to derive OS estimates for nilotinib and dasatinib of 13.0 and 13.4 years respectively (second-line CP). Using the same approach for bosutinib gives an OS of 12.8 years (third-line CP). The base-case in TA251 used a cumulative approach, where OS is equal to the duration of treatments in the pathway. If this method is applied to the TKIs in second-line CML (TA241), OS is reduced to approximately 9.4 and 10.1 years for nilotinib and dasatinib respectively (second-line CP). Similarly, bosutinib OS (third-line CP) is also reduced using this method and a substantial increase in the ICER is seen. **CONCLUSIONS:** There are methodological inconsistencies in NICE's assessments of TKIs for CML. Applying the OS methodology from TA251 to TA241 may have led to nilotinib not being recommended for routine use in the NHS. The impact of new methodologies on previous appraisal results and recommendations should be considered when assessing the validity of a new approach.

CL2

ADDRESSING HETEROGENEITY IN BASELINE RISK OF COPD EXACERBATIONS USING META-REGRESSION

Cope S[1], Buckley F[2], Baldwin M[3], Kraemer M[4], Jansen J[2]

[1]*Mapi, Toronto, ON, Canada,* [2]*Mapi, Boston, MA, USA,* [3]*Novartis Pharmaceuticals UK Limited, Horsham, West Sussex, UK,* [4]*Novartis Pharma AG, Basel, Switzerland*

OBJECTIVES: To evaluate differences across randomized controlled trials (RCTs) concerning interventions for moderate to severe chronic obstructive pulmonary disease (COPD) in terms of baseline exacerbation rates and the association with treatment effects by means of a network meta-analysis (NMA). This NMA is performed based on RCTs evaluating the long-acting bronchodilators indacaterol 75/150/300µg OD, salmeterol 50µg BID, formoterol 12µg BID, tiotropium bromide 18µg/5µg OD, and glycopyrronium bromide 50µg OD. **METHODS:** The rate of moderate or severe exacerbations was extracted from RCTs identified with a systematic literature review. A Bayesian NMA was used to synthesize the treatment effects of the different trials. The association between treatment effects and baseline exacerbation rate with placebo was assessed with a meta-regression model assuming a constant treatment-by-baseline risk interaction term. **RESULTS:** Twenty-four RCTs were included that differed mainly in terms of smoking status, COPD severity, use of inhaled corticosteroids, exacerbation definition, and exacerbation history. Across the RCTs the rate of exacerbations per patient year for patients in the placebo arm ranged from 0.40 to 1.91. Baseline risk was negatively associated with the rate ratios reflecting treatment effects across the RCTs. The coefficient for baseline risk was -0.35 (95% credible intervals: -0.49, -0.18). **CONCLUSIONS:** Based on a NMA of RCTs regarding the efficacy of long-acting bronchodilators in terms of the rate of exacerbations per patient year, baseline risk of exacerbations acts as a significant treatment effect modifier and should be accounted for in the model.

CL3

VALIDATION OF SURROGATE ENDPOINTS IN ADVANCED SOLID TUMOURS: SYSTEMATIC REVIEW OF STATISTICAL METHODS, RESULTS, AND IMPLICATIONS FOR POLICY MAKERS

Ciani O[1], Davis S[2], Tappenden P[2], Garside R[3], Stein K[1], Cantrell A[2], Saad E[4], Buyse M[5], Taylor R[1]

[1]*University of Exeter, Exeter, UK,* [2]*University of Sheffield, Sheffield, UK,* [3]*University of Exeter, Truro, UK,* [4]*Dendrix, Sao Paulo, Brazil,* [5]*IDDI, Louvain-la-Neuve, Belgium*

