17th Biennial European Meeting of the Society for Medical Decision Making Leiden, the Netherlands, June 10–12, 2018

PS1-1 IMPROVING SELECTION OF PATIENTS FOR ENDOVASCULAR TREATMENT OF ACUTE ISCHEMIC STROKE: EXTERNAL VALIDATION OF A CLINICAL DECISION TOOL IN SIX RANDOMIZED CONTROLLED TRIALS

Health Services, Outcomes and Policy Research (HSOP)

Esmee Venema, MD[1], Maxim J.H.L. Mulder, MD, PhD[2], Bob Roozenbeek, MD, PhD[2], Scott Brown, PhD[3], Charles B.M.J. Majoie, MD, PhD[4], Andrew M. Demchuk, MD, PhD[5], Keith W. Muir, MD, PhD[6], Antoni Davalos, MD, PhD[7], Peter J. Mitchell, MD, PhD[8], Serge Bracard, MD, PhD[9], Michael Hill, MD, PhD[5], Philip White, MD, PhD[10], Bruce C.V. Campbell, MD, PhD[11], Francis Guillemin, MD, PhD[12], Jeffrey L. Saver, MD, PhD[13], Tudor G. Jovin, MD, PhD[14], Mayank Goyal, MD, PhD[5], Aad van der Lugt, MD, PhD[15], Ewout W. Steyerberg, PhD[16], Diederik W.J. Dippel, MD, PhD[2] and Hester F. Lingsma, PhD[1], (1)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (2)Department of Neurology, Erasmus University Medical Center, Rotterdam, Netherlands, (3)Altair Biostatistics, St Louis Park, MN, (4)Department of Radiology, Academic Medical Center, Amsterdam, Netherlands, (5)Departments of Clinical Neuroscience and Radiology, Hotchkiss Brain Institute, Cummings School of Medicine, University of Calgary, Calgary, AB, Canada, (6)Institute of Neuroscience & Psychology, University of Glasgow, Queen Elizabeth University Hospital, Glasgow, United Kingdom, (7)Department of Neuroscience, Hospital Germans Trias y Pujol, Barcelona, Spain, (8)Department of Radiology, Royal Melbourne Hospital, University of Melbourne, Melbourne, Australia, (9)Department of Diagnostic and Interventional Neuroradiology, University of Lorraine and University Hospital of Nancy, Nancy, France, (10)Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, United Kingdom, (11)Department of Medicine and Neurology, Melbourne Brain Centre Royal Melbourne Hospital, University of Melbourne, Melbourne, Australia, (12)Department of Clinical Epidemiology, Inserm, University of Lorraine and University Hospital, Nancy, France, (13)Department of Neurology and Comprehensive Stroke Center, David Geffen School of Medicine, University of Los Angeles, Los Angeles, CA, (14)Stroke Institute, Department of Neurology, University of Pittsburgh Medical Center Stroke Institute, Presbyterian University Hospital, Pittsburgh, PA, (15)Deptartment of Radiology, Erasmus University Medical Center, Rotterdam, Netherlands, (16)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands; Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, Netherlands

Purpose:

Benefit of endovascular treatment (EVT) varies between individual patients with acute ischemic stroke. The MR PREDICTS decision tool, previously developed in data from the MR CLEAN trial, predicts outcome with and without EVT based on patient and imaging characteristics. We externally validated this model with data from recent trials, with the ultimate aim to improve selection of patients for EVT.

Method(s):

Individual patient data was derived from six randomized controlled trials within the HERMES collaboration (ESCAPE, REVASCAT, SWIFT-PRIME, EXTEND-IA, THRACE and PISTE). Outcome of the ordinal logistic regression model was the modified Rankin Scale (mRS) score at 90 days after stroke. Treatment benefit was defined as the difference between the predicted probability of functional independence (mRS score 0-2) with and without EVT. Model performance was evaluated according to discrimination (measured with the c-statistic) and calibration. All data from the derivation and validation cohort was combined to update the model coefficients.

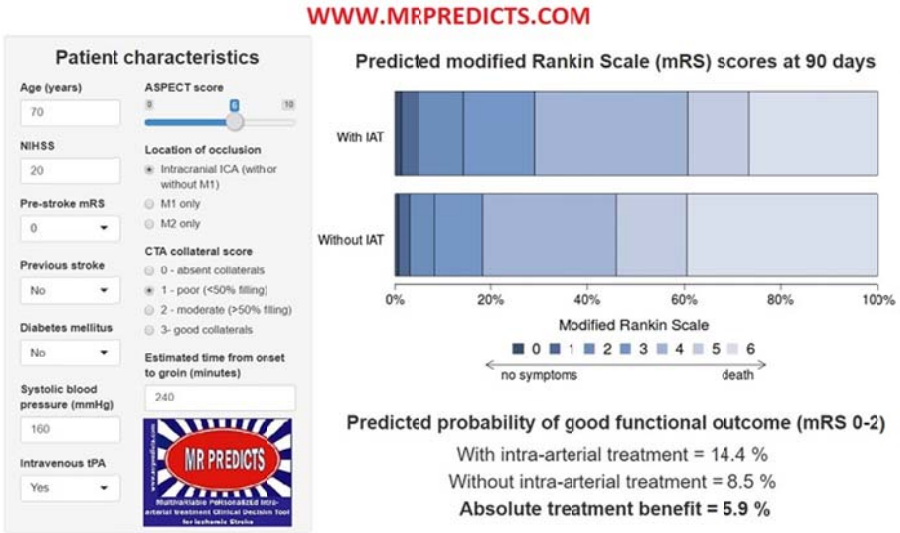
Result(s):

We included 1243 patients in the validation cohort (633 assigned to EVT, 610 assigned to control). The c-statistic was 0.67 (95% confidence interval [CI] 0.65-0.69) for the ordinal mRS and 0.73 (95% CI 0.70-0.76) for functional independence, similar to previous performance. Outcomes were systematically better than predicted (calibration slope 0.89 and intercept 0.52). The observed probability of functional independence was higher than predicted for both treated patients (35% vs 26%) and controls (54% vs 40%), but the observed treatment benefit was similar (19% and 14%). Figure 1 shows a screenshot of the decision tool for use in clinical practice.

Conclusion(s):

MR PREDICTS predicted outcome in a large heterogeneous trial population with discriminative value comparable to other well-known prediction tools. The updated model might be used to support clinical decision making in ischemic stroke by selection of patients for EVT.

Figure 1. Screenshot from the MR PREDICTS decision tool



PS1-2 THE IMPACT OF THE RISING COLORECTAL CANCER INCIDENCE IN YOUNG ADULTS ON THE OPTIMAL AGE TO START SCREENING IN THE US: A MICROSIMULATION ANALYSIS

Health Services, Outcomes and Policy Research (HSOP)

Elisabeth Peterse, MSc[1], Reinier Meester, PhD[1], Rebecca Siegel, MPH[2], Jennifer Chen, MPH[3], Andrea Dwyer, BS[4], Dennis Ahnen, PhD[5], Robert Smith, PhD[2], Ann Zauber, PhD[3] and Iris Lansdorp-Vogelaar, PhD[1], (1)Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, Netherlands, (2)American Cancer Society, Atlanta, GA, (3)Memorial Sloan Kettering Cancer Center, Department of Epidemiology & Biostatistics, New York, NY, (4)University of Colorado Cancer Center, Denver, CO, (5)University of Colorado Cancer Center and Gastroenterology of the Rockies, Denver, CO

Purpose: In 2016, the MISCAN-Colon model was used to inform the US Preventive Services Task Force (USPSTF) colorectal cancer (CRC) screening guidelines, which recommend screening from ages 50 to 75 years for average risk individuals. However, these models did not take into account the increase in CRC incidence below the age of 50 years. In this study, we re-evaluated the optimal age to start screening, age to end screening and screening interval in light of the increase in CRC incidence observed in young adults.

Method(s): We adjusted the simulated lifetime CRC incidence in the MISCAN-Colon model to reflect the observed increase in young onset incidence. In line with the strong birth cohort effect, the current generation of 40-year-olds was assumed to carry forward escalated disease risk as they age. Life-years gained (benefit), the number of colonoscopies (burden) and the ratios of incremental burden to benefit (efficiency ratio) were projected for different screening strategies. Strategies differed with respect to test modality, ages to start screening (40, 45, 50), ages to stop screening (75, 80, 85), and screening intervals (depending on screening modality). We then determined the model-recommended strategies in a similar way as we did for the USPSTF, using similar efficiency ratio thresholds to the previously accepted efficiency ratio of 39 incremental colonoscopies per life-year gained.

Result(s): The life-years gained and the number of colonoscopies for each colonoscopy strategy are plotted in Figure 1. Because of the higher CRC incidence, model-predicted life-years gained from screening increased compared to our previous analyses for the USPSTF. Consequently, the balance of burden to benefit of screening improved, with colonoscopy screening every 10 years starting at age 45 years resulting in an efficiency ratio of 32 incremental colonoscopies per life-year gained.

Conclusion(s): This decision-analytic modeling approach suggests that based on the increase in young-onset CRC incidence, screening initiation at age 45 years has a favorable balance between screening benefits and burden. Screening until age 75 years with colonoscopy every 10 years, fecal immunochemical testing annually, flexible sigmoidoscopy every 5 years, and computed tomographic colonography every 5 years was recommended by the model as these strategies provided similar life-years gained at an acceptable screening burden.

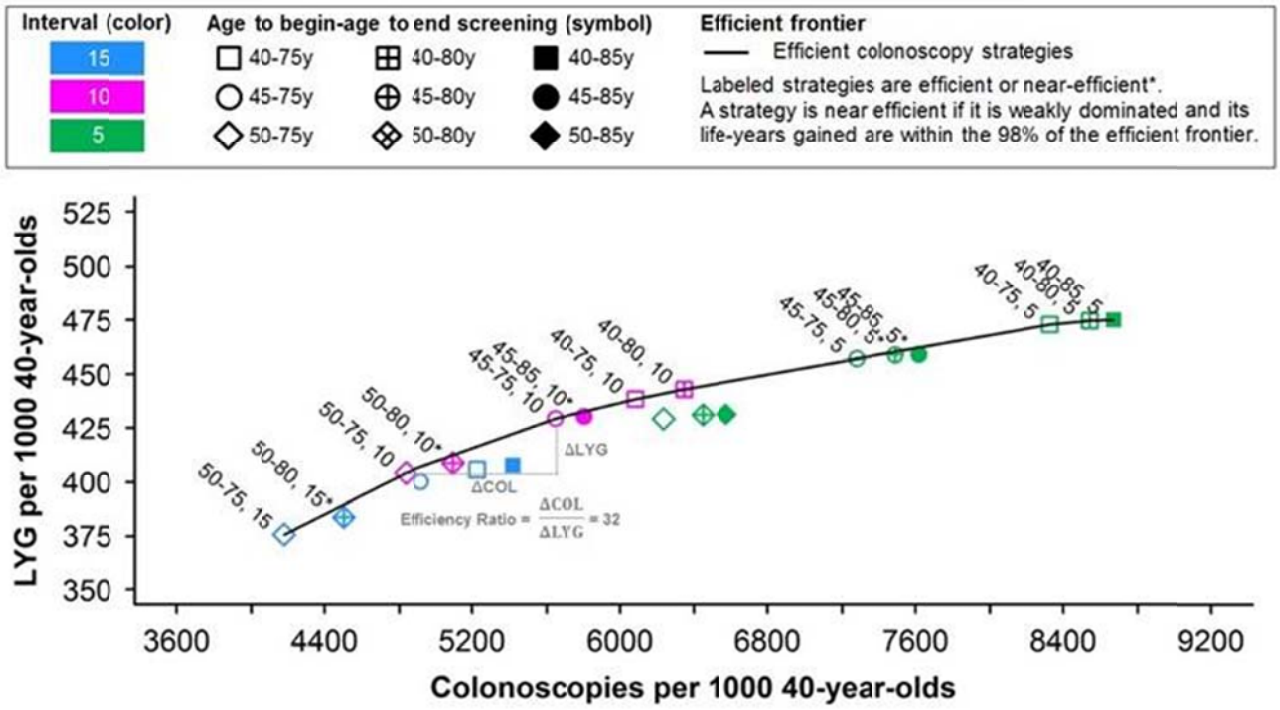


Figure 1. Lifetime number of colonoscopies and life-years gained for the colonoscopy screening strategies evaluated by MISCAN-Colon.

PS1-3 NATURE AND REPORTING CHARACTERISTICS OF SYSTEMATIC REVIEWS FOR UK HEALTH TECHNOLOGY ASSESSMENT: A CROSS-SECTIONAL STUDY

Health Services, Outcomes and Policy Research (HSOP)

Christopher Carroll, PhD, MSc, MA, BA and Eva Kaltenthaler, University of Sheffield, Sheffield, United Kingdom

Purpose: A recent study claimed that increasing numbers of reviews are being published and many are poorly-conducted and reported.[1] Given the role of systematic reviews in healthcare decision-making, especially HTA reviews, the aim of the present study was to assess how well reporting standards in systematic reviews published in 2014 in the UK Health Technology Assessment monograph series compare with the reporting in Cochrane and other "non-Cochrane" systematic reviews from the same year, as reported elsewhere.[1]

Method(s): All relevant UK HTA programme systematic reviews published in 2014 were identified. After piloting of the form, two reviewers each extracted relevant data on conduct and reporting from these reviews. These data were compared with data for Cochrane and "non-Cochrane" systematic reviews from 2014, as published elsewhere.[1] All data were tabulated and summarised.

Result(s): There were 30 UK HTA programme systematic reviews and 300 other systematic reviews, including Cochrane reviews (n=45). Fewer UK HTA reviews covered therapeutic and pharmaceutical topics (53% and 20% respectively) than Cochrane (100% and 51%). The percentage of HTA reviews with required elements of conduct and reporting was frequently very similar to Cochrane and much higher than all other systematic reviews, e.g. availability of protocols (90% for HTAs, 98% for Cochrane and 16% for other non-Cochrane reviews, respectively); the specification of study design criteria (100%, 100%, 79%); the reporting of outcomes (100%, 100%, 78%), quality assessment (100%, 100%, 70%) and other processes; the searching of trial registries for unpublished data (70%, 62%, 19%); reporting of reasons for excluding studies (91%, 91% and 70%) and reporting of authors' conflicts of interests (100%, 100%, 87%). However, HTA reviews compare less favourably with Cochrane and other reviews in the assessment of publication bias.

Conclusion(s): UK HTA systematic reviews are often produced within a specific policy-making context and cover a greater variety of topics than Cochrane reviews. This has implications for timelines, tools and resources. However, they still tend to present standards of conduct and reporting equivalent to "gold standard" Cochrane reviews and superior to systematic reviews more generally. This underlines how systematic reviews produced in the context of HTA, and healthcare decision-making, achieve standards of conduct and reporting not required by many other reviews.

[1] Page et al: Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study, PLOS Medicine, 2016; 13(5): e1002028.

PS1-4 THE REPORTING AND IMPACT OF SURGEONS' EXPERTISE IN RANDOMIZED CONTROLLED TRIALS OF TOTAL HIP AND TOTAL KNEE ARTHROPLASTY: A SYSTEMATIC REVIEW

Health Services, Outcomes and Policy Research (HSOP)

Christopher Carroll, PhD, MSc, MA, BA and Faizan Mahmood, BA, University of Sheffield, Sheffield, United Kingdom

Purpose: The aim of this work was to review how surgeons' expertise, which is a potential confounder of outcomes, is reported in randomized controlled trials (RCTs).

Method(s): A systematic review of surgeons' expertise as reported in RCTs comparing total hip arthroplasty (THA) with hemiarthroplasty (HA), and comparing the different total knee arthroplasty (TKA) techniques of standard parapatellar (MP), midvastus (MV), subvastus (SV) and quadriceps sparing (QS). We searched MEDLINE, EMBASE, Science Citation Index, The Cochrane Library, and Conference Proceedings Citation Index-Science (CPCI-S); reference checking was also performed. Study selection, data extraction, critical appraisal (Cochrane risk of bias tool) and meta-analysis were undertaken by two reviewers.

Result(s): 75 relevant RCTs were identified: 65 RCTs comparing approaches to TKA and 10 for THA. The risk of bias due to surgeons' reported expertise was categorised as low, moderate or high. The majority of RCTs in this sample were assessed as being at high risk of bias on this variable: 42/65 TKA trials (56%) and 5/10 THA trials (50%). Approximately one quarter were deemed to be at low risk of bias: 15/65 TKA trials (23%) and 3/10 THA trials (30%). The reporting of this variable in both comparisons improved from 2009 onwards, e.g. the proportion of low risk of bias RCTs increased from 7% to 35% for the MV vs MP comparison, and from 0% to 53% for SV vs MP; the proportion of trials at high risk of bias also decreased exponentially, from 85% to 53% for MV vs MP, and from 83% to 26% for SV vs MP. Meta-analyses did not demonstrate any meaningful differences between the findings of trials at high risk of bias compared with trials at low/moderate risk of bias, except that the latter did tend to produce findings of relatively greater uncertainty across four of the five outcomes in this sample. This accords with other evidence that better-conducted trials can tend to produce less 'positive' findings.

Conclusion(s): Surgeons' expertise is often poorly reported but has improved since 2009, possibly due to the dissemination of the CONSORT-NPT extension statement. However, it is not adequately captured by any other risk of bias domain and is therefore worthy of assessment. This is the first systematic review to seek to assess both the reporting and impact of this risk of bias variable.

PS1-5 USING BALANCING METHODS TO ESTIMATE INCIDENCE RATES OF DISEASE OUTCOMES IN TYPE 2 DIABETES FOLLOWING ANTIPLATELET TREATMENT FOR SECONDARY PREVENTION OF ACUTE CORONARY SYNDROMES (ACS)

Health Services, Outcomes and Policy Research (HSOP)

Georgios Gkountouras, MSc[1], Rachel Elliott, BPharm Phd FRPharmS[2], Lukasz Tanajewski, PhD[1] and Anthony Avery, PRHO, SHO, GP[3], (1)Division for Pharmacy Practice and Policy, Nottingham, United Kingdom, (2)University of Manchester, Manchester, United Kingdom, (3)School of Medicine, Nottingham, United Kingdom

Purpose:

Routinely collected data can be used to assess effectiveness of interventions in the absence of trial data if confounding is handled appropriately. Balancing techniques are used commonly to address selection bias. In this study, two balancing methods were used to address the imbalance in patient characteristics between different antiplatelet exposure groups, in adults with Type 2 diabetes (T2DM) prescribed antiplatelet treatment for secondary ACS prevention.

Method(s):

A cohort of adults with T2DM prescribed antiplatelets (aspirin-clopidogrel (AC), aspirin-prasugrel (AP), asprin-ticagrelor (AT)) was derived using the UK Clinical Research Practice Datalink (CPRD) linked with English hospital episode statistics (HES). The outcomes of interest were all-cause death, bleeding, myocardial infarction (MI), and stroke. The variables used for balancing were cohort entry year, length of follow-up, general practice location, gender, age, smoking status, alcohol consumption, and number of medicines prescribed.
For balancing purposes, sampling weights were based on the propensity scores of being allocated to any antiplatelet regimen. Propensity scores were generated using two methods [multinomial propensity score (MNPS) and covariate balancing propensity score (CBPS)]. Incidence rates were expressed as events per 100 patient-years.

Result(s):

A cohort of 4784 adults with T2DM was extracted: AC (4449); AP (230); AT (105). The AC group was the largest because this was the first treatment available. Outcome incidence rates (Table 1) suggest AP and AT groups were associated with lower rates of MI, stroke and death than AC. AT was associated with lower bleed rates than AC or AP. Using balancing methods did not alter incidence rates in the AC group for investigated outcomes. AT group incidence rates were reduced post-balancing across all outcomes. For the AP group, balancing generated increased incidence rates for bleeding and death. Reduced stroke and MI rates were observed in the AP and AT with mixed results in terms of side-effects (bleeding).

Conclusion(s):

Using balancing methods allowed for adjustment of characteristics that might affect the type of antiplatelet prescribed, and both methods provided similar changes in incidence rates. Ticagrelor and prasugrel provided equal or better outcomes than clopidogrel, as in clinical trials in the general population, but the magnitude of treatment effect requires further verification, due to low numbers in prasugrel and ticagrelor groups.

PS1-6 PERSONAL VALUE IN PERSONALIZED MEDICINE

Health Services, Outcomes and Policy Research (HSOP)

Mirte Schaafsma[1], Olga Damman, PhD[2], Marleen Jansen, PhD[1], Richelle Singotani[2], Carla van El, Ph.D.[1], Anant Jani, MD, PhD[3], Muir Gray, MD, PhD[3] and Martina Cornel, MD, PhD[1], (1)Department of Clinical Genetics, Amsterdam Public Health research institute, VU University Medical Center, Amsterdam, Netherlands, (2)Department of Public and Occupational Health, Amsterdam Public Health research institute, VU University Medical Center, Amsterdam, Netherlands, (3)Department of Primary Care, University of Oxford, Oxford, United Kingdom

Purpose:

The aim of this study is to explore how personalized medicine can help to study and achieve personal value and optimize patient outcomes.

Method(s):

We explored concepts of both personalized medicine and value based health care in search of a fruitful paradigm for research and care to tailor health care to individual needs and preferences. To study these concepts, literature was gathered on both topics, experts were interviewed, and meetings were organized..

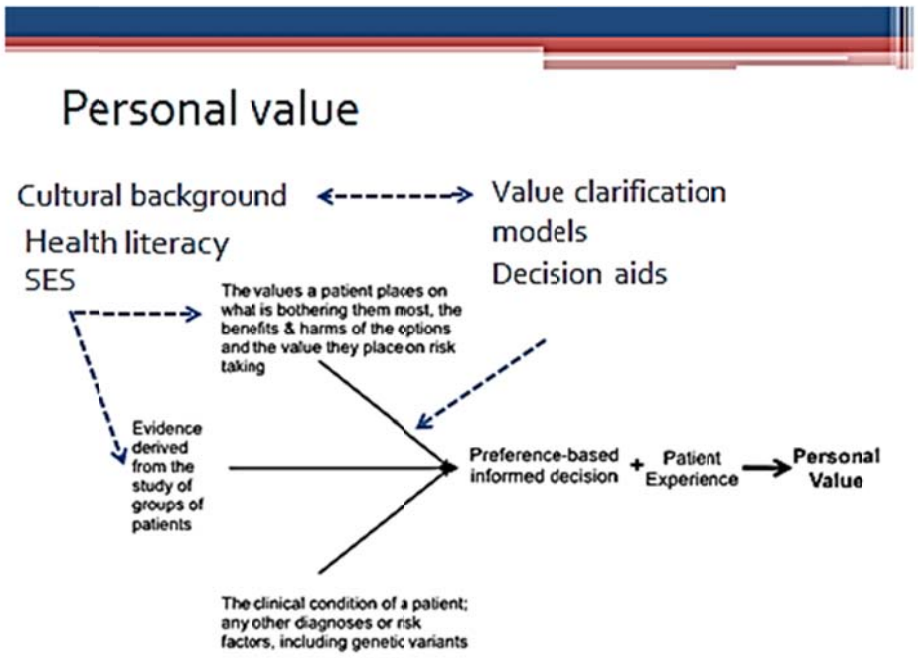
Result(s):

Whereas value based health care strategies focus on outcomes that take individual preferences into account, many definitions of personalized medicine stress tailoring interventions to the biological and genetic make-up of individuals.

We opt for a broad concept of personalized medicine that reaches beyond biological and genomic markers to include cultural, socio-economic and psychological dimensions influencing health and well-being.

Building on the firm basis of evidence based medicine the Triple Value paradigm (Gray and Jani 2016) differentiates between technical, allocative and personal value. Allocative value is determined by how well assets are distributed to different subgroups in a population. Technical value is determined by how well the allocated resources are used for all the people in need in the population, and includes efficiency. Personal value is determined by how well the outcome relates to the preferences and values of each individual, and combines preference-based informed decision making with patient experience.

Personal value can be optimized by studying and clarifying how cultural, psychological and socio-economic factors impact on values and preferences of patients, as well as by taking evidence, the clinical condition and the patient’s genotype into account in decision making.



Adapted from: Gray M and Jani A, Promoting Triple Value Healthcare in Countries with Universal Healthcare. *Healthc Pap,* 2016. **15** (3):p.42-8.

Conclusion(s):

A fruitful strategy for health research and care can be based on a broad concept of personalized medicine addressing genetic, socio-economic, psychological and cultural differences between patients. Such a broad view on personalized medicine will aid in optimizing personal value by ensuring that patients' values and preferences, the available evidence, and the patient's genetic make-up and condition are taken into account in medical decision making.

PS1-7 PREDICTING PERSISTENT POST-CONCUSSIVE SYMPTOMS FOLLOWING MILD TRAUMATIC BRAIN INJURY

Health Services, Outcomes and Policy Research (HSOP)

Maryse C. Cnossen, PhD.[1], Joukje Naalt, van der, PhD.[2], Joke Spikman, PhD.[3], Daan Nieboer[4], John Yue, MD[5], Geoffrey Manley, PhD.[5], Ethan Winkler, PhD.[5], Nicole von Steinbuechel[6], Suzanne Polinder, PhD[1], Ewout W. Steyerberg, PhD[7] and Hester F. Lingsma, PhD[1], (1)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (2)University Medical Center Groningen, Groningen, Netherlands, (3)Groningen University, Groningen, Netherlands, (4)Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, Netherlands, (5)University of California, San Francisco, CA, (6)University of Gottingen, Gottingen, Germany, (7)Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands

Purpose: Traumatic brain injury (TBI) is a leading cause of death and disability worldwide. Approximately 70-90% of all TBIs can be classified as mild TBI (mTBI). Many mTBI patients recover completely during the first weeks to months. However, many develop persistent post-concussion symptoms (PPCS), referring to a constellation of cognitive (e.g. forgetfulness, concentration problems), somatic (e.g. headache, dizziness) and emotional (e.g. irritability, depression) symptoms. Interventions to prevent and reduce PPCS are especially promising when they are personalized and provided early post-injury. Consequently, a valid prediction model for PPCS could support decision-making in this area by identifying patients who might benefit for personalized treatment- or preventive interventions. The objectives of the current study were to externally validate existing prediction models for PPCS, and to develop a new model based the variables including in existing models and the addition of the following complaints at presentation at the emergency department (ED): headache, nausea/vomiting and neck pain.

Method(s): MTBI patients (Glasgow Coma Scale score 13-15) were prospectively recruited from three Dutch level I trauma centers between 2013-2015 in the UPFRONT study. PPCS were assessed using the Head Injury Severity Checklist at six-month post-injury. Patients were classified as having PCS if they reported at least 3 out of 8 symptoms according to the International Classification of Diseases (ICD)-10 diagnosis of PCS. Two published prediction models were examined for calibration and discrimination. The final model, comprising variables of existing models (e. g. demographics, clinical characteristics and two-week symptoms) with the addition of complaints at the ED, was developed with logistic regression and bootstrap validation with 100 repetitions.

Result(s): Overall 591 patients (mean age 51years, 41% female) were included; 241 (41%) developed PPCS. Existing models performed poorly at external validation (AUC: 0.57-0.64, calibration slope below 1 (0.21-0.54). The newly developed model included female sex (OR 1.48, 95%CI [1.01–2.18]), neck pain (OR 2.58,[1.39–4.78]), two-week post-concussion symptoms (OR 4.89,[3.19–7.49]) and two-week posttraumatic stress (OR 2.98,[1.88-4.73]) as predictors. Discrimination of this model was adequate (AUC after bootstrap validation: 0.75).

Conclusion(s): Existing prediction models for PPCS perform poorly. A new model performs reasonably. External validation of this model is warranted before it could be considered for medical decision making in clinical practice.

PS1-9 OLDER KIDNEY PATIENTS' PERCEPTIONS OF A SHARED DECISION-MAKING PROCESS WHETHER OR NOT TO START DIALYSIS

Decision Psychology and Shared Decision Making (DEC)

Wouter Verberne, MD, St Antonius Hospital, Nieuwegein, Netherlands, Karen Prantl, Drs., Dutch Kidney Patients Association (NVN), Bussum, Netherlands and Willem Jan Bos, MD, PhD, Sint Antonius Hospital and Leiden University Medical Center, Nieuwegein and Leiden, Netherlands

Purpose: To assess perceptions from older patients with advanced chronic kidney disease (CKD) of a shared decision-making process on starting dialysis or choosing conservative care.

Method(s): We developed a questionnaire in collaboration with patient representatives and the Dutch Kidney Patients Association to assess patients' perceptions of our shared decision-making process on starting dialysis. We included patients from a non-academic teaching hospital in The Netherlands diagnosed with stage 4/5 CKD and aged ≥70 years, who had chosen either dialysis or conservative care (CC) after careful counselling. A shared-decision making process was initiated when the estimated glomerular filtration rate fell below 20 mL/min/1.73m^2, and included in-depth discussions in which oral and/or written information was given about practicalities, benefits, and risks of dialysis and CC. Patients were counselled by their nephrologist, renal nurse, and social worker, who paid a visit to the patient's home if indicated. Descriptive statistics were used to determine group differences.

Result(s): 99 patients completed the questionnaire (response rate: 77%), of which 75 chose dialysis and 24 conservative care. Mean age was 79.7 years (standard deviation: 5.1 years), 68% was male, 59% lived with a partner, and 73% stated to be religious. Of the 75 patients in the dialysis group, 40 were still in the predialysis phase and 35 were receiving dialysis treatment (25 haemodialysis, 10 peritoneal dialysis). Overall, all patients stated that they were satisfied with the decision-making process (dialysis group: median score 8.0 on a 11-point Likert scale; CC group: 9.0; $p = 0.06$), as well as with their treatment decision (8.0 vs. 9.0; $p = 0.07$). 53% of the patients in the dialysis group considered their own opinion as most important in their treatment choice compared to 95% in the conservative care group ($p = 0.03$). More patients in the dialysis group felt forced to make a decision (32% vs. 1%; $p = 0.011$), or still doubted their treatment decision (17% vs. 0%; $p = 0.03$).

Conclusion(s): Older kidney patients felt satisfied with the shared decision-making process on starting dialysis or choosing conservative care. Those who chose conservative care experienced more autonomy and ownership in the decision-making.

PS1-10 DIFFERENCES IN GLUCOSE CONTROL BY IMMIGRANT GENERATION IN A COHORT OF TYPE 2 DIABETES PATIENTS LIVING IN TYROL/AUSTRIA
Health Services, Outcomes and Policy Research (HSOP)

Wilhelm Oberaigner, Assoc.-Prof. Priv.-Doz. Dr.[1], Sabine Geiger-Gritsch, Dr.[2], Hans-Robert Schoenherr, MD[3], Gerald Bode, MD[3], Christoph Ebenbichler, Prof.[3], Martin Juchum, MD[3], Uwe Siebert, MD, MPH, MSc, ScD[4] and Monika Lechleitner, Prof.[3], (1)UMIT - University for Health Sciences, Medical Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA, Hall i.T., Austria, (2)Department of Clinical Epidemiology of TIROL KLINIKEN, Innsbruck, Austria, Innsbruck, Austria, (3)Diabetes Register Tyrol, Innsbruck, Austria, Innsbruck, Austria, (4)UMIT, Dept. of Public Health, Health Services Research & Health Technology Assessment / Harvard Univ., Dept. Health Policy & Management, Institute for Technology Assessment / ONCOTYROL, Division for HTA, Hall i. T. (Austria) / Boston (USA), Austria

Purpose: For decades, Austria has been a country of immigration although immigration is not in the country's perception (Report by the Migration Council of Austria, 2016). According to official statistics, in 2016 about 22% of the Austrian population had immigration background, the percentage in Tyrol was slightly lower at 20%. As immigration is known to be associated with an elevated risk of diabetes, we aimed at analyzing HbA1c levels as global measure for glucose control by immigrant status including level of immigrant generation.

Method(s): The diabetes register of Tyrol (DRT) was initiated in 2005 and collects demographic, anamnestic and clinical data on patients with diabetes mellitus who undergo healthcare mainly at hospital outpatients departments in Tyrol/Austria. Since 2010, immigrant status was assessed using a questionnaire proposed by Schenk at al (Bundesgesundheitsbl, 2006), which allows categorizing by first to third immigrant generation. The present analysis covers data from 7033 patients with type 2 diabetes for whom immigrant status was assessed. HbA1c in % at last visit was analysed in a logistic regression model for HbA1c $\geq$ 8%. We used multivariate logistic regression to determine odd ratios (OR) and 95% confidence intervals (95% CI), controlling for potential confounders (i.e., age, sex, level of physical activity). Sensitivity analysis was performed using HbA1c as continuous outcome. The analysis was conducted using STATA Version 13.

Result(s): We analyzed a total of 7033 patients with type 2 diabetes. Of all patients, 4425 (62.9%) were non-immigrants, 1273 (18.1%) first generation immigrants, 189 (2.7%) second generation immigrants and 1146 (16.3%) third generation immigrants. Mean age at last visit was 66 for non-immigrant patients and 59/61/69 for first to third generation immigrants respectively. Multivariate logistic regression for HbA1c $\geq$ 8% showed ORs of 1.32 (95% CI 1.15-1.51) for first, 1.00 (95% CI 0.72-01.38) for second, and 0.76 (95% CI 0.65-0.88) for third generation immigrants compared to non-immigrant patients after adjusting for age, sex and level of physical activity. Results were robust when using continuous HbA1c.

Conclusion(s): Our results clearly demonstrate that for the analysis of immigration in diabetes patients, it is necessary to take immigrant generation into account. In our population, glucose control needs to be especially improved for first generation immigrants, whereas third generation immigrants have better glucose level control that non-immigrants.

PS1-11 BETWEEN-CENTER AND BETWEEN-COUNTRY DIFFERENCES IN OUTCOME AFTER ANEURYSMAL SUBARACHNOID HEMORRHAGE IN THE SUBARACHNOID HEMORRHAGE INTERNATIONAL TRIALISTS (SAHIT) REPOSITORY

Health Services, Outcomes and Policy Research (HSOP)

Simone A. Dijkland, MD[1], Blessing N.R. Jaja, MD, PhD[2], Mathieu van der Jagt, MD, PhD[3], Bob Roozenbeek, MD, PhD[4], Ewout W. Steyerberg, PhD[5], Diederik W.J. Dippel, MD, PhD[4], R. Loch Macdonald, MD, PhD[2] and Hester F. Lingsma, PhD[1], (1)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (2)Division of Neurosurgery, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada, (3)Department of Intensive Care, Erasmus University Medical Center, Rotterdam, Netherlands, (4)Department of Neurology, Erasmus University Medical Center, Rotterdam, Netherlands, (5)Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands

Purpose:

Differences in outcome among different settings are common in many diseases. A central asset of the value-based healthcare approach is analyzing whether such differences in outcome may represent differences in quality of care. We aimed to quantify between-center and between-country differences in outcome after aneurysmal subarachnoid hemorrhage (aSAH).

Method(s):

We analyzed data from 5530 aSAH patients from three randomized clinical trials including 170 centers and 22 countries. We used random-effects logistic regression adjusted for patient characteristics to estimate between-center and between-country differences in unfavorable outcome, defined as Glasgow Outcome Scale 1-3 (severe disability, vegetative state or death) or modified Rankin Scale 4-6 (moderately severe disability, severe disability or death) at three months. Between-center and between-country differences were quantified with the median odds ratio (MOR), which is based on the variance of the random effects and can be interpreted as the odds ratio for unfavorable outcome between two randomly selected centers or countries. We analyzed the total database and performed subgroup analysis stratified for continent (Europe, North America and Oceania).

Result(s):

The proportion of patients with unfavorable outcome was 27%. We found substantial between-center differences (MOR=1.26), which were not explained by patient characteristics (adjusted MOR=1.30). Between-country differences were also found (adjusted MOR=1.41). Between-center differences were absent in Oceania and North America (adjusted MOR=1.00), but were present in Europe (adjusted MOR=1.51).

Conclusion(s):

Outcome after aSAH differs between centers and countries, especially in Europe. Since these differences could not be explained by patient characteristics, future research should focus on explanations regarding differential treatment policies and quality of care.

PS1-12 A SIMPLIFIED MODEL OF THE COST-EFFECTIVENESS OF SCREENING IN R: A TEACHING AND RESEARCH TOOL
Quantitative Methods and Theoretical Developments (QMTD)

James O'Mahony, PhD, Trinity College Dublin, Dublin, Ireland and Joost van Rosmalen, PhD, Erasmus MC, University Medical Center, Department of Biostatistics, Rotterdam, Netherlands

Purpose: To demonstrate a simplified pedagogical model of the cost-effectiveness of cancer screening and explain its potential as a teaching and research tool.

Method(s): The models applied in the cost-effectiveness analysis of screening interventions are typically designed to address specific policy questions and consequently are often large and complex. We describe the rationale for employing a lightweight, fully shareable and transparent alternative to such large applied models for the purposes of teaching and methods research. We present an overview of the code of a simplified, discrete-event, microsimulation model of cancer screening coded in R and supported with a Microsoft Excel-based user interface for the specification of input parameters. We demonstrate the components of the model relating to the natural history of disease, test performance and anticipated health gain and healthcare costs.

Result(s): We show how the costs and effects of multiple alternative screening strategies can be simulated in the model. Using the process of comparative statics, we show how the efficient frontier and incremental cost-effectiveness ratios of alternative screening programmes vary with changes in key parameters such as disease incidence and test sensitivity. Furthermore, we demonstrate how the choice of the optimal screening policy for a given cost-effectiveness threshold varies with changes in input parameters. As such, the model provides a tool with which to demonstrate the qualitative relationships between parameters and the optimal policy in a way that is faster and more accessible than employing a full applied model. We go on to describe the potential for further applications and extensions of the model, including probabilistic analysis and calibration.

Conclusion(s): The simplified model provides a transparent and easy-to-use demonstration of the fundamentals of the cost-effectiveness of screening. The model is fully shareable and represents a useful open-source teaching and research tool to enhance methods research in the cost-effectiveness of screening. Most models used in applied research are not fully published, due both to their large size and to concerns about sharing intellectual property. Such incomplete reporting compromises transparency and hinders methods research. Our simplified model avoids these problems with fully-shareable code that can be employed and adapted by anybody. This alternative offers a more appropriate tool for teaching the basics of screening cost-effectiveness and conducting methods research.

PS1-13 SPATIAL DYNAMICS SIMULATION OF PSYCHIATRIC PATIENTS - REHOSPITALISATION SIMULATION FOCUSSING ON DIABETES COMORBIDITY AND REGIONAL HEALTH SERVICES IN INTERNATIONAL CONTEXT
Quantitative Methods and Theoretical Developments (QMTD)

Guenther Zauner, Msc, dwh GmbH, simulation services / DEXHELPP (Decision Support for Health Policy and Planning), Vienna, Austria, Christoph Urach, MSc, Institute for Analysis and Scientific Computing, Technical University Vienna, Vienna, Austria, Martin Bicher, PhD, Technical University Vienna, Vienna, Austria and Niki Popper, PhD, Dexhelpp (Decision Support for Health Policy and Planning), Vienna, Austria

Purpose: The development of the burden of psychiatric diseases is essential for planning purposes of health policy makers. In CEPHOS-LINK (FP7-project; No-603264), several strategies were incorporated to analyse psychiatric re-hospitalisations. Three research questions were defined: **1)** Change of re-hospitalisations in the future, **2)** Theoretical improvement of structure of care on a NUTS3-level, **3)** Possible impact of rising diabetes prevalence. Simulation of the future trends in the total number of patient's behaviour and the costs for these scenarios had to be performed for Austria (AT), Slovenia (SL), and Veneto.

Method(s): We used a simulation approach based on an agent-based population model called **GEPOC**, developed in **DEXHELPP** (FFG, No-843550). The implemented Python framework was extended by several modules for each question. The modules for the other two questions are exchangeable or used together. The parameterization for AT/SL/Veneto was calibrated by pooled claims.

Result(s): The results represent average values out of ten simulation runs to minimize stochastic effects. Pathways of simulated patients are available on microscopic level, meaning that they can be followed over time. For the macroscopic evaluations events of the patients are aggregated up to year 2039 for different patient characteristics. The outcomes show an increase in both index- and re-hospitalisation rates, especially for females and patients with psychotic diagnoses. Looking at NUTS3 regions with highest average driving time to hospitals a change to optimised care can lower the re-hospitalisation rate by nearly 39% in AT, about 20% in SL and in Veneto with actual lowest distance to service only 7.5%. In case of diabetes it has to be noticed that the demographic changes are effecting especially prognosis for AT. Additionally the economic impact of the (re)hospitalisation cases are analysed.

Conclusion(s): All three scenarios show that psychiatric hospitalizations are rising, especially in Austria and Veneto region. The most drastic changes can be assumed to come in the timeframe of the next 10 years for non-psychotic diagnosis. Changing diabetes prevalence also has an impact on psychiatric patients' re-hospitalisation and shows that comorbidities should not be neglected when analysing future development of re-hospitalisation rates.

