multiple courses of TACE is difficult to ascertain since additional courses may be prescribed under a patient-specific treatment protocol or due to non-optimal tumor response. Nonetheless, mean survival after discontinuing TACE was relatively similar regardless of number of treatments received.

PODIUM SESSION I: CONCEPTUAL PAPERS

CP1 ADJUSTING FOR INFLATION IN ECONOMIC EVALUATIONS OF HEALTH TECHNOLOGIES: ARE WE DOING IT WRONG?
Paulden M
University of Toronto, Toronto, ON, Canada

OBJECTIVES: Economic evaluations of health technologies typically require consideration of the time value of money. Conventionally, all costs are represented in 'real terms' by adjusting for inflation. Future costs are then discounted to account for time preference. Although much has been written on the practice of discounting, health economists have paid surprisingly little attention to the issue of appropriately adjusting for inflation. This paper argues that the conventional approach to adjusting for inflation in economic evaluations of health technologies is inappropriate. METHODS & RESULTS: The conventional approach follows the recommendations of the Washington Panel. Costs must be converted to "constant dollars" using a single inflation rate representing the rate of "general price inflation". However, "if the prices of the goods in question change at a rate different from general price levels, this variation should be reflected in the adjustment used". Some analyses therefore use the "Medical Component" of the Consumer Price Index (CPI), or an equivalent measure. Because of this relative similarity within language groups, the conventional approach to be appropriate requires that all costs change at the same rate over time. This is generally not the case - some costs may rise (e.g. pharmaceuticals) at the same time as other costs fall (e.g. personal computers). In particular, products losing patent protection will experience a sudden fall in price out-of-line with general price inflation. A solution is to assign each cost a unique time profile subject to specific market conditions. Rather than applying an inflation rate, future costs are instead estimated using a unique projection model for each cost component. CONCLUSIONS: The conventional approach to adjusting for inflation is inappropriate. A solution to this problem is to select a unique time profile for each cost component. Models routinely used by financial analysts may provide an example for how this projection can be done in practice.

CP2 THE NOTION OF REPRESENTATIVE LANGUAGES IN THE CONTEXT OF TRANSLATABILITY ASSESSMENT
Basse SJ, Martin ML, McCarrier KP
Health Research Associates, Inc., Seattle, WA, USA

BACKGROUND: While current best-practices in PRO development include evaluation of the relative ease of translation for global trial use prior to instrument finalization, methodologies for this translatability assessment (TA) vary greatly. In the proposed approach, representative languages (RLs) are selected to assess the translation difficulty of PRO concepts without the time and cost of evaluating multiple languages with shared characteristics. METHODS: In the genealogical approach employed by linguists, languages sharing a common ancestor that became separated by geographical or socio-political boundaries will evolve in distinct ways, resulting in sets of languages (families) with common linguistic features (e.g. word order, phrasal structure, morphology, lexical, syntactic features). Because of this relative similarity within language groups, efficiency can be gained by assessing translatability with sets of appropriately selected representative languages, which can in turn predict translation problems likely to affect others in their linguistic families. As such, use of appropriate selection is of key importance. RLs should be based both on linguistic properties and other features salient to outcomes research. A family or group of languages may also be defined by shared characteristics that are not purely linguistic in nature. Features such as geographic and cultural (religious/dietary/social) aspects, number and distribution of speakers worldwide, and criteria related to health care utilization or study implementation should be considered in the definition of language families/groups and in the selection of RLs. CONCLUSIONS: Despite differences that undeniably exist between individual languages, limited information can be gained by the repetitive assessment of prospective translation difficulty within groups having similar characteristics. Instead, the use of a representative language to assess translation difficulty for a related group of languages provides greater resource efficiency and more effective application of TA in providing important feedback prior to finalization of newly developed measures.

CP3 PIECEWISE MODELING OF TIME-TO-EVENT DATA WITH FLEXIBLE PARAMETERIZATION OF COVARIATES AND EFFECTS
Ishak EI
UnitedBioSource Corporation, Dorval, QC, Canada

