plotted graphically. Different weightings for each of the attributes were also tested. A further analysis was conducted to test the impact of including the drug cost as an attribute in the aggregate index scores. **RESULTS**: The literature review identified further commonly cited criteria: ‘convenience of administra-
tion’, ‘age of the target population’, ‘quality of life’, and ‘drug innovation’ that were added to the aggregate index scores. In the drugs studied, the Kappa was 0.808 and 0.704 when costs were included and not included, respectively. The standard error of the slope varied from 771.9 to 1143.3 when costs were included and not included, respectively. **CONCLUSIONS**: This quantitative study provided insight into using MCDA and its framework that was proposed by Hughes-Wilson et al. to a range of orphan drugs in different diseases to test the correlation between drug price and aggregated MCDA scores for each product.
addition, the Dutch National Health Care Institute commented on usefulness for decision makers. The agreement was a separate group convened for the event. Experts could comment during a workshop at ISPOR Montreal 2014. RESULTS: 35 Validation techniques were identified and grouped into four categories: conceptual model validation, computerized model validation, data validation and operational validation. Around 30 HE experts commented on this set, which is currently sent out for a final, fifth round. CONCLUSIONS: When filled out by the modellers, AdVISHe (Assessment of the ValIdation Status of Health-economic decision models) supports user models in assessing the validation status of a model. It will be useful as part of a reimbursement dossier, by providing systematic and transparent insight into the validation efforts performed and their results.

PRM60 MODELLING SURVIVAL IN THE PRESENCE OF DIFFERENT MECHANISMS OF ACTION: IPILIMUMAB AND VEMURafenib IN ADVANCED MELANOMA Lee D1, Porter J2, Hervé N1, Hatwell A1

1Freshfields Bruckhaus Deringer, UK, 2Forsyte Ltd, Uxbridge, UK

OBJECTIVES: Traditional indirect comparison treatment methods assume the underlying survival profiles of treatments are similar (i.e. proportional hazards). This assumption is unlikely to hold for the comparison of ipilimumab and vemurafenib. Whereas vemurafenib exhibits improved short-term survival compared with ipilimumab, pooled study data for ipilimumab consistently show that patients achieve durable long-term survival. We present a method to compare across trials with differing survival profiles accounting for follow-on treatments and different baseline characteristics. METHODS: Comparative survival estimates for ipilimumab and vemurafenib were produced using patient-level data from trial CA184-024 for ipilimumab and survival cure rates from trial BIRM-3 (alongside data for vemurafenib). The BIRM-3 vemurafenib overall survival curve was adjusted to account for (a) the effect of second-line ipilimumab (via a funnel-state methodology); (b) patient baseline characteristics; and (c) KIR6.3 and CA184-024, by means of a model (Korn model), constructed to predict the outcomes for dacarbazine-treated patients. The resulting survival estimates were compared with naïve unadjusted survival curve fits, and estimates produced using a hazard ratio meta-analysis (HRMA) from a three-part reimbursement dossier, by providing systematic and transparent insight into the validation efforts performed and their results.

PRM61 HEALTH ECONOMIC MODELS IN ALZHEIMER’S DISEASE: A CRITICAL ASSESSMENT Weiler S1, Droeschel D2, Kaiser K3

1Max Planck Market Access & Pricing Strategy GmbH, Wai am Rhein, Germany, 2University of Freiburg, Freiburg, Germany

OBJECTIVES: Alzheimer’s Disease destroys brain cells, causing problems with memory, thinking, and behavior. A severe enough onset can affect family and social relationships, etc. Consequently, the most basic activities of daily living fail. Tens of thousands of patients with AD have been treated with disease-modifying agents. These options have been introduced and evaluated from a health economic perspective. However, given the specific characteristics of the disease an evaluation of existing models is needed. METHODS: The following databases were searched systematically: PubMed, Health Technology Assessment Database, NHS Economic Evaluation Database, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, DAHTA database, PSYNDEx and PsyCINFO. For the abstracts that met the pre-defined inclusion criteria, full text articles were obtained and evaluated for inclusion in the assessment. RESULTS: After eliminating duplicates the search yielded 219 articles of which another 940 were excluded based on the title selection. Finally 59 articles have been reviewed in full text after abstract review. Out of those articles 39 were deemed to be relevant based on the research question. The majority of models (48%) have been Markov models, other methods being used were various statistical analysis applications, micro-simulation, and discrete-event simulations. Limitations of existing models include the following: Focus on cognitive function as disease progression only, lack of inclusion of correlation between disease progression and other factors (e.g. residential status); lack of complete structure of diagnosis and treatment of disease (e.g. including non-drug treatments). Based on the Drummond checklist for health economic models the quality of models proved generally to be high but the majority of those lack presenting 10% or more of the pathway of the natural history of the disease. CONCLUSIONS: Current models do not allow decision makers optimally characterizing the disease, to better assess the costs and benefits of a wide range of potential interventions. Potential new models need to take the disease characteristics and specifics more appropriately into account.