PS1-14 TWO NOVEL REGISTRY-BASED PREDICTION MODELS FOR OVERALL SURVIVAL IN PATIENTS WITH METASTATIC ESOPHAGEAL OR GASTRIC CANCER

Quantitative Methods and Theoretical Developments (QMTD)

Héctor van den Boorn, MSc.[1], Ameen Abu-Hanna, PhD[1], Emil ter Veer, MSc.[1], Jessy Joy van Kleef, MSc.[1], Florian Lordick, MD, PhD[2], Michael Stahl, MD, PhD[3], Jaffer Ajani, MD, PhD[4], Rosine Guimbaud, MD, PhD[5], Se Hoon Park, MD, PhD[6], Susan Dutton, MSc.[7], Yung-Jue Bang, MD, PhD[8], Nadia Haj Mohammed, MD, PharmD[9], Mirjam Sprangers, PhD[1], Rob Verhoeven, PhD[10], Koos Zwinderman, PhD[1], Martijn van Oijen, PhD[1] and Hanneke van Laarhoven, MD, PhD[1], (1)Academic Medical Center, Amsterdam, Netherlands, (2)University Cancer Center Leipzig, Leipzig, Germany, (3)Kliniken Essen-Mitte, Essen, Germany, (4)University of Texas MD Anderson Cancer Center, Houston, TX, (5)University Hospital of Rangueil, Toulouse, France, (6)Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea, Republic of (South), (7)OCTRU, University of Oxford, Oxford, United Kingdom, (8)Seoul National University Hospital, Seoul National University College of Medicine, Seoul, Korea, Republic of (South), (9)UMC Utrecht, Utrecht, Netherlands, (10)Netherlands Comprehensive Cancer Organisation (IKNL), Eindhoven, Netherlands

Purpose: Prediction models for decision-making in oncology are increasingly being used, but few are available for esophagogastric cancer, particularly in the metastatic setting. The aim of this study is to construct prediction models for overall survival in patients with metastatic esophageal or gastric cancer.

Method(s): Data from patients with metastatic esophageal (N=8670) and gastric (N=4804) cancer diagnosed in the period 2005-2015 were retrieved from the nationwide Dutch cancer registry. Multivariate Cox regression models, extended with treatment interactions, were created to predict overall survival. Multiple imputations were used to handle missing data. Predictor selection was performed via the Akaike Information Criterion (AIC) and was extended by a Delphi consensus among experts in the field of palliative esophagogastric cancer. Validation was performed with an 11-fold temporal validation. Both the concordance-index (c-index) and calibration were used to assess model quality.

Result(s): The Delphi consensus yielded seven important predictors of survival and are shown with the AIC-selected predictors in Table 1. The c-indices show consistent discriminative ability during validation, i.e. 0.71 and 0.68 for respectively the esophageal and gastric cancer models. There is close agreement between predicted and observed survival, with an error of 1.7% and 2.2% for respectively the esophageal and gastric cancer models.

Conclusion(s): The models yield fair discrimination and high calibration levels, and provide a good foundation for further investigation in clinical practice to determine their added value in decision-making.

Table 1: Overview of selected predictors in the esophageal and gastric cancer models (#: selected during Delphi consensus).

Predictor	**Esophageal cancer model**	**Gastric cancer model**
Gender	X	X
Age[#]	X	X
cT-stage	X	X
cN-stage	X	X
Primary tumor location[#]	X	X
Tumor morphology[#]	X	X
Number of distant metastatic sites[#]	X	X
First line treatment type[#]	X	X
Metastasis only in distant lymph nodes	X	X
Liver metastasis[#]	X	
Peritoneal metastasis[#]	X	
Age * First line treatment	X	X
Liver metastasis * first line treatment	X	
Number of distant metastatic sites * First line treatment		X

PS1-15 PREDICTION IN SURVEILLANCE OF BARRETT'S ESOPHAGUS: THE EFFECT OF MULTIPLE MEASUREMENTS OF BIOMARKERS ON THE ESTIMATED NEOPLASTIC PROGRESSION RISK

Quantitative Methods and Theoretical Developments (QMTD)

Carlijn A.M. Roumans[1], Manon C.W. Spaander[2], Iris Lansdorp- Vogelaar, PhD[3], Katharina Biermann[4], Marco J. Bruno[2], Ewout W. Steyerberg, PhD[5] and Dimitris Rizopoulos[6], (1)Department of Public Health, Department of Gastroenterology & Hepatology, Erasmus University Medical Center, Rotterdam, Netherlands, (2)Department of Gastroenterology & Hepatology, Erasmus University Medical Center, Rotterdam, Netherlands, (3)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (4)Department of Pathology, Erasmus University Medical Center, Rotterdam, Netherlands, (5)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands; Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, Netherlands, (6)Department of Biostatistics, Erasmus University Medical Center, Rotterdam, Netherlands

Purpose: Barrett's esophagus (BE) is a premalignant condition, where surveillance is performed periodically to reduce morbidity and mortality related to esophageal adenocarcinoma (EAC). The harm-benefit ratio of this strategy is questionable, because identification of high-risk patients is difficult. To improve risk stratification, additional biomarkers, and their variations between and -in particular- within patients should be taken into account. We aimed to develop a model incorporating all follow-up (FU) measurements of low-grade dysplasia (LGD), p53, and SOX2 to study their predictive performance.

Method(s): In this multicenter prospective cohort study, we included consecutive BE patients from 15 hospitals with at least 0.5 year FU time. The study endpoint was identification of high-grade dysplasia (HGD) or EAC. Data were collected during every FU. Contrary to previous research, we incorporated all available longitudinal measurements of LGD, p53, and SOX2 and linked them to the risk of neoplastic progression using a multivariate joint model. The longitudinal profiles of each marker were estimated with a mixed effects logistic regression model, which were incorporated in a time-varying Cox model afterwards.

Result(s): The median FU time was 7.2 years (IQR 5.4-9.9) of 628 patients included; 48 developed HGD or EAC. If a patient would have only 2 FU moments, one with normal (0) expression of the biomarker, one with aberrant (1), the hazard ratio (HR) of neoplastic progression was 1.2 for LGD (p=0.61), 1.5 for p53 (p=0.004), and 5.0 for SOX2 (p=0.004), annually. With more FU endoscopies, these multiple observations would set the probability of aberrant expression to a value between 0 and 1, with a ditto proportion of the previously mentioned hazard ratios. Dynamic risk profiles of neoplastic progression could be estimated for individual patients during their FU, based on these biomarker patterns.

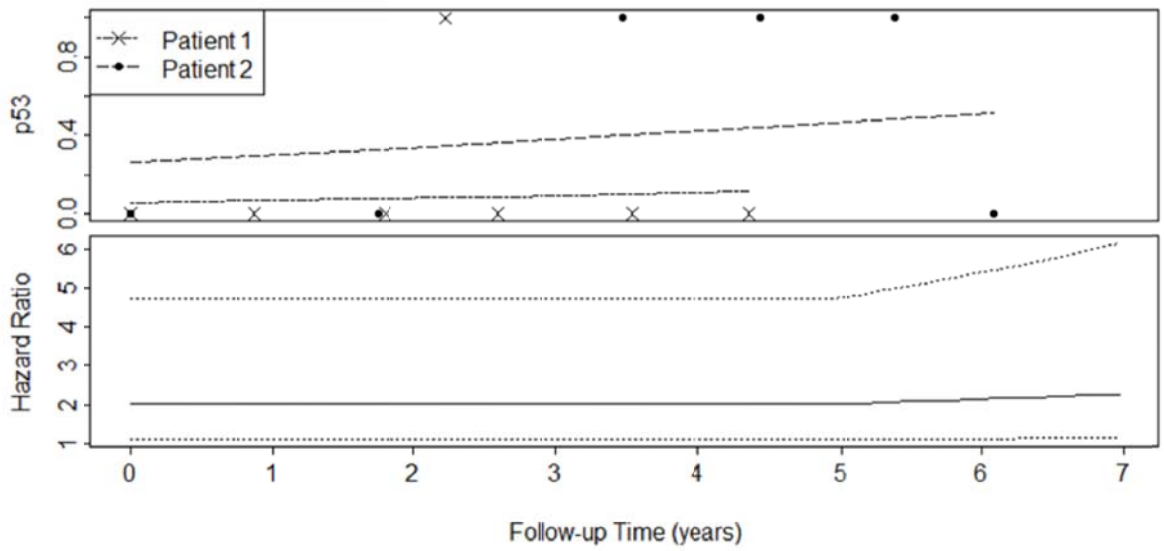


Figure. Two exemplary patients: (top) longitudinal evolution of probability for normal (0) or aberrant (1) expression of p53 in multiple FU moments in time. Symbols are measurements, lines are predictions of the model. P53 measurements are incorporated in a joint model (bottom). The difference in HR between patient 1 and 2 for neoplastic progression (HGD/EAC) with 95% CI is shown, if p53 changes in multiple FU moments in time.

Conclusion(s): The risk of neoplastic progression can be estimated better by p53 and SOX2 than LGD if measurements of all FU endoscopies are taken into account. In the future, this model will provide an updated prediction of neoplastic progression based on new observations in ongoing FU, which will assist in medical decision making. The combination of dynamic observations of biomarkers such as p53 and SOX2 merits further study to improve the identification of high-risk patients in a personalized surveillance program, eventually reducing the burden of surveillance.

PS1-17 STRATEGIC BEHAVIOR AND THE COST-EFFECTIVENESS THRESHOLD
Applied Health Economics (AHE)

Mike Paulden, PhD, University of Alberta, Edmonton, AB, Canada

Purpose: In a health care system with a constrained budget, funding new health technologies displaces other services and results in health losses for other patients. Under existing theory, the cost-effectiveness threshold (λ) is set equal to the incremental cost-effectiveness ratio (ICER) of these displaced services (k). This paper advances existing theory to incorporate strategic behavior, demonstrating that λ should be set below k.

Method(s): The existing theoretical model was extended by incorporating the following assumptions: 1. The threshold is publicly stated and fixed while numerous technologies are appraised; 2. Manufacturers are strategic and 'price to the threshold', resulting in ICERs equal to λ; 3. Each manufacturer has a minimum 'reserve price', and hence 'reserve ICER', needed to supply the technology; 4. Some reserve ICERs lie below k.

Result(s): The optimal λ is between zero and k. If λ is zero, no new technologies are funded, resulting in no net population health benefit. If λ is equal to k, manufacturers strategically price so the ICER equals k; health gains are exactly offset by health losses, such that net population health benefit is again zero. If λ is between zero and k, each funded technology has positive net population health benefit. Within this region, any marginal increase in λ has two counteracting effects: 1. Manufacturers whose reserve ICERs are now met will supply their technologies, increasing net population health benefit; 2. Manufacturers whose reserve ICERs were already met with the lower λ will strategically raise prices, reducing net population health benefit. At the optimal λ, these two effects counteract exactly.

Conclusion(s): When strategic behavior is considered, the optimal λ lies below k. This results in improved population health outcomes compared to using the threshold implied by existing theory. This has important implications for decision making and empirical research into estimating the threshold.

PS1-18 A SYSTEMATIC REVIEW ON COST-EFFECTIVENESS OF CERVICAL CANCER SCREENING IN EUROPE WITH SPECIFIC INTEREST ON RISK-ADAPTED STRATEGIES

Applied Health Economics (AHE)

Gaby Sroczynski, MPH, Dr.PH[1], Artemisa Gogollari, Mag.sc.hum. Dr. med.[2], Emi Naslazi, MPH[1], Nora Pashayan, MD, MSc, MSt, PhD[3], Martin Widschwendter, PhD[4] and Uwe Siebert, MD, MPH, MSc, ScD[5], (1)Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria, (2)Department of Public Health, Health Services Research and Health Technology Assessment, UMIT and ONCOTYROL - Center for Personalized Cancer Medicine, Division of Health Technology Assessment and Bioinformatics, Hall in Tirol, Austria, (3)UCL - University College London, Faculty of Population Health Sciences, London, United Kingdom, (4)UCL - University College London, Department of Women's Cancer, London, United Kingdom, (5)UMIT, Dept. of Public Health, Health Services Research & HTA / Harvard Univ., Dept. Health Policy & Management, Institute for Technology Assessment, Hall i.T. (Austria)/Boston (USA), Austria

Purpose: We systematically reviewed current evidence on long-term effectiveness and cost-effectiveness of cervical cancer screening in Europe with specific interest on risk-adapted strategies.

Method(s): A systematic literature search in relevant databases (Medline/Embase/Cochrane Library/CRD/EconLit) was performed to identify decision-analytic modelling studies assessing the cost-effectiveness of cervical cancer screening strategies in Europe. We summarized study characteristics and results in standardized evidence tables. Outcomes included quality-adjusted life-years (QALY), life-years gained (LYG), and the incremental cost-effectiveness ratios (ICER). Economic results were converted to 2015 Euros using gross domestic product purchasing power parity and the consumer price index.

Result(s): We included 14 studies, comprising analyses of different settings (i.e. vaccinated or non-vaccinated women) mostly for countries with population-based organized screening. All studies reported HPV-based screening to be more effective in terms of patient-relevant outcomes compared to cytology alone independent of HPV vaccination status. Overall, HPV-based screening strategies were considered to be cost-effective at a willingness-to-pay threshold of 50,000 Euro/QALY or LYG conditional on screening intervals of at least three years in non-vaccinated women and at least five years in vaccinated women. HPV screening was mostly accompanied with a triage test for HPV-positive women. Most studies recommended starting screening in general at age 25 using cytology with cross over to HPV-based screening at age 30 years or older. The upper age limit for screening varied with most studies ending screening at age 65 years. While studies consider generic cancer risk aspects to restrict screening algorithms (e.g. tailoring screening start, end and frequency according to age and HPV status), further more personalized screening and follow-up strategies according to a woman's screening and vaccination history and taking into account other risk factors (e.g. life style, smoking) have not been evaluated in the included studies.

Conclusion(s): In conclusion, the evidence from decision-analytic modeling studies suggests that HPV-based screening is more effective compared to cytology alone, and can be considered cost-effective at screening intervals of at least 3 years in non-vaccinated and at least 5 years in vaccinated women. Current risk-tailored screening programs are based on generic restrictions to a specific age and using triage or follow-up tests for HPV-positive women before referring to colposcopy directed biopsy. In future research, predictive biomarker for risk-based management of screen-positives and more personalized screening algorithms should be considered.

PS1-21 A COST-EFFECTIVENESS ANALYSIS OF SYSTEMIC THERAPY IN METASTATIC HORMONE-SENSITIVE PROSTATE CANCER
Applied Health Economics (AHE)

Niranjan Sathianathen, MBBS (Hons)[1], **Fernando Alarid-Escudero, PhD**[2], Karen M. Kuntz, ScD[2], Nathan Lawrentschuk[3], Damien Bolton[3], Declan Murphy[4], Simon Kim[5] and Badrinath Konety[1], (1)University of Minnesota, Minneapolis, MN, (2)Division of Health Policy and Management, University of Minnesota School of Public Health, Minneapolis, MN, (3)University of Melbourne, Department of Surgery, Urology Unit and Olivia Newton-John Cancer Research Institute Austin Health, Heidelberg, Australia, (4)Peter MacCallum Cancer Centre, Parkville, Australia, (5)Case Western Reserve University, Cleveland, OH

Purpose:

The results from randomized trials have demonstrated that both docetaxel and abiraterone acetate when combined with androgen deprivation therapy (ADT) confers a survival benefit compared to ADT alone in newly diagnosed metastatic, hormone-sensitive prostate cancer. While abiraterone may be more effective than docetaxel, it is also more expensive. This study aims to evaluate the effectiveness and costs of docetaxel and abiraterone compared to ADT.

Method(s):

We developed a state-transition model to simulate the natural progression of metastatic prostate cancer. Model parameters were derived from the published literature and through calibration to observed epidemiological data. Following diagnosis, a hypothetical cohort of men could be treated with docetaxel+ADT, abiraterone+ADT or ADT alone. Once disease progresses to the castration-resistant state, treatment with one of the approved therapies in this setting was initiated. We conducted deterministic and probabilistic sensitivity analyses to account for uncertainty in the parameter estimates. Costs were calculated from a health sector perspective and adjusted to 2017 US dollars.

Result(s):

Compared to treatment with ADT alone, docetaxel and abiraterone resulted in discounted quality-adjusted survival gain of 1.7 and 3.1 months, respectively. The projected lifetime discounted costs of ADT, docetaxel and abiraterone was $63,062, $65,569and $256,656, respectively. Using a willingness-to-pay threshold of $100,000 per quality-adjusted life year, treatment with docetaxel and ADT was a cost-effective strategy. The monthly cost of abiraterone would have to be less than $1,323 (base case value = $9,399) for it to be cost-effective at the same willingness-to-pay threshold. The results were sensitive to utility and cost of docetaxel.

Conclusion(s): Immediate treatment with docetaxel and ADT is the most cost-effective treatment option for men with newly metastatic, castrate naïve prostate cancer.

PS1-22 COST EFFECTIVENESS OF ADULT PNEUMOCOCCAL VACCINATION IN US UNDERSERVED MINORITIES AND TOTAL POPULATION: DIFFERENT POLICIES OR BETTER IMPLEMENTATION?

Applied Health Economics (AHE)

Kenneth J. Smith, MD, MS[1], Angela Wateska, MPH[2], Chyongchiou Jeng Lin, PhD[3], Mary Patricia Nowalk, PhD, RD[3] and Richard K. Zimmerman, MD, MPH, MS[2], (1)University of Pittsburgh, Section of Decision Sciences, Pittsburgh, PA, (2)University of Pittsburgh School of Medicine, Pittsburgh, PA, (3)University of Pittsburgh, School of Medicine, Dept of Family Medicine, Pittsburgh, PA

Purpose:

In US adults aged <65 years, pneumococcal vaccination is recommended when immunocompromising or other specific comorbid conditions are present, but only 23% of these patients are currently vaccinated, and less in underserved minorities. The cost-effectiveness of programs to improve vaccination rate or of alternative vaccination policies to increase protection is unclear.

Method(s):

A Markov model estimated the cost-effectiveness of current US adult pneumococcal vaccination recommendations (both pneumococcal polysaccharide vaccine [PPSV] and pneumococcal conjugate vaccine [PCV] in immunocompromised patients aged <65 and only PPSV in those with other high-risk conditions), without and with a program to improve vaccine uptake, and differing policies using alternative comorbidity- or age-based indications in hypothetical 50-year-old cohorts, separately examining US black and total populations. Implementation program-based vaccine uptake improvement was based on a trial examining the 4 Pillars™ Practice Transformation Program in diverse primary care practices. Cohort health status and mortality were modeled using US databases. Age- and comorbidity-specific pneumococcal illness rates came from US CDC data. Vaccine effectiveness (VE) was estimated using trial and Delphi panel data, and differing VE assumption scenarios were examined. Cost and utility data came from medical literature and database sources. A health system perspective was taken, with future costs and effectiveness discounted at 3%/yr. over a lifetime horizon. One-way and probabilistic sensitivity analyses were performed.

Result(s):

In both black and total population cohorts, a program that increased uptake of current adult vaccination recommendations was favored compared to a current recommendation strategy without a program. Recommendations with a program cost $40,300/QALY gained in black populations ($57,000/QALY in the total population) compared to recommendations without a program. Other age- or comorbidity-based vaccination strategies were dominated. Programs were favored, at a $100,000/QALY threshold, unless they improved absolute vaccine uptake by <1.7% (base case: 12.3%) in black populations or <2.8% in the entire population. Less conservative VE assumptions increased implementation program favorability. Probabilistic sensitivity analyses confirmed strong program strategy favorability (82.3-99.8% likelihood at $100,000/QALY) in all scenarios.

Conclusion(s):

Programs to increase adult pneumococcal vaccination uptake are economically reasonable, and more so in US underserved minorities than in the total population. If addressing race-based health disparities is a priority, evidence-based programs to increase vaccination in underserved minorities may be useful.

PS1-24 THE GENERALIZABILITY OF CLINICAL PREDICTIVE MODELS FOR PATIENTS WITH ACUTE CORONARY SYNDROMES: RESULTS FROM INDEPENDENT EXTERNAL VALIDATIONS

Decision Psychology and Shared Decision Making (DEC)

Benjamin S. Wessler, MD[1], Jason Nelson[2], Christine Lundquist, MPH[1], Ben Van Calster[3], David Van Klaveren, PhD, MS[4], Ewout W. Steyerberg, PhD[5] and David Kent[2], (1)Tufts Medical Center, Boston, MA, (2)Predictive Analytics and Comparative Effectiveness (PACE) Center, Tufts Medical Center, Boston, MA, (3)KU Leuven Department of Development and Regeneration, Leuven, Belgium, (4)Predictive Analytics and Comparative Effectiveness Center, Boston, MA, (5)Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands

Purpose: Independent external validations are recommended to understand the trustworthiness of clinical predictions, however most CPMs have never been evaluated in this way. Here we conduct independent external validations of contemporary acute coronary syndrome (ACS) CPMs with a focus on discrimination and calibration in new databases.

Method(s): A systematic review identified CPMs predicting outcomes for patients with ACS. Independent external validations were performed by evaluating model performance using individual patient data from 3 large clinically appropriate publically available clinical trials enrolling patients with ACS. The 'relatedness' of matches between CPMs and databases were characterized based on clinical differences between the validation and derivation populations (e.g. treatments received). CPM performance was evaluated for discrimination and calibration for original and recalibrated models.

Result(s): We screened 234 ACS CPMs and 20 clinically appropriate CPM-trial pairs were identified. Of the 20 matches 9 were distantly related. During external validations, the median c statistic of the original 'as published' CPMs (no recalibration) was 0.70 (IQR, 0.67 to 0.72), with a median decrease of -26.5% (IQR, -34.6% to -21.7%) compared to the c-statistic reported in the derivation cohort. Overall, the median Harrell's E was 0.03 (IQR, 0.02 to 0.05). Recalibration of the intercept led to substantially improved calibration with median change in Harrell's E of -59.7% (IQR -84.9% to -38.7%). Recalibration of the slope had a more modest incremental effect on error (median change in Harrell's E of –11.5% (IQR -27.8% to -2.7%). Performance of CPMs developed on datasets characterized as 'distantly related' rather than 'related' to the validation dataset showed somewhat greater relative decrease in discrimination compared to discrimination reported in their development cohort: -30.3% (-37.7% to -24.8%) – versus -23.3% (IQR, -28.3% to -20.4%). Similarly, measures of bias were also slightly greater for these 'distantly related' models (Harrell's E of 0.05 (IQR, 0.03 to 0.05) versus Harrell's E of 0.02 (IQR, 0.01 to 0.03).

Conclusion(s): For ACS CPMs, models and databases were often incompatible. Independent external validations generally demonstrate that CPMs perform significantly worse than originally presented. Identifiable differences in validation and derivation populations may sometimes partially account for poor CPM performance during external validations. Recalibration of the intercept alone generally corrected most of the prediction error, though recalibration of the slope was also important in a minority of models.

PS1-25 DEVELOPMENT AND EVALUATION OF AN INNOVATIVE COMPUTER-BASED DECISION AID TO ENHANCE INFORMED PARTICIPATION IN COLORECTAL CANCER SCREENING

Decision Psychology and Shared Decision Making (DEC)

Anke Woudstra[1], Evelien Dekker[2], Tom Broens[3], Ellen Smets[4] and Mirjam Fransen[1], (1)Department of Public Health, Amsterdam Public Health research institute, Academic Medical Centre Amsterdam, University of Amsterdam, Amsterdam, Netherlands, (2)Department of Gastroenterology and Hepatology, Academic Medical Centre Amsterdam University of Amsterdam, Amsterdam, Netherlands, (3)Department of Medical Informatics, Amsterdam Public Health research institute, Academic Medical Centre Amsterdam, University of Amsterdam, Amsterdam, Netherlands, (4)Department of Medical Psychology, Amsterdam Public Health research institute, Academic Medical Centre Amsterdam, Amsterdam, Netherlands

Purpose: In the Dutch colorectal cancer screening programme, individuals are expected to make an autonomous, informed decision about screening participation. This is a challenge for screening invitees, especially for those with inadequate health literacy, i.e. the ability to access, understand and use health information. The screening programme is complex and invitees have to make a decision about participation without initial support of a health care professional. Individuals with low health literacy have been shown to experience more uncertainty in their decision than those with adequate health literacy.

Method(s): Following the International Patient Decision Aid Standards (IPDAS) and a framework on health literacy skills for informed decision making, we developed a digital decision aid prototype that is sensitive to individuals with lower health literacy. The decision aid will be qualitatively pre-tested among 30 participants with low and adequate health literacy for accessibility, acceptability and usability in February 2018. After further adaptation on content and design, the decision aid will be pilot-tested among 200 individuals eligible for colorectal cancer screening with varying levels of health literacy in May 2018. Outcome measures are decisional conflict, knowledge, attitude, behavioral intention, anxiety and risk perception.

Result(s): The decision aid contains elements including video, audio and explanations of medical words to improve the comprehensibility of the screening information. A values clarification exercise was developed to support appraisal of the screening information and weigh up potential harms and benefits of colorectal cancer screening for personal relevance.

Conclusion(s): The development of a decision aid is an essential first step in promotion equal opportunities for informed decision making in colorectal cancer screening. The study provides a basis to study the effect on the decision aid on informed decision making among participants with lower health literacy. The findings from the pre-test and the pilot study will be presented at the conference.

PS1-26 DECISION MAKING IN PRIMARY CARE EMERGENCIES: WHEN WOULD GENERAL PRACTITIONERS ATTEND TO CALL OUTS?
Decision Psychology and Shared Decision Making (DEC)

Magnus Hjortdahl, MD and **Peder A. Halvorsen, MD, PhD**, UiT - The Arctic University of Norway, Tromsoe, Norway

Purpose: In Norway general practitioners (GPs) on call are obliged to take part in call outs to pre-hospital emergencies whenever necessary. Whether it is necessary or not is left to the GP's discretion. We aimed to study potential drivers of this decision.

Method(s): In an online survey 1002 GPs were randomised to hypothetical call out alerts and asked whether they would attend. The scenario presented a 60 year old male with acute dyspnoea. We let three attributes vary across different versions of the scenario: Whether or not there were other patients waiting, whether the emergency incident was nearby or distant, and whether the setting was a car accident or acute illness. This yielded 2x2x2=8 versions of the scenario. Each GP saw one version only. We used items from Jackson Personality Inventory to assess the GPs' risk attitude.

Result(s): The GPs were less likely to attend when the emergency was far away, when there were other patients waiting and when the incident was an acute illness rather than a car accident. Across the 8 scenarios the proportion of GPs that would attend to the call out varied from 59 % (non-traumatic dyspnoea, other patients in the waiting room, incident 45 minutes away) to 87% (car accident, no patients in the waiting room, incident 15 min away, chi-square 61, df 7, $p<0.001$). In multivariable logistic regression analysis GPs with long patient lists were less likely to attend, whereas risk seeking GPs and those participating in multidisciplinary team training were more likely to attend. There was a statistically significant interaction between "clinical setting" and "other patients waiting" ($p=0.02$). In the car accident scenario about 80% would attend the callout irrespective of whether there were other patients waiting or not, whereas in the non-traumatic dysponea scenario the GPs were less likely to attend the callout when other patients were waiting (60% versus 78%).

Conclusion(s): Our data suggests that when considering emergency alerts, not only the patient's condition but also the distance to the incident and a crowded waiting room would get weight in the GPs' decisions. Furthermore, team training and risk attitude might influence this decision. If this is true in real life, educational and organisational measures might increase GP participation in call outs to pre-hospital emergencies.

PS1-28 INSUFFICIENT INFORMATION: REDUCING INAPPROPRIATE EXPECTATIONS AND REQUESTS FOR ANTIBIOTICS IN PRIMARY CARE
Decision Psychology and Shared Decision Making (DEC)

Alistair Thorpe, BSc, Miroslav Sirota, PhD, Sheina Orbell, PhD and Marie Juanchich, PHD, University of Essex, Colchester, United Kingdom

Purpose: We test the hypothesis that clinically inappropriate expectations and requests for antibiotics are driven by a lack of information about illnesses and antibiotic effectivity.

Method(s): In a correlational study (Study 1), participants (n = 402) recalled a situation in which they experienced a cold then expressed their beliefs about the illness, beliefs about antibiotics, and their expectations to receive and/or request antibiotics from a family physician. In a between-subjects experiment (Study 2), participants (n = 190) imagined having a cold then expressed their expectations to receive and/or request antibiotics in a 2 (illness information: baseline knowledge vs. viral nature information) × 2 (antibiotic information: baseline knowledge vs. antibiotic effectivity information) design.

Result(s): In Study 1, we found that illness beliefs, antibiotic beliefs, and past antibiotic usage predicted expectations of the physicians' behaviour $F(22,379)=12.20$, $p < .001$, $R^2 = .414$, expectations of antibiotics as a treatment, $F(22,379)=18.65$, $p < .001$, $R^2 = .520$, and the likelihood of requesting antibiotics $F(22,379)=14.62$, $p < .001$, $R^2 = .459$. For antibiotic beliefs, dimensions related to social perception (e.g., social norms $r = .487$, $r = .505$, $r = .493$, $p < .001$) had a greater influence than dimensions assessing knowledge of antibiotic effectivity, appropriate usage, and resistance (with r values for antibiotic knowledge dimensions ranging from -.196 to -.333, $p < .001$). In Study 2, we found that the provision of information about antibiotics reduced expectations of the physicians' prescribing behaviour, $F(1,186)=22.78$, $p < .001$, $\eta^2_p = .11$ and expectations for antibiotics as a treatment, $F(1,186)=7.55$, $p = .007$, $\eta^2_p = .04$. When controlling for specific dimensions of prior antibiotic and illness beliefs, the provision of antibiotic information reduced the likelihood of requesting antibiotics. The provision of information regarding the nature of the illness did not significantly influence expectations or requests.

Conclusion(s): Both information provision and prior beliefs influence expectations and requests for antibiotics. The provision of information regarding the effectivity and side effects of antibiotics decreases, but does not diminish clinically inappropriate expectations and requests of antibiotics. The effect is relatively small and partly depends on individuals' prior beliefs. Interventions attempting to reduce expectations and requests should explore additional social factors that will lead to a synergic effect.

PS1-29 PERCEPTIONS OF SUBSEQUENT SCREENING PARTICIPATION AFTER A FALSE ALARM FOR COLORECTAL CANCER. A QUALITATIVE STUDY
Decision Psychology and Shared Decision Making (DEC)

Pia Kirkegaard, Department of Public Health Programmes, Randers, Denmark, Adrian Edwards, Division of Population Medicine, Cardiff University, Cardiff, United Kingdom and Berit Andersen, Department of Public Health Programmes, Randers Regional Hospital, Randers NØ, Denmark

Purpose:
To explore how screening participants perceive and manage a positive screening result followed by a negative clinical follow-up investigation.

Method(s):
Screening participants with a positive fecal immunochemical test (FIT) result were identified in the Danish national colorectal cancer (CRC) screening program and invited for interviews 1-2 days before follow-up investigation (colonoscopy), and re-interviews 4-6 weeks after, if they had a negative follow-up result, i.e. false alarm for CRC. Thirty people were invited, three declined participation, and five could not be interviewed for practical reasons. Sixteen participants were selected for interview in their own homes, using a semi-structured interview guide, and six participants were interviewed by phone for data saturation purposes. Transcribed data were analyzed thematically, based on an interpretive tradition of ethnography.

Result(s):
The most prominent themes before colonoscopy were symptom appraisal and communication with family or friends. Most participants attributed the positive FIT result showing blood in the stool to pre-existing non-malignant conditions but a few were very worried. Communication strategies included discussions with family or friends about the positive FIT result and the upcoming colonoscopy, or containing information until the colonoscopy had provided the definitive diagnostic result. There was no apparent need for communication with healthcare professionals before the colonoscopy. The most prominent themes after negative colonoscopy were trust, gratitude, and obligation to participate in subsequent screening. Many participants saw empathetic and patient-involving behavior of the medical staff as a cornerstone of their experience with follow-up colonoscopy. They expressed trust in the ability of the medical staff to perform the colonoscopy and in the validity of the result. The screening program was widely portrayed as a common good. Participation in CRC screening was considered by many a moral and social obligation and a means to stay healthy, and there was a significant motivation for subsequent screening.

Conclusion(s):
A positive screening result may cause worry for some participants, potentially to require support, but most participants considered it unconcerning. The screening participants understood that screening implies risks for false alarm, and they accepted this risk as worth taking, including in future screening. A false alarm for cancer might not reduce screening participants' trust in the benefit of screening or deter them from subsequent screening participation.

PS1-30 IMPROVEMENTS IN THE PRIMARY CARE SETTING ON PSA TESTING FOR PROSTATE CANCER

Decision Psychology and Shared Decision Making (DEC)

Julia van Tol-Geerdink, PhD[1], IM van Oort, MD PhD[1], T Bluemer, MD[2] and DM Somford, MD, PhD[3], (1)Radboud UMC, Nijmegen, Netherlands, (2)Organisation of Chronic Primary Care (OCE), Nijmegen, Netherlands, (3)CWZ Hospital, Nijmegen, Netherlands

Purpose:

PSA screening can reduce prostate cancer mortality (by 1 per 1000 screened men), but is can also lead to overdiagnosis and overtreatment with serious side-effects. Once tested, men come into a habit of yearly PSA tests. And once diagnosed, it is not easy to defer treatment. This project aims to improve the decision-making around PSA test in two ways: 1) a decision aid for patient and family doctor to decide whether to test, and 2) after testing, evidence-based guidelines for the family doctor on the need for referral to the urologist and/ or retest intervals.

Method(s):

Urologists and family doctors were involved in the development of the decision aid and the referral/retest guidelines. Sixty-one family doctors were asked to discuss the decision aid with patients who requested a PSA test, and when PSA was tested, to follow the referral-retest guidelines. The intervention was evaluated by questionnaires. The decisions made, patient knowledge and decisional conflict scores were compared before and after the intervention.

Result(s):

To date, 153 questionnaires were received from 177 patients (86% response); 65 before the intervention (control group) and 88 in the intervention group. After the introduction of the decision aid, patients scored higher on knowledge questions (70% vs. 54% correct, $p<.001$) and lower on Decisional Conflict for Value Clarity (25 ±15 vs. 32 ±23, $p=.048$). In addition, more patients deferred the PSA test (59% vs. 28%, $p<0.001$).

Physicians (54% response rate) were also positive. Of them, 88% endorsed the decision aid and 76% used it in their practice. They gave the decision aid a median rating of 8 out of 10. The guideline was endorsed by 85% of the physicians and used in daily practice by 72%. Main reason not to use the decision aid or guideline, had been that they had forgotten.

Conclusion(s):

Evidence-based tools were developed for family doctors 1) for shared decision making with patients on whether or not to test PSA and 2) after tests, on when to refer or retest. These tools yielded better-informed and less conflicted patients and reduced the number of PSA tests.

PS1-31 VALUES CLARIFICATION ABOUT RADIOTHERAPY IN RECTAL CANCER TREATMENT: INDICATIONS OF BENEFITS FOR DECISION PROCESS AND QUALITY OF LIFE

Decision Psychology and Shared Decision Making (DEC)

Arwen H. Pieterse, PhD[1], Marleen Kunneman, PhD[2], Wilbert B. van den Hout, PhD[1], Monique C.M. Baas-Thijssen, RN[3], Corrie A.M. Marijnen, MD PhD[3] and Anne M. Stiggelbout, PhD[1], (1)Medical Decision Making, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands, (2)Knowledge and Evaluation Research (KER) Unit, Mayo Clinic; and Medical Decision Making, Dept of Biomedical Data Sciences, Leiden University Medical Center, Rochester, MN, (3)Dept of Radiotherapy, Leiden University Medical Center, Leiden, Netherlands

Purpose:

To assess whether a values clarification method (VCM) helps newly-diagnosed rectal cancer patients to make trade-offs in deciding about pre-operative radiotherapy.

Method(s):

In a cohort of rectal cancer patients, we offered an adaptive conjoint analysis-based VCM to a subset of patients (VCM-group). We audiotaped consecutive encounters with radiation oncologists in which the decision to undergo pre-operative radiotherapy or not was discussed, and used a study-specific coding scheme to code patient treatment-related preferences. Patients were asked to complete questionnaires pre-encounter (T0), immediately post-encounter (T1), and at six (T2) and 12 (T3) months follow-up to assess: socio-demographics and anxiety (STAI-6) (T0), values clarity (Decisional Conflict Scale; T1), regret (Decision Regret Scale; T2-3), presence of treatment harms (faecal incontinence, sexual problems; EORTC QLQ-CR29) and their impact (Incontinence Impact Questionnaire, Sexual Distress Scale) (T2-3). Patients were interviewed about their valuation of treatment outcomes (Time Trade-Off) and those in the VCM-group about the VCM. Two raters coded the audiotapes for expression of treatment-related preferences. Missing questionnaire data were multiply imputed.

Difference between the patients in the VCM versus care-as-usual group regarding: 1) expression of treatment-related preferences, 2) utility for treatment harms, 3) evaluation of the decision making process, and 4) impact of treatment harms at follow-up were tested using median tests or using linear regression analyses correcting for baseline differences, as appropriate.

Result(s):

N=135 patients participated in the full study; additionally N=38 had only their encounter taped. The patients in the VCM (N=33) versus care-as-usual (N=102) group did not significantly differ on baseline characteristics, except that the VCM group comprised more women (45% *vs*. 26%, *P*=0.04). In the VCM (N=33) versus care-as-usual (N=140) group significantly more often a treatment-related preference was expressed. the VCM (N=33) versus care-as-usual (N=102) group did not significantly differ in values clarity, nor in presence of incontinence or sexual problems at six or 12 months follow-up. The VCM (N=33) versus care-as-usual (N=102) group experienced significantly lower regret and less impact of treatment harms at six-months follow-up; differences were in the same direction but not significant at 12 months.

Conclusion(s):

Being explicitly invited to think about treatment benefits/harms seems beneficial to foster the expression of preferences and in living with treatment harms.

PS1-33 FROM INDIVIDUAL TO POPULATION PREFERENCES: RANDOM UTILITY MODELS VS DIRICHLET DISTRIBUTION

Patient and Stakeholder Preferences and Engagement (PSPE)

Tommi Tervonen, PhD, Evidera, London, United Kingdom, Francesco Pignatti, MD, European Medicines Agency, London, United Kingdom and Douwe Postmus, PhD, University Medical Center Groningen, Groningen, Netherlands

Purpose: To assess variability and convergence of three parametric approaches for estimating the average treatment benefit-risk trade-offs.

Method(s): Individual-level trade-offs from a three-attribute patient preference study (n=560) were used in a simulation study to assess measurement error of conditional and random parameter logit models applied to individual choice data, and Dirichlet distribution applied to individual trade-off data. Discrete choice experiments (DCE) were conducted for sample sizes from 20 to 500, where in each one subjects drawn from the empirical distribution were simulated to answer 5-16 DCE questions. Conditional and random parameter logit models were fit to the answer data, and a Dirichlet distribution estimated for the trade-off weights that were originally obtained with 6 adaptive pair-wise questions. Minimum sample size to achieve convergence was calculated as Mean Squared Error (MSE) <0.01 (p=0.05), where error was defined as Euclidean distance of the obtained estimate and the estimate with the full sample (n=560, 16 questions). Logit model variability was additionally assessed with coefficient p-values.

Result(s): Conditional logit model converged to <0.01 MSE with a minimum of 6 questions and sample size 200. Random parameter logit did not converge to <0.01 MSE in these experiments, whereas Dirichlet distribution converged to <0.01 MSE with sample size 40, and to <0.001 MSE (p=0.05) with sample size 180. Conditional logit model had statistically non-significant (p>0.2) estimates for the least important attribute in 3-6 out of 20 instances with 6 questions, when sample size varied between 200-300.

Conclusion(s): Logit models could not reliably estimate population preferences in this low-dimensionality data set when the same number of questions (six) was used in the simulated DCE as was originally used for collecting trade-off data using an adaptive method. Dirichlet distribution is a promising alternative for describing population preferences based on individual preference estimates, outperforming simulation of random utility models with this data set. Future research should assess generalizability of these results with other data sets, and in simulation studies.

Disclaimer: The views presented here are those of the authors and not of their organizations.

PS1-34 HEALTH-RELATED QUALITY OF LIFE AFTER TRAUMATIC BRAIN INJURY: DERIVING A VALUE SET FOR THE QOLIBRI-OS AND QOLIBRI IN 3 EUROPEAN COUNTRIES

Patient and Stakeholder Preferences and Engagement (PSPE)

Daphne C. Voormolen, MSc[1], Suzanne Polinder, PhD[1], Nicole von Steinbuechel[2], Mark Oppe, PhD[3] and Juanita A. Haagsma, PhD[1], (1)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (2)University of Gottingen, Gottingen, Germany, (3)EuroQol Research Foundation, Rotterdam, Netherlands

Purpose: The Quality of Life after Brain Injury (QOLIBRI) is an instrument specifically developed to measure health-related quality of life (HRQoL) of individuals after traumatic brain injury (TBI). The QOLIBRI Overall Scale (QOLIBRI-OS) is an indexed version of the complete scale. The objective of our study was to derive value sets for the QOLIBRI-OS/QOLIBRI in the Netherlands (NL), United Kingdom (UK) and Italy (IT) by using Discrete Choice Experiments (DCEs) in the general population. These value sets allow calculation of utility weights for TBI health states measured with the QOLIBRI-OS/QOLIBRI. These data can ultimately be used in economic evaluations, Comparative Effectiveness Research (CER) and decision-making in the field of traumatic brain injury (TBI).