OBJECTIVES: Licensing and reimbursement of anticancer drugs should rely on evidence from patient-relevant endpoints such as overall survival (OS). Nevertheless, evidence from surrogate endpoints may also be useful, as it may expedite the regulatory approval and coverage decisions of new therapies. It is therefore essential that candidate surrogate endpoints be properly validated. However, there is no consensus on statistical methods for such validation and on how the evidence thus derived should be applied by policy makers. **METHODS:** We review meta-analyses of therapeutic interventions against advanced solid tumours published until December 2012 that quantified the statistical association between progression-free survival (PFS) or time-to-progression (TTP) and OS. We assessed the suitability of the two surrogates using three current surrogate validation frameworks: Bucher's framework, the German Institute of Quality and Efficiency in Health Care's (IQWiG) framework and the Biomarker-Surrogacy Evaluation Schema (BSES3). **RESULTS:** Thirty-one meta-analyses were included which employed a variety of statistical methods to assess surrogate validity. The strength of the association between PFS or TTP and OS was generally low. The level of evidence (observation-level vs. treatment-level) available supporting an association between PFS or TTP and OS varied considerably by cancer type, by evaluation tools and was not always consistent even within one specific cancer type. **CONCLUSIONS:** Not in all solid tumours the treatment-level association between PFS or TTP and OS has been investigated. According to the IQWiG's framework, only PFS achieved acceptable evidence of surrogacy in metastatic colorectal and ovarian cancer treated with cytotoxic agents, whereas in no indication did the two candidate endpoints achieve good evidence of surrogacy according to BSES3. Our study emphasises the challenges of surrogate-endpoint validation and the importance of building consensus on appropriate statistical techniques to examine surrogacy and on the development of evaluation frameworks for policy makers.

CL4

FRAMEWORK FOR EVIDENCE ASSESSMENT BASED ON GRADE AND APPLICATION TO HPV VACCINATION IN MALES IN THE EUROPEAN HEALTH CARE CONTEXT

Siebert U[1], Sroczynski G[2], Baker P[3], Borget I[4], Castellsagué X[5], Chapman R[6], von Knebel-Doeberitz M[7], Mortensen GL[8], La Torre G[9]

[1]*UMIT – University for Health Sciences, Medical Informatics and Technology / ONCOTYROL / Harvard University, Hall i. T./ Innsbruck / Boston, Austria,* [2]*UMIT – University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria,* [3]*Men's Health Consultant, Brighton, UK,* [4]*Institut Gustave Roussy, Villejuif, France,* [5]*L'Hospitalet de Llobregat, Barcelona, Spain,* [6]*Evidera, London, UK,* [7]*University of Heidelberg, German Cancer Research Center, Heidelberg, Germany,* [8]*AnthroConsult, Aarhus C, Denmark,* [9]*Sapienza University of Rome, Rome, Italy*

OBJECTIVES: To develop and apply an extended framework for evidence assessment based on the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach using the example of male human papilloma virus (HPV) vaccination Europe. **METHODS:** A pan-European multidisciplinary expert group was established to develop an extended GRADE framework that includes explicit assessment of cost-effectiveness, medical needs, and patient aspects, ethical and social issues. Using an expert panel process, we assessed the feasibility of using this framework by applying it to male HPV vaccination in Europe. Studies were assessed using the specific framework tools; results and feasibility were discussed; and consensus was achieved through a modified Delphi method. **RESULTS:** We identified three advisory committees (ACIP/USA; NACI/Canada; STIKO/Germany) using GRADE for vaccines assessment. Institutions handled data beyond vaccine efficacy and safety differently and did not formally grade economic evidence. We adopted the grading methodology of ACIP for the key factor 'Benefits and Harms' and developed modules for grading evidence type and quality of economic evaluations ('Economic Evaluation') and for systematically assessing epidemiology, disease burden and unmet medical needs, as well as ethical, social and patient aspects ('Values and Preferences'). The feasibility test demonstrated that all framework components were feasible in the case of HPV vaccination. Overall evidence type for cost-effectiveness was low with uncertainty in results. Cost-effectiveness was best, when all HPV-related diseases and outcomes were included and when assuming low coverage in females and lower vaccine prices. **CONCLUSIONS:** The GRADE approach is applicable in assessing vaccinations and was successfully applied to HPV vaccination in males. The assessment of benefits and harms can be extended by explicit assessment of the evidence on cost-effectiveness and other key factors including unmet medical needs, and ethical, social and patient aspects. This extended framework can better inform policy- and decison makers.