Projection of time-to-event distributions is necessary to obtain accurate estimation of life expectancy, or prediction of event times for economic models. Parametric survival analysis techniques are typically used, and can represent a broad range of shapes. In some cases, however, the best distributional fit fails to capture the variation in hazards over the entire time span, or it provides acceptable fit to the data but yields clinically implausible projections (e.g., constant hazard of death). More flexible techniques, like piecewise exponential models, can overcome these issues but remain generally underused. In piecewise models, the time axis is divided into contiguous segments with a common parametric distribution assumed within each segment, but values of the parameters are allowed to vary. In addition to greater flexibility, this framework allows inclusion of time-dependent parameters and time-dependent effects. Two important considerations are the number and placement of divisions on the time axis, and the choice of the common distribution. Examination of the cumulative and log-cumulative hazard plots can assist with these issues. For instance, the number/placement of divisions for a piecewise-exponential model could be determined visually such that the points within each division of the cumulative hazard plot follow a linear pattern. The same can be done with log-cumulative hazard function for a piecewise-Weibull model. Although piecewise-exponential models can be made progressively more flexible by increasing the number of segments to capture even very complex patterns, the hazard for the last segment can be limiting for selection. Thus, models based on Weibull distributions may be more appropriate, and possibly achieve similar fit with fewer segments. The overall selection in these cases involves using optimization strategies (e.g., grid search for placement of divisions) and use of fit statistics to select distributions.

CP4 MEASURING HEALTH OUTCOMES IN THE ABSENCE OF RIGOUR: WILLFUL IGNORANCE OR DELIBERATE MALPRACTICE?
Finn P
University of Leeds, Leeds, UK

RESULTS: In economic evaluations of health technologies undertook by regulatory agencies across the world. As a consequence, health economists now play a critical part in generating the evidence base used to determine both access to and the price of treatment. No matter the complexity of any evaluation, every project needs to describe and value the benefits of health care interventions. The computation of an ICER depends totally on the capacity to quantify marginal changes in health outcomes. The orthodoxy of using QALYs as the outcome of capturing such outcomes via the use of generic health status measurement systems (for example HUI or EQ-5D) together with their corresponding social preference weights. The requirement that the values of the general population constitute the "correct" perspective is one element of the health economics creed. A second dictates that the "worth" of a health outcome shall be expressed in terms of utility - a concept that lacks a defined unit of measure or any agreed standard evaluation method. A regrettable fit to the data is a "reasonable" approach. Health economists recognise the non-commensurability of Standard Gamble (SG) and Time Trade-Off (TTO) methods their public posture generally belies this contradiction. The status of the QALY as a useful metric of health benefit/loss has been fatally compromised by the failure of the scientific community to agree on a single method for determining the quality-adjustment factor. The preparedness of health economists to ignore this gap in their armamentarium runs counter to the rational practice of science. This paper challenges the intellectual deadweight of traditional health economics, specifically in regards to the measurement of health outcomes. Examples of defective practice drawn from Canadian and UK HTA reports will be used to illustrate the conceptual issues raised in this paper.
challenged by payers and regulatory authorities to develop evidence describing the burden of illness and justifying the payer investment.

HC2 HEALTH CARE EXPENDITURES AND OUTCOME AMONG ELDERLY CANCER PATIENTS WITH CANCER

Eli Lilly and Company USA /Australia, West Ryde, Australia

OBJECTIVES: Determine the relationship between health care expenditures and outcomes with cancer, compared to those with cancer and depression, after controlling for demographic, socio-economic, access to care and other health status variables. METHODS: Cross-sectional data on 4766 adults from multiple years (2006, 2007, 2008, and 2009) of the nationally representative Medicare household survey, Medical Expenditure Panel Survey (MEPS) were used. Cancer and depression was identified from Medical conditions file. Dependent variables consisted of total, inpatient, outpatient, emergency room, prescription drugs and other expenditures. OLS on logged dollars and generalized linear models with log-link were performed. All analyses accounted for the complex survey design of the MEPS. RESULTS: Overall, 14% of individuals with cancer reported having depression. Among individuals with cancer and depression the average health care expenditures were $18,401 compared to $12,091 among those without depression. After adjusting for demographic, socio-economic, access to care and other health status variables, those with depression had about 32% greater total expenditures compared to those without depression. Expenditures for every type were higher among individuals with depression compared to those without depression. Individuals with cancer and depression were more significantly more likely to use hospital rooms (AOR = 1.46) and prescription drugs (AOR = 3.56) compared to their counterparts without depression. CONCLUSIONS: Among adults with cancer, those with depression had higher health care utilization and expenditures compared to those without depression. Policy efforts to reduce excess health care expenditures associated with depression may include screening for depressive symptoms and preventing major depression, timely depression treatment once depression is detected.