Method(s): A stepwise procedure was used to develop value sets for the QOLIBRI-OS/QOLIBRI. First, a Rasch analysis was performed on QOLIBRI-OS/QOLIBRI data of 1000 TBI patients that was collected six months post-injury. With this Rasch analysis the number of instrument items was reduced to five items. Second, a selection was made of approximately 400 health states that covered the entire spectrum of severity to be used in the health state valuation task. Third, a panel of judges, consisting of members of the general public of NL, UK and IT evaluated the selected health states. Data were collected using a web-based survey with 16 DCE questions. Lastly, DCE answers were analysed using a conditional logistic regression and the utilities and weights were calculated by using a multinomial slope model.

Result(s): The data analyses resulted in estimates of the values for all potential health states according to the responses for the selected health states. A total of 11.759 respondents completed the questionnaire. To check if respondents understood the DCEs, we calculated the sum scores per health state, subtracted health state B from A and analysed the probability of respondents choosing A (Fig.1). A utility score for the QOLIBRI-OS/QOLIBRI health states is generated from the DCE responses. All these results are still preliminary, because not all data analyses have been finalized yet.

Conclusion(s): We conclude that by transforming the QOLIBRI-OS/QOLIBRI into utility scores we have paved the way for application in CER and in summary measures of population health, which informs decision-makers on the best interventions and strategies for TBI patients.

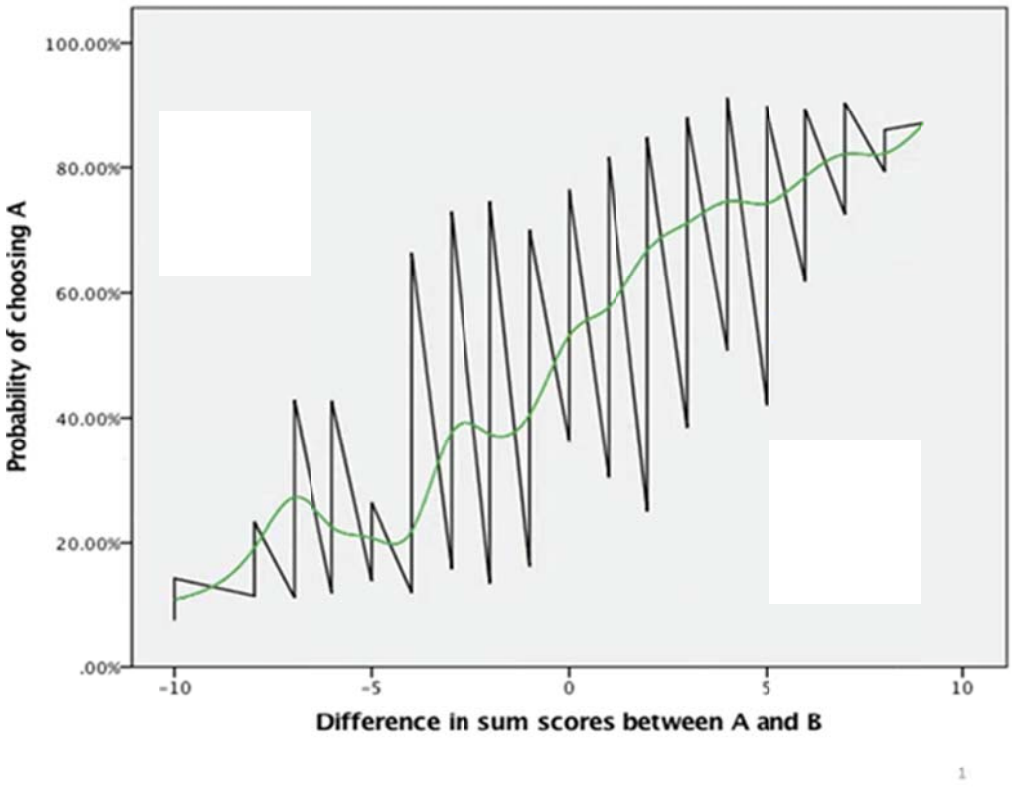


Figure 1. The probability of choosing health state A when looking at the difference in sum scores between health state A and B.

PS1-35 PSYCHOLOGICAL FACTORS AFFECTING PATIENT PREFERENCES: REVIEW OF PSYCHOLOGICAL MEASURES USED IN PATIENT PREFERENCES AND DECISION STUDIES

Patient and Stakeholder Preferences and Engagement (PSPE)

Selena Russo, PhD[1], Ulrik Kihlbom, PhD[2], Cathy Anne Pinto, PhD[3], Chiara Jongerius, MSc[4], Flavia Faccio[4], Silvia Francesca Maria Pizzoli[4], Rosanne Janssens[5], Gwenda Simon, MA PhD[6], Marie Falahee, MA PhD[6], Isabelle Huys, MSc, PhD[7], Esther de Bekker-Grob, PhD[8], Jorien Veldwijk, PhD[9], Douwe Postmus, PhD[10], **Dario Monzani, PhD**[11] and Gabriella Pravettoni, PhD[4], (1)Behavioural Sciences Unit, Kids Cancer Centre, Sydney Children's Hospital,, Randwick, Australia, (2)Centre for research ethics and bioethics Uppsala University, Uppsala, Sweden, (3)Merck Research Laboratories, Center for Observational and Realworld Evidence, Rahway, NJ, (4)Applied Research Division for Cognitive and Psychological Science, IEO European Institute of Oncology, Milan, Italy, (5)Department of Pharmaceutical and Pharmacological Sciences, University of Leuven, Leuven, Belgium, (6)Rheumatology Research Group Institute of Inflammation and Ageing (IIA) University of Birmingham, Birmingham, United Kingdom, (7)KU Leuven, Leuven, Belgium, (8)Erasmus University Rotterdam, Rotterdam, Netherlands, (9)Erasmus Choice Modelling Centre, Erasmus University , Rotterdam, Netherlands, (10)University Medical Center Groningen, Groningen, Netherlands, (11)Applied Research Division for Cognitive and Psychological Science, IEO European Institute of Oncology, Milan, Italy, Milan, Italy

Purpose: To date, research has been mainly focussed on how to elicit patient preferences, with less attention on why patients form certain preferences. A systematic review was conducted to assess what psychological instruments are currently used and which psychological constructs are known to have an impact on patient preferences including the formation of preferences and preference heterogeneity.

Method(s): A systematic literature search was conducted using PubMed, MEDLINE, Embase, PsycINFO, and Google Scholar to identify studies that used a validated instrument to assess psychological variables in relation to patients' preferences. The articles identified by the search were screened according to the following inclusion criteria: i) Studies including the development, validation, evaluation or use of an instrument measuring psychological dimensions in relation to patients' preference elicitation or health-related decisions; ii) The instruments undertook a validation process and present psychometric information on validity and reliability; iii) Quantitative studies; iv) Focus on human beings; v) Full text available; vi) English language; vii) Studies published after 1980. The results of this systematic review were supplemented with psychological constructs and instruments identified by a panel of international experts.

Result(s): A total of 56 instruments measuring 29 constructs were identified, of which 30 instruments resulted from the systematic review and 26 from the expert panel. Health literacy, numeracy, patient activation and health locus of control influence patients' preferences and decisions and measurements of those constructs have shown consistent results. The measurements of self-efficacy, health-beliefs and of risk propensity are promising fields of study, but the number and consistency of existing results is not yet satisfactory. Measurements of personality traits, anxiety and depression offer inconclusive or inconsistent results.

Conclusion(s): Results outline some initial recommendations. Based on the empirical evidence demonstrating an influence of health literacy, numeracy and patient activation on the formation of preference, it is advisable to incorporate instruments in future patient preference studies that measure these constructs. Within the category of psychological constructs that could explain preference heterogeneity, health locus of control is a strong predictor of decisions in several healthcare contexts and is useful to consider when designing patient preference study. Future research should continue to explore the association of psychological constructs with preference formation and heterogeneity to build on these initial recommendations.

PS2-1 STRIVING FOR A SOCIETAL PERSPECTIVE: AN ANALYTICAL FRAMEWORK FOR THE ECONOMIC EVALUATIONS OF INTERVENTIONS WITH COSTS AND EFFECTS FALLING ON MULTIPLE SECTORS AND DECISION MAKERS

Health Services, Outcomes and Policy Research (HSOP)

Simon Walker, MSc[1], Susan Griffin, PhD[2], Miqdad Asaria, PhD[1] and Mark Sculpher, PhD[2], (1)Centre for Health Economics, University of York, York, United Kingdom, (2)University of York, York, United Kingdom

Purpose:

Conventional application of cost-effectiveness analysis assumes a single healthcare budget used with the objective to improve health. In reality, many decisions are more complex - impacting budgets held by varied decision-makers, whose mandates differ, and whose interests extend beyond health. A naive 'societal perspective' does not address situations where heterogenous priorities of multiple decision-makers conflict, and adding costs across multiple budgets conceals important differences in opportunity cost. This study presents how economic evaluation can be conducted to inform and impartially input into deliberations of decision-makers in such contexts.

Method(s):

A framework was developed to capture the impact of interventions on multiple outcomes and multiple budgets, allowing for competing objectives of decision makers (including various notions of equity). We set out an "impact inventory" which captures an intervention's effect on different individuals in terms of both direct impacts and opportunity costs. We then illustrate alternative approaches for aggregating impacts, the societal value judgements these imply, and their impact on ranking of interventions.

Result(s):

The "impact inventory" sets out a series of assessments for an analyst to make when performing evaluation to inform a range of public sector decision-makers. These assessments reflect the institutional arrangements of public sector bodies (e.g. funding allocation) as well as their valuations of different outcomes and any equity considerations. The approach provides information on cost-effectiveness from a range of perspectives, and facilitates deliberation between different decision makers when there are conflicting objectives by making the trade-offs and value judgements explicit. The social impact inventory of the second Cost-Effectiveness Panel is a special, restricted case of this more general approach

Conclusion(s):

Through appropriate assessment of the evidence and by clearly distinguishing the points at which value judgements feed into the evaluation process, assessments of interventions with multiple outcomes and costs falling across different sectors can reliably be informed.

PS2-2 EFFICIENT ASSIGNMENT IN REGRESSION DISCONTINUITY DESIGN: VALIDATION IN TWO LARGE RANDOMIZED TRIALS

Health Services, Outcomes and Policy Research (HSOP)

Nikki van Leeuwen, MSc[1], Hester F. Lingsma, PhD[2], Daan Nieboer[1] and Ewout W. Steyerberg, PhD[3], (1)Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, Netherlands, (2)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (3)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands; Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, Netherlands

Purpose:

The regression discontinuity (RD) design is a quasi-experimental design to study effectiveness of treatments. In the RD design treatment is assigned to a subset of patients based on a baseline variable; e.g. older patients receive treatment while younger patients do not. Although RD might provide valid treatment effect estimates, it requires substantially more patients then an RCT to obtain the same reliability. We aimed to compare different assignment approaches to increase the statistical efficiency of the RD design.

Method(s):

We used data from two RCTs (CRASH and GUSTO) to resemble RD with different strategies to assign treatment, using variables that correlated weakly and strongly with the outcome under study (14-day and 30-mortality respectively). The treatment effect was analysed with local logistic regression and logistic regression with restricted cubic spline and polynomial adjustment. To assess the difference in statistical efficiency, standard errors were compared between the different treatment assignment strategies. The gain in efficiency was expressed in number of patients needed in reference to an RCT.

Result(s):

Preliminary analysis showed that the treatment effect estimates had lower standard errors (SE 0.09 and 0.08) when the assignment variable poorly correlated with the outcome, compared to assignment based on variables strongly correlating with outcome (SE 0.19 and 0.21 for CRASH and GUSTO respectively). Up to 1.2 times as many patients in RD based on an assignment model poorly correlating with outcome, compared to 5.5 to 7.7 times as many patients in RD based on an assignment model highly correlating with outcome, were needed to obtain the same efficiency as in an RCT.

Conclusion(s):

The inefficiency of RD design may be partly remedied by using an assignment variable that poorly correlates with outcome, but depends more on the method of analysis.

PS2-3 EFFECTS OF SHARED DECISION-MAKING COMPARED TO USUAL CARE FOR PROSTATE CANCER SCREENING AND TREATMENT: A SYSTEMATIC REVIEW AND META-ANALYSIS OF RANDOMISED CONTROLLED TRIALS
Health Services, Outcomes and Policy Research (HSOP)

Nahara Martinez-Gonzalez, Stefan Neuner-Jehle, Andreas Plate, Stefan Markun, Thomas Rosemann and Oliver Senn, Institute of Primary Care, University of Zurich and University Hospital of Zurich, Zurich, Switzerland

Purpose: Shared decision-making (SDM) is highly recommended for medical decisions regarding prostate cancer (PC) screening and treatment. We synthesized the evidence on the effectiveness of SDM compared to usual care.

Method(s): We comprehensively and systematically searched for primary peer-reviewed and grey literature of randomised controlled trials (RCTs) conducted in primary and specialized care that compared SDM to usual care for PC. We included and critically appraised eligible RCTs published in English. We calculated the unadjusted study-specific and pooled relative risks (RR) or standardized mean differences (SMD), and (where applicable) performed random-effects meta-analyses**.**

Result(s): After evaluating the full-text of 270 reports, eight RCTs comprising 2,825 patients met the inclusion criteria. The studies were conducted mainly in North America and represent diverse interventions. *PC screening* was addressed in four RCTs. Meta-analyses (of two studies at most) showed knowledge improvement with SDM (SMD 0.23, 95%CI 0.02 to 0.43), but SDM was not different to usual care in reducing patient participation rates (RR 1.03, 95%CI 0.90 to 1.19) or decisional conflict (SMD -0.04, 95%CI -0.23 to 0.15; SMD -0.05, 95%CI -0.24 to 0.14). There were other fifteen outcome measures of interest. In 40% (n=6/15) of the individual trial estimates, SDM reduced physicians' tendency for PC screening, and improved the accuracy of patients' perception of lifetime-risks and men's views towards screening. The other 60% showed no significant differences between SDM and usual care for other outcomes. *PC treatment* was addressed in four RCTs. The studies reported twenty-four outcome measures of interest. Of these, 37.5% individual trial estimates showed that SDM significantly improved knowledge, perception of being informed, and patient-perceived quality of life (QoL) at four weeks after the intervention. The other 62.5% individual trial estimates resulted in no significant differences between groups for other outcomes. Overall, the studies are generally small and prone to high risk of bias. There is an important dearth of data on the effects of SDM on disease-relevant endpoints such as symptoms and mortality.

Conclusion(s): SDM may positively influence men's knowledge of PC screening and treatment. SDM for PC treatment may have a positive short-term effect on patient-perceived QoL. The evidence on SDM effectiveness is limited mainly by the absence of disease-specific and clinically relevant outcome measures.

PS2-4 PREDICTION MODELS FOR FUNCTIONAL OUTCOME 6-MONTHS AND 1-YEAR AFTER TRAUMA

Health Services, Outcomes and Policy Research (HSOP)

Leonie de Munter, MSc[1], Suzanne Polinder, PhD[2], Cornelis van de Ree, MD[1], Nena Kruithof, MSc[1], Ewout W. Steyerberg, PhD[3] and Mariska de Jongh, PhD[1], (1)ETZ, Xperiment Trauma TopZorg, Tilburg, Netherlands, (2)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (3)Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands

Purpose: Trauma patients often perceive an impaired health status after trauma. This study aims to develop a prediction model for health status in the general trauma population, based on 6-months and 12-months outcome. The models should incorporate easily accessible predictors.

Method(s): The Brabant Injury Outcome Surveillance (BIOS-study) is a multicenter prospective observational cohort study. Adult trauma patients were included from August 2015 through November 2016 if they were admitted to one of the ten hospitals in the county Noord-Brabant, the Netherlands. Outcome measures were the EuroQol-5 dimensions (EQ-5D) and the Health Utilities Index (HUI) 6 months and 1 year after trauma. Summary scores were calculated for all outcome measures. Possible predictors were pre-injury health status, injury severity, patient characteristics and frailty pre-injury (measured with the Groningen Frailty Index). All potential predictors were assessed with univariable linear regression. Predictors were included in the multivariable model if $p<0,2$. The model performances were assessed with Nagelkerke R-square (R^2).

Result(s): A total of 2,106 patients and 1,938 patients were used to develop a prediction model for, respectively, 6 months and 1 year functional outcome. Pre-injury health status and frailty pre-injury were the strongest predictors for functional outcome in the general trauma population. Age, comorbidity, social economic status, functional capacity index, injury severity score, length of stay in hospital were also included in the multivariable prediction models. The model explained 52% of the variance for eq5D-utility ($R^2=0.52$) and 45% of the variance was explained for HUI 2 or HUI 3 index ($R^2=0.45$).

Conclusion(s): To our knowledge, these are the first models to predict functional outcome 6 months and 1 year after trauma. The models seem promising for predicting health status of trauma patients in the western society. However, future research is recommended to externally validate the models.

PS2-5 VALIDATION, CALIBRATION, AND CLINICAL USEFULNESS OF FINDRISC, AUSDRISK, AND ADA RISK SCORES FOR DETECTING 5-YEAR INCIDENT TYPE 2 DIABETES IN TEHRAN LIPID AND GLUCOSE STUDY (TLGS)
Health Services, Outcomes and Policy Research (HSOP)

Mojtaba Lotfaliany, MD[1], Brian Oldenburg, PhD[1], Fereidoun Azizi, MD[2] and **Davood Khalili, MD PhD**[3], (1)Non-Communicable Disease Control Unit (NCDU), Center for Health Equity, School of Population and Global Health, University of Melbourne, Melbourne, Australia, (2)Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran (Islamic Republic of), (3)Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran (Islamic Republic of)

Purpose:

A number of non-invasive diabetes risk score were developed; very few were implemented in practice. Three of these risk scores namely FINDRISK, AUSDRISK, and ADA risk score utilized in different clinical trials and showed very promising results. Current study aimed to validate, calibrate, and assess clinical usefulness of FINDRISC, AUSDRISK, and ADA risk scores for detecting 5-year risk of incident type 2 diabetes in the Tehran Lipid and Glucose Study (TLGS), a long-term ongoing community-representative cohort study with a triennial basis data collection conducted in the capital city of Iran, Tehran.

Method(s):

Data from 2682 participants in phase 3 of TLGS (2005-2007) who were followed in phase 5 of TLGS (2011-2013) was used to assess the performance of the diabetes risk scores. To avoid over-fitting in calibration estimates, the re-calibrated and calibrated-in-large models were derived using the data from 2594 participants in phase 1 (1999-2001) of study who were followed in phase 3 (2005-2007). Separate logistic regressions were fitted in the phase 1 data-set to derive “calibration-in-large” (calibrating the model intercept) and “re-calibrated” (calibrating the model intercept and the slope) models for each risk score. To evaluate the performance of the models, area under the receiver operating characteristic curve (AUC), and the Hosmer-Lemeshow chi-squared were estimated for each risk score. Clinical usefulness of the models were assessed by plotting net benefit fraction for each risk score. Maximum benefit was derived by assessing the net benefit fraction at the threshold equal to 5-year cumulative incidence rate of type 2 diabetes.

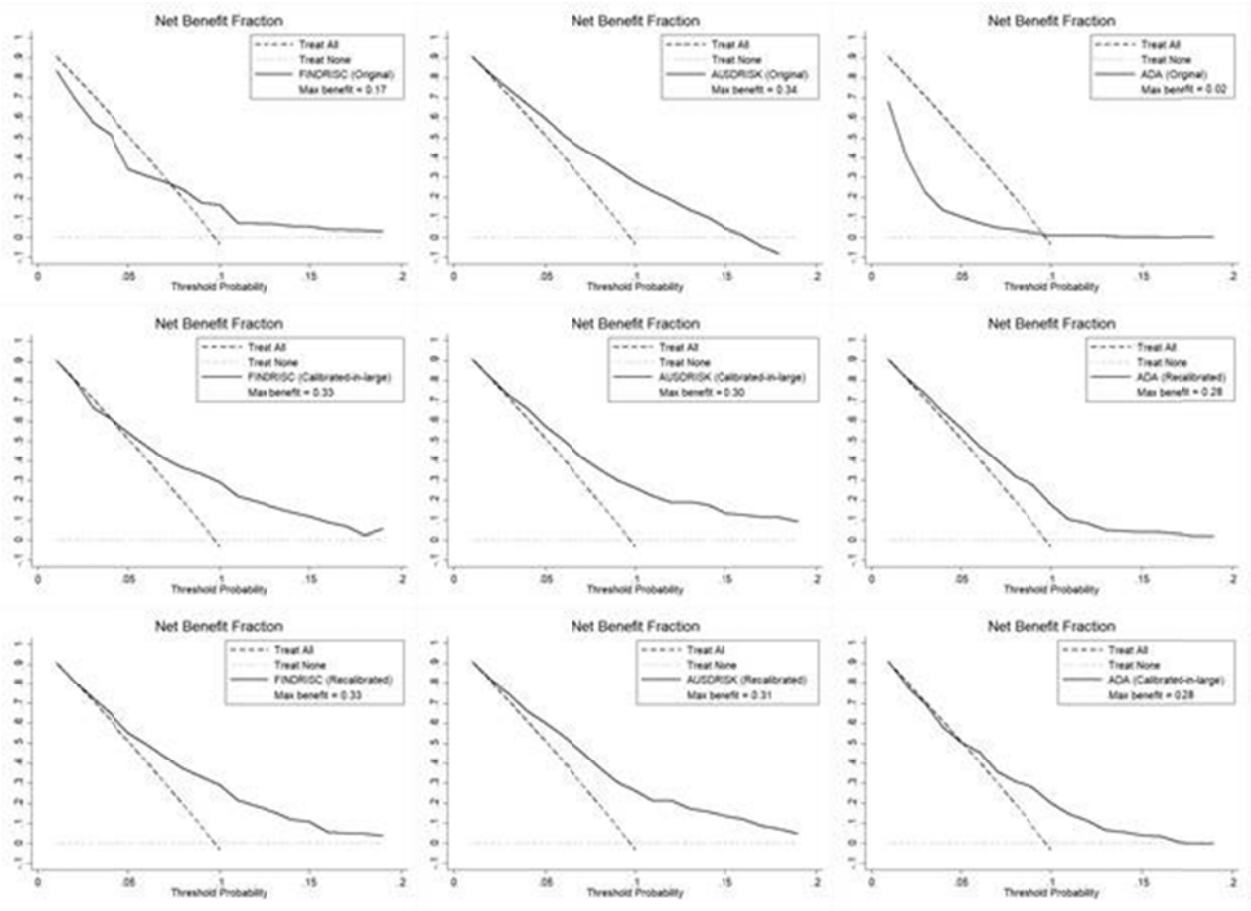
Result(s):

Cumulative incidence of type 2 diabetes was 9.7%. Regarding the discrimination power of the models, AUSDRISK had the highest AUC (0.722). Regaining the calibration, among original risk scores, AUSDRISK with $Chi^2 = 175.4$, among the calibrated-in-large risk scores, FINDRISC with $Chi^2=50.1$, and among re-calibrated risk scores, FINDRISC with $Chi^2=14.0$ had the best performance. Among original risk scores, AUSDRISK with maximum net benefit fraction of 0.34, among the calibrated-in large risk score, FINDRISC with maximum net benefit

fraction of 0.33, and among recalibrated risk scores, FINDRISC with maximum net benefit fraction of 0.33 had the best usefulness (Figure).

Conclusion(s):

Original model for AUSDRISK and calibrated-in-large model for FINDRISC with threshold risk of 10% is recommended to be used for screening of high risk individuals in Iranian population.



PS2-6 DECISION AIDS AND CONSECUTIVE TREATMENTS: THE CASE OF MULTIPLE SCLEROSIS

Decision Psychology and Shared Decision Making (DEC)

Ana Manzano, PhD[1], **Hilary Bekker, PhD**[2], Ieva Eskyte, PhD[1], Sue Pavitt, BSc, PhD[1], Helen Ford[3], George Pepper[4], David Meads, PhD[5], Jeremy Chataway[6], Klaus Schmierer[7] and Edward Webb[8], (1)University of Leeds, Leeds, United Kingdom, (2)Leeds Institute of Health Sciences School of Medicine - University of Leeds, Leeds, United Kingdom, (3)Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom, (4)Shift.ms, Leeds, United Kingdom, (5)Leeds Institute of Health Sciences, Leeds, United Kingdom, (6)University College of London, London, United Kingdom, (7)Barts MS, Blizard Institute, London, United Kingdom, (8)Leeds Institute of health Sciences, Leeds, United Kingdom

Purpose: To develop a discrete choice experiment (DCE) task for application to a patient decision aid supporting people with relapsing remitting multiple sclerosis (PwRRMS) make disease modifying treatment (DMT) choices across an escalating and de-escalating treatment course.

Method(s): Mixed-methods including: 1) A critical literature review to generate a framework for conceptualising the complex context within which decision making takes place; 2) Focus groups methods (n=4) including PwRRMS (n=5), neurologists (n=8) and specialist nurses (n=4) exploring preferences for different DMTs; 3) Semi-structured interviews with PwRRMS (n=30) to explore factors explaining decision making to start a treatment regimen, switch and/or stop, timing of decision points, and preferences for treatment; a) an on-line DCE using a think-aloud method to pilot the task with PwRRMS.

Result(s): Across relapse remitting multiple sclerosis some PwRRMS only ever choose one treatment, and others move through several consecutive treatments escalating and de-escalating associated risks. Making these treatment choices is emotionally demanding, either as a first step onto the treatment pathway when people's understanding and experiences of treatment is impoverished, and/or switching when worrying about the rate of progression of MS. There was evidence that some PwRRMS preferred to defer decision making rather than make a choice about a riskier treatment. There were also misunderstanding about treatment outcomes and increasing risks of treatment options along the disease continuum. There seemed to be a significant need for information to support PwRRMS switching between treatment regimens, both in terms of the treatment options, their burden and implications for understanding a fluctuating but progressive disease state.

Conclusion(s):

These data illustrate a complexity between treatment decision making and progressive disease states not currently recognised by patient decision aid development and evaluation guidance. Patient decision aids supporting PwRRMS will need to help support trade-offs across the disease trajectory.

PS2-8 SYSTEMATIC OVERVIEW OF PATIENT-REPORTED OUTCOMES FOR MYELODYSPLASTIC SYNDROMES

Health Services, Outcomes and Policy Research (HSOP)

Igor Stojkov, MPharm[1], Ursula Rochau, MD, MSc[2], Conrads-Frank Annette, PhD[2], Helena Borba, PhD[3], Karin Koinig, PhD[4], Marjan Arvandi, MS[2], Fabio Efficace, MD, PhD[5], Reinhard Stauder, MD, PhD[4] and Uwe Siebert, MD, MPH, MSc, ScD[6], (1)UMIT - University for Health Sciences, Medical Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA, Hall i.T., Austria, (2)Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (3)Department of Pharmacy, Pharmaceutical Sciences Postgraduate Research Program, Federal University of Parana, Curitiba, Brazil, (4)Department of Internal Medicine V (Hematology and Oncology), Innsbruck Medical University, Innsbruck, Austria, (5)Health Outcomes Research Unit, Gruppo Italiano Malattie Ematologiche dell'Adulto (GIMEMA), Rome, Italy, (6)UMIT, Dept. of Public Health, Health Services Research & Health Technology Assessment / Harvard Univ., Dept. Health Policy & Management, Institute for Technology Assessment / ONCOTYROL, Division for HTA, Hall i. T. (Austria) / Boston (USA), Austria

Purpose: Myelodysplastic syndromes (MDS) are a heterogeneous group of clonal myeloid neoplasms typically observed in aged populations. Beside clinical parameters, patient-reported outcomes (PROs) are an additional asset in the disease management. We aimed to provide an overview of quality-of-life measurement instruments (QoL-MI) as a multidimensional source of PROs, from which patients will select the MDS core PROs in a subsequent survey. The core PROs will contribute to a new MDS-core outcome set (MDS-COS).

Method(s): To identify studies applying QoL-MI in MDS, we performed a comprehensive systematic literature search (PubMed, Cochrane Library, Scopus, Web of Science). We excluded studies measuring solely one health domain or studies without an available English version of the QoL-MI. Studies were screened by at least two independent reviewers (IS/HB/UR) and the observed QoL-MI were summarized in an evidence table. The domains and items within each of the observed QoL-MI were extracted and categorized into potential PROs by a group of six experienced researchers and clinical experts.

Result(s): Our first search resulted in 2863 studies; 81 studies were included. Overall, we identified twelve generic, six cancer-specific and two MDS-specific QoL-MI. The most commonly used instruments were EORTC QLQ-C30 (26 studies), FACIT-An (17 studies), SF-36 (16 studies), QoL-E (10 studies), and FACIT-BMT (7 studies). These five instruments composed 68% of the total 112 QoL-MI, whereas the MDS-specific QoL-MIs were only applied in 10%. After the experts' categorization, the following 40 PROs were derived: shortness of breath, weakness, need to rest, fatigue, sleep disturbances, emotional wellbeing, depression, fear of MDS progression or transformation to leukemia, body image, loss of independence, physical activity, ability to work/activities of daily living, basic mobility, memory difficulties/cognition, speaking difficulties/speech-language problems, transfusion dependence, medication use, fear of side effects of treatment, pain, headache, colds/infections, urinary incontinence, hearing

problems, eye problems, impatience, tremor, skin problems, loss of appetite, loss of weight, vomiting/nausea, change in sense of taste, defecation/change in digestion, financial difficulties, relationship with friends/relatives/partner, sexuality/sexual activity, confidence in health care services, disease knowledge, general health, hospital dependence, and general quality of life.

Conclusion(s): Overall, QoL assessment in MDS patients was rare and MDS-specific QoL-MIs were scarce. To incorporate the patients' perspective in the first MDS-COS, we will initiate a Delphi selection process for essential PROs, including clinical experts and MDS patients.

PS2-9 TO BRIDGE OR NOT TO BRIDGE: MODELLING PERIPROCEDURAL ANTICOAGULATION MANAGEMENT

Health Services, Outcomes and Policy Research (HSOP)

Simon van der Pol, PharmD[1], Maartje Jacobs, PharmD[2], Karina Meijer, MD, PhD[1], Margriet Piersma, MD[3], Robert Tieleman, MD, PhD[2], Maarten Postma, PhD[4] and Marinus van Hulst, PharmD, PhD[2], (1)University Medical Center Groningen, Groningen, Netherlands, (2)Martini Hospital, Groningen, Netherlands, (3)Certe Thrombosis Service Groningen, Groningen, Netherlands, (4)University of Groningen, Groningen, Netherlands

Purpose: For atrial fibrillation (AF) patients receiving vitamin K antagonists (VKAs), careful management of anticoagulation is important around surgical procedures to minimize the stroke and bleeding risks. If the VKA needs to be stopped periprocedurally to reduce the risk of bleeding, a decision needs to be made whether to bridge this period with a low-molecular weight heparin (LMWH). We aimed to develop a model to compare two periprocedural strategies for AF patients that have to interrupt VKA treatment: administering a LMWH or forgoing bridging therapy.

Method(s): A probabilistic Markov model was developed to simulate both a bridge and a non-bridge cohort of AF patients periprocedurally. Modelled events were based on the clinically used CHA_2DS_2-VASc and HAS-BLED stroke and bleeding prediction rules. To predict strokes, INR values were considered. Quality-adjusted life expectancy, based on the beforementioned clinical endpoints, was the main outcome considered.

Result(s): The base case analysis shows that bridging anticoagulation increases the bleeding rate, but reduces the stroke rate. Bridging may be beneficial for patients with a CHA_2DS_2-VASc scores of 6 or higher and HAS-BLED scores of 0 to 2. For expected shorter periods to reach therapeutic INR, bridging therapy is less likely to be beneficial.

Conclusion(s): For patients at high risk of bleeding, bridging anticoagulation Is not likely to be beneficial. For patients at high risk of stroke and low risk of bleeding, bridging anticoagulation may result in additional quality adjusted life years. INR management is an important factor to consider periprocedurally when making the decision whether to bridge.

PS2-11 CONSISTENCY BETWEEN SCORE AND GLOBORISK PREDICTION MODELS FOR CARDIOVASCULAR DISEASE MORTALITY IN A WESTERN ASIAN COUNTRY

Quantitative Methods and Theoretical Developments (QMTD)

Noushin Fahimfar, MD, PhD candidate of Epidemiology[1], Akbar Fotouhi, MD, PhD[1], Mohammad Ali Mansournia, MD, PhD[1], Reza Malekzadeh, MD[2], Fereidoun Azizi, MD[3], Marjan Mansourian, PhD[4], Sadaf G.Sepanlou, MD, PhD[2], Farzad Hadaegh, MD[5], Mohammad Hassan Emamian, MD, PhD[6], Hamid Roohafza, MD[4], Hossein Poustchi, MD, PhD[2], Akram Pourshams, MD[2], Ewout W. Steyerberg, PhD[7] and Davood Khalili, MD PhD[5], (1)Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran (Islamic Republic of), (2)Digestive Diseases Research Center, Digestive Diseases Research Institute, Tehran University of Medical Sciences, Tehran, Iran (Islamic Republic of), (3)Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran (Islamic Republic of), (4)Isfahan Cardiovascular Research Center, Isfahan Cardiovascular Research Institute, Isfahan University of Medical Sciences, Isfahan, Iran (Islamic Republic of), (5)Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran (Islamic Republic of), (6)Ophthalmic Epidemiology Research Center, Shahroud University of Medical Sciences, Shahroud, Iran (Islamic Republic of), (7)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands; Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, Netherlands

Purpose: To find the consistency between SCORE and Globorisk, prediction models developed from pooled cohort studies in Europe and America, respectively, to predict cardiovascular mortality in Iranian population.

Method(s): We included 24427 participants (11187 men) aged 40 to 80 years from four population-based cohort studies in Iran. Updating for the baseline survival and the slope of the linear predictor were applied. The models' discrimination was assessed using C-index and calibration plots were used to compare the predicted probability to the observed risk. The agreement between risks, predicted by the models, was assessed using intraclass correlation coefficient (ICC) index and Bland-Altman plot. Kappa index was also used to evaluate the agreement of the models at the threshold probability of 5%. Decision curve analysis was used to compare the clinical usefulness of the models in a wide range of thresholds.

Result(s): The 10-year observed risk were 0.042 (95%CI: 0.037-0.048) and 0.021 (0.018-0.025) in men and women, respectively. The average predicted risk in men were 0.044 and 0.046 for SCORE and Globorisk, respectively. These values were 0.022 and 0.026 in women. Discrimination power of the models were 0.784 (95%CI: 0.756-0.811) for SCORE and 0.793 (0.766-0.820) for Globorisk in men; the corresponding values in women were 0.778 (0.742-0.814) and 0.792 (0.757-0.828), respectively. Based on decision curve analysis, the models' net benefits were the same in men, but with the superiority of the Globorisk in risks higher than 7%. The Kappa index at the threshold probability of 5% was 0.92 (95% CI: 0.91-0.93) and 0.86 (0.85-0.87) in men and women, respectively. In line with kappa, the ICC was estimated as 0.96 (95%CI: 0.95-0.96) in men and 0.89 (0.88-0.89) in women. Bland-Altman plot graphically showed good agreement between the models.

Conclusion(s): Although SCORE and Globorisk functions have been derived in different cohorts with different methodologies, both recalibrated models have a good agreement to estimate the cardiovascular mortality risk in Iranian population.

PS2-13 EARLY IDENTIFICATION OF MAJOR TRAUMA: EXTERNAL VALIDATION AND COMPARISON OF PROGNOSTIC PREHOSPITAL TRAUMA MODELS IN 167,043 PATIENTS

Quantitative Methods and Theoretical Developments (QMTD)

Charlie Sewalt, BSc[1], Esmee Venema, MD[1], Eveline Wiegers[1], Dennis den Hartog, MD, PhD[2], Fiona Lecky, MD, PhD[3] and Hester F. Lingsma, PhD[1], (1)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (2)Department of Surgery, Erasmus University Medical Center, Rotterdam, Netherlands, (3)School of Health and Related Research, University of Sheffield, Sheffield, United Kingdom

Purpose:

Patients with major trauma could benefit from treatment in a dedicated trauma center, but early identification of major trauma (Injury Severity Score (ISS) > 15) remains difficult. Over the past decades, several prehospital trauma models for predicting mortality have been developed. The aim of this study is to externally validate existing prognostic models for trauma patients and to assess their ability to predict mortality and major trauma in the prehospital setting, to ultimately improve prehospital decision making.

Method(s):

Prediction models for validation were identified in the literature through a systematic search. For the validation we used data from the Trauma Audit and Research Network (TARN) from the UK, and selected all trauma patients transported with emergency medical services (EMS) or helicopter emergency services (HEMS) between 2013 and 2016. Patients transferred in from another hospital were excluded. The outcome measures were in-hospital mortality and major trauma (ISS>15). The performance of the models was assessed in terms of discrimination and calibration using the C-statistic and calibration slope and intercept. We assessed overtriage and undertriage rates of major trauma, defined as the proportion of false positive and false negative patients, for all prediction models.

Result(s):

A total of 167,043 patients were included to validate six prediction models. Discrimination ranged from 0.603 (95% confidence interval 0.597-0.610) for MGAP to 0.806 (0.802-0.809) for mREMS in predicting in-hospital mortality and varying from 0.591 (0.588-594) for mREMS to 0.737 (0.734-0.739) for KTS in predicting major trauma. The calibration slopes for in-hospital mortality ranged from 0.414 to 0.725. The RTS model was the best calibrated model for in-hospital mortality (calibration slope 0.725 and intercept 0.136). Triage of major trauma was most accurate with the MGAP model (undertriage rate 29.8% and overtriage rate 53.0%).

Conclusion(s):

This study shows that most prehospital trauma models perform reasonable in predicting in-hospital mortality but are inadequate to identify major trauma patients. None of these prediction models met the maximal undertriage and overtriage rates (5% undertriage and an overtriage rate of 25-35% according to the American College of Surgeons). To improve pre-hospital decision making in trauma patients, future studies should focus on which patient groups benefit from treatment in a major trauma center and on the development of tools to identify these patients in the pre-hospital setting.

PS2-14 METHODS FOR INCLUSION OF CORRELATED PARAMETERS IN DETERMINISTIC SENSITIVITY ANALYSES

Quantitative Methods and Theoretical Developments (QMTD)

Peter Gal, MSc, Evidera Inc., Budapest, Hungary and Mustafa Oguz, PhD, Evidera Inc., London, United Kingdom

Purpose:

Inputs of economic models with parameter uncertainty are recommended to be included in deterministic sensitivity analyses (DSA) to assess model outcomes' sensitivity to key factors. In DSA, model parameters are tested at the lower and upper bounds of their estimated confidence intervals (CI). When a group of parameters are correlated, their joint extreme values are not well defined. This study offers rigorous and practical methods to establish meaningful bounds that take into account the correlation structure of the parameters.

Method(s):

The first method sets each parameter in a separate DSA scenario to the boundary of their own CI. All other correlated parameters are set to their expected values conditional on the set value of the selected parameter. The conditional expected values may be calculated by closed formulas given their means and covariance matrices.

The second proposed method sets all correlated parameters in one common scenario. To define the relevant boundary of their joint distribution, first a context is defined with a relevant outcome measure that allows the ranking of sets of multiple parameters. A typical such context and outcome measure could be the mean survival time for parametric survival curves. Then, for the DSA scenario the parameter sets corresponding to the low and high quantile values of the calculated measures are selected. Depending on the context, the boundary parameter sets may be directly calculated or, more generally, they may be identified by simulation.

Result(s):

Our first method is recommended when separate scenarios for each correlated parameter are required. Typical such examples are regression coefficients for utility, cost, or risk equations. Our second method allows for the joint setting of boundary values for a set of correlated parameters, and their inclusion in common DSA scenarios. This could be the recommended method when setting the individual parameters separately to their marginal distribution extreme values would imply unreasonable scenarios that is often the case for parametric survival curves.

Conclusion(s):

The prevalent current practice for correlated model parameters with uncertainty is to either omit them from DSA or include them without properly accounting for their covariance structure. We propose two practical methods that allow for more complete and credible sensitivity analyses by incorporating known information about the parameters covariances.

PS2-15 SHOULD UNCERTAINTY BE INCORPORATED IN PREFERENCE-SENSITIVE DECISION SUPPORT? INVITATION TO DEBATE
Quantitative Methods and Theoretical Developments (QMTD)

Mette Kjer Kaltoft, PhD, University of Southern Denmark, Odense, Denmark, Jesper Bo Nielsen, PhD, Research Unit for General Practice, Odense, Denmark and Jack Dowie, PhD, London School of Hygiene and Tropical Medicine, London, United Kingdom

Purpose: In the dominant paradigm, researchers establish the uncertainties, along with the expected average outcomes, and hand over the task of dealing with the separated outputs – e.g. Means and Credible Intervals (CIs) - to the decision maker. Whether individual or group, the decision maker is expected to make the necessary trade-off as part of the final deliberative 'making up of mind'. There is no attempt to synthesise, transparently and systematically, the joint means and uncertainties of the performance ratings for all options on all criteria. The essence of an MCDA-based person-centred decision support tool (DST) is that all required value judgments are elicited from the decision maker *inside* the tool, at or near the point of decision. Leaving them analytically unsupported in the mean-uncertainty trade-off task is unacceptable and may constitute 'symbolic violence'. Sensitivity analyses and stochastic displays merely re-present the task in a different way.

Method(s): The method involves treating the means and CIs of all outcomes and process considerations as separate criteria. The decision maker has their importance weights for mean and uncertainty elicited for each criterion separately, so the resulting trade-offs are criterion-specific and may differ between life expectancy and quality of life, for example. The Expected Value (EV) score for each option within the MCDA-based DST incorporates this trade-off. (The uncertainty reported in a study cannot be regarded as an indicator of its technical quality, which should be separately graded.)

