A complex individual patient simulation model (URPDS Outcome Model version 1.3) was used with quality adjusted life years (QALY) and cost of complications as model inputs. To reduce this order uncertainty, 200 patients were simulated for each input combination selected. ANN simulation meta-models using a sample of 200 individual runs were developed and cross-validated to approximate the original simulation data without any specific input-output functional relationships. As a consequence, any number of input parameters. Performance was compared with a Gaussian Process (GP) meta-model, and a valid and better predictive meta-model was then used for PSA. RESULTS: From ANN meta-models, the mean absolute percentage error (defined as positive difference between the predicted and true output divided by the range in true output) was 3.8% for costs and 1.4% for QALYs compared with 5.1% and 2.1% in GP meta-models. The distribution of errors was approximately symmetric and zero means that mean costs and QALYs for an intervention are unlikely to be affected by the small inconsistencies associated with this approach. CONCLUSIONS: ANN performs better predicting the probability than GP meta-models in estimating costs and QALYs from the UKPDS outcome model. A PSA carried out using the ANN meta-model demonstrated the potential for ANN in analysing complex health economic models.

DA2 A CHOICE THAT MATTERS: COMPARING METHODS OF DATA SYNTHESIS IN COST-EFFECTIVENESS MODELLING

Vos P, Al M, Oppe M, Rutten-Van Mólen K
IMTA, Erasmus University, Rotterdam, The Netherlands

OBJECTIVES: Different methods of meta-analysis on parameter models can lead to different outcomes of cost-effectiveness (CE) modeling. As the “true” CE is unknown, it is unclear which method performs best. We compared different methods of meta-analysis with regards to the underlying “true” CE outcome. METHODS: In a simulation study we constructed two patient populations and their treatments (true vs. observed) with events and a progressive lethal disease. T-tests from these populations, comparing two treatments, varying the number of trials, trial sizes and between-study heterogeneity in scenarios. From each trial, utilities, transition and event probabilities, risk-differences and log-risk-ratios were estimated. These parameters were synthesized using frequentist fixed-effects (FE) and random-effects (RE), Bayesian fixed-effects (BFE) and random-effects (BRE) models. A CE model was filled and probabilistic sensitivity analysis was performed. We repeated this trial sampling, leading to 1000 sets of health economic outcomes for each scenario. We compared methods of meta-analysis on bias and coverage, the percentage of draws that the “true” outcome lies in the confidence interval. RESULTS: Even in the most heterogeneous scenario, biases were limited to approximately 5%, and similar for all methods, but small biases in individual treatments occasionally led to biases up to 30% in the difference between arms. FE models consistently have lower coverage than BFE. With homogeneous trials, all methods have coverage above 80% for all outcomes. BRE has coverage higher than 99% for all outcomes, regardless of heterogeneity. With heterogeneity, RE methods perform better than FE and BRE has a lower coverage compared to BFE. All methods, even with heterogeneous trials, have 100% coverage around the ICER. CONCLUSIONS: BFE or BRE models are preferred in all situations, as they are more conservative. However, insight in the real level of heterogeneity is important, as using BFE without heterogeneity will overestimate uncertainty.

DA3 THE POWER OF ASSUMPTIONS

Van Hout BA, Stevens JW
University of Sheffield, Sheffield, UK

OBJECTIVES: To develop a method which increases the potential to find statistically significant differences in costs and effects when a trial is powered using a sample size calculation method that allows for treatment assignment only. Larger sample sizes gave less biased results, while a higher correlation between independent variables resulted in more biased estimates of the treatment and subgroup effect. Correlation between the independent variables and the subgroup variable did not lead to biased results. CONCLUSIONS: The results show the feasibility and validity of the PS in subgroup analyses when analyzing registry databases and indirect comparisons in economic evaluations.

PODUM SESSION I: LATEST INSIGHT IN THE ESTIMATION OF PRODUCTIVITY COST: BETTER DESCRIBING THE SOCIETAL VALUE

PC1 THE USE AND PERFORMANCE OF PRODUCTIVITY SCALES TO EVALUATE PRESENTEEISM IN MOOD DISORDERS

Despréglé N1, Danchenko N2, François C3, Lersberg B4, Drummond M4
1Insitut Pasteur, Paris, France; 2Institut de Recherche de Santé et de Sécurité au Travail, Paris, France; 3Opuniment, London, England, France; 4Optuniment, Uxbridge, Middlesex, UK

OBJECTIVES: Mood disorders are associated with a high societal cost, mainly due to productivity loss and in particular presenteeism. The latter should therefore be measured with the most appropriate tool. The objective is to review the use of ten instruments in mood disorders and to provide recommendations about the most appropriate instruments according to the situation. METHODS: A systematic review was conducted using PubMed focusing on ten instruments: Endicott Work Productivity Scale (EWPS), Health & Labour Questionnaire (HLQ), WHO Health and Work Performance Questionnaire (HPQQ), Health and Work Questionnaire (HWOQ), LAM Employment Absence and Productivity Scale (LEAPS), Sheehan Disability Scale (SDS), Stanford Presenteeism Scale (SPS), Work and Health Interview (WHI), Work Limitations and Questionnaire (WLQ). A value of 1 means the best estimates of the treatment and subgroup effect. Correlation between the independent variables and the subgroup variable did not lead to biased results. CONCLUSIONS: The results show the feasibility and validity of the PS in subgroup analyses when analyzing registry databases and indirect comparisons in economic evaluations.

SIMULATION STUDY

van Eeren H1, Spreuwenberg MD2, van Manen JG3, de Rooij M4, Stijnen T5
1Erasmus University Medical Center, Rotterdam, The Netherlands; 2Maastricht University, Maastricht, The Netherlands; 3Vrije Universiteit Medical Centre, Amsterdam, The Netherlands; 4Department of Medical Statistics & Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands

OBJECTIVES: The use of registry databases and indirect comparisons has become important in health economic evaluations. Lack of randomization could lead to selection bias due to pretreatment differences between patients. To avoid selection bias, the propensity score method (PS) (Rosenbaum & Rubin, 1983) is often applied. However, average treatment effects can vary within different subgroups. It is yet unclear how to perform subgroup analyses when the propensity score method is applied. METHODS: A Monte Carlo simulation is conducted to test the performance of different forms of the PS in subgroup analysis. RESULTS: PSs differ in whether the variables included in the PS are indicators of the subgroup and were related to treatment assignment, to outcome or related to both assignment and outcome. Furthermore the PS is estimated in two ways, primarily on treatment assignment only and secondly on a combination of the treatment assignment and subgroup variable. These PSs were used as adjustment in a regression model. Simulations are accomplished for 18 different settings varying sample size, correlation between independent variables and correlation between independent variables and subgroups. RESULTS: The PS without inclusion of the variable for subgroups, but with inclusion of variables related to outcome, is the most appropriate. The PS should be included as a covariate in a regression model together with the variable for subgroups as covariate, where the PS is based on treatment assignment only. Larger sample sizes gave less biased results, while a higher correlation between the independent variables resulted in more biased estimates of the treatment and subgroup effect. Correlation between the independent variables and the subgroup variable did not lead to biased results. CONCLUSIONS: The results show the feasibility and validity of the PS in subgroup analyses when analyzing registry databases and indirect comparisons in economic evaluations.