HC3 DISCREPANCIES BETWEEN FDA APPROVAL AND CMS COVERAGE FOR DRUGS AND DEVICES

Chambers BJ, Neumann PJ
Tufts Medical Center, Boston, MA, USA

OBJECTIVES: Following FDA approval, medical technology must still gain Centers for Medicare and Medicaid Services (CMS) coverage before it can be reimbursed. However, the two agencies use approval processes based on different evidentiary standards. We identified the type and nature of discrepancies between FDA approval and CMS national coverage determinations (NCDs) for drugs and devices. METHODS: We used the Tufts Medical Center NCD database, which contains detailed information on 165 NCDs since 1999. For each device or Part B drug considered in an NCD (1999-2011) (n=69), we searched the FDA website to identify the approved indication. We classified CMS coverage as: more restrictive than FDA approval, i.e., conditions were placed on coverage beyond the FDA-approved label, equivalent to FDA approval, or less restrictive than FDA approval, i.e., CMS covers off-label indications. Further, we categorized conditions on CMS coverage as: "patient-related", e.g., restricted to patients with certain comorbidities or characteristics; "place in therapy", e.g., tied to use as second-line therapy; or "technology-related", e.g., restricted to a particular application of the drug or device. RESULTS: CMS has covered FDA-approved drugs or devices taken through the Medicare NCD process in 80% of cases (55/69). For CMS covered drugs and devices (n=55), coverage was more restrictive in 32 cases (58%), equivalent to FDA approval in 8 cases (14%), and less restrictive in 7 cases (13%). Most common coverage restrictions were patient-related (78%), e.g., laparoscopic gastric banding to treat obesity is covered for patients suffering from an obesity-related comorbidity, and place in therapy (38%), e.g., coverage for extracorporeal immunoadsorption is covered for rheumatoid arthritis patients who have failed three disease-modifying antirheumatic drugs (DMARDs). In roughly one third of cases, CMS placed multiple restrictions on coverage. CONCLUSIONS: CMS coverage determinations are generally more restrictive than corresponding FDA approval. CMS often restricts coverage to patients with the most severe disease.

HC4 HOSPITALIZATION COSTS AND OUTCOMES AMONG ELDERLY CANCER PATIENTS IN THE UNITED STATES

Eli Lilly and Company, Indianapolis, IN, USA, Eli Lilly and Company USA /Australia, West Ryde, Australia

OBJECTIVES: To assess the patient-, hospital-, and discharge-level characteristics associated with hospitalization among elderly cancer patients in the United States (US). Hospitalization outcomes (length of stay [LOS], total charges, and mortality) among elderly patients with cancer were also studied. METHODS: A cross-sectional descriptive analysis of the 2009 Healthcare Cost and Utilization Project (HCUP) database was conducted. Patients were identified based on diagnosis (any-listed) of cancer using Clinical Classification Software (CCS). A control group of patients without cancer were identified by matching on age and gender. The study sample was conducted using FREUD-SURF procedures in SAS v9.2. RESULTS: In 2009, a total of 3,325,174 (weighted) hospitalizations occurred among elderly patients with cancer in the US. Elderly cancer patients had higher total hospital charges ($39,406 vs. $37,756), longer LOS (5.7 days vs. 5.4 days), and higher mortality (4.8% vs. 3.6%) as compared to those without cancer. A greater proportion of hospitalizations among cancer patients occurred in teaching hospitals (44.1% vs. 38.9%, p<0.001). In terms of location, a greater proportion of hospitalizations for cancer patients occurred in hospitals located in urban areas in comparison to those without cancer (88.1% vs. 84.7%, p<0.001). Total charges for hospitalizations among elderly patients with prostate cancer (average LOS=4.9 days), lung cancer (average LOS=6.1 days), and breast cancer (average LOS=4.7 days) were roughly $16.1, $16.0, and $16.0 billion, respectively. Mortality rates during hospitalization were the highest for those with pancreatic (10%), liver (9.7%), and lung cancer (8.7%). CONCLUSIONS: Elderly patients with cancer had significantly greater hospitalization burden as compared to those without cancer. Hospital mortality rates were the highest for elderly patients with pancreatic, liver, and lung cancer, respectively.