Result(s): Three proof of method examples are provided, each of which can be explored interactively online (https://ale.rsyd.dk, enter 1406 as survey ID). They comprise a hypothetical example with two options/three considerations/six criteria; an empirical 6/2/4 example on bone medications; and a 6/4/8 example on anti-retroviral regimes for HIV patients.

Conclusion(s): In one view setting up mean and uncertainty as separate criteria involves double counting, the calculation of the mean having already 'synthesised' the uncertainty in the distribution. In the alternative view, the mean calculation, while reflecting the uncertainty in the distribution, has left it to be addressed by the decision maker. Their preference-based weighting of mean and uncertainty can be entered into a new EV calculation, leading to a preference-sensitive result. A method to implement this is shown to be feasible. Whether it should be followed is opened for debate.

PS2-16 IN SEARCH OF SWEDEN'S COST-EFFECTIVENESS THRESHOLD

Applied Health Economics (AHE)

Jonathan Siverskog, MSc and Martin Henriksson, PhD, Linkoping University, Linkoping, Sweden

Purpose:

For Sweden, and many other countries, there is no empirical estimate of a cost-effectiveness threshold representing the marginal cost at which healthcare is generating health. This is a problem, since without one we do not know whether reimbursement decisions, by displacement, are expected to increase or decrease population health. The purpose of this study is to investigate the viability of estimating the threshold for healthcare in Sweden using publicly available data on aggregate expenditure and mortality.

Method(s):

In this study, we analyse the relationship between years of life lost per capita and health care expenditure per capita. We use publicly available data with annual frequency to construct a nation-level time series for 1970-2016 and a panel for 2003-2016, including data for twenty regional councils. The time series is used to estimate a vector autoregressive model, where we test for cointegration and causality using the Johansen and Granger tests, respectively. The panel data set is analysed using two-stage least squares regression. We consider a number of instrumental variables reflecting exogenous shocks to councils' costs and labour market conditions to address the issue of endogeneity.

Result(s):

Our time series approach reveals a negative long-run relationship between healthcare expenditure and life years lost. However, the results indicate that decreases in mortality are causing increases in expenditure, rather than other way around, which might be interpreted as longevity driving healthcare costs.

Analysis of our panel data set points to a positive relationship between years of life lost and healthcare expenditure, which is consistent with higher spending in regions with an older population, or simply poorer health. When instrumenting for expenditure, we find a negative relationship, which translates to a marginal cost per life year of 367 thousand SEK in 2016. This estimate varies wildly with the set of instruments, however, and none of the instruments considered are sufficiently strong to generate reliable or significant estimates.

Conclusion(s):

There are many different channels through which healthcare expenditure and mortality influence each other. Our regression analyses reveal some of these, but fail to uncover the one relevant for the estimation of a cost-effectiveness threshold. Therefore, in lack of some ingenious (and exogenous) instrument, our provisional conclusion must be that Sweden's cost-effectiveness threshold is not to be found using aggregate level data.

PS2-17 COST-EFFECTIVENESS OF PROBLEM SOLVING THERAPY IN OUTPATIENT REHABILITATION FOR STROKE
Applied Health Economics (AHE)

Anne Berkhoff, MSc[1], Marieke Visser, PhD[2], Adriaan van't Spijker, PhD[3], Jan van Busschbach, PhD[3], Gerard Ribbers, MD, PhD[4] and Majanka Heijenbrok-Kal, PhD[2], (1)Rijndam Rehabilitation, Rotterdam, Netherlands, (2)Rijndam Rehabilitation, Erasmus University Medical Center, Rotterdam, Netherlands, (3)Erasmus University Medical Center, Rotterdam, Netherlands, (4)Rijndam Rehabilitation, Rotterdam, Erasmus University Medical Center, Rotterdam, Netherlands

Purpose: To study the cost-effectiveness of Problem Solving Therapy (PST) in outpatient stroke rehabilitation in comparison to usual (UC) alongside a randomized clinical trial (RCT).

Method(s): During out-patient rehabilitation participants with stroke (18-75 years old) were included in an RCT with one year follow-up. Patients were randomly assigned to an 8-week group intervention consisting of add-on PST or UC. Alongside this RCT a cost-effectiveness study was performed from the health care perspective. HRQoL was measured with the EQ-5D-5L and costs were calculated based on the the Trimbos and iMTA questionnaire (Tic-P). Measurements were performed pre- and post-intervention, and after 6 and 12 months follow-up. Linear mixed models were used to estimate trajectories of effects and costs over time adjusted for baseline differences between the groups. Incremental cost-effectiveness ratios (ICER's) were calculated, expressed as costs per Quality Adjusted Life Year (QALY) gained. The uncertainty around the ICER was determined by bootstrapping (5000 iterations).

Result(s): In total, 138 patients were included in the RCT, of which 72 were assigned to PST and 66 to UC. No significant differences were found between PST and UC for HRQoL (PST: 0.813 vs UC: 0.806, p=0.539) or total healthcare costs (PST: €11,571 vs UC: €9025, p=0.118) over total follow-up time. The mean cost of PST was €462 per patient. At baseline and directly post-intervention, mean health care costs were significantly higher in the PST group than in the UC group (PST baseline: €4114 (SD 311) vs UC baseline: €3141 (SD 320), p<0.01; PST post-intervention: €3356 (SD 318) vs UC baseline: €2288 (SD 320), p<0.01), but health care costs were equal at 12 months follow-up (PST: €925 (SD 325) vs UC: €954 (SD 339), p=0.61). The ICER of PST versus UC was 104,528 €/QALY gained (95 CI:62,093-226,064), indicating that PST was not cost-effective.

Conclusion(s): PST was not cost-effective in comparison with usual care. However, results showed a larger decrease in health care consumption and costs during 6 months follow-up in the PST group compared with UC. These results suggest that PST could enhance recovery after discharge.

PS2-18 COST-UTILITY OF EARLY BREAST CANCER SURVEILLANCE IN SURVIVORS OF ADOLESCENT HODGKIN'S LYMPHOMA
Applied Health Economics (AHE)

Lauren Tessier, HBSc, Institue of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada, Jill Furzer, MA, Insitute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada, David Hodgson, MD, MPH, FRCPC, Princess Margaret Cancer Centre & The University of Toronto, Toronto, ON, Canada, Cecilia Cotton, PhD, The University of Waterloo, Waterloo, ON, Canada, Paul Nathan, MD, MSc, FRCPC, The Hospital for Sick Children & The University of Toronto, Toronto, ON, Canada, Sumit Gupta, MD, PhD, FRCPC, Division of Haematology/Oncology, Hospital for Sick Children, Toronto, ON, Canada, Peter C. Coyte, PhD, University of Toronto, Toronto, ON, Canada and Petros Pechlivanoglou, Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, ON, Canada

Purpose:
Adolescent females treated for Hodgkin's Lymphoma (HL) are at an increased risk of breast cancer (BC). The cost-effectiveness of high-risk BC surveillance in this population is unknown. The purpose of this study was to evaluate the cost-utility of seven distinct high-risk BC surveillance strategies using magnetic resonance imaging (MRI) or mammography, or their combination, beginning at various starting ages, relative to the current standard of care (SOC) in Ontario, annual mammography and MRI beginning at age 30.

Method(s):
A discrete event simulation model was used to simulate the life history of a theoretical cohort of 25-year-old females treated for HL in adolescence. The model utilized BC incidence and mortality data from the Surveillance, Epidemiology, and End Results (SEER) Program, Canadian all-cause mortality data, and a tumour growth algorithm from a previously published study. Screening, diagnostic and treatment costs were derived from various published sources specific to Ontario. Assumptions included that: all females were BC-free at simulation start; surveillance stopped after age 75; and BC risk persisted until age 80. We estimated BC incidence and mortality, life expectancy, quality-adjusted life-years (QALYs), healthcare costs, and the relative cost-utility among screening strategies.

Result(s):
Relative to the SOC, annual MRI initiated at age 25 and annual MRI initiated at age 25 with a switch to annual mammography at age 50 both resulted in average cost savings of $174 and $1,825 dollars per woman screened, respectively, over the lifetime. Additionally, MRI initiated at age 25 resulted in an increase of 0.03 QALYs, while switching to mammography at age 50 lead to a QALY gain of 0.01.

Conclusion(s):
Current high-risk breast cancer surveillance guidelines do not constitute the most cost-effective strategy in survivors of adolescent HL. Annual MRI from ages 25 to 75 or annual MRI from ages 25 to 49 with a switch to annual mammography from ages 50 to 75 both represent less costly and more effective options.

PS2-19 IMPROVING IDENTIFICATION OF PATIENT SUBGROUPS VIA A STEPWISE COVARIATE SELECTION APPROACH

Quantitative Methods and Theoretical Developments (QMTD)

Diego Saldana Miranda, Ph.D.[1], Srinivasa Rao Pelluri, M.Sc.[2], Valery Risson, Ph.D. MBA[1] and Gorana Capkun, Ph.D.[1], (1)Novartis Pharma A.G., Basel, Switzerland, (2)Novartis Healthcare Pvt. Ltd., Hyderabad, India

Purpose:

Patient subgroup identification methodologies previously described[1,2] rely on the separate modelling of each study arm's effect with a fixed set of covariates. We describe a stepwise approach for covariate selection based on each predictor's contribution to the observed treatment difference.

Method(s):

Following approaches described previously for continuous, binary, and time-to-event outcomes[1,2], two separate regularized generalized linear models are fitted on the study drug and control arms using a training dataset. The resulting models are used to predict trial outcomes on a separate validation dataset. This results in two predicted outcomes per patient, one with the study drug and one with the control treatment. The difference in predicted outcomes results in a score which can be used to retain patients that benefit the most from the study drug. A cross validation approach in which the training and validation sets are resampled several times and the threshold for patient selection increased progressively is used to obtain an average treatment difference curve as well as confidence intervals. The resulting curve can be used to select appropriate score thresholds.

The approach is further extended to count outcomes via the use of Poisson regression, and a stepwise covariate selection method based on the treatment difference curves is used to obtain an optimal set of covariates. This extended approach is applied on masked data from a clinical trial, where the primary outcome to be optimized is a count outcome.

Result(s): The use of the covariate selection approach increased the resulting treatment difference curve compared to the use of the full set of covariates, resulting in a significant increase in the observed treatment difference for the enriched subpopulation. Additional sensitivity analyses through bootstrapping showed that the improvement of the effect was present in 99% of the resampled results.

Conclusion(s):

The proposed methodology shows promising results and deserves consideration for future use in medical decision making.

[1]Zhao L, Tian L, Cai T, Claggett B, Wei LJ. EFFECTIVELY SELECTING A TARGET POPULATION FOR A FUTURE COMPARATIVE STUDY. J Am Stat Assoc. 2013;108(502):527-539.

[2]Li J, Zhao L, Tian L, et al. A predictive enrichment procedure to identify potential responders to a new therapy for randomized, comparative controlled clinical studies. Biometrics. 2016;72(3):877-87.

PS2-20 COST-EFFECTIVENESS OF PROPHYLACTIC HYSTERECTOMY IN FIRST-DEGREE FEMALE RELATIVES WITH LYNCH SYNDROME OF PATIENTS DIAGNOSED WITH COLORECTAL CANCER IN THE UNITED STATES: A MICRO SIMULATION STUDY

Applied Health Economics (AHE)

Maaike Alblas, MSc[1], Elisabeth Peterse, MSc[1], Mengmeng Du, PhD[2], Jennifer Chen, MPH[2], Ann Zauber, PhD[2], George W. Mutter, MD, PhD[3], Ewout W. Steyerberg, PhD[4], Nikki van Leeuwen, MSc[1] and Iris Lansdorp-Vogelaar, PhD[1], (1)Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, Netherlands, (2)Memorial Sloan Kettering Cancer Center, Department of Epidemiology & Biostatistics, New York, NY, (3)Harvard Medical School Boston, Department of Pathology, Boston, MA, (4)Leiden University Medical Center, Department of Biomedical Data Sciences, Leiden, Netherlands

Purpose:

In the past decade, universal testing of all colorectal cancers for Lynch Syndrome (LS) has become standard practice with the aim of identifying first-degree relatives (FDR) with Lynch. Patients with LS have substantially higher risk of colorectal cancer and endometrial cancer compared to the general population. The practice of universal testing for LS and offering FDR with LS intensive screening for colorectal cancer has proven to be (cost-)effective. In this study, we developed a microsimulation model for endometrial cancer to evaluate the additional cost-effectiveness of also offering prophylactic hysterectomy to female FDR with LS.

Method(s):

We developed the MISCAN Endometrial model, which incorporates a natural history model for the development of hyperplasia with and without atypia into endometrial cancer. The age specific incidence of endometrial cancer was calibrated to SEER from the Surveillance, Epidemiology and End result Program (SEER) from 2009 to 2013. The progression of hyperplasia to endometrial cancer was based on literature and differed between hyperplasia without atypia and hyperplasia with atypia. We simulated a population of 10 million women with an age range matching that of FDR with LS in a Dutch study of universal testing of LS in colorectal cancer. Subsequently, we used the MISCAN-Endometrial model to estimate life-years gained and the resulting incremental cost-effectiveness ratio from offering this cohort prophylactic hysterectomy from age 40 (after child-bearing age). This strategy was compared to no prophylactic hysterectomy in FDR.

Result(s):

In the absence of prophylactic hysterectomy in FDR, the number of endometrial cancer deaths was estimated to be 111 per 1,000 women. Total associated costs for the treatment of endometrial cancer were $7.3 million. Offering prophylactic hysterectomy to FDR after age 40 would reduce the number of endometrial cancer deaths to 0.77 per 1,000 women (-99%) and increase the costs (treatment of endometrial cancer and prophylactic hysterectomy) to $16.3 million (+113%). The resulting incremental cost-effectiveness ratio of performing prophylactic hysterectomy in FDR was $5,566, i.e. well below accepted thresholds.

Conclusion(s):

Offering prophylactic hysterectomy after age 40 to female FDR with LS in a program of universal testing for LS in colorectal cancers seems highly cost-effective. However, future research is necessary to incorporate the impact of hysterectomy on quality of life in the analysis.

PS2-22 COST-EFFECTIVENESS OF EARLY DETECTION AND PREVENTION STRATEGIES FOR OVARIAN CANCER: A SYSTEMATIC REVIEW
Applied Health Economics (AHE)

Gaby Sroczynski, MPH, Dr.PH[1], Artemisa Gogollari, Mag.sc.hum. Dr. med.[2], Felicitas Kuehne, MSc[3], Lara Hallson, MPH, Dr.[2], Nora Pashayan, MD, MSc, MSt, PhD[4], Martin Widschwendter, PhD[5] and Uwe Siebert, MD, MPH, MSc, ScD[6], (1)Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria, (2)Department of Public Health, Health Services Research and Health Technology Assessment, UMIT and ONCOTYROL - Center for Personalized Cancer Medicine, Division of Health Technology Assessment and Bioinformatics, Hall in Tirol, Austria, (3)Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (4)UCL - University College London, Faculty of Population Health Sciences, London, United Kingdom, (5)UCL - University College London, Department of Women's Cancer, London, United Kingdom, (6)UMIT, Dept. of Public Health, Health Services Research & HTA / Harvard Univ., Dept. Health Policy & Management, Institute for Technology Assessment, Hall i.T. (Austria)/Boston (USA), Austria

Purpose: The objective was to systematically review cost-effectiveness studies evaluating early detection and prevention strategies for ovarian cancer in various populations.

Method(s): A systematic literature search for decision-analytic modeling studies evaluating the cost-effectiveness of early detection and/or prevention strategies for ovarian cancer was performed in relevant databases (Medline / Embase / Cochrane Library / CRD / EconLit). Study characteristics and results including life-years gained (LYG), quality-adjusted life-years (QALY), and incremental cost-effectiveness ratios (ICER; in cost/LYG or QALY) were summarised in standardised evidence tables. Economic results were converted into Euros (index year 2015) using gross domestic product purchasing power parity and the consumer price index.

Result(s): A total of twenty-four studies were included differing in terms of target population, discount rate, perspective and evaluated strategies. The studies evaluated (1) genetic testing for mutations and preventive measures for diagnosed mutation carriers, (2) early detection of ovarian cancer and (3) risk reducing prevention measures for women at an increased risk of developing ovarian cancer. Genetic testing for mutations and prevention for diagnosed mutation carriers achieved ICERs/ICURs of 10,000-42,000 Euro/LYG (8,000-70,000 Euro/QALY). Early detection of ovarian cancer utilising the CA125-based longitudinal Risk of Ovarian Cancer Algorithm (ROCA) in women from the general population aged 50 achieved ICURs of 9,500-80,000 Euro/QALY for multimodal screening with a risk-adjusted algorithm. The results were sensitive in terms of test performance characteristics, test costs and test frequency. Risk-reducing interventions in premenopausal women, yielded ICERs between 700 Euro/LYG for women with a 10% lifetime risk of developing ovarian cancer to 47,500 Euro/LYG for women with 2% lifetime risk of developing ovarian cancer (6,000 to 52,000 Euro/QALY); in postmenopausal women respective ICERs ranged from 2,000 to 47,000 Euro/LYG (2,000 to 757,000

Euro/QALY). For mutation carriers risk-reducing surgical prevention strategies were reported to be cost-saving or achieved ICERs below 2,600 Euro/LYG (below 16,000 Euro /QALY).

Conclusion(s): Based on our findings, both early detection and preventive surgery in women at increased or high risk of developing ovarian cancer can be considered effective and cost effective. For women at average risk, multimodal screening based on the ROCA from age 50 upwards can be considered cost-effective.

PS2-23 DEVELOPING A DECISION AID FOR A MENTAL HEALTH PROBLEM: LESSONS LEARNED

Decision Psychology and Shared Decision Making (DEC)

Wilma Otten, PhD, Leanthe van Harten, MSc and Fieke Pannebakker, PhD, TNO the Netherlands Organisation for Applied Scientific Research, Leiden, Netherlands

Purpose: A decision aid for choosing a treatment for patients with conversion disorder was developed based on a recent Dutch Standard of Care. During the development we learned several lessons that seem specific for a mental health problem compared to our experiences with decision aids for somatic health problems.

Method(s): The decision aid was developed by writers of the Standard of Care (LvH, FP) and an expert on decision aids (WO). Extra input was derived from interviews with patients, close relatives of patients, and health care professionals. The development and intermediate products were discussed in four workgroups consisting of patients, relatives and professionals. A mockup of the interactive digital decision aid was built to test this aid with the end users.

Result(s): The following points arose from working with a specific mental health problem. (a) The very limited empirical evidence regarding effective treatments promoted the use of patient experiences. (b) Using professional experiences was difficult because of vested interest in specific therapies. (c) Patients not only wanted a decision aid, but also an aid how to communicate with their health care professional. (d) Describing the treatment in such a way that patients could really imagine what the therapy involved proved a real challenge. For example, cognitive behavioral therapy as a treatment depends on the specific application to conversion disorder by a specific therapist. (e) Addressing the therapeutic alliance between client and therapist.

Conclusion(s): Decision aids for mental health problems are certainly feasible and appreciated. Some problems need more solutions, like using experiences of patients and health care professionals. Also informing patients about the nature of treatments remains a challenge in the case of mental health.

This work is financed by Netwerk Kwaliteitsontwikkeling GGz (the Netherlands)

PS2-24 CONCEPTUAL MODELS OF SHARED DECISION MAKING: A SYSTEMATIC REVIEW OF THE LITERATURE

Decision Psychology and Shared Decision Making (DEC)

Hanna Bomhof-Roordink, MSc, Fania R. Gaertner, PhD, Anne M. Stiggelbout, PhD and Arwen H. Pieterse, PhD, Medical Decision Making, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands

Purpose:

To 1) provide an up-to-date overview of conceptual models of shared decision making (SDM), including their empirical foundation and type of decision, 2) give insight in the prominence of elements and actors within the models, and 3) integrate elements into components.

Method(s):

Three electronic databases (PubMed, Embase, Cochrane) were systematically searched for peer-reviewed articles published from inception up until June 21, 2016. Pairs of raters independently screened titles and abstracts, and then full-texts of all articles retrieved. One rater extracted and analysed the data, another checked it.

Result(s):

The search yielded 1765 records, from which 22 conceptual models described in 22 different articles from 18 different first authors were included for this review. Ten models were developed based on empirical data gathered with the purpose to inform the model, and twelve were based on analytical thinking. Fourteen models were developed for a specific setting, such as psychiatry or oncology. Nine models focused on treatment decision making; for half of the models, the authors did not state the type of decision of interest. The models identified criteria, stages, elements, steps, or skills, often in a certain sequence, but most without further clarification of relationships. We identified 52 different elements in the 22 models. Six elements previously identified by Makoul and Clayman (2006) were part of at least 11 (50%) models: List options, Benefits/risks, Patients' preferences, Check/clarify patient understanding, Make or explicitly defer decision, and Learn about the patient. For each of the 52 elements, models differed in whether an actor was stated or not. Overall, clinicians had a role regarding 46 elements in one or more models, compared to 30 elements for patients. We integrated the elements into 25 components, some describing explicit phases in SDM processes, others relating to more general behaviours. The frequency of components and the frequency of the patient role in them differs markedly.

Conclusion(s):

This systematic review provides an up-to-date overview of key elements within conceptual models of SDM. It shows that models show some similarities, but large heterogeneity exists. We urge researchers to elucidate relationships between elements in future models.

PS2-25 HOW IS CERVICAL CANCER SCREENING INFORMATION COMMUNICATED IN UK WEBSITES?
Decision Psychology and Shared Decision Making (DEC)

Yasmina Okan, PhD, Centre for Decision Research, Leeds University Business School, Leeds, United Kingdom, Samuel G. Smith, PhD, Leeds Institute of Health Sciences, University of Leeds, Leeds, United Kingdom and Wandi Bruine de Bruin, PhD, Centre for Decision Reserch, Leeds University Business School, Leeds, United Kingdom

Purpose: Patients increasingly search for information online. Web-based information about cervical cancer screening plays an important role in women's screening decisions. We therefore investigated whether UK websites about cervical cancer screening (1) contain key information about benefits and harms of screening, possible screening outcomes, and cervical cancer risks, (2) present probabilistic information using formats recommended in the risk communication literature, (3) include appeals for participation and/or informed decision making.

Method(s): We identified 14 UK websites through Google, by entering common search terms pertaining to cervical cancer screening. We coded website content using an established checklist of 16 items, including benefits and harms of screening, possible screening outcomes, and cervical cancer risks. We also examined whether the format of any probabilistic information involved verbal quantifiers, numbers, or graphical displays, as well as whether risk reduction was communicated in relative vs. absolute terms. Finally, we coded whether websites included appeals for participation and/or statements concerning informed decision making.

Result(s): We report on three main findings. First, benefits and harms of screening and possible screening outcomes were mentioned frequently: Discussed benefits focused on risk reduction for developing cervical cancer (in 71% of websites), whereas the harms mentioned included overdiagnosis and overtreatment (86%), pain/discomfort related to the cytology test (86%), and the possibility of false negatives (71%) and false positives (57%). In contrast, the risks of developing cervical cancer or dying from cervical cancer were not discussed in any website. Second, risk reduction for developing cervical cancer was typically presented numerically, in all cases in relative terms. In contrast, quantifications of harms were mostly verbal. Graphical displays were only included in two websites, depicting possible screening outcomes (e.g., likelihood of abnormal results). Finally, appeals for participation were present in 86% of websites, with 42% of these also referring to informed choice concerning screening.

Conclusion(s): The existing heterogeneity in the information included in UK websites about cervical cancer screening may compromise women's informed decision making. Probabilistic information was often not conveyed in formats known to facilitate understanding. Recommendations to avoid use of verbal quantifiers alone, to present absolute rather than relative risk reductions, and to use graphical displays were often not met. Designing websites that adhere to such recommendations may support informed cervical screening uptake and avoid potentially harmful misunderstandings.

PS2-27 DO PEOPLE WITH A DIFFERENT REGULATORY FOCUS MAKE DIFFERENT DECISIONS ABOUT COLORECTAL CANCER-SCREENING PARTICIPATION?

Decision Psychology and Shared Decision Making (DEC)

Linda Douma, MSc, Amsterdam Public Health research institute, VU University Medical Centre, Amsterdam, Netherlands, Ellen Uiters, PhD, National Institute for Public Health and the Environment (RIVM), Bilthoven, Netherlands and Danielle Timmermans, PhD, Department of Public and Occupational Health, Amsterdam Public Health research institute, VU University Medical Center, Amsterdam, Netherlands

Purpose: Previous studies have shown that having a promotion focus (e.g. wanting to achieve a state of being healthy) or prevention focus (e.g. wanting to avoid getting ill) can affect people's health-related decisions and behaviour by emphasising aspects and information that seem relevant in light of what they want to achieve. However, this issue has not yet been researched regarding colorectal cancer (CRC) screening. With our study, we aimed to examine the relationship between people's focus and their CRC screening participation, as this could provide insights for supporting people in making this complex decision.

Method(s): An online survey was carried out among a sample of first-time CRC screening invitees (1282 respondents, response rate 49%). We assessed people's promotion and prevention focus, focus on the advantages or disadvantages of CRC screening, screening participation and main considerations (e.g. cancer is a serious illness) concerning their screening decision.

Result(s): Generally, CRC screening participants score higher on both a promotion and prevention focus than non-participants. A dominant focus was present among non-participants, namely a prevention focus. CRC screening participants show a dominant focus on the advantages of CRC screening. The relationship between people's focus and their screening participation is partially mediated by people's main considerations concerning CRC screening.

Conclusion(s): People's focus differs between CRC screening participants and non-participants. To optimally support people in making their screening decision, the provided information should appeal to their different foci/frames of mind regarding this decision. Future research could examine what this in reality means for developing information materials.

PS2-28 DECISION SUPPORT IN THE CHOICE BETWEEN NEOBLADDER AND STOMA AFTER BLADDER REMOVAL

Decision Psychology and Shared Decision Making (DEC)

Julia van Tol-Geerdink, PhD[1], DM Somford, MD, PhD[2], C Wijburg, MD PhD[3] and JA Witjes, MD PhD[1], (1)Radboud UMC, Nijmegen, Netherlands, (2)CWZ Hospital, Nijmegen, Netherlands, (3)Rijnstate Hospital, Arnhem, Netherlands

Purpose:

After bladder removal for cancer, patients face a difficult choice between a stoma and a bladder replacement (neobladder). Both have pros and cons. To date, the decision is mainly based on patient and hospital characteristics. A stoma is usually chosen for fragile patients, for women and in low volume hospitals. The decision should, however, also reflect the patients personal preference, as recommended by current guidelines. For this purpose, we developed and tested a decision aid (DA).

Method(s):

Urologists, specialized nurses and patients were involved in the development of the DA, according to IPDAS criteria. Subsequently, the DA was used and evaluated by new patients. Their opinion on content and use were assessed by questionnaire. In addition, knowledge and participation were compared between patients with DA and patients without.

Result(s):

In total, 47 patients were asked to participate and 41 gave informed consent (87%). Two were non-eligible because of tumor characteristics and two did not respond. In all, 37 questionnaires were received; 12 from the control group and 25 from the DA group.

The DA made all patients participate actively in the decision making; none of the patients in the DA group reported that (only or mainly) the doctor had decided vs. 25% of patients in the control group ($p=0.011$). The DA group answered 96% of knowledge questions correct and the control group 83% ($p=0.066$). Patients were positive about the content of the DA. All patients reported that they "seldom or never" found the content complicated or words difficult to understand. Moreover, all stated that the presentation of the pros and cons of the options was balanced.

Patients received the DA on the day of diagnosis (38%), within one week after diagnosis (19%) or later (43%). Satisfaction with the timing differed between these groups (resp. 88%, 100% and 33% satisfied, $p=0.018$) .

Conclusion(s):

The DA was evaluated positively and made all patients participate actively in the choice between stoma and neobladder. As for the use in daily practice; receiving the DA on the same day as the cancer diagnosis is not considered too early by most patients. On the other hand, most patients who receive it after more than a week would have preferred to receive it earlier.

PS2-29 RESEARCH PARTICIPANTS EXPERIENCES OF SHORT-TERM MENTAL DISTRESS AFTER RECEIVING CARDIOVASCULAR RISK INFORMATION
Decision Psychology and Shared Decision Making (DEC)

Asa Grauman, MSc Public Health, PhD student, Centre for Research Ethics & Bioethics, Uppsala university, Uppsala, Sweden, Mats G Hansson, Professor of Biomedical Ethics, Centre for Research Ethics & Bioethics, Uppsala University, Uppsala, Sweden, Arvid Puranen, M.D., Sunderby Hospital, Norrbotten County Council, Luleå, Sweden, Stefan James, M.D. Professor of cardiology, Uppsala Clinical Research center, Uppsala University, Uppsala, Sweden and Jorien Veldwijk, PhD, Erasmus Choice Modelling Centre, Erasmus University , Rotterdam, Netherlands

Purpose: To investigate psychological outcomes after receiving cardiovascular risk information in men and woman 50-64 years old.

Method(s): This study was conducted within the Swedish CardioPulmonary BioImage Study (SCAPIS). Randomly selected individuals from the general Swedish population underwent thorough health examinations and received written reports with test results related to cardiovascular risk factors. Individuals with abnormal findings were referred to either primary health care centers or specialized hospital care. Psychological factors were assessed on the first day of the study and again three months later, and included worry of having a heart attack, Short form health survey 12 (SF-12) and Hospital and Anxiety Scale (HADS). Secondary outcomes included knowledge about cardiovascular risk factors and intention to change life style. Change in outcomes was tested by analyzing differences between baseline and follow-up. Subgroup analysis was carried out to see whether change in psychological factors differ among those who were referred or labelled with new diagnose. Associations were tested using univariate and multivariable statistical methods.

Result(s): Baseline data was available for 612 individuals, and 434 (71 %) of those individuals also completed the questionnaire after three months. Level of worry at baseline was associated with family history of heart attack, high cholesterol, abdominal obesity, perceived stress, health literacy level and perceived risk of having a heart attack. There was an increase in worry after three months (p=0,01). After performing subgroup analysis, the increase in worry was only significant for those referred to hospital or diagnosed with coronary arteriosclerosis. The same participants also increased their mental distress accordingly to the SF-12 mental score and the HADS anxiety score. There were no significant differences in remaining outcomes after three months.

Conclusion(s): Mental distress increased three months after receiving cardiovascular risk information for individuals diagnosed with coronary arteriosclerosis. The results of this study therefore indicate that risk assessment should take psychological factors in consideration.

PS2-30 VISUAL ATTENTION WHEN USING INFORMATION IN A PATIENT DECISION AID: AN EYE-TRACKING STUDY AMONG OLDER CANCER PATIENTS
Decision Psychology and Shared Decision Making (DEC)

Olga Damman, PhD[1], Julia Van Weert, PhD[2], Jozua Murris, Msc[3] and Danielle Timmermans, PhD[1], (1)Department of Public and Occupational Health, Amsterdam Public Health research institute, VU University Medical Center, Amsterdam, Netherlands, (2)University of Amsterdam, Amsterdam, Netherlands, (3)VU University, Amsterdam, Netherlands

Purpose: To assess older cancer patients' visual attention to core information in a patient decision aid, and to test differences between exposure to a 'complex' table versus an adapted 'simple' table (in an attempt to reduce cognitive effort for older patients).

Method(s): Eye-tracking experiment with 18 participants (≥ 65 yrs), all (former) cancer patients. Participants viewed information about pros and cons of two equivalent treatment options for early stage lung cancer: surgery and stereotactic ablative radiotherapy (SABR). Eye tracking assessed participants' dwell times related to three types of information (1) treatment outcomes (e.g., 5 yrs survival rates); (2) side-effects (e.g. risk of fatigue); (3) procedures (e.g., removal of lymph glands and associated certainty that can be obtained that tumor is gone). A mixed between-within subjects ANOVA was conducted with relative dwell time as dependent variable, type of information (outcomes, side-effects, procedures) as within-subjects variable and type of table (complex, simple) as between-subjects variable.

Result(s): In both tables, participants attended most to information about treatment outcomes (relative dwell time complex table 201 ms; relative dwell time simple table 269 ms). The main effect of type of information on relative dwell time approached significance ($p = .055$, $\eta p2 = .282$). Relative dwell times were significantly shorter for procedures compared to both treatment outcomes ($p < .001$, $\eta p2 = .033$) and side-effects ($p = .019$, $\eta p2 = .407$). No significant interaction effect was found between type of information and type of table ($p = .596$, $\eta p2 = .030$), nor was there a significant main effect of type of table ($p = .117$, $\eta p2 = .208$). Participants exposed to the 'simple' table did overall spend more time per word compared to those exposed to the 'complex' table, but this difference was not significant.

Conclusion(s): The least amount of visual attention was spent on the procedures of treatment options by older cancer patients. Simplifying the typical complex table did not seem to alter this. An important question is whether this information about procedures is actually considered relatively less important by patients, as it seems to include relevant details on the certainty that can be obtained that the tumor is gone. It may be that participants simply did not focus on it because it was presented at the bottom of the table.

PS2-32 HOW TO PRESENT RISK IN PATIENT PREFERENCE STUDIES: A SCOPING REVIEW

Decision Psychology and Shared Decision Making (DEC)

Elisabeth Furberg, PhD, Centre for Research Ethics & Bioethics, Uppsala, Sweden, Asa Grauman, MSc Public Health, PhD student, Centre for Research Ethics & Bioethics, Uppsala university, Uppsala, Sweden, Cathy Anne Pinto, PhD, Merck Research Laboratories, Center for Observational and Realworld Evidence, Rahway, NJ, Richard Hermann, MD, MPH, AstraZeneca, Wilmington, DE, Rosanne Janssens, Department of Pharmaceutical and Pharmacological Sciences, University of Leuven, Leuven, Belgium, Esther de Bekker-Grob, PhD, Erasmus University Rotterdam, Rotterdam, Netherlands, Jorien Veldwijk, PhD, Erasmus Choice Modelling Centre, Erasmus University , Rotterdam, Netherlands and **Ulrik Kihlbom, PhD**, Centre for research ethics and bioethics Uppsala University, Uppsala, Sweden

Purpose: Most patient preference studies include risk attributes. For the validity of the results, it is crucial that the respondents understand these attributes. The aim of this scoping review was to provide an overview ways to present risk in patient preference studies.

Method(s): The search query included descriptions of 'patient' and 'risk communication' to get all relevant papers. PubMed, PsycINFO, SCOPUS and Cochrane databases were then consulted followed by a snowballing search for the inclusion of additional papers. After the removal of duplicates, the papers were screened with the following exclusion criteria: (i) not in English, (ii) no full text available, (iii) published before 1980, (v) not a research study, (vi) does not evaluate preference methods, (vii) does not focus on risk communication in a health-care setting, and (viii) does not focus on humans.

Result(s): In total 143 individual studies and 10 reviews were included. The risk formats described and tested in the literature were grouped into three main categories: (i) numerical formats, (ii) verbal formats, and (iii) visual formats. A fourth category "communication style", was added to capture relevant information influencing the interpretation of the other formats. In summary, the results reported in the included studies showed that:

- Risk was better understood using numbers than words such as "low", "medium" or "high";
- Absolute risk was better understood than relative risk;
- Frequencies were preferred over percentages;
- Adding visual formats to numerical or verbal formats increased understanding;
- Individualization of the risk information was positively perceived, but there was insufficient evidence to conclude that individualization enhances understanding;

The results of comparing different formats within the categories were inconstant.

Conclusion(s): Even though very few of the found studies focussed on patient preferences studies, we may conclude that such studies may benefit from combinations of risk formats as well as individualization when optional. E.g. in explorative preference studies, through direct interaction with patients during pilot (exploratory) testing, the risk format can be adjusted and

further tailored to the needs of the patient population. The ambiguous results of comparisons of different sub-categories may relate to contextual factors, such as characteristics of the respondents, the kind of risk and severity of risk involved.

Acknowledgement: This work received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking PREFER grant n° 115966

PS2-33 PATIENT-CENTERED BENEFIT-RISK DECISION-MAKING AND THE ROLE OF EDUCATIONAL TOOLS AND PSYCHOLOGICAL INSTRUMENTS IN PREFERENCE ELICITATION- YEAR 1 OF IMI PREFER

Patient and Stakeholder Preferences and Engagement (PSPE)

Ulrik Kihlbom, PhD[1], Selena Russo, PhD[2], Dario Monzani, PhD[3], Geert Vander Stichele, PharmD[4], Sarah Verschueren, PhD[4] and Cathy Anne Pinto, PhD[5], (1)Centre for research ethics and bioethics Uppsala University, Uppsala, Sweden, (2)Behavioural Sciences Unit, Kids Cancer Centre, Sydney Children's Hospital,, Randwick, Australia, (3)Applied Research Division for Cognitive and Psychological Science, IEO European Institute of Oncology, Milan, Italy, Milan, Italy, (4)MindBytes, Gent, Belgium, (5)Merck Research Laboratories, Center for Observational and Realworld Evidence, Rahway, NJ

Purpose: While researchers have developed a variety of methods for eliciting patient preferences, there has been less focus on psychological instruments or educational tools that can help determine why patients form certain preferences and make certain decisions. If preferences are to be accurate expressions of trade-offs patients make when having properly understood benefit, harms and other treatment attributes, then education and psychological measurements will be crucial. As a task in PREFER, a 5-year Innovative Medicines Initiative (IMI), we aimed to identify, describe and assess the clinical feasibility of using scenario-based interactive tools (SBITs), and profile psychological variables that can affect the construction, elicitation, and interpretation of preferences.

Method(s): As a first step, scoping and systematic reviews were performed. Based on that information, motivational factors and appraisal criteria were developed for selecting specific psychological measurements and educational features.

Result(s): Based on a systematic review of the available empirical evidence, 29 psychological constructs were identified with a theoretical basis for influencing preferences and decision-making. These constructs were sorted in three classes depending on the strength of evidence. Distinctions between psychological constructs that can be regarded as explanatory and those that can be regarded as formative were drawn.

A scoping review on educational tools revealed that SBITs have largely not been developed and evaluated in a healthcare setting. Moreover, it was recognized that SBITs can be constructed with a wide variety of features. As such, preliminary guidelines have been developed to chart how the characteristics of a particular case study relate to the selection of specific educational features.

Conclusion(s):

There are some psychological instruments for which there are satisfying evidence and may explain preference heterogeneity or determine representativeness of the sample population. Further evidence is required to determine how useful other instruments are. SBITs are possibly useful in patient preference studies because they may: (1) stimulate a more personal and subjective view and reflection about treatment options, (2) help motivate participants and (3)

help participants overcome cognitive challenges, inherent to many preference methods; but, more research is needed.

Acknowledgement: This work received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking PREFER grant n° 115966 and benefited from contributions from IMI2 PREFER Work Package 2 participants.

PS2-34 ENGAGING STAKEHOLDERS IN THE DESIGN OF A DECISION AID TO IMPROVE SECONDARY PREVENTION AFTER STROKE

Patient and Stakeholder Preferences and Engagement (PSPE)

Talya Porat, PhD, Iain Marshall, Euan Sadler, Miguel Vadillo, Christopher McKevitt, Charles Wolfe and Vasa Curcin, King's College London, London, United Kingdom

Purpose: To engage stroke survivors and other stakeholders in the identification and design of a decision support system (DSS) to improve secondary prevention after stroke.

Method(s): We engaged a range of stakeholders (N=37), including stroke survivors and carers (N=11), health and social care professionals (N=16), commissioners and policy makers (N=6) and researchers (N=4), to explore their needs and to collaboratively design a DSS to improve secondary prevention after stroke. Stakeholders participated in group meetings (including using nominal group techniques such as priority setting and consensus building), focus groups and interviews. The design of the DSS was underpinned by theories of shared decision making and behaviour change techniques.

Result(s): Eight themes (needs from a DSS) were identified in the meetings and interviews: 1) involve stroke survivors in decisions concerning their treatments; 2) communicate risk in a meaningful way; 3) compare individual's perceived risk with their actual predicted risk; 4) personalise treatments to help control stroke risk factors; 5) display the effectiveness of recommended treatments; 6) address concerns about treatment and barriers to adherence; 7) support continuity of care; and 8) identify individuals at high risk of recurrent stroke. These themes informed the collaborative design of DOTT (Deciding on Treatments Together), a decision aid to be used in primary care during clinical consultations between the healthcare professional and stroke survivor, aiming to facilitate shared decision making on treatments to reduce recurrent stroke risk. DOTT will be integrated with the electronic health records system, displaying the stroke survivor's predicted risk of having a recurrent stroke (based on rules generated from the South London Stroke Register), and propose the most effective personalised treatments to reduce the risk (based on the stroke survivor's characteristics), with their relative benefit.

Conclusion(s): Engaging stakeholders throughout the design process ensures a system that is in line with stroke survivors' and healthcare professionals' needs. DOTT has the potential to reduce stroke recurrence by adopting a data-driven patient-centred approach, leading to stroke survivors making more informed decisions, taking ownership for the treatment decisions, improving their adherence to the agreed management plan and thus reducing their risk of a further stroke. While DOTT currently targets stroke only, the approach and the design features could be used for a range of chronic diseases requiring long-term management.