CARDIOVASCULAR DISEASE OUTCOMES RESEARCH STUDIES

CV1

CHALLENGES IN MODELLING THE COST EFFECTIVENESS OF INTERVENTIONS IN CARDIOVASCULAR DISEASE

Burgers LT, Redekop WK, Severens JL

Erasmus University Rotterdam, Rotterdam, The Netherlands

OBJECTIVES: Modelling is essential in performing economic evaluations for various reasons. For example, modelling is necessary if extrapolation of short-term or intermediate results to long-term outcomes is required and numerous strategies need to be evaluated without direct evidence. However, modelling inherently poses challenges that need to be dealt with since models always represent a simplification of reality. The aim of this study is to identify and analyse the challenges in modelling the cost-effectiveness of cardiovascular disease (CVD) interventions. **METHODS:** A questionnaire was sent to 40 corresponding authors (systematically selected) of recent model-based economic evaluations of CVD interventions published in high-impact cardiovascular, health economics and general medical journals. Respondents were asked to provide their own challenges and also rank the importance of challenges identified using a pilot version of the questionnaire distributed to 7 experienced researchers. Furthermore, we analysed the discussion sections of the papers to identify unmentioned challenges. Solutions, if available, were based on input from the respondents and the recommendations of the ISPOR-SMDM task force. **RESULTS:** The systematic literature search identified 1720 potentially relevant articles. The limit of 40 authors was reached after screening 294 titles and abstracts. Beside the challenge of lack of data, preliminary results show that it was difficult to obtain a sufficiently valid, precise and accurate cost-effectiveness estimate due, for example, to interrelating clinical outcomes or extrapolating from surrogate outcomes. Both challenges often exist in CEAs evaluating CVD prevention strategies. **CONCLUSIONS:** The preliminary results of this study showed examples of CVD modelling challenges encountered during studies published in high-impact journals. Modelling guidelines do not provide sufficient assistance in resolving all challenges but it is probably unrealistic to expect this. Some of the reported challenges are specific to the type of intervention and disease, but most challenges are present in all types of interventions and diseases.

CV2

APPLICATION OF BEHAVIOURAL ECONOMICS TO THE UNDERSTANDING OF ADHERENCE: DOES AN INDIVIDUAL'S TIME PREFERENCE INFLUENCE ADHERENCE TO MEDICATIONS?

Fargher EA, Morrison V, Plumpton CO, Hughes DA

Bangor University, Bangor, UK

OBJECTIVES: There is general support that individual time preference affects health-related behaviours. People with a high, positive time preference value their immediate health higher than future health, even if presented with extreme scenarios of intertemporal choice. We hypothesised that adherence to medication requires trade-offs between immediate and delayed health benefits. Patients with lower time preference rates may be more adherent to medication as they place a higher value on the future benefits of adherence. **METHODS:** Hypertensive adult patients across Europe were invited to complete a web-based survey that had been translated and piloted. Patients' time preference was assessed (4-items) to calculate individual discount rates (%) in both short term (3-years) and medium term (6-years). Medication adherence was measured using the Morisky questionnaire (primary analysis) and the Medication Adherence Report Scale (MARS, secondary analysis). Sample size calculation, based on 5% one-sided confidence, assuming 30% non-adherence with Morisky measure indicated n=323 per country. Missing data were imputed using multiple imputation in STATA. The significance of the association with adherence was assessed using the Wald test statistic. **RESULTS:** 969 patients completed the questionnaire across England, Wales and Hungary, 79% of possible responses were observed. Short and medium term time preference rates in England, Wales and Hungary were in the expected directions, but the relationship was not statistically significant. Based on Morisky adherence - Wales (short): adherent 8.7%, non-adherent 9.4% (p=0.541); (medium): adherent 4.7%, non-adherent 5.0% (p=0.611). England (short): adherent 7.8%, non-adherent 9.5% (p=0.163); (medium): adherent 3.7%, non-adherent 4.5% (p=0.095). Hungary (short): adherent 19.0%, non-adherent 18.2% (p=0.504); (medium): adherent 8.9%, non-adherent 8.6%