POD I: HEALTH TECHNOLOGY ASSESSMENT STUDIES

HT1 INTER-COUNTRY VARIABILITY IN COVERAGE DECISIONS FOR ORPHAN DRUGS: CRITERIA DRIVING HTA RECOMMENDATIONS IN SIX COUNTRIES

Nicola F, Karavos P
London School of Economics and Political Science, London, UK

Inter-country variability in access to orphan drugs across countries has been highlighted in a number of studies. Understanding the reasons driving coverage decisions is a way forward in identifying areas where HTA methods may be improved. OBJECTIVES: Objectives are three-fold: a) to establish a methodology and framework for systematic comparison of HTA processes across countries; b) to identify the criteria driving HTA recommendations for a sample of orphan drugs; and, c) to understand the reasons for diverging recommendations and propose ways to minimize these. METHODS: All common orphan drug-indication pairs were identified in six countries (England, Scotland, France, Sweden, Canada and Australia) between 2001 and 2012 were selected. Agreement levels in HTA outcomes between countries were measured using Cohen’s kappa scores. Thematic analysis, by creating an NIH-9 coding manual, was conducted to systematically compare each compound. Reasons for diverging HTA outcomes were differentiated based on whether they are a consequence of country-specific considerations or of the HTA process, and ranked by frequency of occurrences. RESULTS: Fourteen orphan drug-indication pairs were retrieved. Agreement in HTA outcomes was poor (κ = [0.5, 0.3]). Eight drug-indication pairs appraised by at least four HTA bodies were analyzed, five of which received positive outcomes. Preliminary results suggest that in four of five cases, reasons for diverging recommendations were a consequence of the HTA process. Examples of non-homogeneous assessments include: lack of appropriate primary endpoint, lack of long-term data, evidence not reflecting clinical practice, orphan status or unmet clinical need. CONCLUSIONS: Preliminary results identify the criteria driving the assessments and reasons why they result in diverging HTA outcomes, enabling a better understanding of these processes by elucidating the expectations and value judgments from HTA bodies, particularly on the orphan status, and identifying areas where more consensus on what constitutes appropriate HTA methodologies is needed. Final results will quantify these criteria in a systematic manner.

HT2 AGENCY AGREEMENT IN HEALTH TECHNOLOGY ASSESSMENT STUDIES: SUBMISSIONS AND DECISIONS

Jaksa A, Demb J, Ho YS, Daniel K
Context Matters, Inc., New York, NY, USA

OBJECTIVES: HTA agencies often review the same drugs for the same/similar indications. How often do agencies agree on their reimbursement decisions? Previous research has compared reimbursement recommendations for the same drugs) for a limited number of agencies, but studies have rarely focused on more than 2 agencies. We collect and analyze a large number of health technology assessments from several countries to explore how often the agencies agree on their reimbursement decisions. METHODS: The data covered five agencies that make reimbursement decisions: NICE, SMIC, PBAC, HAS and CADTH’s Common Drug Review. Our analysis only included decisions for drugs that were reviewed by at least two agencies, producing a total of 195 reviews. RESULTS: There was generally a high level of agreement among all pairs of agencies, ranging from 56% (PBAC; CADTH) to 91% (NICE; HAS). It is important to note that within the sample of drugs reviewed, all agencies issued positive recommendations at very high rates – all but CADTH issued positive recommendations for more than 80% of the drugs reviewed. This fact alone would produce high levels of agreement, even if agencies’ recommendations were statistically independent. Actual agreement rates observed were close to those implied by independence. CONCLUSIONS: Agencies agree on their reimbursement decisions quite often, but at rates close to those implied by their high overall positive recommendation rates alone. Our analysis highlights the need to focus on identifying the determinants of agencies’ high rates of agreement.

HT3 ISSUES IN THE SELECTION OF COMPARATORS FOR REGULATORY AND HTA SUBMISSIONS

Delazer AM1, Murray JP2, Klopcin MN1, Johnston JA1, Grainger D2
1 Eli Lilly and Company, Indianapolis, IN, USA, 2 Eli Lilly and Company USA /Australia, West Ryde, Australia

There was generally a high level of agreement between all pairs of agencies, ranging from 56% (PBAC; CADTH) to 91% (NICE; HAS). It is important to note that within the sample of drugs reviewed, all agencies issued positive recommendations at very high rates – all but CADTH issued positive recommendations for more than 80% of the drugs reviewed. This fact alone would produce high levels of agreement, even if agencies’ recommendations were statistically independent. Actual agreement rates observed were close to those implied by independence. CONCLUSIONS: Agencies agree on their reimbursement decisions quite often, but at rates close to those implied by their high overall positive recommendation rates alone. Our analysis highlights the need to focus on identifying the determinants of agencies’ high rates of agreement.