PS2-35 "THEY EXPLAIN LUNG TRANSPLANT LIKE IT'S CHANGING A CAR PART": ENRICHING DECISION MAKING ABOUT INTUBATION AND LUNG TRANSPLANT USING PATIENT NARRATIVES FOR ADULTS WITH CYSTIC FIBROSIS

Patient and Stakeholder Preferences and Engagement (PSPE)

Melissa Basile, PhD[1], Eugenia Boye-Codjoe, MPH[1], Janice Wang, MD[2], Katherine Henthorne, LMSW[2], Danielle Rodriguez-Madera[1], Denis Hadjiliadis, MD[3], Jessica Goggin, MAS, RN[4], Jorge Huamantla[1], Victoria Shaffer, PhD[5] and **Negin Hajizadeh, MD, MPH**[6], (1)Northwell Health, Manhasset, NY, (2)Northwell Health, Lake Success, NY, (3)University of Pennsylvania School of Medicine, Philadelphia, PA, (4)University of California, San Diego, La Jolla, CA, (5)University of Missouri, Columbia, MO, (6)Hofstra Northwell School of Medicine and Northwell Health, Manhasset, NY

Purpose: Preliminary results from analysis of narrative interviews about intubation and lung transplant among adults with Cystic Fibrosis (CF) or their caregivers, for use in a web-based decision aid for CF advance care planning.

Method(s): Individuals were recruited from 3 study sites, and through an ad placed on the Cystic Fibrosis Foundation's Community Voice webpage seeking individuals to share stories about intubation or lung transplant. We asked open-ended questions covering participants' experiences with intubation and/or lung transplant including: a) process (events leading up to, during, and after the procedures); b) information seeking for decision making about intubation/lung transplant; and c) feelings/emotions (did they make the 'right' decision, would they do anything differently, and what would they tell someone else considering intubation or lung transplant?). Interviews: averaged 64 minutes, were audio recorded and transcribed for qualitative analysis.

Result(s): N=15 participants were recruited (10 CF adults and 5 family members); individual narratives focused on: intubation N=1, lung transplant N=8, both N=6; 3/5 family member participants had CF loved-ones who had died. Preliminary analysis showed a diverse array of experiences: most people never discussed the possibility of needing intubation for respiratory failure with a clinician. Several people provided details about being intubated - one person described it like "a railroad track going down my throat", while others could not recall the experience. One family member discussed her loved-one's chronic rejection after transplant and her decision to withdraw life support; two people described waking up after transplant and breathing for the first time with new lungs, but most expressed frustration at not being extubated for several days/weeks afterwards; some participants recalled the moment they went in for transplant, while others did not recall anything due to sedation and intubation in the days prior to transplant; most participants expressed sadness when thinking about their donors. Finally, some participants experienced almost no post-transplant complications, while others continue having multiple complications including rejection and undergoing a second double-lung transplant.

Conclusion(s): CF patient intubation and lung transplant narratives focused on personal-level experiences which may not come up during usual clinician-patient communications. Therefore patient narratives representing a wide array of outcomes can be a vital source of information to include in patient decision aids, alongside of prognostic estimates and risk/benefits comparisons of treatment options.

PS2-36 EARLY VALUE ASSESSMENT OF IMAGE-GUIDED SURGICAL TECHNOLOGIES USED IN ONCOLOGY

Patient and Stakeholder Preferences and Engagement (PSPE)

Melanie Lindenberg, MSc[1], Janine A. van Til, PhD[2], Valesca Retel, PhD[1], Theo Ruers, MD PhD[1] and Wim van Harten, MD PhD[3], (1)Netherlands Cancer Institute, Amsterdam, Netherlands, (2)University of Twente, Enschede, Netherlands, (3)University of Twente and the Netherlands Cancer Institute, Amsterdam, Netherlands

Purpose: Surgical procedures are increasingly executed with technologies, such as robotics, image guidance, and tracking techniques. This can have serious financial impact on investments and operational costs. Early stage healthcare technology assessment (HTA) can help in positioning promising innovations. This study - part of a large HTA project - aimed to rank and assess the expected value of three image-guided surgical technologies that can be used in different oncologic interventions, to steer research and development.

Method(s): We used an Analytic Hierarchy Process, a multiple criteria decision analysis technique, to rank the three technologies and a combination of those. The analysis focused on the following interventions: resection of lymph nodes (LN) within the pelvis, and resection of rectal, tongue, liver, and breast tumors. The criteria set was based on the Core HTA model, literature and was discussed with clinicians. The set contained 16 criteria categorized in four domains: effectiveness, efficiency, technology and organizational. The technologies were defined with the technical developers. First, clinicians weighted the criteria by pair-wise comparisons in an interview. Second, they compared the technologies pair-wise per criteria and intervention. Per intervention, weights of the criteria and score of the technology were combined to calculate the expected value of each technology and rank accordingly. Finally a sensitivity analysis was conducted.

Result(s): Twenty-three clinicians participated in the weighting interviews and 16 in the scoring interviews. The value of the technology in all interventions was mainly described by the performance on the effectiveness domain (59%-66%) comprising e.g. risk on recurrence and complications. In 4/5 interventions, the combination received the highest score (22%-38%). Per intervention the technologies were ranked as follows: in rectal and tongue tumors, optical imaging (20%;19%) and navigation (19%;18%) were ranked second and third respectively. In resecting liver tumors, navigation was ranked second (24%), the other technologies (augmented reality (AR), optical imaging) scored close to usual care (14%;11%). In LN removal the navigation was ranked first (29%) followed by the combination (28%) and AR (16%).

Conclusion(s): Clinicians see great potential in the combination of technologies but its value is mainly explained by one of the three technologies. This analysis gained insight in the expectations of clinicians regarding the technologies to steer further technical development. Clinical research is necessary to show the true added value of these technologies.

PS3-1 A FRAMED FIELD EXPERIMENT ON THE EFFECT OF FEEDBACK ON ANTIBIOTICS PROVISION IN PEDIATRICS

Health Services, Outcomes and Policy Research (HSOP)

Kerstin Eilermann[1], Katrin Halstenberg[2], Ludwig Kuntz[3], Kyriakos Martakis[2], Bernd Roth[2] and Daniel Wiesen[3], (1)Cologne Graduate School and Department of Healthcare Management, University of Cologne, Cologne, Germany, (2)Children's and Adolescents' Hospital, University Hospital of Cologne, Cologne, Germany, (3)Department of Healthcare Management, University of Cologne, Cologne, Germany

Purpose: To analyze the causal effect of expert feedback in form of a simple benchmark on individual pediatricians' antibiotic therapy decisions.

Method(s): In a randomized, controlled framed field experiment pediatricians decided on length of first-line antibiotic therapies for 40 stylized pediatric routine cases (in total, 8,760 decisions were made by 73 subjects). In the intervention group, subjects received feedback in form of an expert benchmark (average length of antibiotic therapy). Pediatricians' individual characteristics were elicited in a post-experimental questionnaire comprising items for demographics, clinical characteristics, personality traits, and risk attitudes.

The expert benchmark was elicited in an online survey, in which directors of pediatric departments in Germany (n=20) were asked for their recommendations on length of antibiotic therapy for the cases used in the experiment. In order to verify the experts' opinion as a suitable benchmark for appropriate length of antibiotic therapies, the decisions from the survey were – separately for each case – compared with recommendations on antibiotic therapy from national and international guidelines.

Result(s): Pediatricians exposed to feedback, which allowed for a comparison of their own aggregated length of therapy with an expert benchmark, significantly reduced chosen length of antibiotics therapy. The announcement of feedback did not significantly affect pediatricians' choices. Pediatricians' clinical experience, personality traits, and risk attitudes significantly affected their decisions on length of antibiotic therapy.

Conclusion(s): Pediatricians' antibiotics provision is significantly affected by expert feedback. Our findings make a strong case for the inclusion of individual feedback in antibiotic stewardship programs to addressing inappropriate antibiotics provision. Pediatricians' risk attitudes, personality traits, and professional experience explain the heterogeneity in decisions on length of antibiotic therapies.

PS3-2 IMPROVEMENT OF THE PERFORMANCE OF SURVIVAL PREDICTION IN THE AGEING TRAUMA POPULATION

Health Services, Outcomes and Policy Research (HSOP)

Leonie de Munter, MSc[1], Nancy ter Bogt, PhD[2], Suzanne Polinder, PhD[3], Charlie Sewalt, BSc[3], Ewout W. Steyerberg, PhD[3] and Mariska de Jongh, PhD[4], (1)ETZ, Xperiment Trauma TopZorg, Tilburg, Netherlands, (2)Network Emergency Care Euregio, Enschede, Netherlands, (3)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (4)Netwerk Emergency Care Brabant, Tilburg, Netherlands

Purpose: The overestimation of survival predictions in the ageing trauma population result in negative benchmark numbers in hospitals that mainly treat elderly patients. The aim of this study was to develop and validate a modified Trauma and Injury Severity Score (TRISS) with minimal adjustments for accurate survival prediction in the ageing trauma population.

Method(s): This retrospective study was conducted with data from Network Emergency Care Brabant and Network Emergency care Euregio. Missing values were imputed. New prediction models were created in the development set (n=15,530), including age (continuous or categorical) and Anesthesiologists Physical Status (ASA). The models were externally validated in a validation set (n=15,504). Subsets were created based on age (≥75 years) and the presence of hip fracture. Model performance was assessed by Nagelkerke R^2, discrimination and calibration. A final model was created based on both datasets.

Result(s): No differences were found between the baseline characteristics of the development dataset and the validation set. The inclusion of ASA in the prediction models showed significant improved discriminative abilities in the two subsets (e.g. 0.522 [95% CI: 0.461, 0.583] vs 0.735 [95% CI: 0.689, 0.781] for elderly patients with hip fracture). The final model showed high agreement between observed and predicted survival.

Conclusion(s): The includion of ASA and age as continuous variable in survival prediction is a simple adjustment of the TRISS methodology to improve survival predictions in the ageing trauma population. A new model is presented, through which patients with isolated hip fractures could be included in the evaluation of trauma care.

PS3-4 ROUTINE DATA BASED ANALYSES OF MORTALITY OF PATIENTS WITH FRACTURES DUE TO FALLS LESS THAN THREE METERS - A COMPARATIVE APPROACH

Health Services, Outcomes and Policy Research (HSOP)

Guenther Zauner, Msc, dwh GmbH, simulation services / DEXHELPP (Decision Support for Health Policy and Planning), Vienna, Austria, Ingrid Wilbacher, PhD, Main Association of Austrian Security Institutions, Vienna, Austria, Niki Popper, PhD, Dexhelpp (Decision Support for Health Policy and Planning), Vienna, Austria, Florian Endel, Technical University Vienna, Institute for Analysis and Scientific Computing, Vienna, Austria and Christoph Urach, MSc, Institute for Analysis and Scientific Computing, Technical University Vienna, Vienna, Austria

Purpose: Within DEXHELPP (Decision Support for Health Policy and Planning: Methods, Models and Technologies based on Existing Health Care Data, FFG, No-843550) fractures due to falls less than three meters for the elderly are identified as one big health problem resulting in increased mortality. In this project a comparison of:

(1) real world mortality,
(2) Charlson index based on hospitalisation, and
(3) Charlson index based on medical treatment

is performed.

Method(s): We used linked claims data in the Austrian health care sector. We analysed ICD10 main diagnosis subgroups of fracture categories for the years 2006 and 2007. A half year wash out phase is used to identify incident cases.
We used a probabilistic matching of ATC codes (Anatomical Therapeutic Chemical Classification) to ICD10 diagnostic groups to calculate the Charlson index using the description of Quan et al. In parallel this method is applied for up to six additional hospital diagnoses and performed a comparison of the results of the implemented methods.

Result(s): We identified in total 32.343 patients older than 60 years with main fracture ICD10 diagnosis and followed up for five years for mortality identification purpose. During the first year 4.596 persons died. Low and moderate disease severity have a vastly reduced mortality rate. For low severity it is only about 1/3 and for moderate severity 2/3 compared to high disease severity with 21% annual mortality. The integration of Charlson index works and gives an adequate forecast method. Especially combination of intramural based index and medical treatment-based strategy are promising but have to be set up individually especially depending on age and gender.

Conclusion(s): The realised work represents the actual Austrian situation. Additional analysis including longer wash-out phase seems promising. Due to the fact, that a high number of persons is affected, the socio-economic impact is high. The results reflect the findings from international studies and could be used for prevention strategy planning and evaluation models.

PS3-5 PREDICTING LENGTH OF STAY AFTER A ROAD TRAFFIC ACCIDENT WHILE ACCOUNTING FOR COMPETING RISKS AND TIME-DEPENDENT VARIABLES

Health Services, Outcomes and Policy Research (HSOP)

Griet Van Belleghem, Msc[1], Stefanie Devos, PhD[1], Ives Hubloue, PhD MD[2], Door Lauwaert, RN[2], Karen Pien, MD[2], Ronald Buyl, PhD[1] and Koen Putman, PhD[1], (1)Vrije Universiteit Brussel, Jette, Belgium, (2)Universitair ziekenhuis brussel, Jette, Belgium

Purpose: While concerns about length of hospital stay are one of the most prominent questions patients have during hospitalization after trauma, the limited research on prediction modelling show only moderate performance. The aim of this study is to use a modelling technique based on the cox proportional hazards model to estimate the length of stay considering competing endpoints and time-dependent variables.

Method(s): The analyses were performed on an individual linked database of all hospitalized road traffic victims in the period 2009-2011 (N=64 377). The technique 'dynamic prediction by landmarking in competing risks' was used to estimate duration until discharge as it controls for bias induced by including time-dependent variables in a competing risks model. The competing endpoint is dying during hospitalization. Landmark points (LM) were set every 3 days until day 21. The model performance was checked based on C-statistic and martingale residuals.

Result(s): The supermodel behaved consistently over the different LM's. The model showed a good model calibration on predictions during the first 30days. The less severe a patient is injured the higher the probability of being discharged. Injured body-regions will in general lead to a lower probability of being discharged except for injuries to the face. The influence of the location of an injury on length of stay varies over the different LM's.

Conclusion(s): The use of dynamic predictions by landmarking in competing risks based on administrative data is feasible and leads to a decent model fit.

PS3-6 DO CANCER BIOMARKERS MAKE TARGETED THERAPIES COST-EFFECTIVE? A SYSTEMATIC REVIEW IN METASTATIC COLORECTAL CANCER

Health Services, Outcomes and Policy Research (HSOP)

Kelly Seo and John Cairns, London School of Hygiene and Tropical Medicine (LSHTM), London, United Kingdom

Purpose:

This study aims to determine the impact of predictive biomarkers on the cost-effectiveness of targeted therapies in the context of metastatic colorectal cancer.

Method(s):

A literature search was performed using Medline, Embase, EconLit, NHSEED. Economic evaluations of biomarker-guided therapies were searched from 2000 until October 2017. Targeted therapies granted a marketing authorization with companion biomarkers by the European Medicines Agency (EMA) or US Food and Drug Administration (FDA) were included in the literature search. Studies were selected based on the pre-defined PICOS framework (population, intervention, comparator, outcome, study type). Methodological quality was assessed with the Quality of Health Economic Studies (QHES) instrument. A selection of included studies and a final QHES score per study were resolved between two assessors.

Result(s):

45 studies were included in this review. Of these, 16 studies evaluated the intrinsic value of cancer biomarkers, whereas the remaining studies focused on assessing the cost-effectiveness of corresponding drugs. Most studies indicated favourable cost-effectiveness of predictive biomarkers for targeted therapies in mCRC. Some studies even reported that biomarkers were cost-effective, while their corresponding therapies were not cost-effective. A considerable number of economic evaluations were conducted in pre-defined genetic populations and thus, often failed to fully capture the biomarker's clinical and economic values. An average QHES score of 74 was found.

Conclusion(s):

Cancer biomarkers for targeted therapies in mCRC were found to be improving the cost-effectiveness of targeted therapies. However, it did not necessarily make their corresponding therapies cost-effective. The companion biomarkers shown to reduce therapy costs however, their savings were not sufficient enough to make corresponding agents cost-effective. Meanwhile, evaluation of predictive biomarkers was often restricted to the cost of tests not on their clinical values or biomarker prevalence.

PS3-7 IN-HOSPITAL MORTALITY VARIABILITY FOR CHARLSON COMORBIDITY INDICES USING ADMINISTRATIVE DATA
Health Services, Outcomes and Policy Research (HSOP)

Alberto Freitas, PhD, Joao Vasco Santos, MD, Mariana Lobo, MSc and Cristina Santos, PhD, CINTESIS and MEDCIDS, Faculty of Medicine of the University of Porto, Porto, Portugal

Purpose: The Charlson Comorbidity Index (CCI) is one of the most widely used methods for summarizing comorbidities into a score in hospital episodes by means of administrative data (International Classification of Diseases – ICD – codes). Each associated comorbidity has a weight, based on the adjusted risk of mortality within one year. Higher CCI scores may also be associated with increased length of stay, hospital costs and readmissions. Since the original CCI proposed weights, others studies have suggested different comorbidity weights. This study aims to compare and analyze trends for the CCI using different approaches (different weights for each comorbidity), and their relation with in-hospital mortality by disease groups.

Method(s): This study included data from adults (18 years-old and over) hospitalizations in mainland Portuguese public hospitals between 2000 and 2014. Three different approaches were used to weight: the original by Charlson/Deyo et al. (1987/1992), the proposed by Schneeweiss et al. (2003), and the revision by Quan et al. (2011). The Agency for Healthcare Research and Quality's Clinical Classifications Software (CCS) for ICD-9-CM was used to group the principal diagnosis. Receiver Operating Characteristic (ROC) curves were calculated as discrimination measures.

Result(s): Circa 10 million episodes were included in this analysis. In general, the AUC was higher considering the original Charlson/Deyo weights, when compared to the other approaches. Considering specific disease groups (CCS), with at least 1000 deaths, the biggest differences among approaches were found for "Chronic ulcer of skin" with 60.4% (95% CI [59.1-61.7]) for Charlson/Deyo, 63.0% (95% CI [61.7-64.4]) for Schneeweiss, and 55.2% (95% CI [53.8-56.6]) for Quan.

Conclusion(s): Important differences among weights in different scenarios (different diseases groups) exist and need to be properly understood in order for a proper selection, use and interpretation of the CCI. Yet, results highlight and increased suitability of this index over overs years, using administrative data.

Acknowledgments to: The Project "NORTE-01-0145-FEDER-000016" (NanoSTIMA), financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF); the Portuguese Ministry of Health's Authority for Health Services (ACSS I. P.) for providing access to national hospitalizations data.

PS3-9 FACTORS INFLUENCING THE UPTAKE OF A CHOOSING WISELY RECOMMENDATION BY ORTHOPAEDIC SURGEONS

Health Services, Outcomes and Policy Research (HSOP)

Tessa Rietbergen[1], Ron Diercks, MD, PhD[2], Ieke van der Wel[1], Elske van den Akker, PhD[1], Rob Nelissen, MD, PhD[3], Enrike van der Linden, MD, PhD[3], P.J. Marang-van de Mheen, PhD[1] and **Leti van Bodegom-Vos, PhD**[1], (1)Leiden University Medical Center, Department of Biomedical Data Sciences, section Medical Decision Making, Leiden, Netherlands, (2)University Medical Center Groningen, Department of Orthopaedics, Groningen, Netherlands, (3)Leiden University Medical Center, Department of Orthopaedics, Leiden, Netherlands

Purpose:

To assess which factors are associated with the uptake of the Choosing Wisely recommendation “No arthroscopic surgery and no MRI for patients of 50 years and older with knee degenerative knee complaints"™ (in short: CW-recommendation) by Dutch orthopaedic surgeons.

Method(s):

We performed an internet-based survey among 421 Dutch orthopaedic surgeons, from which 243 (58%) responded. This survey included questions about background characteristics, uptake of the CW-recommendation (4-point scale), and forty factors possibly influencing the uptake of the CW-recommendation (4-point scale). Factors were based on literature and interviews with orthopaedic surgeons, and classified according to the framework of Grol. Spearman rank correlations were used to identify factors associated with the uptake of the CW-recommendation. Next, as individual factors may be related to each other we included factors that were significantly associated with the uptake of the CW-recommendation into a multivariable logistic regression model ($p<0.05$). For this analysis, we dichotomized the answers on the uptake of the CW-recommendation and factors because of few observations in some cells. Factors increasing the uptake of a CW recommendation were classified as facilitators, factors decreasing the uptake as barriers.

Result(s):

Among respondents, 200 OS (82%) reported to adhere to the CW-recommendation for arthroscopic surgery and 197 (81%) for MRI. In univariate analyses, 17/40 factors were significantly associated with the uptake of the CW-recommendation for arthroscopic surgery, and 16/40 with the uptake of the CW-recommendation for MRI. In multivariate analyses, 6/17 factors were independently significantly associated with the uptake of the CW-recommendation for arthroscopic surgery (see table 1). Four of these independently significantly associated factors acted as facilitators and two factors as barriers for the uptake of the CW-recommendation for arthroscopic surgery. The uptake of the CW-recommendation regarding MRI was significantly independently associated with 3/16 factors. One of these factor acted as facilitator and two as barriers for the uptake of the CW-recommendation for MRI.Â

Conclusion(s):

Barriers and facilitators were mostly found on the individual professional level, related to the awareness of, belief in and attitude towards the CW-recommendation of orthopaedic surgeons. Strategies to improve the uptake of the CW-recommendation among orthopaedic surgeons should at least be targeted at these factors. Â

Table 1. Significant independent associations between factors (barriers/ facilitators) for the uptake of a CW-recommendation regarding arthroscopic surgery and MRI and follow-up of this CW recommendation by orthopaedic surgeons.

Framework level (GroI)	Arthroscopic surgery β (95% CI) (B/F)	MRI β (95% CI) (B/F)
Innovation		
I agree with the CW-recommendation	4.66 (2.30 till 7.02)(F)	2.69 (1.30 till 4.69)(F)
Individual professional		
Awareness		
I am aware of the CW-recommendation regarding arthroscopic surgery/ MRI	5.00 (1.76 till 8.26)(F)	-
Beliefs		
I belief in added value of arthroscopic surgery/ MRI	-1.88 (-3.10 till -0.66)(B)	-0.91 (-1.76 till -0.08)(B)
Attitude		
I feel I do not need to follow the CW-recommendation, because of my clinical experience	-1.83 (-2.86 till -0.81)(B)	-0.87 (-1.67 till -0.08)(B)
Behavioural routines		
I search effectively to the latest evidence and guidelines about diagnosis and treatment of knee complaints	1.34 (0.20 till 2.49)(F)	-
Social context		
My colleagues all follow the CW-recommendation	1.36 (0.22 till 2.52)(F)	-

B = Barrier, mentioned factor hampers the follow-up of the CW-recommendation
F= Facilitator, mentioned factor facilitates the follow-up of the CW-recommendation

PS3-10 ECONOMIC IMPLICATIONS OF DISPARITIES IN CERVICAL CANCER SCREENING

Health Services, Outcomes and Policy Research (HSOP)

Vivian Chia-Rong Hsieh, PhD, Department of Health Services Administration, China Medical University, Taichung, Taiwan

Purpose: The Taiwan Health Promotion Administration (HPA) currently funds a population-based Pap smear test for females aged 30 or over. However, the uptake of this screening strategy is still low. This research aims to estimate the forgone health and economic benefits associated with disparities in subsidized cervical cancer screening in Taiwan.

Method(s): Using population-based National Health Insurance (NHI) administrative databases, cancer and death registries from 2012-2015, differential uptake of subsidized annual Pap Smear test and prevalence of cervical cancer due to socioeconomic disparities were calculated. Socioeconomic classifications considered were income level, occupational group and area of residence. From an economic perspective, we then investigated levels of associated health care spending (laboratory tests, outpatient, inpatient, emergency care) in females with and without screening, as well as across social groups. Finally, potential years of life lost (PYLL) are estimated to illustrate the economic consequences of inequalities in cervical cancer screening.

Result(s): From our results, however, the overall screening rate is still lower than 30%, with a slight decrease annually. By income level, those in the bottom 20% (lowest income quintile) have the lowest screening rate of ~22%, higher prevalence rate (252.4/100,000 in 2015) than higher income females, and highest disease-specific mortality rates (12.0/100,000 in 2015). By occupation type, workers with contracted or temporary jobs have the highest screening rate of approximately 31%, even higher than that of fixed-term, permanent workers (~27%). The former is also observed with highest cancer prevalence rates. Those in eastern and remote areas have the highest hospital admission rates (52.1/100,000 in 2015). Unexpectedly, PYLL due to cervical cancer is projected to be highest in the middle income group.

Conclusion(s): Our results suggest that economic costs due to disparities in cervical cancer screening is considerable. Differential uptake of subsidized Pap Smear test across socioeconomic groups generated a subsequent impact on the disease-related health care costs, service utilization and mortality rates which were seen highest in the marginalized populations.

PS3-11 SIMULATION OF DETECTION OF OLIGO-RECURRENCES IN NON-SMALL CELL LUNG CANCER

Quantitative Methods and Theoretical Developments (QMTD)

Henri Wolff, MSc[1], Leonie Alberts, MSc[2], Lisanne Kastelijn, PhD[3], Sherif El Sharouni, PhD[4], Franz Schramel, PhD[2], Rein Vos, PhD[5] and Veerle MH Coupe[1], (1)VU University Medical Center, Amsterdam, Netherlands, (2)St. Antonius Hospital, Nieuwegein, Netherlands, (3)Ziekenhuis Gelderse Vallei, Ede, Netherlands, (4)University Medical Centre Utrecht, Utrecht, Netherlands, (5)Maastricht University, Maastricht, Netherlands

Purpose: Oligo-recurrent disease is defined as one to five recurrences detected after a disease free interval and treatment of the primary tumour with curative intent. Oligo-recurrences in non-small cell lung cancer (NSCLC) are often also treated with curative intent, but additional detected recurrences are common in this group of patients. A prediction model can be used to understand the mechanisms of false-positive oligo-recurrence detection (see figure).

Method(s): We developed a simulation model that generates 10.000 patients. A Gompertz-growth model determines the amount of metastases that have reached the visible phase at the time of a scan (see figure). The model simulates a CT surveillance schedule for the asymptomatic detection of recurrences at fixed time points. In addition, each patient also has a chance of unscheduled symptomatic detection of recurrences. Symptomatic detection depends on the number of metastases that have reached detectable size, and a patient-specific probability that one metastasis becomes symptomatic.

The model generates variation in the patient specific parameters: growth rate, chance to become symptomatic, and size ratio of recurrences. The generated patients are classified into the subgroups true oligo-recurrence, false-positive oligo-recurrence, and poly-recurrence.

Result(s): The model was able to reproduce recurrence ratios from medical literature. The amount of false-positive oligo-recurrences is most significantly determined by the frequency of follow-up scans per year. False-positive oligo-recurrences have significantly more symptomatic detections than true oligo-recurrences. Differences between patient groups in growth rate and size ratio were insignificant.

Conclusion(s): The model is hypothesis generating, but predictions can easily be verified in clinical practice. Better understanding of the mechanisms behind metastatic disease can help improve both follow-up rationale and patient selection for treatment of oligo-recurrences, leading to improved cost effective treatments.

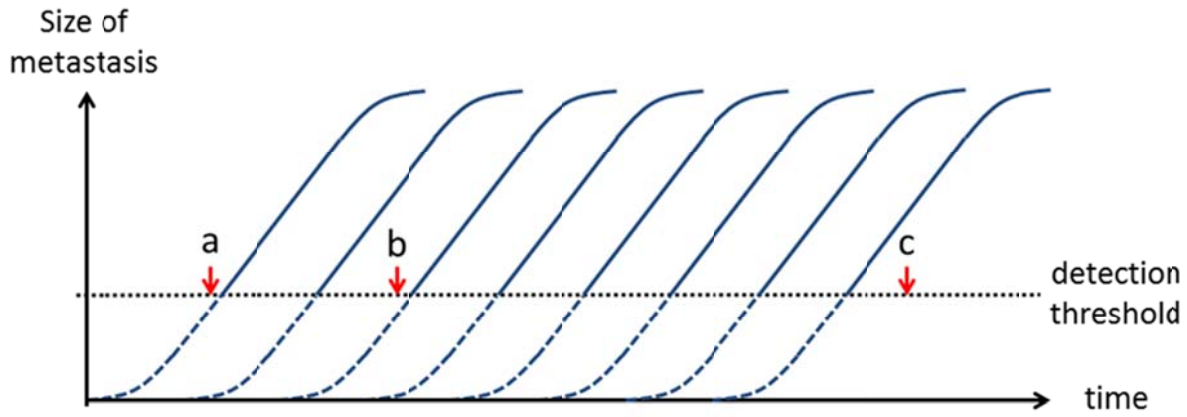


Figure: Metastases (blue lines) become visible on a scan when they grow above the detection threshold. Before this time they remain in the 'invisible phase' (dashed lines). If a scan occurs at time 'a', all metastases are invisible, and the patient would be considered to be 'recurrence free'. At time 'b', two recurrences would be visible on the scan. With 6 more invisible metastases present, we define this as a false-positive oligo-recurrence. At time 'c', 6 recurrences would be visible on the scan, which is defined as a poly-recurrence.

PS3-12 USE OF STATED-REVEALED PREFERENCE MODELS FOR A MULTISTAKEHOLDER APPROACH

Quantitative Methods and Theoretical Developments (QMTD)

Christine C. Huttin, PhD, ENDEPUSresearchinc and University Aix Marseille, Cambridge, MA

Purpose:

The paper proposes the development of a stated-revealed preference model in health care. It follows a series of methodological papers compiled in the book "economics and medical decision making" (Huttin, May 2017) and two methodological papers on congruence of frames for conjoint surveys and design developments for econometric modeling (Huttin, THC, March and Oct, 2017) . The epidemiological and economic model uses multiple levels of disaggregation (Huttin, 2007, 2010, and 2017). As the inclusion of MCDAs for calibration of structural economic models is under researched. In the R&D ecosystem, to integrate individual level evidence that physicians and patients can use, new PRCTs may be used, as well as practice changing studies based on enough evidence (e.g. Mindact trial with patient stratification balancing high/low genomic risks and high/low clinical risks , Hudis and Dickler, 2016).

Method(s): A conceptual framework is proposed, based on the compilation of papers. The current methodological development involves two steps: Step 1 A Stated choice experimental design, based on the Lens model and revised versions, involving physician and patients in shared decision making process. Unobserved physicians' preference heterogeneity is represented in a stated preference model, with an inclusion of effects of financial and economic information due to physician or practice economics, treatment costs of drug and devices, and patient economics (cost cognitive cues) .Step 2: A disease econometric model , including a partition of medical practices based on physicians and practices' characteristics is designed and tested on diabetes type II (Huttin and Wong; Huttin and Atwood) It helps to identify potential cut off points and critical thresholds where stated data can complement billing information on complex cost sharing mechanisms.

Result(s): the framework is presented on the case of three groups of judges: physicians, patients and payers; different statistical runs are compared for disease econometric models with and without the partitioning of medical practices, controlling for IT diffusion.Identification of cut off points are presented using the existing literature and comparing with analytical datasets used so far.

Conclusion(s): The feasibility of a joint estimation combining stated and revealed data elements on cost of care is discussed, based on an estimation of a joint model using maximum likelihood methods and analysis of variances.

PS3-13 STUDY PROTOCOL OF A CAUSAL COMPARATIVE EFFECTIVENESS ANALYSIS AND DATA AVAILABILITY IN THE AUSTRIAN SETTING: THE CASE OF WHEN TO START STATIN TREATMENT

Quantitative Methods and Theoretical Developments (QMTD)

Felicitas Kuehne, MSc[1], Beate Jahn, PhD[2], Annette Conrads-Frank, PhD[3], Marvin Bundo, MD[1], Marjan Arvandi, MS[4], Florian Endel, BSc[5], Niki Popper, PhD[6], Christoph Urach, MSc[7], Gottfried Endel, MD[8], Michael Gyimesi, MSc[9], Goodarz Danaei, Sc.D.[10], Ankur Pandya, PhD[11], Thomas A. Gaziano, MD, MSc[12] and Uwe Siebert, MD, MPH, MSc, ScD[13], (1)Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (2)UMIT - University for Health Sciences, Medical Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA, Hall i.T., Austria, Hall i.T., Austria, Austria, (3)UMIT - University for Health Sciences, Medical Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA, Hall i.T., Austria, (4)Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (5)DEXHELPP, Vienna, Austria, (6)DEXHELPP, Institute for Analysis and Scientific Computing, Technical University Vienna, Vienna, Austria, (7)Institute for Analysis and Scientific Computing, Technical University Vienna, Vienna, Austria, (8)Main Association of Austrian Social Security Institutions, Vienna, Austria, (9)Austrian Public Health Institute, Vienna, Austria, (10)Department of Global Health and Population, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, (11)Center for Health Decision Science, Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Boston, MA, (12)Center for Health Decision Science, Department of Health Policy and Management, Harvard T.H. Chan School of Public Health/ Division of Cardiovascular Medicine, Brigham and Women's Hospital, Boston, MA, (13)UMIT, Dept. of Public Health, Health Services Research & HTA / Harvard Univ., Dept. Health Policy & Management, Institute for Technology Assessment, Hall i.T. (Austria)/Boston (USA), Austria

Purpose:

Real world evidence (RWE) is currently widely debated as source to gain information outside the artificial setting of clinical trials. Common challenges occurring with RWE are confounding, missing/misclassified data, no clear treatment assignment, dynamic treatment regimens, and switching. Within DEXHELPP, a large project on decision support tools for health-policy planning, the utilization of several databases within Austria is being endorsed.

The aim of this project was to develop a causal (counterfactual) approach for analyzing such datasets using the research question when-to-start statin treatment to prevent cardiovascular disease and to provide a set of evidence/variables that are needed to answer this question and assess the availability of these data.

Method(s):

We generate a study protocol following the "target trial" approach comparing four starting-statin-treatment strategies, i.e., no statin treatment, start when the ESC-SCORE exceeds 1%, 5%, and 10%. We describe the data structure needed for the causal assessment and file a data request to assess the data availability in the Austrian setting.

Result(s):

Individuals between 40 and 75 years of age without an acute myocardial infarction (MI) or stroke enter the study at the time they first exceed the risk-threshold of 1% and are followed up for 15 years. Potential time-independent and time-dependent confounding and selection bias are assessed using directed acyclic graphs (DAGs). Patients are replicated and assigned to each treatment arm. A per protocol analysis is applied censoring patients at time of protocol violation. Inverse probability of censoring weighting accounts for informative censoring and time-dependent confounding. Austrian hospital data including information on diagnoses, prescriptions, age, and gender are available. However, information on the outpatient services, continuous information on laboratory values, behaviors, and measurements are lacking. This information would need to be indirectly estimated using diagnoses data or data on filed prescriptions. The main challenge is the availability of repeated measurements on time-dependent confounders during the study time horizon.

Conclusion(s):

DAGs and a protocol following the "target trial" approach are important tools to guide study design and analytic approach in deriving causal effects from big data. This approach can guide the decision on whether or not an outcomes study will be successful and inform further evidence generation.

PS3-14 A BAYESIAN APPROACH TO EXPLORE AND IMPROVE METHODOLOGIES FOR THE IDENTIFICATION AND ESTIMATION OF SUBGROUP EFFECTS IN CLINICAL TRIALS

Quantitative Methods and Theoretical Developments (QMTD)

Claudia Geue, PhD[1], Neil Hawkins, PhD[1], Richard Grieve, PhD[2], Jeremy Oakley, PhD[3], Monica Hernandez, PhD[4] and James Carpenter, DPhil[2], (1)University of Glasgow, Glasgow, United Kingdom, (2)London School of Hygiene and Tropical Medicine, London, United Kingdom, (3)University of sheffield, Sheffield, United Kingdom, (4)University of Sheffield, Sheffield, United Kingdom

Purpose:

To compare statistical methods for: i) inference regarding subgroup (interaction) effects amongst a set of candidate variables, and ii) estimation of the magnitude of accepted subgroup effects. Simulations will help to inform the development of a fully Bayesian framework for subgroup analysis, accounting for effects of multiple testing and incorporating a framework for eliciting prior judgment from clinical experts.

Method(s):

Fixed, random and mixture models were compared. The fixed interaction effects model estimates subgroup interaction effects independently. The random interaction effects model treats the subgroup interaction terms as a random effect providing shrunken estimates. The degree of shrinkage will depend on the extent to which the observed variation is compatible with random chance. However, the model does assume exchange-ability of all putative subgroup effects. The mixture model allows the subgroup interaction terms to represent a mixture of: i) zero and ii) non-zero values. The models were compared in simulation studies where the true interaction effects are known. The number of candidate binary subgrouping variables and the number and magnitude of interaction effects were varied. The performance of the models was assessed by estimating sensitivity (true positive rate) measuring the ability to correctly identify subgrouping variables, and specificity (true negative rate) measuring the ability to correctly identify variables which do not identify subgroups. These were estimated across a range of thresholds for accepting a subgroup effect. These were allowed to vary between the methods. Bias was measured using Mean Absolute Deviation (MAD) of interaction effect estimates across all candidate variables and across estimated interaction effects for accepted subgroups.

Result(s): ROC curves were produced, plotting sensitivity against specificity for the three different models. Preliminary results provide evidence that the fixed effects model performs best in terms of inference, however the random and mixed effects models seemed to perform better with respect to estimation in accepted subgroups.

Conclusion(s): The underlying loss function needs to be considered when selecting both the threshold for acceptance of a subgroup effect and the statistical model used for analysis.

PS3-15 A SYSTEMATIC REVIEW ON COST-EFFECTIVENESS OF EARLY DETECTION AND PREVENTION STRATEGIES FOR ENDOMETRIAL CANCER
Applied Health Economics (AHE)

Gaby Sroczynski, MPH, Dr.PH[1], Artemisa Gogollari, Mag.sc.hum. Dr. med.[2], Annette Conrads-Frank, PhD[1], Martin Widschwendter, PhD[3], Nora Pashayan, MD, MSc, MSt, PhD[4] and Uwe Siebert, MD, MPH, MSc, ScD[5], (1)Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria, (2)Department of Public Health, Health Services Research and Health Technology Assessment, UMIT and ONCOTYROL - Center for Personalized Cancer Medicine, Division of Health Technology Assessment and Bioinformatics, Hall in Tirol, Austria, (3)UCL - University College London, Department of Women's Cancer, London, United Kingdom, (4)UCL - University College London, Faculty of Population Health Sciences, London, United Kingdom, (5)UMIT, Dept. of Public Health, Health Services Research & HTA / Harvard Univ., Dept. Health Policy & Management, Institute for Technology Assessment, Hall i.T. (Austria)/Boston (USA), Austria

Purpose: The purpose of this review was to systematically review current evidence evaluating the long-term effectiveness and cost-effectiveness of endometrial cancer early detection and prevention strategies in asymptomatic women.

Method(s): A systematic literature search on decision-analytic modelling studies assessing the cost-effectiveness of early detection and/or prevention strategies for endometrial cancer was performed using relevant electronic databases (Medline/Embase/Cochrane Library/CRD/EconLit). We summarized study characteristics and results in standardized evidence tables. Outcomes included quality-adjusted life-years (QALY), life-years gained (LYG), and incremental cost-effectiveness ratios (ICER). Economic results were converted to 2015 Euros using gross domestic product purchasing power parity and the consumer price index.

Result(s): We included seven studies evaluating early detection and prevention strategies for endometrial cancer in asymptomatic women with different cancer risk profiles. The strategies included: (1) Family history based genetic testing for germline mutations and prevention strategies for diagnosed mutation carriers, (2) endometrial cancer screening in women with different risk profiles, (3) preventive measures including prophylactic surgery in women at high risk for endometrial cancer. Genetic testing for Lynch syndrome in women 25-40 years old with a family history of endometrial cancer, followed by prevention strategies yielded ICERs below 40,000 Euro/QALY. One study evaluating annual screening with an hypothetical serum biomarker panel including prolactin reported ICERs between 19,000 and 39,000 Euro/LYG for women at increased risk for cancer (e.g. obese with BMI of 30 kg/m2, age 45-80; Tamoxifen usage up to 5 years, age > 60), and 58,000 Euro/LYG for women at average risk for cancer (general population, age 50), all compared to no screening as the next non-dominated strategy. Annual screening strategies with transvaginal ultrasound or biopsy were reported to be dominated in all analyses. Another study in obese women revealed ICERs above 82,000 Euro/LYG for cancer prevention using contraceptives or annual screening using biopsy.

Prophylactic surgery in asymptomatic women with Lynch syndrome at age 40 years achieved ICERs below 11,000 Euro/QALY.

Conclusion(s): The results of our study suggest that in asymptomatic women at high risk for endometrial cancer, genetic testing and preventive strategies for mutation carriers can be considered cost effective. Annual screening with a set of biomarkers, if validated, may have also the potential to be cost effective in women at increased cancer risk.

PS3-16 EARLY HEALTH TECHNOLOGY ASSESSMENT OF TISSUE ENGINEERED HEART VALVES IN ELDERLY PATIENTS
Applied Health Economics (AHE)

Simone Huygens, MSc, Erasmus MC, Rotterdam, Netherlands, Johanna Takkenberg, MD, PhD, Department of Cardiothoracic Surgery, Erasmus University Medical Center, Rotterdam, Netherlands and Maureen Rutten-van Molken, PhD, institute of Medical Technology Assessment, Rotterdam, Netherlands

Purpose: In-situ tissue engineering provides a promising method to create heart valves with potential to grow, repair and last a lifetime. They have potential to limit or even eliminate the disadvantages of existing heart valve substitutes (e.g. use of anticoagulants with subsequent risk of bleedings and reoperations due to valve degeneration). We aim to investigate the potential cost-effectiveness of surgical aortic valve replacement with tissue engineered heart valves (TEHV) compared to bioprostheses in elderly patients (>70 years).