PRM62 APPROACHES USED TO MODEL THE RELATIONSHIP BETWEEN PROGRESSION-FREE SURVIVAL (PFS) / TIME-TO-PROGRESSION (TTP) AND OVERALL SURVIVAL (OS) WITHIN HEALTH ECONOMIC MODELS OF CANCER THERAPIES Rafia R., Ward S.E.

University of Sheffield, Sheffield, UK

OBJECTIVES: Within health economic models of metastatic cancer therapies assumptions regarding the relationship between progression-free survival (PFS) and overall survival (OS) are typically required, notably when OS data are immature or unavailable. A review was undertaken to identify the methods that have been used within health economic models regarding this relationship and to identify the methods that are recommended for the data used. RESULTS: 41 relevant articles were available, PFS/OS data were not available or immature. METHODS: All NICE technology appraisals in the advanced and/or metastatic cancer setting completed by December 2013 were reviewed. The review included all relevant appraisal documents publicly available on the NICE website containing information on the methods used and/or rationale for the approach taken to model the relationship between OS and PFS/TPP within the health economic model. This included the sponsor submission and updated analyses, the independent Assessment Report, and other reports/analyses in relation to the appraisal process. RESULTS: In those instances where OS data were immature or not available, PFS/TPP was typically assumed to be a valid surrogate of OS. Justification for this approach was incrementally reported. In some cases a sensitivity analysis with a quantification of the assumed relationship was informed by published evidence and/or expert judgement. In some cases attempts were made to explore the potential impact of this relationship in sensitivity analysis. CONCLUSIONS: The methods and/or approach for the approach model the relationship between OS and PFS/TPP in health economic models has been inconsistently reported and justified. Whilst some health economic models attempted to quantify this relationship, further transparency is required. A consensus needs to emerge on the most appropriate approaches to be used within health economic models to quantify this relationship, specifically when OS data are not available or immature and to identify the circumstances when such approaches may be most relevant.

PRM63 COMPARISON OF METHODS TO ESTIMATE HEALTH STATE UTILITIES IN METASTATIC BREAST CANCER (MBC) Hudgens S1, Briggs A2, Trembly C2, Forsythe A1, Lloyd A4

1Clinical Outcomes Solutions, Tucson, AZ, USA, 2University of Glasgow, Glasgow, UK, 3Kesei Inc, Quincy MA, USA, 4University of Sheffield, UK

OBJECTIVES: Patient-level utility values for different stages of MBC and toxicities commonly associated with chemotherapy regimes are useful for health economic assessments. Three methods to estimate utilities exist when direct utility data are not available: using ‘utility mapping’ from existing disease-specific scales, vignette studies that describe the health states; or derivation of preference-based measures from an existing condition-specific scale. This study compares utility estimates in MBC utilizing the above methods. METHODS: Based on data from a phase 3 clinical trial in MBC (N=1102) utility mapping was conducted using a published regression algorithm to convert the EORTC QLQ-C30 questionnaire to the EQ-5D utility. Mean utility values were estimated for relevant health states: stable disease (SD), tumor response (TR), and disease progression (DP). RESULTS: Three vignette studies that describe the health state utility values in MBC were compared. The 5-year A557 utility estimates ranged from 0.561 to 0.593 for intervention one, 0.562 to 0.594 for intervention two and 0.563 to 0.595 for intervention three. The difference in costs ranged from €561 to €912 for intervention one, €739 to €1350 for intervention two and €1410 to €1618 for intervention three. The 5-year cost-effectiveness ratios (ICERs) for the most comprehensive intervention, intervention four, was €17,000/QALY for two models, €25,000+28,000/QALY for three models.