Method(s): The potential cost-effectiveness of TEHV compared to bioprostheses was estimated with a lifetime patient-level simulation model. Different scenarios with varying performance and costs of TEHV were explored including 25.000 patients per run. In the best case scenario, we assumed that there is no valve dysfunction and risks of endocarditis (due to reduced sensitivity to inflammation), stroke, bleeding, and valve thrombosis (due to less thrombogenicity of the valve and no need for postoperative anticoagulants) reduce to the level of the general population. In the worst case scenario, we assumed an 25% increase in valve dysfunction and endocarditis and equal risks on other events compared to bioprostheses. In both scenarios we varied the costs of TEHV: 50% less, equal, or twice the costs of bioprostheses.

The clinical input parameters for the comparator treatment were derived from a national cardiac surgery database, meta-analysis, and other published literature. Cost input parameters were based on health insurance claims data and utilities were based on patient questionnaires (EQ-5D-5L) and published literature.

Result(s): The results are presented in the table.

	Best case scenario			**Worst case scenario**		
	ΔQALY	ΔCosts (€)	ICER (€/QALY)	ΔQALY	ΔCosts (€)	ICER (€/QALY)
Price TEHV (50% of bioprosthesis price)	0.41	-3.762	TEHV dominates	-0.05	-601	11.112
Price valves equal	0.41	-2.664	TEHV dominates	-0.05	679	Bioprosthesis dominates
Price TEHV (2 times bioprosthesis price)	0.41	-470	TEHV dominates	-0.05	3.240	Bioprosthesis dominates

Conclusion(s): The potential cost-effectiveness of TEHV in the aortic valve position in elderly patients is promising. In the best case scenario, the use of TEHV instead of bioprostheses results in QALY gains and cost reductions, also when TEHV are twice as expensive as bioprostheses. When TEHV can be produced at a lower cost than bioprostheses, the worst case scenario shows there is a loss of QALYs accompanied by net savings of €601.

PS3-17 COST-UTILITY ANALYSIS USING EQ-5D-5L DATA: DOES HOW THE UTILITIES ARE DERIVED MATTER?

Patient and Stakeholder Preferences and Engagement (PSPE)

Fan Yang, PhD, Centre for Health Economics, University of York, York, United Kingdom, Nancy Devlin, PhD, Office of Health Economics, London, United Kingdom and Nan Luo, PhD, Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

Purpose:

NICE recommends using crosswalk algorithm and the UK 3L value set to generate utilities from 5L data, although an English 5L value set is available. In this study, we explored how the two approaches would affect cost-effectiveness estimates for seven countries (England, Canada, the Netherlands, China, Japan, South Korea and Singapore) where both value sets are available.

Method(s):

Three Markov models assessing the cost-effectiveness of hemodialysis (HD) versus peritoneal dialysis (PD) for patients with end-stage renal disease (ESRD) were used to compare the incremental QALYs estimated using utilities derived from the 5L value set approach (5L) and the crosswalk approach (c5L). Utilities of HD and PD were estimated using primary 5L data collected from 150 patients. Direction and magnitude of the difference in 5L and c5L based incremental QALYs were examined for each country. The mathematic functions for estimating incremental QALYs were used to examine the relationship between utility and incremental QALYs.

Result(s):

In the model for non-diabetic ESRD patients, incremental QALYs based on 5L (range: 1.744 to 2.004) were similar to those based on c5L (range: 1.764 to 1.982). In the model for diabetic ESRD patients, incremental QALYs based on 5L (range: 1.454 to 1.633) were generally higher than those based on c5L (range: 1.365 to 1.568). In the model for all ESRD patients, incremental QALYs based on 5L (range: 0.290 to 0.480) were generally lower than those based on c5L (range: 0.315 to 0.493).

In all three models, 5L (c5L) generated more incremental QALYs when value sets of China (South Korea) were used. Such patterns were not observed for other counties. The magnitude of the differences in 5L and c5L based incremental QALYs varied with the value sets used. The largest and smallest differences were observed when value sets of China (range: 0.086 to 0.196) and the Netherlands (range: -0.009 to 0.023) were used, respectively. In all models, incremental QALYs was a positive linear function of both utility of PD and difference in utilities of PD and HD.

Conclusion(s):

The value set and crosswalk approaches could lead to different cost-effectiveness results when 5L data are used to estimate QALYs. The direction and magnitude of the difference vary with value sets used and may not be deduced from difference in utilities alone.

PS3-18 NEGOTIATING THE RELATIONSHIP BETWEEN PRICE AND VOLUME: IMPLICATIONS FOR ICER CALCULATIONS IN HETEROGENEOUS TREATMENT POPULATIONS

Applied Health Economics (AHE)

Helen O'Donnell, MPharm, MSc[1], Laura McCullagh, BSc (Pharm), PhD.[2], Michael Barry, MB, PhD[1] and Cathal Walsh, PhD[3], (1)National Centre for Pharmacoeconomics, Dublin 8, Ireland, (2)Trinity College Dublin, Dublin 8, Ireland, (3)University of Limerick, Limerick, Ireland

Purpose: Restricting reimbursement to subgroups of the licensed population and engaging in price negotiations with producers are two methods payers employ to optimise the use of new technologies. The purpose of this study is to describe a new method for incremental cost-effectiveness ratio (ICER) calculations in these scenarios.

Method(s):

In scenarios where an intervention is cost-effective in one subgroup and fails to be cost-effective in another, a decision may be made to restrict reimbursement to the cost-effective subgroup. However if it is profit maximising to do so, producers may be willing to negotiate a lower price for the intervention for reimbursement in the full population to generate increased sales. Traditionally ICER calculations would only consider the incremental costs and gains of the new intervention in the remaining subgroup. We propose a broader incremental comparison where the additional cost-savings from the reimbursed subgroup (that would be generated by a price reduction) are included in the calculation of the ICER of the new intervention for the other subgroup to reduce the incremental costs.

The implications of the new method are illustrated through the use of a simulation study. We propose that a new medicine "DrugA" is available to treat two diseases, 'Yellow' and 'Blue' at a cost of €13,000/patient /year. We assume a greater rate of disease progression and a larger treatment effect in 'Yellow' compared to 'Blue'. A Markov model with parameters compatible with these assumptions was generated. An incremental cost-effectiveness analysis was conducted versus standard of care. We assume a cost effectiveness threshold of €45,000/Quality Adjusted Life Year (QALY).

Result(s): At a cost of €13,000/patient/year the drug is cost-effective in 'Yellow' disease (ICER=€43,305/QALY) but not cost-effective in 'Blue' (ICER=€56,091/QALY). It is assumed DrugA is then reimbursed for 'Yellow' but price negotiations take place where the price would fall to €11,000 if reimbursement was extended. Under the traditional method DrugA is still not cost-effective in 'Blue' (ICER=€47,607/QALY). However under the new method when the cost-offsets from treatment of 'Yellow' are included, DrugA would be cost-effective and considered for reimbursement. (ICER=€38,807/QALY)

Conclusion(s):

The results show that the novel method has to potential to change conclusions regarding the cost-effectiveness of interventions. Use of this method should be considered when price negotiations take place in tandem with cost-effectiveness analysis.

PS3-19 COST EFFECTIVENESS OF BREAST CANCER SCREENING - A SYSTEMATIC REVIEW OF EUROPEAN ECONOMIC MODELS WITH FOCUS ON OVERDIAGNOSIS AND OVERTREATMENT

Applied Health Economics (AHE)

Nikolai Muehlberger, Priv.-Doz., DVM, MPH, UMIT - University for Health Sciences, Medical Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA, Hall i.T., Austria, Gaby Sroczynski, MPH, Dr.PH, Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria, **Artemisa Gogollari, Mag.sc.hum. Dr. med.**, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT and ONCOTYROL - Center for Personalized Cancer Medicine, Division of Health Technology Assessment and Bioinformatics, Hall in Tirol, Austria, Beate Jahn, PhD, UMIT - University for Health Sciences, Medical Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA, Hall i.T., Austria, Hall i.T., Austria, Austria, Nora Pashayan, MD, MSc, MSt, PhD, UCL - University College London, Faculty of Population Health Sciences, London, United Kingdom, Ewout W. Steyerberg, PhD, Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands; Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, Netherlands, Martin Widschwendter, PhD, UCL - University College London, Department of Women's Cancer, London, United Kingdom and Uwe Siebert, MD, MPH, MSc, ScD, UMIT, Dept. of Public Health, Health Services Research & HTA / Harvard Univ., Dept. Health Policy & Management, Institute for Technology Assessment, Hall i.T. (Austria)/Boston (USA), Austria

Purpose: Recent research suggests higher estimates of overdiagnosis and overtreatment and has started a new controversy about the benefits, harms and cost effectiveness of breast cancer (BCa) screening. Our objective was to review economic studies evaluating BCa screening in the European health care setting with specific focus on the consideration of overdiagnosis- and overtreatment-related harms.

Method(s): Relevant databases (Medline/Embase/Cochrane Library/CRD/EconLit) were systematically searched for decision-analytic modeling studies evaluating the cost effectiveness of BCa screening strategies in the European health care context. Study characteristics, methodological details and results were extracted into standardized evidence tables. Economic results, including incremental cost-effectiveness ratios (ICER) expressing cost per life year (LY) or quality-adjusted life year (QALY) gained, were converted to 2015 Euros using the GDP-PPP and CPI. To evaluate the consideration of overdiagnosis- and overtreatment-related harms, we investigated whether overdiagnosis was modeled (explicitly or implicitly) and long-term treatment-related disutilities (also accounting for overtreatment-related harms) were incorporated.

Result(s): Twenty five studies evaluating BCa screening were reviewed. Fifteen studies included screening strategies resembling currently established screening programs for women at average

risk, such as biennial or triennial mammography screening between the ages of 50-70. In all 15 studies, the ICERs of currently established breast cancer screening strategies fall far below 30,000 Euros/QALY or LY gained, which is considered to be cost-effective in most European countries. Fourteen of the 15 studies modeled the natural history of BCa in the absence of screening, which is an implicit way to consider overdiagnosis (lifetime risk of overdiagnosis = BCa incidence with screening – BCa incidence without screening). Despite that, none of the studies reported model predictions for overdiagnosis. Five of these 14 studies focused exclusively on screening effects on life expectancy, which is insufficient to account for the mainly nonfatal harms by overdiagnosis and overtreatment. Nine of these 14 studies considered effects on quality of life, including long-term disutility due to treatment. However, as most studies apply an averaged utility in the post-treatment phase, it was difficult to judge whether all relevant harms were included.

Conclusion(s): Economic studies suggest that BCa screening is effective and cost-effective in the European settings. However, it is difficult to judge whether overdiagnosis- and overtreatment-related harms are sufficiently accounted for. Further studies should make these important aspects more transparent.

PS3-20 FUTURE COSTS OF CANCER - PREDICTIONS BASED ON REGISTER DATA AND INNOVATION TRENDS

Applied Health Economics (AHE)

Christoffer Bugge, Msc, University of Oslo, Oslo, Norway, Erik Magnus Saether , PhD, Oslo Economics, Oslo, Norway and Ivar S. Kristiansen, MD, PhD, MPH, Department of Health Management and Health Economics, University of Oslo, Oslo, Norway

Purpose: The rising financial burden of cancer on health-care systems worldwide has led to the increased demand for evidence-based research on which to base planning and budget decisions. Models are well fit to predict future incidence and prevalence of cancer and health care costs in order to aid health care managers in allocating resources. The aim of this study is to model the future cost of cancer in Norway based on administrative register data and innovation trends.

Method(s): Based on real world register data, statistical predictions and clinical expertise we develop a prediction model for future costs of cancer in the period from 2017 to 2034. The health predictions is based on NORCAN (Cancer Registry of Norwegian). The treatment predictions are based on cancer stage, first/second/third/fourth line treatment, drug costs and administration costs. Health care utilization, productivity losses are based on register data while QALY gains were taken from the scientific literature. We decompose the key cost drivers and predict their future development. We use clinical experts to predict the consequences of innovations and how the future care delivery will take place.

Result(s): The cancer related costs of the health care, except nursing and care services are estimated at NOK 15 billion in 2014 (US$1.00=NOK8.50). Here, specialist health care comprised nearly 90 percent of the total costs. The costs of anti-cancer drugs were NOK 1,7 billion, with an annual growth of 26 percent (real price adjusted). Total societal costs (excluding value of lost life years) were estimated to NOK 40 billion. The value of life years lost were estimated to nearly NOK 100 billion. The prediction model will predict the future cost development for the key components.

Conclusion(s): Use of register data and prediction model can assist policy makers in preparing for future cancer challenges and more efficient cancer care.

PS3-23 REPRODUCTIVE DECISION SUPPORT FOR COUPLES AT RISK FOR HEREDITARY CANCER: AN EFFECT EVALUATION
Decision Psychology and Shared Decision Making (DEC)

Kelly Reumkens, Marly Tummers, Joyce Habets, Christine de Die and Liesbeth van Osch, Maastricht UMC+, Maastricht, Netherlands

Purpose: A predisposition for hereditary cancer is usually autosomal dominant, implying that there is a 50% risk of transmitting the mutation to offspring. This knowledge may evoke challenging decision-making processes among couples at risk for hereditary cancer and child wish. This project aims to support carrier couples during reproductive decision-making by the use of an online patient decision aid.

Method(s): The decision aid was developed according to the IPDAS guidelines and focuses on three main reproductive options that enable carrier couples to have a genetically related child: 1. natural conception without genetic testing, 2. natural conception with prenatal diagnosis, and 3. preimplantation genetic diagnosis (PGD). After completing a needs assessment, a usability test and a pilot study, the effectiveness of the decision aid is currently assessed in a nation-wide one group pretest-posttest study among couples with hereditary cancer and child wish. Main outcomes (decisional conflict, knowledge and decision self-efficacy) are assessed before use of the decision aid (T0), immediately after use (T1) and at two weeks (T2) and three months (T3) after use of the decision aid using paired samples t-tests. P-values of <0.05 were considered to indicate statistical significance.

Result(s): Preliminary short-term results of 100 participants at T0, T1, and T2 and long-term results of 79 participants at T3 indicate a significant decline in mean decisional conflict scores (all *p's <0.001*). Furthermore, couples' knowledge regarding available reproductive options was significantly increased (all *p's <0.001*). Couples' decision self-efficacy did not significantly increase at T1 ($p=0.124$) and T2 ($p=0.055$). At T3 decision self-efficacy increased significantly ($p<0.001$).The decision aid was overall highly appreciated by the target group and evaluated with a mean score of 8.4 (scale 1-10). The mean time spent using the decision aid was 24 minutes (range 2-104 minutes) and participants viewed on average 20 out of 36 pages.

Conclusion(s): Preliminary findings will be extended with data on a total of >80 participants at T3, including informed choice outcomes. In anticipation of these findings, we will strive for structural implementation of the decision aid in oncogenetic counselling of couples with hereditary cancer and child wish. Ultimately, it is expected that the decision aid will enable end-users to make an informed decision, which may lessen the negative psychological impact of decision making on couples' daily life and wellbeing.

PS3-24 EVALUATION OF A DECISION AID FOR TREATMENT DECISION MAKING IN PATIENTS WITH URINARY TRACT SYMPTOMS DUE TO BENIGN PROSTATIC HYPERPLASIA

Decision Psychology and Shared Decision Making (DEC)

Fieke van der Wijden, MD[1], Maarten Cuypers, MSc[2], Marieke de Vries, PhD[3], **Isabel de Angst, MD**[1], Romy R.E.D. Lamers, MD[1], Harm van Melick, MD, PhD[4], Jeltje de Beij, MD[5], Dennis Oerlemans, MD, PhD[6], Cees van de Beek, MD[7] and Paul J.M. Kil, MD, PhD[1], (1)Elisabeth-Tweesteden Hospital, Tilburg, Netherlands, (2)Radboud University Medical Center, Nijmegen, Netherlands, (3)Radboud University, Nijmegen, Netherlands, (4)St. Antonius Hospital, Nieuwegein, Netherlands, (5)Zuyderland Medical Center, Heerlen, Netherlands, (6)Laurentius Hospital, Roermond, Netherlands, (7)Maastricht University Medical Center, Maastricht, Netherlands

Purpose: Compared to evidence on treatment decision aids (DAs) for prostate cancer, evidence on treatment DAs for men with lower urinary tract symptoms due to benign prostatic hyperplasia (LUTS/BPH) is limited. Therefore, we evaluated the benefit of a newly developed web-based DA on treatment preferences and actual received treatments. We also evaluated the DA usability.

Method(s): Between July 2016 and January 2017, patients who consulted the urologist because of LUTS/BPH, were invited to use the DA. Patients were eligible for inclusion, if they had to choose between watchful waiting/lifestyle advices and medication or between (continuing) medication and surgery. Treatment preferences and responses on values clarification exercises (VCEs) were extracted from the DA. Furthermore, usability of the DA among patients and healthcare professionals was assessed using questionnaires.

Result(s): A total of 126 patients completed the DA and were included for analyses. Although 51% of patients did not had an initial preference, they were all able to indicate a final treatment preference after DA use. Among 80% of the DA users with initial preference, treatment preference did not change after DA use. In 79% the final treatment preference was in concordance with the actual received treatment. Most statements of the VCE discriminated well between final treatment preferences in DA. Overall, the content of the DA and its usefulness was well accepted by patients and their healthcare professionals.

Conclusion(s): Most patients with LUTS/BPH received treatments in concordance with their final treatment preference after DA use. The DA succeeds in clarifying patients' values and preferences. Therefore, our findings suggest that the implementation of this DA could support the decision-making process for both patients and healthcare professionals, in either eliciting or confirming their final treatment preference.

PS3-25 COMMUNICATING DETAILED INFORMATION ABOUT COLORECTAL CANCER SCREENING TO CITIZENS WITH LOWER EDUCATIONAL ATTAINMENT USING AN ELECTRONIC DECISION AID: A QUALITATIVE STUDY
Decision Psychology and Shared Decision Making (DEC)

Pernille Gabel, MD, PhD Fellow[1], Mette Bach Larsen, PhD, MScPH[1], Pia Kirkegaard, PhD, MA[1], Adrian Edwards[2] and Berit Andersen[1], (1)Department of Public Health Programmes, Randers Regional Hospital, Randers NØ, Denmark, (2)Division of Population Medicine, Cardiff University, Cardiff, United Kingdom

Purpose: Compared to average educational attainment citizens, citizens with lower educational attainment (LEA) less frequently take up colorectal cancer (CRC) screening, and to a lesser degree read and understand conventional screening information. The information needs of LEA citizens ranges from a clear recommendation to elaborating information. Some decision aids (DAs) are designed to support informed decision making about CRC screening participation, but none embraces diversion in information needs. The aim of this study was to develop such a DA tailored to LEA citizens.

Method(s): A prototype of the DA was developed based on the IPDAS guidelines along with LEA citizens' information needs. The online DA presented information in steps. Values clarification questions were included and answers summarized in a choice-barometer on the last page. Statistics were presented in both relative and absolute numbers. Both user testing, peer review and field testing were conducted using focus group and telephone interviews and email correspondences with LEA citizens and healthcare professionals. Data was analyzed using thematic analysis.

Result(s): The citizens found the DA easy to understand and the text of suitable length. They easily and intuitively navigated around the DA, and stated, that they felt encouraged to think about benefits and harms of CRC screening without being overloaded with information.

Conclusion(s): This DA represents a new way of communicating detailed information about CRC screening to LEA citizens, enabling citizens to make value-based decisions. Further, this work might serve as an inspiration when developing information material in other screening programs.

PS3-26 DEVELOPMENT AND VALIDATION OF PERSONALISED PREDICTION TO ESTIMATE FUTURE RISK IN PATIENTS WITH ASTHMA; USING CLINICAL PARAMETERS AND EARLY TREATMENT RESPONSE

Decision Psychology and Shared Decision Making (DEC)

Suzanne Boer[1], Persijn Honkoop[1], Rik Loijmans[2], Jiska Snoeck-Stroband[1], Gerben ter Riet[2], Tjard Schermer[3] and Jacob Sont[1], (1)Leiden University Medical Center, Leiden, Netherlands, (2)Academic Medical Centre, Amsterdam, Netherlands, (3)Radboud University Medical Centre, Nijmegen, Netherlands

Purpose:

At regular patient visits for asthma the risk of future adverse outcomes should be assessed. Therefore, we aimed to develop, and validate, a risk prediction score for level of future risk, including patient characteristics and information on early treatment response.

Method(s):

We used data of 304 adult asthmatics from a 12-month primary care RCT with three-monthly assessments. With logistic regression we modelled the association between level of future risk and patient characteristics including early treatment response. Future risk was defined as Asthma Control Questionnaire (ACQ) score ≥ 1.5 at 12 months or the experience of at least one exacerbation during the final six months. Predictive performance was assessed by AUROC and HL-test. We developed a risk prediction score based on the regression coefficients.

Result(s):

Performance of the risk prediction score improved taking into account data on early treatment response (AUROC 0.84) compared to a model containing only baseline characteristics (AUROC = 0.78). The score includes six easy to obtain predictors; sex, ACQ-score and exacerbations in the previous year at baseline and at 1st follow-up ACQ, smoking status and exacerbations in the previous three months (indicating early treatment response). External validation yielded an AUROC of 0.77. The risk prediction score classified patients into three risk groups, low (absolute risk 11.7%), intermediate (47.0%) and high level of risk (72.7%).

Conclusion(s):

We developed and externally validated a risk prediction score, quantifying both level of current asthma control and the guideline-defined future risk. Patients' individual risk can now be estimated in an easy way, as proposed but not specified, by asthma management guidelines.

PS3-27 CARDIOVASCULAR RISK COMMUNICATION BASED ON A VISUAL DECISION AID

Decision Psychology and Shared Decision Making (DEC)

Anders Elkaer Jensen[1], Jens Soendergaard[1], Rodney T. Jackson[2] and **Jesper Bo Nielsen, PhD**[1], (1)Research Unit for General Practice, Odense, Denmark, (2)Section of Epidemiology and Biostatistics, Auckland, New Zealand

Purpose: Less than 50% of patients diagnosed with hypertension and treated in general practice have reached a blood pressure within the recommended level of the national guideline. Compliance is a main problem for these patients, but efficient tools for increasing patient compliance are sparse. Our objective is to evaluate the risk-prediction and risk-communication tool: "Your Heart Forecast", to see if it can improve patient compliance, health literacy and empowerment. The software is based on visual illustrations of risk levels and functions interactively, so that patients can change their risk profiles in the program and thereby modulate the visual display of their predicted average risk.

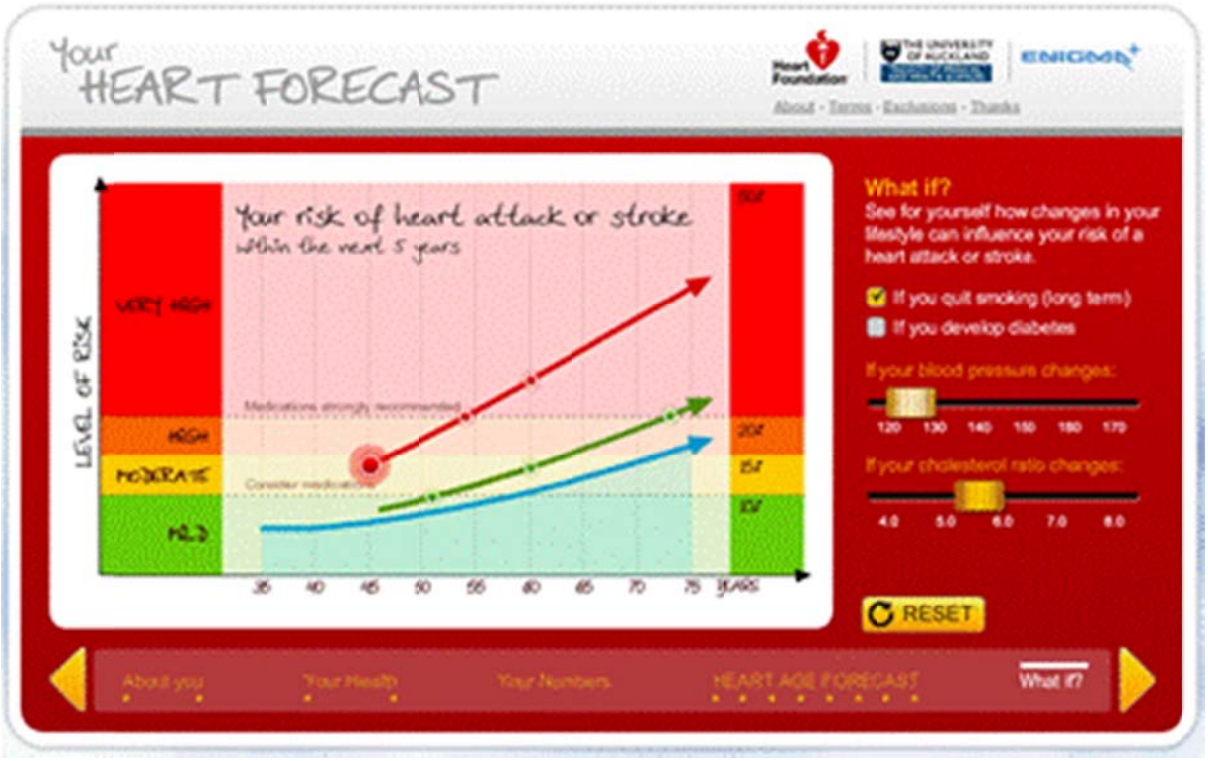


Figure 1. Risk of heart attack or stroke within the next 5 years for male, aged 45 years, smoker, BP = 140/85 (present risk – red line; ideal risk – blue line; predicted risk following suggested change in risk factors – green line).

Method(s): Patients will be followed in a cluster-randomised controlled trial in the setting of general practice using surveys at inclusion and after 6 and 12 months. Besides questionnaires we will measure the participants blood pressure as a hard outcome and draw data from the rich source of registers available in Denmark (socio-economy, prescription database, disease

registers). After 6 months, we will conduct qualitative interviews with a subgroup of patients from the intervention group.

Result(s): Expectedly we will find that the use of Your Heart Forecast can lower patients' blood pressure and increase their compliance, health literacy and empowerment. We expect to show an increase in general health literacy and patient empowerment measured by Patient Activation Measure(PAM13).

Conclusion(s): We expect that this software can improve patient compliance and thereby be an efficient tool to implement in the national blood pressure control program in Denmark. We hypothesize that the cost of using this programme is less than expenses for hospitalisation due to complications and comorbidity to hypertension in non-compliant patients.

PS3-30 MEASURING SHARED DECISION MAKING IN ONCOLOGY: DEVELOPMENT OF THE ISHARE QUESTIONNAIRES

Decision Psychology and Shared Decision Making (DEC)

Hanna Bomhof-Roordink, MSc, Fania R. Gaertner, PhD, Anne M. Stiggelbout, PhD and Arwen H. Pieterse, PhD, Medical Decision Making, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands

Purpose:

To 1) determine the construct of shared decision making (SDM) for oncology, 2) develop a patient and oncologist questionnaire to measure SDM, and 3) determine their content validity and comprehensibility.

Method(s):

The construct was defined based on a systematic literature review of conceptual models of SDM, a newly-developed model of SDM in oncology that was based on a large qualitative study among stakeholders, and input from five international SDM-experts. We selected the elements most typical for SDM from these sources, and formulated items for each element. International SDM-experts, cancer patients and oncologists rated content validity in an online questionnaire. We assumed a formative measurement model, i.e., the items are not necessarily closely associated. Item reduction then cannot be conducted based on factor analysis, so we asked cancer patients in an online field-test to indicate the most important item per element, to inform further item reduction. Comprehension of the items was tested in cycles of cognitive interviews with cancer patients and oncologists.

Result(s):

We identified 17 elements based on the literature, the newly-developed model and expert opinion, and developed 153 corresponding items. Content validity was rated on item level in two cycles by five SDM-experts and 12 cancer patients and we arrived at a selection of 19 elements and 66 items. Subsequently, oncologists rated content validity on element level. Combining these results with those from the field-test among 131 cancer patients, we narrowed the selection down to 14 elements and 24 items. Next, cognitive interviews were held with nine patients and five oncologists, and items were removed, adapted and tested further. The final patient and oncologist questionnaires each contain the same 15 items, three of which explicitly assess patient behavior, and phrased according to the target group. A six-point response format was chosen, ranging from ‘Not done at all’ to ‘Done completely’.

Conclusion(s):

This study provides a short patient and oncologist questionnaire to measure SDM in oncology, based on a clearly defined construct and a thorough development process. It assesses both oncologist and patient behaviors, and focuses on the SDM-process rather than on one consultation.

PS3-31 USING A SHORT GRAPH LITERACY SCALE TO PREDICT PRECURSORS OF HEALTH BEHAVIOR CHANGE

Decision Psychology and Shared Decision Making (DEC)

Yasmina Okan, PhD, Centre for Decision Research, Leeds University Business School, Leeds, United Kingdom, Eva Janssen, PhD, Maastricht University, Maastricht, Netherlands, Mirta Galesic, PhD, Santa Fe Institute, Santa Fe, NM and Erika A. Waters, PhD, Washington University School of Medicine, St. Louis, MO

Purpose: Visual displays can facilitate risk communication and promote better health choices. Their effectiveness in predicting risk comprehension is influenced by graph literacy. Yet, it is unclear how graph literacy relates to other cognitive, affective, and conative precursors of health behavior change. We constructed a new 4-item scale to examine this question in a probabilistic US sample.

Method(s): Initial scale construction was conducted in a laboratory, and involved selecting the 4 items that had the most favorable psychometric properties from the validated 13-item graph literacy scale by Galesic & Garcia-Retamero (2011). For validation, the 4 items were used to predict understanding of health risk information in probabilistic national samples in Germany and the US. The new 4-item scale was then administered in a separate study involving a probabilistic US sample comprised of 47% racial/ethnic minorities and 46% with limited formal education. Participants viewed a risk ladder displaying hypothetical risk calculator results for diabetes, stroke, heart disease, and colon cancer. Next, they completed items assessing key cognitive, affective, and conative precursors of behavior change described in theories of health behavior.

Result(s): Graph literacy was associated in theoretically expected ways with all cognitive precursors assessed. Specifically, it was positively associated with risk comprehension, message acceptance, and response efficacy (i.e., believing that engaging in a given behavior will reduce risk), but negatively associated with cognitive risk perception and perceived severity of the diseases. Although numeracy had an independent contribution for most factors, graph literacy accounted for unique variance in most cases. Results for affective precursors overall mirrored those for cognitive precursors, although numeracy was a stronger predictor for some affective factors (e.g., affective risk perceptions). Moreover, graph literacy (but not numeracy) predicted some key conative outcomes (e.g., lower defensive processing).

Conclusion(s): Our data suggest that the new 4-item scale is a fast and psychometrically valid method for measuring objective graph literacy. Additionally, we documented the first evidence that graph literacy predicts key cognitive, affective, and conative precursors of health behavior change. Overall, our findings highlight the theoretical and practical relevance of graph literacy. We discuss implications for health risk communication practice and research and mechanisms underlying the association of graph literacy with the different precursors.

PS3-33 INFORMING THE DESIGN OF A NEW PREFERENCE WEIGHTED VALUE FRAMEWORK FOR ONCOLOGY: A MIXED METHODS APPROACH

Patient and Stakeholder Preferences and Engagement (PSPE)

Camilla Baba, PhD[1], Andrew Briggs, DPhil[2], Alison Lipitz-Synderman, PhD[2], Susan Chimonas, PhD[2], Anna Kaltenboeck, PhD[2], Emma McIntosh, PhD[1] and Peter Bach, MD[2], (1)Health Economics and Health Technology Assessment, University of Glasgow, Glasgow, United Kingdom, (2)Memorial Sloan Kettering Cancer Center, New York, NY

Purpose:

As a response to rising oncology therapy costs, multiple value frameworks are emerging. However, involvement in their design and conceptualisation from economists has been limited and none have been based on preference weightings as legitimate indicators of value for included health gain, nor have the effects of contextual factors on preference weightings been adequately differentiated and assessed. This study uses qualitative methods to identify oncology treatment attributes and contextual factors that informs the design of a discrete choice experiment (DCE) to elicit values and develop a value framework with relative preference weightings for future decision-makers.

Method(s):

In January/February 2017, three focus groups were conducted with cancer patients, oncology physicians and nurses in New York. Following nominal group technique (NGT) methodology, the groups identified and prioritised cancer therapy treatment and delivery attributes. Qualitative thematic analysis of transcripts was conducted in order to identify treatment-related attributes and contextual factors. These attributes and factors were employed in the design of a discrete choice experiment.

Result(s):

A total of 30 attributes were identified. Cancer therapy health gains (efficacy and toxicity) were the first priority across focus groups. Thereafter, groups differed in their priorities. Patients (n=8) highlighted the importance of evidence quality, existence of alternative treatments, the long-term adverse effects, how well established the treatment is and treating oncologist/centre reputation. Nurses (n=10) focused on administration harms, communication and treatment innovation. The physicians (n=6) prioritised carers and family inconvenience and burden, costs to patients, functional outcomes, and the societal costs of the treatment with reference to the disease burden to be addressed. Following a thematic analysis, a conceptual framework was developed whereby treatment-related attributes were distinguished from the contextual factors in which treatment takes place, and categorized by whether they were health-related, cost-related and non-health-related, A DCE was then designed to estimate the preference weights for treatment attributes, framing trade-offs under varying contextual attributes to assess their effect on preference weights.

Conclusion(s):

This work assessed the viability of NGT to comprehensively identify attributes related to treatment decision-making among different stakeholders, and to distinguish treatment-specific attributes from contextual factors that may influence preference weights.

PS3-34 AN OVERVIEW OF HEALTH UTILITY VALUES IN THYROID CANCER, THYROID NODULES, THYROID AUTONOMY AND INTELLECTUAL DISABILITY RELATED TO IODINE DEFICIENCY

Patient and Stakeholder Preferences and Engagement (PSPE)

Vjollca Qerimi, MPharm[1], Monika Buchberger, MSc[2], Sibylle Puntscher, PhD[3], Conrads-Frank Annette, PhD[2], Uwe Siebert, MD, MPH, MSc, ScD[4] and Ursula Rochau, MD, MSc[2], (1)UMIT - University for Health Sciences, Med. Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA / Faculty of Pharmacy, University of Skopje, Macedonia, Hall i. T., Austria, (2)UMIT - University for Health Sciences, Medical Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA, Hall i. T., Austria, Austria, (3)Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (4)UMIT, Dept. of Public Health, Health Services Research & Health Technology Assessment / Harvard Univ., Dept. Health Policy & Management, Institute for Technology Assessment / ONCOTYROL, Division for HTA, Hall i. T. (Austria) / Boston (USA), Austria

Purpose: To give an overview on published health utility values (HUV) for health states that can be associated with iodine deficiency (ID) including thyroid cancer, thyroid nodules, thyroid autonomy, or intellectual disability.

Method(s): We conducted a comprehensive systematic literature search in PubMed/MEDLINE, Tuft's CEA Registry, and Cochrane to identify studies reporting on HUV related to thyroid cancer, thyroid nodules, thyroid autonomy, or intellectual disability. We used evidence tables to systematically extract and summarize health utility parameters based on disease, population, type of preference-based questionnaire and original source (if applicable), methods for utility estimation when reported, and HUV.

Result(s): Overall, we identified 1326 studies. After title/abstract and full-text screening, 32 studies were included. These studies report HUV related to health states in thyroid cancer (n=16), thyroid nodules (n=11), and intellectual disability (n=5). No study was identified reporting HUV for thyroid autonomy.

Health utilities were derived with different outcome measurements including SF-36, SF-6D, EQ-5D, time-trade-off questionnaires, standard gamble, Health Utility Index (HUI), and expert opinion. In several studies, such as cost-utility analyses, utilities were taken from other published literature (e.g., quality-of-life studies). However, in some of the related original sources, data were originally collected for medical conditions unrelated to the health states of interest.

The identified utilities for thyroid cancer were dependent on treatment strategies and adverse events and ranged between 0.60-0.80 (without treatment), 0.71-0.85 (pre-ablation), 0.60-0.80 (post-ablation), 0.70-0.85 (post-radioiodine treatment), 0.20-0.99 (with complications related to treatment), and 1 (surveillance after hemithyroidectomy or total thyroidectomy/disease free). The utilities for thyroid nodules were reported for health states such as disease-free after

hemithyroidectomy or total thyroidectomy (around 0.9), recurrence (in a range of 0.5-0.6), and pre- and post-ablation (in a range of 0.5-0.8). Utility values for intellectual disability ranged from 0.5-0.8 depending on the underlying disorder (e.g., congenital hypothyroidism, childhood meduloblastoma, meningitis, subclinical hypothyroidism).

Conclusion(s): We identified several studies reporting HUV for different health states of thyroid cancer, thyroid nodules, and intellectual disability related to iodine deficiency. In the future, the identified utilities can be used in decision-analytic modeling to evaluate the benefit-harm ratio of prevention programs for iodine deficiency disorders. However, deciding on the HUV for the relevant health states requires more detailed analysis on the population, the different treatment options and the stage of the disease.

PS3-35 AN OVERVIEW OF HEALTH-RELATED QUALITY OF LIFE DATA IN MULTIPLE MYELOMA AND MAPPING OF HEALTH UTILITY VALUES
Patient and Stakeholder Preferences and Engagement (PSPE)

Vjollca Qerimi, MPharm[1], Igor Stojkov, MPharm[2], Sibylle Puntscher, PhD[3], Beate Jahn, PhD[4], Andrea Manca, PhD, MSc[5], Bernhard Holzner, PhD[6], Eva Gamper, PhD[6], Wolfgang Willenbacher, MD[7], Roman Weger[8], Lucia Neppl[8], Dan Greenberg, PhD[9], Uwe Siebert, MD, MPH, MSc, ScD[10] and Ursula Rochau, MD, MSc[11], (1)UMIT - University for Health Sciences, Med. Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA / Faculty of Pharmacy, University of Skopje, Macedonia, Hall i. T., Austria, (2)Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (3)Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (4)UMIT - University for Health Sciences, Medical Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA, Hall i.T., Austria, Hall i.T., Austria, Austria, (5)The University of York, Centre for Health Economics, United Kingdom/ Department of Population Health, Luxembourg Institute of Health, Luxemburg, York, United Kingdom, (6)Department of Psychiatry and Psychotherapy, Medical University of Innsbruck, Innsbruck, Austria, Innsbruck, Austria, (7)Internal Medicine V, Hematology and Oncology, Medical University, Innsbruck, Austria, (8)Oncotyrol – Center for Personalized Cancer Medicine, Innsbruck, Austria, Innsbruck, Austria, (9)Department of Health Systems Management, Ben-Gurion University of the Negev, Be'er Sheva, Israel, (10)UMIT, Dept. of Public Health, Health Services Research & HTA / Harvard Univ., Dept. Health Policy & Management, Institute for Technology Assessment, Hall i.T. (Austria)/Boston (USA), Austria, (11)Department of Public Health, Health Services Research and Health Technology Assessment UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i. T., Austria

Purpose: The goal of our study is to give an overview on published health-related quality of life (HRQoL) data in patients with multiple myeloma (MM) and to assess whether the identified data could be used to derive health utility values (HUV) using mapping methods.

Method(s): We updated a systematic literature search conducted in PubMed/MEDLINE and performed an additional comprehensive search in the Cochrane Central Register of Controlled Trials (CENTRAL), Tufts CEA Registry, Web of Science, EQ-5D and European Organization for Research and Treatment of Cancer (EORTC) database to identify studies reporting on HRQoL or HUV in patients with MM derived from the EQ-5D and EORTC questionnaires (QLQ-C30 or QLQ-MY20). We used evidence tables to systematically extract and summarize HRQoL data and HUV from each study based on type of questionnaires used, study population/stage of the disease, number and timing of follow-ups, treatment combinations, and reported functional scores or utility values.

Result(s): We initially identified 847 studies reporting HRQoL data and HUV. After two phases of screening, 30 studies were included in further analyses. All included studies applied the

EORTC QLQ-C30 questionnaire to report HRQoL data. Six studies reported health utility parameters derived from the EQ-5D questionnaire and 18 studies presented HRQoL data derived from the EORTC QLQ-MY20 questionnaire. Based on the stage of the disease reported, studies were classified in: non-interventional (n=1); 1st-line treatment including induction therapy and autologous stem cell transplantation (ASCT) (n=9); 1st-line treatment without ASCT (n=8); ≥ 2nd line-treatments including induction therapy and ASCT (n=3); ≥ 2nd line-treatments (n=1); maintenance therapy (n=2); relapse/refractory treatment (n=5); and non-specified (n=1). Treatment combinations included induction therapy before ASCT, bortezomib, thalidomide, lenalidomide, carfilzomib, and/or pamidronate alone or combined with melphalan, prednisone, and/or dexamethasone. Only 14 studies reported complete functional scores and may be used for mapping to predict HUV for comparable patient-population data sets.

Conclusion(s): We identified several studies reporting on HRQoL data and HUV in patients with MM derived from EQ-5D and EORTC, some of which reported HRQoL or HUV for a defined treatment combination and some reported values of patients in aftercare who had received several functional scores from EORTC questionnaires, those data sets are not suitable to use them in mapping methods to derive HUV.

PS3-36 FACTORS INFLUENCING ANTIBIOTIC CONSUMPTION BEHAVIOUR - AN EXPLORATIVE STUDY

Patient and Stakeholder Preferences and Engagement (PSPE)

Mirko Ancillotti, MA[1], Stefan Eriksson, PhD[1], Jorien Veldwijk, PhD[2], Jessica Nihlen Fahlquist, PhD[1], Dan Andersson, PhD[3] and Tove Godskesen, RN, PhD[4], (1)Centre for Research Ethics & Bioethics, Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden, (2)Erasmus Choice Modelling Centre, Erasmus University , Rotterdam, Netherlands, (3)Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden, (4)Department of Health Care Sciences, Ersta Sköndal Bräcke University College, Stockholm, Sweden

Purpose: To identify the factors promoting and thwarting judicious antibiotic use in the Swedish population and their implications for communication strategies.

Method(s): Data were collected through focus group discussions with 23 Swedes from the general population. Recruitment was performed through a site-based approach and with purposive sampling, aiming for maximum variation in gender, age, and education level for the groups' composition. The Health Belief Model was used as a theoretical framework and transcripts were analyzed with qualitative content analysis.

Result(s): Antibiotic resistance (AR) was identified by participants as a health problem with potentially terrible consequences. The severity of the problem was perceived more strongly than the actual likelihood of being affected by it. Metaphors such as climate change were abundantly employed to describe AR as a slowly emerging problem that is somehow creeping up on us.

There was a tension between individual (egoistic) and collective (altruistic) reasons for engaging in judicious behaviour. In discussions, a need for empowerment through good health communication from authorities and family physicians was stressed. To have more knowledge of antibiotics and AR would increase perceived individual self-efficacy and promote judicious behaviour. Another identified cue to action was the awareness of participating in international efforts to curb AR.

Conclusion(s): While it can be considered a basic requirement, just informing the public is insufficient; comprehension of the attitudes and public health behaviour influencing AR is vital for tailoring campaigns, enabling health education programmes to convey relevant, effective messages, and thus induce behaviour modification.

Knowledge about antibiotic consumption and resistance, as well as values such as altruism and trust in the health care system, has significant influence on perceptions of individual responsibility and on behaviour. These factors should therefore be emphasised in health communication. To instead frame AR as a slowly emerging disaster risks diminishing the perception of being susceptible to AR and related health issues.

PS4-1 ARE WE TREATING THE RIGHT PATIENTS? FROM TRIALS TO GUIDELINES TO CLINICAL PRACTICE

Health Services, Outcomes and Policy Research (HSOP)

Eveline Wiegers, BSc[1], Esmee Venema, MD[2], Diederik W.J. Dippel, MD, PhD[3] and Hester F. Lingsma, PhD[2], (1)Department of Public Health, Erasmus University Medical Center Rotterdam, Rotterdam, Netherlands, (2)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (3)Department of Neurology, Erasmus University Medical Center, Rotterdam, Netherlands

Purpose: In the last years, seven randomized controlled trials proved the benefit of endovascular treatment (EVT) compared with best medical care in selected acute stroke patients with large artery occlusion. These trials used different selection criteria which led to uncertainty of treatment effect in specific subgroups. We aim to investigate how the available evidence is translated into international guidelines and the adherence to these guidelines in clinical practice.

Method(s): Different databases were searched to investigate differences and similarities in recommendations between guidelines for acute stroke care that were updated after the results of the EVT trials. The MR CLEAN Registry, containing information of all patients treated with EVT in the Netherlands between March 2014 and June 2016 (n = 1607), was used to investigate guideline adherence.

Result(s): In total, six guidelines were selected for further analysis. Not all guidelines provided guidance in the treatment of specific subgroups of patients such as younger age or mild strokes. Although the same trials were the main source for all guidelines, different recommendations were made regarding treatment of patients with a distal occlusion and patients with a time since onset longer than 5 hours. From the 1607 patients in total, only 620 (38.6%) patients would have been treated if all selected guidelines were adhered to. Having an occlusion of the M2 segment or a small infarct were the most common reasons for not being treated conform guidelines, i.e. patients with these characteristics were treated with EVT in clinical practice while the guidelines do not recommend so.

Conclusion(s): Even in the presence of seven relatively homogeneous RCTs, the translation from evidence to guidelines is not clear-cut. Many patients that were treated with EVT in the Netherlands would not have been treated if the guidelines were followed. Our results implicate that guidelines should be updated and refined in order to guarantee an unambiguous policy in the decision which stroke patients should receive EVT. Future research is necessary to investigate why physicians decide to deviate from guidelines.

PS4-3 DOES SMOKER PROFILE DETERMINE ADOPTION OF A DECISION AID TO IMPROVE EVIDENCE-BASED CESSATION SUPPORT UPTAKE?

Decision Psychology and Shared Decision Making (DEC)

Thomas Gultzow, Master of Science[1], Eline Smit, PhD[2], Raesita Hudales, Master of Science[1], Carmen Dirksen, PhD[3] and Ciska Hoving, PhD[1], (1)Maastricht University, Maastricht, Netherlands, (2)Universiteit van Amsterdam, Amsterdam, Netherlands, (3)Maastricht UMC+, Maastricht, Netherlands

Purpose:

To identify smoker profiles based on decision-related cognitions and to investigate the influence of these profiles on smokers' intention to use an online decision aid to choose evidence-based smoking cessation support tools.

Method(s):

A cross-sectional study involving 200 Dutch smokers with the intention to cease smoking within six months was conducted in January 2018. Hierarchical cluster analysis based on health locus of control and decision-making style was employed to classify smokers into groups. Subsequently, clusters were compared on demographics, smoking behavior, smoking-related cognitions and intention to use the decision aid by employing independent-samples t-tests, Mann-Whitney U tests and linear regression analyses.

Result(s):

Two clusters were identified; intenders and non-intenders. Intenders scored significantly higher on the dependent, avoidant and regret decision-making style, while scoring lower on the spontaneous and intuitive decision-making style, as well as, scoring lower on chance locus of control and powerful others locus of control. The time elapsed since a last cessation attempt was shorter. And while their autonomous motivation to stop smoking was lower, their attitude regarding cessation support was higher. Cluster membership statistically significantly predicted intention to use a decision aid in the future, $F(1, 198) = 9.931$, $p < .002$, adj. $R^2 = .048$.

Conclusion(s):

Smokers that are more interested in a future decision-aid aimed at helping them in their decision to use evidence-based smoking cessation support tools seem to be people that are less likely to make intuitive and spontaneous decisions. They seem to be more decision-making averse overall, while also be more dependent on others to make a choice. Furthermore, they tend to be prone to regret their choices. These results suggest that decision-aids aimed at facilitating the decision of smokers regarding a future cessation attempt should be designed to address the unique characteristics of the identified subgroups.

PS4-4 ACCEPTABILITY AND EFFECTS OF A LUNG CANCER SCREENING (LUCAS) DECISION AID

Health Services, Outcomes and Policy Research (HSOP)

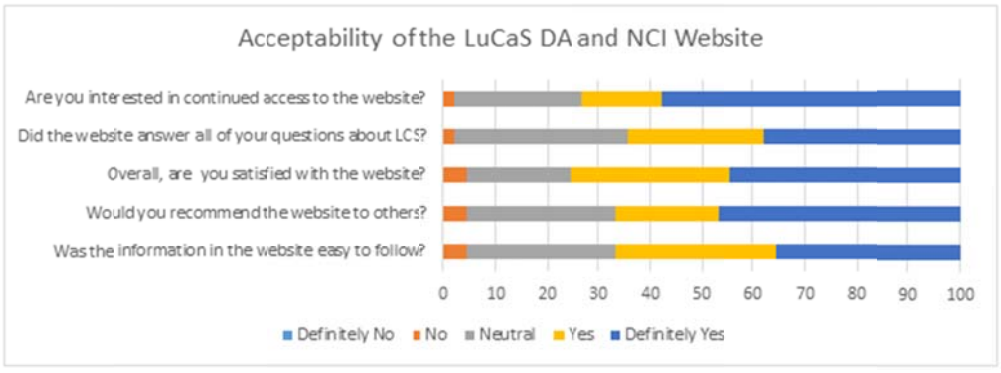
Margaret M. Byrne, PhD[1], Meghan Johnson[2] and Jamie L. Studts, PhD[2], (1)Moffitt Cancer Center, Tampa, FL, (2)University of Kentucky College of Medicine, Lexington, KY

Purpose: We evaluated the efficacy and acceptability of the novel LuCaS DA aid in a pilot randomized controlled trial with the US NCI website information on lung cancer screening (LCS) as a control.

Method(s): Individuals who are at a higher risk of lung cancer due to cigarette smoking (n=50) were recruited in Miami FL and Lexington KY. Participants were randomized to reviewing either the LuCaS DA (n=23) or the US NCI lung cancer screening website (n=27), and completed 3 surveys (baseline, 2-weeks post enrollment, and 2-months post enrollment). Information was collected on demographics, LCS knowledge, LCS Decisional Conflict, acceptability of the DA overall and specific DA components, including a conjoint exercise aimed at clarifying participant preferences.

Result(s): Participant demographics did not vary by randomization group, and included: average age of 52.6 (SD 5.1); 79% female; 62% White; 16% Hispanic; 34% High school education or less and 24% College graduate; 53% Employed; and 88% with some health care insurance; 28% reported health as Excellent or Very Good, and 36% as Fair or Poor. Intervention acceptability reported at the 2-week post survey for both the LuCaS DA and US NCI website was high as measured on a 1-5 scale (see graph) with no differences between the interventions. Conjoint acceptability questions included: "Did the conjoint exercises make sense to you?" with 42% responding "Very much"; and "How much did the conjoint exercises help you explore your personal preference?" with 42% again responding "Very much". Objective LCS knowledge was low but improved slightly after the intervention. Specifically, only half of participants correctly answered at least 1 CT scan knowledge or LCS elibilitiy question at baseline, but this percentage increased to two-thirds at the 2-week survey. Decisional conflict as measured by the Decision Conflict Scale was lower at 2-week post survey than at baseline for all participants and for both the DCS overall score and for all subscales.

Conclusion(s): The LuCaS DA aims to improve individual's ability to make an informed, preference-sensitive decision about lung cancer screening. Our pilot RCT shows that the LuCaS DA is at least as acceptable and useful as the US NCI lung cancer screening site for individuals who might be considering LCS.



PS4-5 INFECTIOUS DISEASE EMERGENCY DEPARTMENT SERVICE UTILIZATION AND DISEASE BURDEN AMONG ELDERLY PEOPLE IN NURSING HOMES

Health Services, Outcomes and Policy Research (HSOP)

Meng-Lun Hsieh, Department of Medical Education, Taipei Veterans General Hospital, Taipei, Taiwan and Vivian Chia-Rong Hsieh, PhD, Department of Health Services Administration, China Medical University, Taichung, Taiwan

Purpose: With the rapidly increasing rate of aging in Taiwan, we wish to investigate the extent of infectious disease (ID) emergency visits, hospitalizations and associated health care costs among nursing home elderlies in the country.

Method(s): Here we conduct an ecological study with economic analysis using a population-based emergency department (ED) sample from 2002-2013. From Taiwan's claims-based National Health Insurance Research Database (NHIRD), we identified adults aged 65 or older who has visited ED with a primary diagnosis of the following IDs: tuberculosis (ICD-9-CM codes: 010-018), upper respiratory infection (URI) (465), pneumonia (486), cellulitis (528.3) and urinary tract infection (UTI) (599.0). Rates of ED visits were first calculated annually and then stratified by age group, sex, area of residence and nursing home residents (yes or no). We then examined ID-specific ED visits, subsequent hospital admissions and deaths, and computed their associated total health care expenditure (sum of costs from ED visits and subsequent hospital admissions) for analysis. Chi-squared test and student's T test were used to test statistical differences between study groups.

Result(s): From 2002-2013, we observed an overall increase in the rate of ED visits among elderlies in Taiwan, from 2,565.5 per 100,000 in 2002 to 3,941.5 per 100,000 in 2013. Pneumonia seemed to be the leading cause of ED visits (2.83%), hospital admissions (5.77%) and deaths (17.4%), followed by UTI (2.11% of ED visits, 2.36% admissions, 10.8% deaths). Average total health care cost per capita was also highest for pneumonia-related ED visits (New Taiwan Dollar, NT$ 91,166±113,183). Among our selected sample, 24.6% (35,825 of 145,886) were nursing home residents. Elderly from nursing homes were observed with statistical significantly higher proportions of ID-related ED visits when compared to non-nursing home elderly: tuberculosis (2.52% vs 1.31%, $P<0.001$), URI (14.1% vs. 11.3%, $P<0.001$), pneumonia (44.1% vs. 13.6%, $P<0.001$), and UTI (45.0% vs. 13.5%, $P<0.001$). Significantly higher ID-related health care costs were also incurred by nursing home elderlies relative to non-nursing home counterparts ($P<0.001$).

Conclusion(s): In general, increasing health care burden attributed to ID among the elderly was observed. Common infectious diseases in nursing home elderly are shown to be more prevalent than in non-nursing home elderly. Findings of this study is critical for addressing health care issues in the elderly as rate of aging continues to accelerate.

PS4-6 RANDOMIZE EVERYONE: CREATING VALID INSTRUMENTAL VARIABLES FOR LEARNING HEALTH CARE SYSTEMS

Health Services, Outcomes and Policy Research (HSOP)

Tor Tosteson, Sc.D., Todd MacKenzie, PhD, John Batsis, MD, MS, Adam Pearson, MD, MS, James Bernat, MD and Jonathan Lurie, MD, MS, Geisel School of Medicine at Dartmouth, Lebanon, NH

Purpose:

The overarching objective of *Randomize Everyone* is to enhance ongoing comparative effectiveness assessment of existing clinical treatments by introducing comprehensive randomization mechanisms and ethically sound procedures for patient informed consent in learning health care systems and their EHRs.

Method(s):

Patient-clinician decision making requires reliable evidence on the comparative effectiveness of alternative treatments. Efforts to improve health care systems have included a large investment in establishing electronic health record systems (EHR). Large volumes of well-organized health care records will provide a resource for conducting comparative effectiveness research. However, observational studies based on EHR data are subject to confounding due to treatment selection bias. Incorporating randomization widely can address this problem, but it is likely that considerable nonadherence will occur. The proposed use of instrumental variables will allow for valid estimates of comparative effectiveness under these circumstances.

Informatics methodologies will be developed for alerting physicians to instances where treatments are being used that are designated for randomization. The system will generate and record the randomized treatment assignments. Modules will be developed for specifying clinical and patient outcomes to be extracted from the EHR and for assessing adherence. Statistical analytics will be integrated into the system for estimating treatment effects based on instrumental variable methods and for automated monitoring for process measures. Appropriate methods for missing data will be implemented.

Focus groups from two demonstration clinics will be presented with structured examples of treatment comparisons and mechanisms of informed consent. A Stakeholder Board including experts in clinical research ethics, clinicians, patients, and institutional stakeholders will be empaneled to evaluate specific proposals for randomizations.

Result(s):

The new methodologies will be demonstrated in two centers at Dartmouth: the Spine Clinic and the Weight and Wellness Clinic using protocols incorporating the recommendations of the Stakeholder Board and designed to be amended for additional randomizations as a mechanism for expanding *Randomize Everyone.*

Conclusion(s):

The new methodologies will benefit patients by providing a system for continually improving comparative effectiveness evidence to support treatment decisions in learning health care systems.

PS4-8 MEASURING PHYSICAL ABILITY AND FUNCTIONAL STATUS THROUGH PATIENT REPORTED OUTCOMES (PROS)

Health Services, Outcomes and Policy Research (HSOP)

Joshua Biber, MS, MBA[1], Anne Thackeray, PhD, MPH[2], Craig Gale, MS[1], Heidi Cozart[1], Howard Weeks, MD[3] and Rachel Hess, MD, MS[4], (1)University of Utah Medical Group, Salt Lake City, UT, (2)University of Utah, Physical Therapy & Athletic Training, SAlt Lake City, UT, (3)University of Utah Health, Salt Lake City, UT, (4)University of Utah, Department of Population Health Sciences, Division of Health System Innovation and Research, Salt Lake City, UT

Purpose: Identify functional status of a patient related to their ability to complete certain tasks as measured through PROMIS Physical Function and improve interpretability of the PROMIS Physical Function score.

Method(s): We have implemented PROs in over 70% of our ambulatory encounters. As part of our assessment, patients are asked to complete the PROMIS Bank v1.2-Physical Function (PF) computer adaptive test (CAT). Functional status is then identified using a t-score. To improve interpretability of these scores, we examined the association of the t-score with specific items and item responses. Using the last question the patient received, the answer to that question, and the t-score generated, we determined the frequency for which the t-score corresponded to the specific item and items response compared to the total frequency of the t-score.

Result(s): From 2015 to 2017, 90,382 PF scores were collected throughout our healthcare system. The highest score reported was a 73.35, which was seen in 314 of the scores. This corresponded to a question of, "Does your health now limit you in participating in active sports such as swimming, tennis, or basketball?" with a response of, "Not at all" in 98.7% of the scores. The lowest score reported was a 15.40, which was seen in 161 of the scores. This corresponded to a question of, "Are you able to transfer from a bed to a chair and back?" with a response of, "Unable to do" in 100% of the scores. The most common score was a 49.76, which was seen in 2,311 of the scores, or 2.6% of the time. This corresponded to a question of, "Does your health now limit you in doing heavy work around the house like scrubbing floors, or lifting or moving heavy furniture?" with a response of, "Very little" in 99.3% of the scores. Similar evaluation was completed for different scoring ranges.

Conclusion(s): As expected, there is consistency to the relationship between final t-score and the answer to last question delivered by the CAT. Looking at the last question, answer, and resulting score, clinicians, patients, and researchers can identify specific limitations in functional ability. This description may allow for a more informative use of the data, improved communication of PROs between patients and providers, and the establishment realistic functional benchmarks for improvement after intervention.

PS4-9 VALUE-BASED STROKE CARE: FEASIBILITY OF COMPREHENSIVE OUTCOME MEASUREMENTS IN THE STROKE COHORT OUTCOMES OF REHABILITATION (SCORE) STUDY, AND CHANGES IN OUTCOMES OVER TIME
Health Services, Outcomes and Policy Research (HSOP)

Iris Groeneveld, PhD[1], Paulien Goossens, MD, PhD[1], Felicie van Vree, MSc[1], Jorit Meesters, PhD[2], Henk Arwert, MD[2], Radha Rambaran Mishre[2], Thea P.M. Vliet Vlieland, MD, PhD[3] and **Leti van Bodegom-Vos, PhD**[4], (1)Rijnlands Rehabilitation Center, Leiden, Netherlands, (2)Sophia Rehabilitation, Den Haag, Netherlands, (3)Leiden University Medical Center, Leiden, Netherlands, (4)Leiden University Medical Center, Department of Biomedical Data Sciences, section Medical Decision Making, Leiden, Netherlands

Purpose:

Accurate measurement of outcomes is a key element in Value Based Health Care (VBHC). This study aimed to explore the feasibility of administering a comprehensive set of patient-reported outcome measures (PROMS), based on a stroke-specific VBHC framework, in stroke patients in inpatient rehabilitation, and analyse the changes in outcomes over time.

Method(s):

The ongoing Stroke Cohort Outcomes of REhabilitation (SCORE) study includes consecutive stroke patients admitted to two inpatient rehabilitation facilities in the Netherlands. In accordance with the Standard Set for Stroke from the International Consortium for Health Outcomes Measurement (ICHOM; Salinas J et al, *Stroke* 2016), which is based on the VBHC concept, PROMS were assessed at 3-4 months post stroke. Additional assessments were conducted upon admission and 6 and 12 months thereafter. PROMS included a) physical functioning and b) communication (Stroke Impact Scale, SIS; Stroke and Aphasia Quality of Life Scale, SAQOL-39); c) societal participation (Utrecht Scale for Evaluation of Rehabilitation, USER-P); d) mood (Hospital Anxiety and Depression Scale, HADS); e) fatigue (Fatigue Severity Scale, FSS); and f) general health (Euroqol-5D, EQ5D). Sociodemographic and clinical characteristics, including the Barthel Index, were registered. Feasibility was defined as participation and attrition rates. Paired t-tests were conducted to analyse changes (mean, 95% CI) over time.

Result(s):

Of 485 patients invited over 29 months, 295 (60.8%) participated. Mean age was 60.5 (SD 12.3), 173 (58.6%) were male, 222 (75.5%) had an ischemic stroke, and median Barthel Index was 16.0. The attrition rates at 3 and 12 months were 2% (n=6) and 16.6% (n=49), respectively. Attrition was due to death (n=6), health problems (n=6), withdrawal (n=30), or loss to follow-up (n=7). Measures of physical functioning, societal participation, mood and general health improved during admission and/or during follow-up, whereas communication and fatigue did not change over time.

Conclusion(s):

In this inception cohort of stroke patients in rehabilitation the feasibility of the administration of comprehensive set of stroke-related PROMs based on the ICHOM framework was demonstrated. Although only administration at 3-4 months is recommended, the assessments at 6 and 12 months appeared valuable, as some PROMS were more responsive than others. In future research, more efforts into the improvement of the initial recruitment rate and the selection of appropriate measures to demonstrate the added value of rehabilitation are needed.

PS4-10 THE USE OF MARKOV MODELS TO DISENTANGLE DIFFERENCE IN PATTERNS OF FUNCTIONING FOR YOUTH AND ADOLESCENTS WITH ADHD: RESULTS FROM THE MTA STUDY

Quantitative Methods and Theoretical Developments (QMTD)

Roel Freriks, MSc[1], **Qi Cao, PhD**[1], Jurjen van der Schans, PhD[1], Annabeth Groenman, PhD[2], Pieter Hoekstra, MD, PhD[2], Erik Buskens, MD, PhD[2], Maarten Postma, PhD[1] and Jochen Mierau, PhD[1], (1)University of Groningen, Groningen, Netherlands, (2)University Medical Centre Groningen, Groningen, Netherlands

Purpose: Delinquency is an important predictor and effect of Attention Deficit Hyperactivity Disorder (ADHD) in youth and adolescents, especially for boys. Previous research shows little variation in health outcomes among different treatments for ADHD. We therefore used delinquency as a proxy of societal functioning to disentangle the observed difference in patterns of behavior between four treatment groups. These patterns are predicted with a Markov model.

Method(s): For this study we use the National Institute of Mental Health MTA cohort, a 14-month randomized treatment study of 579 children with ADHD, ages 7 to 10 years, with naturalistic follow-ups for up to 16 years after baseline. The MTA Study contains a second cohort of 258 children without ADHD, the Local Normative Comparison Group (LNCG), recruited 2 years after baseline and who were thereafter followed up at similar intervals as the children with ADHD. Children of the MTA cohort were randomly assigned to one of four treatment modes: 1. Routine community care (the control group), 2. Intensive medication management only, 3. Intensive behavioral treatment only, and 4. Combination treatment. We converted a five-point delinquency scale into three health states; high functioning, mild to moderate functioning, and severe functioning. Severe functioning was included as an absorbing state to estimate the corresponding survival curves, since previous research showed severe delinquent behavior in youth is extremely persistence. Subsequently, we developed a continuous-time Markov model based on 2-year observational data of the ADHD cohort to predict the 10-year effectiveness of the four treatments.

Result(s): Good validation results were depicted between our model and the observed data. The predicted average probability of not reaching the severe functioning state were respectively 0.855, 0.802, 0.813, and 0.876 for the treatment groups we considered. We tested the robustness of delinquency as outcome variable. Both in level and persistence we found a statistically significant difference between the ADHD and non-ADHD cohort ($P < 0.01$).

Conclusion(s): Using delinquency as a proxy for the development of societal functioning revealed interesting difference in patterns among the four treatment groups for clinicians/practitioners. This study shows that Markov models can be used to predict these patterns.

PS4-11 USING EXPERT ELICITATION TO CAPTURE MODEL UNCERTAINTIES: METHODS FOR POOLING EXPERTS' PRIORS

Quantitative Methods and Theoretical Developments (QMTD)

Dina Jankovic, Msc, University of York, York, United Kingdom

Purpose:

Expert elicitation refers to formal processes for quantifying experts' beliefs about uncertain quantities, typically as probability distributions. Elicitation has been used to characterise uncertainty in cost-effectiveness decision models when other sources of data are not available. Its use to date has been limited, partly due to scepticism around the credibility of elicited priors.

Elicitation is generally conducted with multiple experts to minimise bias and ensure representation of experts with different perspectives. Investigators can subsequently evaluate the contribution of each expert and use it to weight their priors, so that experts who are thought to contribute more are given more say. This is referred to as opinion pooling. Different methods for deriving weights exist; they vary in their assumptions on what determines experts' contribution. The choice of pooling method can affect the resulting estimates of uncertainty, yet it is not clear which method is optimal.

This paper aims to develop a set of guiding principles for choosing the optimal method for opinion pooling.

Method(s):

Two targeted literature searches were conducted. First, literature in elicitation, forecasting and cognitive sciences was searched to understand what determines experts' contribution to the overall estimate of uncertainty (the aggregate prior). The second targeted search was conducted to identify all existing methods for scoring experts' performance. Finally, the existing scoring methods were critiqued on their ability to capture the characteristics believed to affect experts' contribution. The findings were used to propose a set of guiding principles for choosing the optimal method for opinion pooling.

Result(s):

Experts' contribution is believed to depend on their intrinsic characteristics (field-specific experience, quantitative skills, impartiality) and their similarity to other experts in the sample. Methods for capturing experts' contribution broadly fall into one of two categories: 1) assigning a score to experts' field-specific experience (e.g. based on the number of years worked in a relevant field), or 2) scoring experts' ability to predict parameter values by eliciting their priors on seed parameters (that the investigator knows the value of, but the expert does not). Furthermore, the second category includes multiple methods for scoring experts' priors; the methods vary in the value they place on experts' accuracy and precision.

Conclusion(s):

The paper provides a framework for deciding on methods for opinion pooling and highlights their limitations.

PS4-12 MODELING PERSONALIZED ADJUVANT TREATMENT IN EARLY STAGE COLON CANCER (PATTERN)

Quantitative Methods and Theoretical Developments (QMTD)

Gabrielle Jongeneel[1], Marjolein JE Greuter[1], Felice N van Erning[2], Miriam Koopman[3], Cornelis JA Punt[4], Geraldine R Vink[5] and Veerle MH Coupe[1], (1)VU University Medical Center, Amsterdam, Netherlands, (2)Netherlands Comprehensive Cancer Organisation, Eindhoven, Netherlands, (3)University Medical Center Utrecht, Utrecht, Netherlands, (4)Academic Medical Center Amsterdam, Amsterdam, Netherlands, (5)Netherlands Comprehensive Cancer Organization, Utrecht, Netherlands

Purpose: To develop a decision model for the evaluation of different selection strategies for adjuvant treatment in stage II colon cancer patients.

Method(s): A Markov cohort model with one-month cycle length and time horizon of 50 months was developed. Five health states were included; diagnosis, recurrence, post-surgical death (PSD), death other causes (DOC) and death of colon cancer (DCC) (Figure 1). The cohort starts in the state diagnosis. From there, patients can transition to PSD (transition 1), DOC (T2) or recurrence (T3). These three transitions were considered competing events. From recurrence, patients can transition to DCC (T4) and DOC (T5). Data for model quantification were provided by the Netherlands Cancer Registry. The cohort consisted of 2,211 patients diagnosed with early stage colon cancer between 2002-2008 who did not receive adjuvant therapy. 59.6% of the population was aged >70, had predominantly a T3N0M0 stage tumor (89.6%) and less than 10 lymph nodes examined (52.0%). Within 50 months of follow-up, 312 recurrences and 638 deaths were observed. The model was parametrized using parametric survival models adjusted for age, TNM stage and number of lymph nodes examined for the transitions 2-5. Probabilities for transition 1 were based on the assumption that all deaths within 90 days were PSDs. For treatment effect, the transition from diagnosis to recurrence was adjusted based on a 1.2 hazard ratio as reported in a trial. A limitation of the model is the lack of subgroup-specific treatment effects. As an example, three selection strategies were evaluated in which allocation of adjuvant treatment was based on A) TNM stage, B) number of lymph nodes examined and C) both TNM stage and number of lymph nodes examined.

Result(s): According to model predictions, the number of recurrences after 50 months in 1,000 patients was highest in selection strategy A (92) compared to strategy B (86) and C (85). Also the number of deaths due to colon cancer was highest in selection strategy A (42) compared to strategy B (39) and C (39).

Conclusion(s): We developed a model that can be used for evaluating selection strategies for the allocation of adjuvant treatment in stage II colon cancer patients. In future studies, the model can be used to estimate long-term health gain and cost-effectiveness of novel, molecular-based selection strategies.

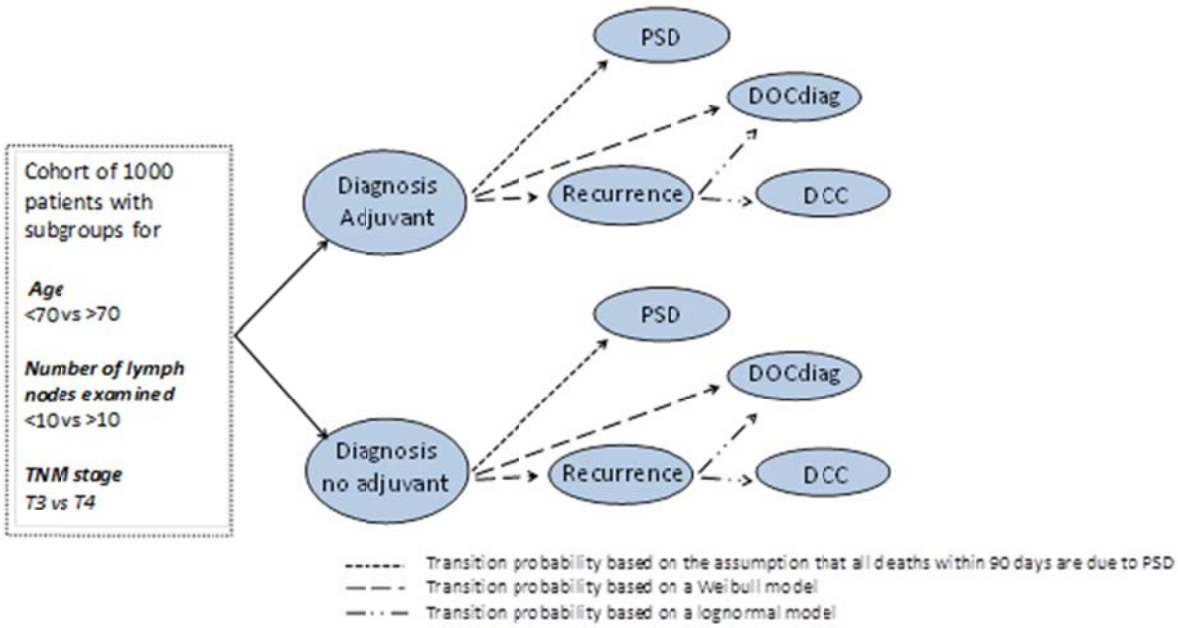


Figure1. Structure of the Personalized Adjuvant TreaTment in EaRly stage coloN (PATTERN) model.

PS4-13 A NEW SOFTWARE TOOL TO COMPUTE THE EXPECTED VALUE OF SAMPLE INFORMATION - AN APPLICATION TO THE HOMEHEALTH INTERVENTION

Quantitative Methods and Theoretical Developments (QMTD)

Anna Heath, Rachael Hunter and Gianluca Baio, University College London, London, United Kingdom

Background: The aim of Expected Value of Sample Information (EVSI) is to use the available evidence on the cost and effectiveness of a new intervention compared to current practice to calculate the economic value of specific trial designs, including the optimum sample size. Despite the potential benefits of using the EVSI, practical applications have been limited due to computational difficulties and ease of interpretation of results.

Purpose: The aim of this work is to develop a software tool that can easily calculate the EVSI and present the results to stakeholders. We have used the HomeHealth feasibility study used to test the capabilities of the new tool and how it might be used in a real world setting.

Method(s): The tool has been developed in R with an interactive R Shiny interface to present the results of the EVSI analysis. The EVSI calculations are based on "moment matching", a method that reduces their computational burden. The HomeHealth intervention was evaluated using a Bayesian cost-effectiveness model. The EVSI is calculated for two distinct study designs – one focusing on costs and one on the primary outcome.

Result(s): The EVSI calculations demonstrate that a clinical trial for the HomeHealth intervention has economic value. It highlights the need to collect cost data alongside the primary outcome of the trial. The R Shiny interface (https://egon.stats.ucl.ac.uk/projects/EVSI/) aids the presentation and interpretation of the results and allows it to be used as part of an application for further funding.

Conclusion(s): Our tool improves the ease with which EVSI can be calculated and presented for trials going for funding for a full RCT. Future work includes making the tool more user friendly and evaluating how trial methodologists and funders interpret the results.

PS4-15 ECONOMIC EVALUATION OF CULPRIT LESION ONLY PCI VERSUS IMMEDIATE MULTIVESSEL PCI IN ACUTE MYOCARDIAL INFARCTION COMPLICATED BY CARDIOGENIC SHOCK: THE CULPRIT-SHOCK TRIAL
Applied Health Economics (AHE)

Jose Robles-Zurita, PhD[1], Andrew Briggs, DPhil[1], Zahidul Quayyum, PhD[2], Keith Oldroyd, MD[3], Uwe Zeymer, MD[4], Steffen Desch, MD[5], Suzanne de Waha, MD[6] and Holger Thiele, MD[7], (1)University of Glasgow, Glasgow, United Kingdom, (2)Queen Mary University London, London, United Kingdom, (3)West of Scotland Regional Heart and Lung Centre, Golden Jubilee National Hospital, Glasgow, UK, United Kingdom, (4)Klinikum Ludwigshafen and Institut für Herzinfarktforschung, Ludwigshafen, Germany, Germany, (5)University Heart Center Lübeck, University Hospital Schleswig-Holstein, Lübeck, Germany, GA, Germany, (6)University Heart Center Lübeck, University Hospital Schleswig-Holstein, Lübeck, Germany, Germany, (7)Department of Internal Medicine and Cardiology, University of Leipzig - Heart Center, Leipzig, Germany, Germany

Purpose:

This paper presents economic evaluation of two revascularization strategies to treat patients with acute myocardial infarction and multivessel coronary artery disease complicated by cardiogenic shock: a) culprit vessel only (CO) percutaneous coronary intervention (PCI), with potentially subsequent staged revascularisation, and; b) immediate multivessel (MV) PCI.

Method(s):

A cost-effectiveness analysis (CEA) and cost utility analysis (CUA) used the 30-day follow-up data for 706 patients randomised in the international multicentre randomised controlled trial, CULPRIT-SHOCK. The CEA considered the composite primary endpoint of death or renal-replacement therapy. The CUA estimated the incremental cost per additional quality-adjusted life years (QALYs) and the incremental net monetary benefit (NMB) using a 1-year and a lifetime horizon. Information from the trial (i.e. clinical impact, quality of life and costs) and previous literature were used to populate a decision analytic model. A health system and societal perspective were considered where the cost for each revascularisation strategy were derived from Germany's national health service, initially, and from other participating countries, in a sensitivity analysis.

Result(s):

At 30 days, the composite primary endpoint of death and renal-failure was lower in the CO-PCI group (45.9%) than in the MV-PCI group (55.4%) (relative risk: 0.83; 95% confidence interval [CI], 0.71 to 0.96; P=0.01). The need for an economic analysis is justified given the consequences for healthcare costs. For example, there were more staged PCI of non-culprit lesions in the CO-PCI arm (17.4%, as opposed to 2.3% in the MV-PCI group), while there were more patients with renal replacement therapy in the MV-PCI group (16.4 % versus 11.6 %). Results of the economic evaluation will be presented once fully analysed.

Conclusion(s):

This economic evaluation will provide detailed health economic analysis further supporting medical decision making regarding potential treatment strategies for patients with acute myocardial infarction and multivessel disease complicated by cardiogenic shock. The economic evaluation within the pragmatic multicountry trial design enriches the external validity of the results.

PS4-16 THE IRRELEVANCE OF ICERS
Applied Health Economics (AHE)

Mike Paulden, PhD, University of Alberta, Edmonton, AB, Canada

Purpose: The incremental cost-effectiveness ratio (ICER) is the most commonly used summary statistic in cost-effectiveness analyses of health technologies. Although the ICER has received criticism, particularly due to its undesirable statistical properties, it remains popular with decision makers. This may be because of its perceived simplicity compared to alternative measures such as 'net benefit', and the ability for decision makers to calculate ICERs without specifying a cost-effectiveness threshold. We demonstrate that the perceived advantages of ICERs are misplaced, and that many common approaches for interpreting ICERs are fundamentally flawed.

Method(s): In depth review of methods to calculate and interpret ICERs and measures of 'net benefit', followed by a review of how each measure is used in practice by decision makers. We also propose a novel means for representing 'net benefit' on the cost-effectiveness plane, allowing for easy interpretation by decision makers.

Result(s): The ICER is more complex to calculate than measures of 'net benefit' and cannot be interpreted in the absence of a cost-effectiveness threshold, nullifying its perceived advantages. Common practices, such as interpreting lower ICERs as implying that a technology is 'more cost-effective', are fundamentally flawed; a lower ICER may be associated with higher, lower, or equal 'net benefit', so is uninformative as to whether a technology is more or less cost-effective. Common uses for ICERs, including their interpretation without a threshold and in sensitivity analyses (e.g. "tornado diagrams"), are entirely unsupported theoretically and may result in misleading findings.

Conclusion(s): If the cost-effectiveness threshold is not known then ICERs cannot be used to inform decisions. If the threshold is known then measures of 'net benefit' are easier to calculate and interpret. It follows that decision makers should take steps to abandon ICERs and instead adopt measures of 'net benefit' to inform their decisions.

PS4-17 SURVEYING THE COST-EFFECTIVENESS OF THE TWENTY PROCEDURES WITH THE LARGEST PUBLIC HEALTH SERVICES WAITING LISTS IN IRELAND: IMPLICATIONS FOR IRELAND'S COST-EFFECTIVENESS THRESHOLD

Applied Health Economics (AHE)

Tse Chiang Chen, MSc, Dane Wanniarachige, MA, Siofra Murphy, Katie Lockhart and James O'Mahony, PhD, Trinity College Dublin, Dublin, Ireland

Purpose: To inform a reconsideration of Ireland's current cost-effectiveness threshold of €45,000/quality-adjusted life-year (QALY) by assessing the cost-effectiveness of interventions demonstrating considerable unmet need as evidenced by large waiting lists within the Irish public health system.

Method(s): Waiting list data for inpatient and day-case procedures in the Irish public health system were obtained from the National Treatment Purchase Fund. The twenty interventions with the largest number of individuals waiting for inpatient and day-case care were identified, using waiting list size as a marker for the greatest unmet need. The academic literature was searched to obtain cost-effectiveness estimates for each intervention. Preference was given to sources from Ireland, followed by the UK and other high-income countries. Cost-effectiveness estimates from foreign studies were adjusted for differences in currency, purchasing power parity and inflation. These estimates were then compared to Ireland's current threshold.

Result(s): Of the top 20 waiting list procedures, 17 had ICERs below €45,000/QALY, 14 fell below €20,000/QALY, and 10 fell below €10,000/QALY. These 17 cost-effective interventions account for over 71% of the overall outpatient and day-case lists in Ireland. Only one procedure had an ICER above the current threshold. Two procedures had ICERs reported for different patient and indication groups that lay either side of the threshold.

Conclusion(s): Ireland is one of the few countries worldwide to have an explicit threshold. This threshold, however, has been criticized for not being based on evidence of opportunity cost of interventions forgone. Our findings indicate that many interventions exhibiting considerable unmet need have ICERs below the threshold. This suggests that population health could be enhanced by lowering the cost-effectiveness threshold to reduce expenditure on new interventions and divert resources to address the unmet need for existing, highly cost-effective interventions. An evidence-informed revision is required in order to ensure the threshold is consistent with its theoretical basis of the opportunity cost of other interventions foregone. A limitation of this study was the difficulty in matching specific procedures from waiting lists with ICER estimates from the literature. Nevertheless, our study represents a useful demonstration of a novel concept of combining routinely-compiled waiting lists with existing cost-effectiveness evidence to inform cost-effectiveness thresholds. This represents a tractable alternative to other recent research to estimate thresholds based on opportunity cost.

PS4-18 MISSING DATA IN TRIAL-BASED COST-EFFECTIVENESS ANALYSIS: A REVIEW OF THE CURRENT PRACTICE

Applied Health Economics (AHE)

Baptiste Leurent, London School of Hygiene & Tropical Medicine, London, United Kingdom, Manuel Gomes, PhD, University College London, London, United Kingdom and James Carpenter, DPhil, London School of Hygiene and Tropical Medicine, London, United Kingdom

Purpose:

Cost-effectiveness analyses (CEA) conducted alongside randomised trials provide key evidence for informing health care decision making, but missing data poses substantive challenges. There have been a number of recent developments in methods and guidelines to address missing data. We aimed to review the extent of, and methods used to address, missing data in trial-based cost-effectiveness analysis

Method(s):

We conducted a review of trial-based CEA published in the Health Technology Assessment journal between January 2013 and December 2015. This journal includes all evaluations funded by the UK National Institute for Health Research Health Technology Assessment Programme (NIHR HTA), providing details of the analysis methods and results. We extracted key information on the extent of missing data, and the methods used to address missing data in the primary and sensitivity analyses. We then critically reviewed our findings, and made recommendations to improve practice.

Result(s):

Fifty-two eligible trials were identified. Nearly all of them had missing data, with a median proportion of participants with complete cost-effectiveness data of 63% (IQR 47% to 81%). The most common approach for the primary analysis was to restrict analysis to those with complete data (43%), followed by multiple imputation (30%). Half of the studies conducted sensitivity analyses for missing data, typically using a limited range of assumptions. Only two (4%) considered possible departures from the missing-at-random assumption.

Conclusion(s):

Missing data remain a major concern for CEA conducted alongside randomised trials. In spite of its limitations, restricting analysis to the subset of complete records is the most common approach. Further improvements are needed, and should focus on i) limiting the extent of missing data, ii) choosing a method valid under contextually plausible assumptions for the primary analysis, and iii) conducting sensitivity analyses to departures from the missing-at-random assumption.

PS4-19 PREDICTING OVERALL SURVIVAL AND PROGRESSION-FREE SURVIVAL FOR EXTENSIVE STAGE SMALL-CELL LUNG CANCER TREATED WITH PROPHYLACTIC CRANIAL IRRADIATION WITH OR WITHOUT THORACIC RADIOTHERAPY

Applied Health Economics (AHE)

Simone Rauh[1], Marjolein Greuter[1], Harm van Tinteren[2], Paul Cobussen[3], Ben Slotman[3] and Veerle Coupe[1], (1)VU University Medical Center, Department of Epidemiology and Biostatistics, Amsterdam, Netherlands, (2)The Netherlands Cancer Institute, Department of Statistics, Amsterdam, Netherlands, (3)VU University Medical Center, Department of Radiation Oncology, Amsterdam, Netherlands

Purpose: To inform a cost-effectiveness evaluation model, we aimed to predict overall survival (OS) and progression-free survival (PFS) for patients with extensive stage small-cell lung cancer (ES-SCLC) treated with prophylactic cranial irradiation (PCI) with or without thoracic radiotherapy (TR).

Method(s): Weibull models were fitted to predict OS and PFS for 495 patients with ES-SCLC treated with chemotherapy and PCI, included in an RCT concerning TR treatment (NTR1527). The same predictors were considered in the OS and PFS models using forward selection (based on Akaike's criterion: $p<0.157$): age, gender, TR treatment, intrathoracic disease, WHO performance status, number of metastases, and presence of liver and/or bone metastases. Interaction terms between treatment and other predictors were evaluated ($p<0.1$). By design, information on metastases was missing for 46% of the patients. These missings were multivariately imputed, generating 46 imputed sets. Estimates were pooled using Rubin's rules. Model discrimination was assessed by area under the curve (AUC).

Result(s): Outliers in OS and PFS had a large influence on predictor estimates, resulting in overestimated survival rates up until 24 months and underestimated rates afterwards. Therefore, OS and PFS were censored at 24 months. Age≤70 years, female gender, TR treatment, interaction between female gender and treatment, and presence of liver and/or bone metastases were statistically significant predictors of longer OS. Except for age, the same predictors were statistically significant predictors of longer PFS. However, to prevent impossible combinations of a longer predicted PFS than predicted OS, age was also included in the PFS model. Median AUC for 2-year survival over the imputed datasets was 0.72 (IQR:0.69-0.74) and 0.78 (IQR:0.74-0.80), respectively, for OS and PFS. Figure 1 shows OS (Fig. 1A) and PFS (Fig. 1B) stratified for the strongest predictor, presence of liver and/or bone metastases.

Conclusion(s): Age, gender, TR treatment, and presence of liver and/or bone metastases predicted OS and PFS. These models will be combined with data on health utilities and costs to evaluate the cost-effectiveness of PCI+TR vs PCI only. In addition, based on the predictors in the Weibull model, cost-effectiveness will be evaluated for specific subgroups of patients.

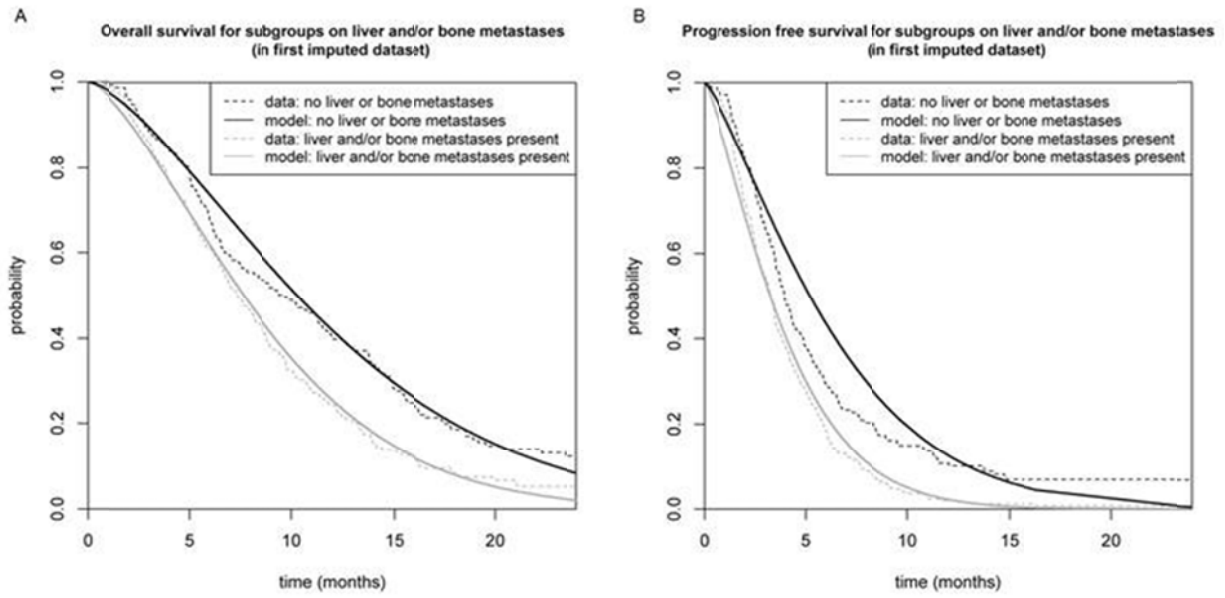


Figure 1: Observed and predicted OS and PFS for subgroups: no liver or bone metastases versus liver and/or bone metastases present – in the first imputed dataset

PS4-20 ASSESSMENT OF CURRENT USE OF BUDGET IMPACT ANALYSES FOR THE EVALUATION OF CANCER SCREENING PROGRAMS
Applied Health Economics (AHE)

Beate Jahn, PhD[1], Jovan Todorovic, MD[1], Marvin Bundo, MD[1], Gaby Sroczynski, MPH, Dr.PH[1], Annette Conrads-Frank, PhD[2], Ursula Rochau, MD, MSc[1], Jagpreet Chhatwal, PhD[3], Dan Greenberg, PhD[4], Josephine Mauskopf, PhD[5] and Uwe Siebert, MD, MPH, MSc, ScD[6], (1)Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (2)UMIT - University for Health Sciences, Medical Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA, Hall i.T., Austria, (3)Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, (4)Department of Health Systems Management, Ben-Gurion University of the Negev, Be'er Sheva, Israel, (5)RTI Health Solutions, RTI International, Durham, NC, (6)UMIT, Dept. of Public Health, Health Services Research & HTA, Harvard Univ., Dept. of Health Policy & Management, Institute for Technology Assessment, Hall i.T./Boston, MA

Purpose: Budget impact analyses (BIA) describe the changes in short-term intervention-related and disease-related costs of new health care technologies. BIA are increasingly required for budgetary planning by decision-makers. Our study systematically reviewed published BIA for cancer screening programs, specifically applied methods, and the degree to which international BIA guidelines are followed in studies evaluating cancer screening programs.

Method(s): A systematic literature search was conducted in MEDLINE and EconLit for BIA evaluating cancer screening programs, published in English language during 2010-2016. Standardized evidence tables were used to extract study characteristics as outlined in the ISPOR BIA Task Force Guidelines including cancer type, model structure and assumptions, definition of population size/characteristics, perspective, time horizon, included costs, source of epidemiologic and clinical data, consideration of health impact, validation, and uncertainty analysis.

Result(s): Ten studies were identified evaluating screening for cancer of breast (n=3), colorectal (n=2) cervical (n=2), lung (n=1), prostate (n=1) and the skin (n=1). Applied model designs varied from different types of decision-analytic models (60%) to simple cost calculators (40%). Simple cost calculators are recommended by ISPOR guidelines as long as important conditions are credibly captured. Complex multiple-cohort models were mainly applied in the reviewed studies that combined cost-effectiveness and BIA (n=3). The time horizon ranged from one year (n=6) to 20 years (n=3). BIA framework should allow for calculating shorter and longer time horizons. Reporting both, annual results and cumulative results may be useful for planning annual expenditures and for showing the net effects on total costs over time because of delays in offsetting disease costs savings. However, not all studies provided this information. All studies included direct condition-related costs and two studies additionally included indirect cost. Health impact on patients was reported in 40% of the studies. Uncertainty analysis was limited in most of the studies. Only 40% of the studies validated their results.

Conclusion(s): Our review highlights a considerable variability in the extent to which the BIA studies evaluating cancer screening programs followed recommended guidelines. For example, most BIA failed to report projections beyond one-year or to report the model validation. To ensure high quality and sound decision support, best practice recommendations should be followed more rigorously in all key aspects.

PS4-21 AN INTEGRATED HEART FAILURE PROGRAMME MANAGEMENT IN THE BASQUE COUNTRY

Applied Health Economics (AHE)

Igor Larranaga[1], Arantzazu Arrospide[1], Ricardo San Vicente[2], Aitziber Barandiaran[2], Adolfo Ruiz de Austri[1], Idoia Larranaga[2], Itziar Perez[3] and Jesus Millas[1], (1)Integrated Healthcare Organization Alto Deba, Arrasate, Spain, (2)Integrated Healthcare Organization Goierri Alto-Urola, Zumarraga, Spain, (3)Integrated Healthcare Organization Goierri Alto-Urola, Integrated Healthcare Organization Goierri Alto-Urola, Spain

Purpose: Heart failure is characterized by frequent transitions in time to instability states, which decrease patient quality of life and generate great resource consumption. The objective of the programme is to implement and evaluate an integrated health and social intervention for heart failure in the Integrated Health Organization of Alto Deba in Spain.

Method(s): The target population are patients with at least one hospitalization in the previous year due to heart failure. The control group is set as patients identified until 2016 and the intervention group as patients identified after the intervention began in 2017. The study design is based on the Plan-Do-Check-Act continuous improvement approach. First, a discrete event simulation model is used to predict the evolution of resource consumption and perform a budget impact analysis to 2020 (Plan). Second, an intervention based on primary care empowerment and focused on patient monitoring and self-care is implemented in 2017 (Do). Third, implementation and effectiveness indicators are analysed statistically to compare 2018 and 2016 and verify if the objectives are achieved or not (Check). Finally, depending on the result of the analysis, the programme is maintained or adjusted (Act).

Result(s): Around 200 patients have been identified in 2016 and 2017. Increase the stable phase in those patients prevents their decompensation, generating at the same time a reduction in the resource consumption rate and its associated costs.

Conclusion(s): In order to evaluate systematically the program, is necessary to measure, not only the effectiveness of the intervention, but also its implementation. For that reason quantified objectives have to be established in the planning phase to have a reference in the evaluation phase.

PS4-22 THE IMPACT OF DISEASE LABELS ON DISEASE EXPERIENCE IN PATIENTS WITH PROLONGED INCURABLE CANCER: A QUALITATIVE STUDY
Decision Psychology and Shared Decision Making (DEC)

Veerle Busink[1], Vincent Ho[2], Donald van Tol[3] and **Hilde Buiting**[1], (1)Netherlands Cancer Institute, Amsterdam, Netherlands, (2)Comprehensive Cancer Center the Netherlands (IKNL), Utrecht, Netherlands, (3)University Medical Center Groningen, Groningen, Netherlands

Purpose:

Advances in medicine have resulted in prolonged disease trajectories, increasingly even in those with incurable cancer. These prolonged disease trajectories have induced discussions about the 'right' medical terminology to be used, and the impact of medical terminologies on patients' well-being. We examined the impact of disease labels on disease experience in patients with prolonged incurable cancer.

Method(s):

Qualitative study based on participant observations and 33 conversations at the day-care unit in two stages covering a period of 2 years (2015-2017). We asked patients about the incurable nature of their disease, using open-ended questions. By explicating the underlying reason of this specific study we sometimes broached the topic of chronicity too. We however especially looked at disease-labeling in an indirect way to avoid possible bias.

Result(s):

Patients ascribed different disease labels to themselves, either on purpose or not. They varied in their preference regarding specific disease labels (e.g. 'chronic', 'palliative', 'stable') in their communication about the incurable nature of their disease. Patients' use of disease-labels seemed to be related in how they coped with their situation. Patients showing a positive or neutral mood were often more comfortable with the label 'chronic', whereas patients with a more negative mood more often felt they could become trapped in a definition until the end of their life. Some patients preferred not to label their disease at all, as they felt it would make them 'less human'. If healthcare professionals used different disease-labels (same situation, same moment), this could distress patients. Although some labels used by healthcare professionals bothered patients, their impact on disease perception seemed low as patients often decided to stick to their own, more 'positive' disease label. Patients who were more willing to perceive themselves as a 'patient' tended to receive more support from their environment.

Conclusion(s):

Patients' explicit use of disease labels can be part of their coping strategy, which can be associated with maintaining identity. Healthcare professionals' use of disease labels, on the other hand, may cause confusion. In a world where the diagnosis of incurable cancer is not a direct death sentence anymore, appropriate use of disease labels by healthcare professionals is warranted: to explore how such labels can contribute to patient's well-being as well as the decision-making process.

PS4-23 THE DOCTOR-PATIENT RELATIONSHIP IN GENERAL PRACTICE. HOW QUALITY OF COMMUNICATION, TRUST, AND EPISTEMOLOGICAL BELIEFS ABOUT MEDICINE AFFECT SATISFACTION AFTER MEDICAL CONSULTATIONS
Decision Psychology and Shared Decision Making (DEC)

Serena Petrocchi[1], Paola Iannello[2], Flavia Lecciso[3], Alessandro Antonietti[2] and Peter Schulz, PhD[4], (1)Università della Svizzera italiana, Lugano, Switzerland, (2)Catholic University of the Sacred Heart, Milan, Italy, (3)Università del Salento, Lecce, Italy, (4)Institute of Communication and Health, University of Lugano (Università della Svizzera italiana), Lugano, Switzerland

Purpose: analyze: generalized (GR) and dyadic reciprocity (DR) of quality of communication and trust in doctor-patient relations; the associations among patients' quality of communication, trust, epistemological beliefs and satisfaction.

Method(s): Participants were 11 GPs (M_{age} = 54.16, SD = 12.28, 7 men) and their 149 patients (M_{age} = 47.48, SD = 9.88, 62.4% women; doctor-patients range 1-30, M = 14, SD = 9.3). After a consultation, doctors and patients independently completed questionnaires on quality of communication (Campbell et al., 2007) and trust (Dugan et al., 2005). Patients completed a questionnaire on their epistemological beliefs about medicine (Kienhues & Bromme, 2012) and satisfaction.

Result(s): MLM modelling provided estimation of reciprocity (quality of communication: B_{GR} = .17, SE = .07, Z = 2.31*, B_{DR} = -.78, SE = .15, Z = -5.25***; trust: B_{GR} = .15, SE = .05, Z = 2.51*, B_{DR} = -.65, SE = .12, Z = -4.27***). Hierarchical regression analysis [F(7 141) = 41.32, p < .0001, R2adj. = .66, Step 2] showed that doctor's years of experience, t = -2.11*, B = -.11, quality of communication skills, t = 2.32*, B = .16, and trust, t = 9.55***, B = .68, were associated with patients' satisfaction after the visit (controlling for patients' age, gender, general health, and number of visits). Test of simple mediation yielded significant effects of: a) stability of epistemological beliefs on patients' trust in doctor, b = .38* (SE = .16); b) patients' trust in doctor on satisfaction, b = .62*** (SE = .09). Indirect effect of stability of beliefs on satisfaction did not reach the significance.

Conclusion(s): Estimation of the generalized reciprocity showed that a physician who perceives high quality of relation/trust tends to have patients who perceive high quality of relation too. As dyadic reciprocity, if a physician perceives high quality of relation/trust with a particular patient (more than with other patients), that patient perceives low quality of relation with that physician (more than the physician's other patients). Regression and mediation analyses suggested that the patients' evaluation of quality of communication and trust in their doctors after a consultation, together with stability of epistemological beliefs, contributed to the patients' satisfaction over and above their age, gender, and number of visits.

PS4-25 DEVELOPMENT OF A WEB-BASED PATIENT DECISION AID FOR DISEASE-MODIFYING DRUGS IN MULTIPLE SCLEROSIS

Decision Psychology and Shared Decision Making (DEC)

Ingrid Kremer, MSc[1], Mickael Hiligsmann, PhD[2], Silvia Evers, PhD, LLM[2] and Peter Jongen, PhD, MD[3], (1)Maastricht University, Maastricht, Netherlands, (2)Department of Health Services Research, Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, Netherlands, (3)MS4 Research Institute, Nijmegen, Netherlands

Purpose: To describe the development of a patient decision aid for disease modifying drugs (DMDs) in multiple sclerosis (MS). The patient decision aid is based on the principles of multi criteria decision analysis. In this approach, each decision is described according to a number of characteristics. A weight is given to these characteristics according to their importance in the decision. weights are compared to the performance of each alternative option on the characteristics to rank the options from most to least preferred.

Method(s): The patient decision aid for MS elicits the patient's preferences for characteristics of DMD treatment and compares these preferences to the treatment options available for the patient. First, focus groups and a best-worst scaling were conducted with patients to identify and prioritize the characteristics which need to be included in the patient decision aid. Next, a literature review was conducted focused on the identified treatment characteristics to collect the best evidence available for the thirteen DMDs now available for patients. The prototype of the decision aid is currently being developed. Usability and comprehensibility of the decision aid will be evaluated during the alpha pilot testing. Neurologists, MS nurses and MS patients (N=20) will be asked to go through the decision aid and interviews will be performed afterwards to identify any issues in use and understandability of the questions, information and results.

Result(s): Ten DMD characteristics were identified that patients regarded as most important for their decisions and were thus included in the patient decision aid. Examples of these characteristics are: effect of the drug on relapse rate, quality of life, risks on adverse events, safety and ease of use. In the literature review estimates of the performance of treatment options were found based on RCTs and observational studies, but a substantial amount of missing evidence was identified. These gaps will be filled by using expert opinions.

Conclusion(s): We are developing a patient decision aid that closely reflects MS patients' needs when making a choice between DMD treatment options, as the selection of information included in the patient decision aid was based on the patient's perspective. Moreover, the patients preferences are explicitly elicited and used in the decision aid to rank the treatment options from most to least important for the individual patient.

PS4-26 THE EVALUATION OF OVARIAN CANCER PATIENTS' PERCEPTION OF THE MEDICAL INFORMATION GIVEN AFTER THE PRIMARY DIAGNOSIS
Decision Psychology and Shared Decision Making (DEC)

Judith den Ouden, MD[1], Regina The, Ir.[2] and Anne van Altena, MD, PhD[1], (1)Radboud university medical center, Nijmegen, Netherlands, (2)ZorgKeuzeLab, Delft, Netherlands

Purpose: To determine the perception on the information given by the doctor about treatment options after the primary diagnosis in patients with advanced stage ovarian cancer.

Method(s): In four hospitals in the Netherlands patients in follow up after primary treatment for advanced stage ovarian cancer were asked to participate. A total of 50 patients were asked to fill out a questionnaire. The questionnaire contained general questions about the patient, the given information on treatment options, and the process of the decision-making. Furthermore, in depth interviews were performed in 10 patients on the same topic.

Result(s): First results from the interviews show that most patients did not feel they had a real choice in therapeutic options. The majority of patients underwent treatment consisting of surgery and chemotherapy without realizing the severity of the diagnosis or the prognosis. In retrospect they would have wanted to know what to expect during and after treatment; i.e. more information on the possible side effects of the therapy and when to seek medical help. The decision making process differed in the study group; some patients rely heavily on their doctors opinion while others want to be fully informed and start to look for information by themselves, to make sure that they are aware of all their options. Results from the questionnaires will follow.

Conclusion(s): The Dutch working group on gynaecologic oncology agreed that there is need for a patient decision aid (DA) to support patients with advanced stage ovarian cancer to make a well-informed decision about their treatment. Previous studies show that a DA effectively supports shared decision-making. The majority of patients are diagnosed in an advanced stage of the disease and face poor prognosis. Before developing a DA we wanted to gather more insight in patients perception of the information that is currently provided and their specific needs. Patients indicated the need for reliable information about all available treatment options for their situation. Furthermore, patients expressed the need to know how the treatment will influence their daily life, and therefore want more specific information about what to expect from possible side effects during and after treatment. This information will be used in the development of a Dutch DA for patients with advanced stage ovarian cancer that can be used in clinical practice.

PS4-27 VARIATIONS IN DISCUSSION RATE ON COLORECTAL CANCER (CRC) SCREENING AMONG PREVIOUSLY UNTESTED PATIENTS IN PRIMARY CARE AND DECISION TAKEN AFTER DISCUSSION: RESULTS FROM A SYSTEMATIC DATA COLLECTION IN THE SWISS SENTINELLA NETWORK

Decision Psychology and Shared Decision Making (DEC)

Yonas Martin[1], Emanuele Prati[1], Leonhard Braun[1], Charles Dvořak[2], Kali Tal[1], Nikola Biller-Andorno[3], Jean-Luc Bulliard[4], Jacques Cornuz[5], Kevin Selby[6] and Reto Auer[5], (1)Institute of primary health care (BIHAM), University of Bern, Bern, Switzerland, (2)President Sentinella Network, Bern, Switzerland, (3)Institute for Biomedical Ethics and History of Medicine, UZH, Zurich, Switzerland, (4)Institute of Social and Preventive Medicine, University of Lausanne, Lausanne, Switzerland, (5)Department of Ambulatory Care and Community Medicine, University of Lausanne, Lausanne, Switzerland, (6)Northern California Division of Research, Kaiser Permanente, Oakland, CA

Purpose: Only 46% of patients who visit their primary care physicians (PCP) in Switzerland are up-to-date with CRC screening, but many patients are never offered the choice of screening. Patients who discuss screening with PCPs may be more likely to get screened. We aimed to determine the following: the proportion of 50-75-year-old patients eligible for screening, who discussed screening with their PC, who chose to be screened, which method they chose, and how many refused screening. We also wanted to describe variation in care between PCPs.

Method(s): We invited 130 PCP practices in the Sentinella Network, a representative sample of PCPs across all regions of Switzerland. Each physician systematically collected data on 40 consecutive patients, including demographic data, data on previous CRC testing (colonoscopy/fecal occult blood testing [FOBT]/other), contra-indications for screening, if a discussion took place, and the decision made (screening method, patient's refusal).

Result(s): We included 92/129 eligible PCPs in our analysis. Mean physician age was 55; 77% were men. PCPs collected data on 3,677 patients (mean age 63; 51% female). Nearly half (46%) had been tested for CRC (FOBT to colonoscopy ratio: 0.1). Half these patients were eligible for a discussion (no previous testing and no contra-indication for testing; N=1,739/3,472). Of these, PCPs discussed CRC screening with 51% (N=883/1,739). After excluding patients with risk factors or symptoms that suggested CRC (N=107), 61% (N=477/776) opted for screening (54% colonoscopy; 45% FOBT), 29% refused, 6% neither refused nor planned a screening. Decision information was missing for 3%. After the consultation, 60% (N= 2,070/3,472) had been tested previously or decided to be tested. Discussion rate varied widely between PCPs (quartile (Q1;31%, Q3:77%), as did choice of test (FOBT/colonoscopy ratio Q1:0, Q3:0.7).

Conclusion(s): Within a systematic data collection, we found PCPs could discuss CRC screening with half of eligible patients. Discussion rates varied widely between PCP practices, but when a discussion took place, it increased the number of patients who decided for screening. Future studies should determine if implementing shared decision making for CRC reduces variation in care.

PS4-28 HOW COUNSELORS COMMUNICATE UNCERTAINTY ABOUT MULTIGENE PANEL TESTING FOR CANCER. AN OBSERVATIONAL STUDY
Decision Psychology and Shared Decision Making (DEC)

Niki Medendorp[1], Marij Hillen, PhD[1], Pomme van Maarschalkerweerd[1], Cora Aalfs[1], Anne M. Stiggelbout, PhD[2] and **Ellen Smets**[1], (1)Academic Medical Center, Amsterdam, Netherlands, (2)Medical Decision Making, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands

Purpose: Multigene panel testing is increasingly used to improve the identification of an inherited cancer predisposition. Panel tests may yield uncertainty, for example about the meaning of variants. Counselees need to be fully informed about these uncertainties, to make an informed decision about whether or not to perform such a test. Until now, it is unknown how uncertainty concerning multigene panel testing is communicated to counselees. Therefore, we investigated the manner and variety in which counselors discuss uncertainty regarding multigene panel tests, using simulated patients (SPs). Additionally, we investigated which characteristics of counsellors are associated with their communication.

Method(s): Counsellors of all eight genetic centers in the Netherlands were instructed to discuss a multigene panel test with a SP. Using a script, SPs represented a counselee visiting for initial cancer genetic counselling who had had multiple cancer types. All consultations were videotaped and double-coded by two coders independently, using a coding scheme based on previous qualitative studies. All utterances of uncertainty by counsellors, framing and initiative of these utterances, and their verbal responses on scripted uncertainty expressions by the SP were coded. Three weeks prior to the consultation, counsellors filled out a questionnaire assessing their demographics, confidence in their own ability to communicate uncertainty, tolerance to uncertainty, attitudes towards shared decision making (SDM) and perceived social norm in how uncertainty should be communicated. Immediately after the consultation, both counsellors and SPs filled out a questionnaire assessing their satisfaction with communication, perceived degree of SDM and session realism.

Result(s): Currently, we have videotaped 21 consultations out of a 30 planned, involving 10 clinical geneticists and 11 genetic counselors, 4 male, and with a mean age of 41 years. Results on the different types and frequency of uncertainties communicated by counsellors, their reactions to uncertain expressions by SPs and associations between counselor characteristics and the manner of communication will be presented.

Conclusion(s): This study provides insight into how uncertainty regarding multigene panel testing is communicated, and whether counsellor characteristics matter. The results will inform future studies on how the discussion of uncertainty affects counselees and how it may be improved.

PS4-29 REVIEW ON EXISTING DECISION AIDS SUPPORTING SHARED DECISION MAKING IN COLORECTAL CANCER SCREENING

Decision Psychology and Shared Decision Making (DEC)

Marvin Bundo, MD[1], Beate Jahn, PhD[1], Gaby Sroczynski, MPH, Dr.PH[1], Ursula Rochau, MD, MSc[1], Marieke de Vries, PhD[2], Bernhard Streicher, PhD[3], Nikolai Muehlberger, Priv.-Doz., DVM, MPH[1], Silivia Angerer, PhD[1], Herbert Tilg, Dr.[4] and Uwe Siebert, MD, MPH, MSc, ScD[5], (1)Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (2)Institute for Computing and Information Sciences (iCIS), Intelligent Systems group, Radboud University, Nijmegen, Netherlands, (3)Social- and Personality Psychology Working Group, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (4)Internal Medicine Clinic I, University of Innsbruck, Innsbruck, Austria, (5)UMIT, Dept. of Public Health, Health Services Research & HTA, Harvard Univ., Dept. of Health Policy & Management, Institute for Technology Assessment, Hall i.T./Boston, MA

Purpose: Most of the tools that support shared decision making (SDM) in colorectal cancer (CRC) screening do not always provide patients with adequate and comprehensible health information (e.g., expected benefits and harms from evidence-based decision-analytic models or meta-analyses). The aim of this study was to identify existing decision aids (DA) that support SDM in CRC screening and review their strengths and limitations

Method(s): First, a literature review was conducted in PubMed from 2000 to 2017 to identify existing DA that support SDM in CRC screening. Second, we searched the Internet to find additional DA designed by clinics or national public health institutes. Then we extracted information on the format of the DA, presented benefits and harms of the screening strategies, numerical and graphical presentations, question prompt lists (QPL) and risk stratification (RS). International guidelines were utilized to evaluate the DA.

Result(s): We identified the following formats: brochures, picture cards, risk calculators and websites. These decision aids provide a description of screening benefits (e.g., CRC cases avoided, CRC deaths avoided) and harms (e.g., perforation, bleeding). Benefits are often presented in detail, sometimes accompanied by graphical representations. However, often there is unbalanced balance-harm reporting. For some DA, numerical figures are either lacking or are incomplete. QPL that would help to structure conversations about health decisions or RS tools are often missing. The most helpful DA are those that present balanced benefits-harm information, comprehensible numerical figures, graphical representations, well-structured QPL and RS. For example, the attractively designed brochures for CRC by Cancer Aid in Austria also do not include information on potential harms and lack graphical representations. IQWIG in Germany developed CRC screening information brochures that include a full range of benefits and harms. However, for an interactive process of SDM, the material is limited regarding patient preference based information and graphical representations.

Conclusion(s): A detailed description of benefits using simple language and graphical representations are usually present in DA and facilitate SDM. However, the description of harms, RS and QPL, which would make SDM more balanced and patient-centered, are often missing or suboptimal . In a further step, our results will be utilized to design a decision aid tool for CRC screening using information from our CRC screening model that will improve patient-centered care in Austria and patient satisfaction.

PS4-30 PATIENT-CLINICIAN COMMUNICATION AND ADHERENCE TO ORAL CHEMOTHERAPY

Decision Psychology and Shared Decision Making (DEC)

Jennifer Elston Lafata, PhD, Daniel Gratie, PharmD and Bobbie Nguyen, Division of Pharmaceutical Outcomes and Policy, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC

Purpose: As oral chemotherapy increasingly is used in cancer care, the challenge of how to support patients in their use has taken on increased urgency. We summarize the literature on patient-clinician communication and patients' adherence to curative oral chemotherapy, and explore whether effective communication is occurring within oncology practices.

Method(s): We used the terms adherence, non-adherence, communication, treatment satisfaction, discordance, barriers, discontinuation and the names of specific oral chemotherapy agents to search for interventional and observational studies of patient-clinician communication and patients' use of curative oral chemotherapy. Literature available before December 28, 2017 was searched via PubMed. Two research team members independently reviewed abstracts to determine studies eligible for full review. We also searched reference lists to identify additional relevant studies. We excluded studies focused on metastatic disease treatment, hormone/endocrine therapy, and those that did not measure patient-clinician communication and adherence. We additionally identified relevant patient-oncologist conversations by searching a database of audio-recorded and transcribed office visits from a national physician panel. We searched transcripts electronically for keywords relevant to oral chemotherapy discussions among patients diagnosed with cancer. Two research staff independently reviewed audio-recordings and transcripts, coding for the presence or absence of key communication behaviors as identified from the systematic review.

Result(s): The initial search resulted in the identification of 279 articles, of which 121 were retained after title review. After reading abstracts, we retrieved the full text for 32 articles. Studies focused on a diversity of communication behaviors, including discussions of medication side effects and treatment goals. Studies also used a diversity of means by which to measure medication adherence, including patient-self report and pharmacy refills. Most audio-recorded office visit conversations were void of key communication behaviors thought to be supportive of patients' medication adherence.

Conclusion(s): Consistent with other systematic reviews of patient-clinician communication in other clinical contexts, we found that the existent literature included few studies and employ a diversity of methods. Furthermore, among our national sample of patient-oncologist office visit conversations, we found that office visit conversations often fall short of what could be considered optimal. As the use of oral chemotherapy continues to expand within oncology care, it is important that oncologists and other clinicians understand how they can support patients' adherence to life-saving oral chemotherapy treatment.

PS4-31 INTRODUCING CLINICAL DECISION MAKING IN UNDERGRADUATE TRAINING: EXPERIENCES FROM COUNTRIES WORLDWIDE

Decision Psychology and Shared Decision Making (DEC)

Rina Triasih, Dr, MMed(Paed), Sardjito hospital/Faculty of Medicine, Universitas Gadjah Mada, Yogyakarta, Indonesia, Luc Debaene, MD, University of Antwerp, Antwerp, Belgium, Rodrigo Henriquez, MD, Facultad de Ciencias de la Salud, Universidad de Las Américas, Quito, Ecuador., Quito, Ecuador, Paulin Basinga, PhD, Bill and Melinda Gates foundation. Nigeria country office., Lagos, Nigeria and **Jef Van den Ende, MD, PhD**, University of Antwerp, Wilrijk, Belgium

Purpose:

In the seventies Pauker and Kassirer translated clinical decision making into everyday clinical work with their concept of the therapeutical threshold, later extended to the test and test-treatment threshold. Kahneman and Tversky stressed the importance of heuristics for clinical reasoning, pushing Bayes' theorem more or less backstage. In the eighties Sackett caused a revolution bringing clinical epidemiology to clinicians, advocating replacing sensitivity and specificity with likelihood ratio's and predictive values with posttest probability.
Up to the nineties, clinical reasoning training was left to individual clinical professors often contradicting each other. Clinical reasoning was conceived 'a natural talent'. Building on the abovementioned paradigm shift, we integrated these ideas in a clinical reasoning model first developed at the Institute of Tropical Medicine, Antwerp, Belgium.

Method(s):

Together with first steps in undergraduate training in Belgium, field experiences concerned Ecuador, Bolivia, Laos and Rwanda, where workshops were organized for practicing clinicians. After several years feedback made it clear that teaching should start in universities, before heuristic biases and wrong conceptions of probability and thresholds become the 'gold reasoning standard' in routine clinical work. From 2002 on, we set up undergraduate courses in Rwanda, Burkina Faso, Ethiopia, Laos, Ecuador and more intensively in Belgian Universities. In Indonesia we started by first training trainers and supervisors, anticipating undergraduate introduction.

Result(s):

Looking back over these 30 years some observations can be made. First, mere teaching does not work. Clinical reasoning can not be taught, it can only be learned in interactive workshops. Second, training should be started in undergraduate years. Third, we should first teach future trainers, with the adagium 'that they should know ten times more than what they teach'. Last, and probably most important: if students master the theory well and are later confronted with a supervisor or colleague who does not master the theory nor speaks the language, conflicts and confusion arise on both sides, and deception on the student's side are the consequences. Therefore all clinicians should be involved in the overall concept.

Conclusion(s):

Conclusions are crystal clear: undergraduate training in clinical reasoning is mandatory, future trainers should be trained thoroughly, and all clinicians should at least master the basic concepts, and should be able to demonstrate them in daily work.

PS4-33 DO'S AND DON'TS FOR INFORMED CONSENT AROUND CHILDBIRTH: PERSPECTIVE OF PATIENTS AND HEALTHCARE PROFESSIONALS
Patient and Stakeholder Preferences and Engagement (PSPE)

Bo Hacking[1], Dana Huppelschoten[2], Annemijn Aarts[3] and **Jeroen van Dillen**[3], (1)Maasziekenhuis Pantein, Nijmegen, Netherlands, (2)Catharina Ziekenhuis Eindhoven, -, Netherlands, (3)Radboudumc Nijmegen, -, Netherlands

Purpose: According to Dutch medical law informed consent is necessary for every medical procedure. Little is known about what information women want during delivery. The purpose of this study was to evaluate patients and healthcare professionals' preferences concerning informed consent for interventions during labor using a non-validated questionnaire.

Method(s): Women in the postpartum department and obstetric healthcare professionals of an academic hospital in the Netherlands during the period of the 15th of June to the 31th of July 2017 were included. Using semi-structured interviews with healthcare professionals and women in the postpartum department we selected six interventions for the development of the questionnaire. The questionnaire was send to 81 patients and 141 healthcare professionals and contained questions concerning informed consent on six interventions during labour: vaginal examination, cardiotocography (CTG), internal fetal monitoring, postpartum oxytocin injection, episiotomy and assisted vaginal delivery.

Result(s): 57 Patients (70,4%) and 58 healthcare professionals (41,1%) completed the questionnaire. Most patients (72,4%) and healthcare professionals (78,6%) mentioned that information concerning al six interventions should be given before delivery. Patients and healthcare professionals preferred to have this information when the intervention was indicated (50,6% resp. 44,5%). Patients preferred not explicitly asking for informed consent (38,1%), while healthcare professionals preferred implicit informed consent (52,3%). Patients mentioned significantly more often that giving consent for an intervention is not necessary compared to healthcare professionals (OR 12,64 (95%CI 7,30-21,87)).

Conclusion(s): Healthcare professionals should present information about the content, indication, risks and alternatives of an intervention, in keeping with the patient's preferences for receiving medical information and the obligation of informed consent for healthcare professionals. Patients want to be informed concerning interventions during labor, preferably during labor when the intervention is indicated. Patients seem to find explicit informed consent less important in comparison to healthcare professionals.

PS4-35 ESTIMATING THE IMPACT OF DIRECT-TO-CONSUMER ADVERTISING ON PREFERENCES FOR PHARMACY ONLY MEDICINES

Patient and Stakeholder Preferences and Engagement (PSPE)

Stephen Goodall, PhD, Rosalie Viney, PhD and Deborah Street, PhD, University of Technology Sydney: Centre for Health Economics Research and Evaluation, Sydney, Australia

Purpose: This project aims to investigate whether a discrete choice experiment could estimate the impact of direct to consumer advertising on pharmacy only medicines.

Method(s): A DCE was developed to investigate preferences about the management of a common health problem, cold sores. The treatment options included pharmacist only medicines (requiring a brief consultation with a pharmacist) and over the counter products. A novel aspect of the DCE was the inclusion of a short 30 second TV-style information advertisement that promoted one of the products included in the DCE. Respondents were allocated to one of three arms: Arms 1 and 2 were shown the advertisement before they answered the DCE task. Arm 3 was a control arm in which the advertisement was shown at the end of the survey, after the DCE task. The DCE task included 16 choice sets, respondents were asked to choose between two products and no treatment. Each product was described by a product type (which was either a named brand or a generic brand), and by product attributes (frequency and duration of use, effectiveness, pharmacist recommendation and cost). The responses were analysed using a generalised multinomial logit and the willingness to pay for each products was estimated.

Result(s): In total, 1295 consumers and 501 Pharmacists participated in the study. The consumer respondents were stratified in terms of whether they were experienced cold sore sufferers or not. Respondents who viewed the advertisement were more likely to seek treatment from a pharmacy. The pharmacist's recommendation had a positive impact on the consumers' choice of product. The advertisement increased awareness of both the advertised product and generic alternatives. The impact of the advertisement was largest amongst consumers who were not experienced cold sore sufferers. These findings suggest the advertisement raised awareness of the pharmacy services and increased their knowledge about treatment options available. The advertisement did not have a significant impact on the recommendations made by pharmacy professionals.

Conclusion(s): The project demonstrates the feasibility of using a DCE to estimate the impact of direct to consumer advertising of pharmacy only medicines. The results will inform decisions about how such advertising is regulated in the Australian setting.

PS4-36 DIFFERENCES BETWEEN VIDEO AND DESCRIPTION IN DISCRETE CHOICE EXPERIMENTS: UPPER LIMB PROSTHETIC MOTION

Patient and Stakeholder Preferences and Engagement (PSPE)

Leslie Wilson, PhD[1], David Szeto, PharmD[2], Matthew Garibaldi, Prosthetist, MS[2] and Fatema Turkistani, Ph.D.[1], (1)University of California, San Francisco, San Francisco, CA, (2)UCSF, San Francisco, CA

Purpose: The Food and Drug Administration released a guidance to allow inclusion of patient preference information for approval decisions for new devices, especially when preference sensitive. 70% of patients with limb loss adopt but don't use their prosthetic devices. Concurrently prosthetic innovation is rapid, resulting in uncertainty about patient's choices, which could be explored by discrete choice measures. Uncertainty still exists, however, around the best methods for presenting discrete choice measures to patients. Our purpose is to compare the use of video presentation of the complex movements involved in prosthetic benefits, with more traditional picture based presentation of a choice based conjoint measure of two prosthetic advances: osseointegration and neuroelectric control.

Method(s): We used a modified meta-ethnographic approach, including concept synthesis, and interviews to construct two choice-based conjoint instruments comparing risks and benefits of upper limb prosthetic choices, one which includes three short imbedded videos to describe prosthetic motion, sensation, and capabilities, and another with descriptions only. The two measures were presented exactly the same except for the video. Using our 20 patient pilot sample of patients with upper limb loss, we compare preference choices between these two presentation approaches.

Result(s): Responses are generally stronger when video is used compared to descriptions only. A 'complete lack of independence in cooking dinner' is equal to a 50% risk of 'serious but treatable infection' (0.06). Patients expressed almost twice the preference for motion ability with 8 grips, full strength, and fluid motion (.09 vs .05), and almost 3 times the aversion to only 2 grips, little strength, and choppy motion (-.15 vs -.06) when shown in video compared to descriptions. The choices with video affected the choices of attributes without video, with 'risk of complete prosthetic failure for 5 years' and 'how connected the prosthetic feels as part of the body' selected less often.

Conclusion(s): Video can be useful for an attribute that needs to convey a complex concept such as hand and arm motion. Although, video can affect the strength of the choices, it did not appear to affect the preference order of levels within an attribute.

1A-1 IMPROVING A PATIENT DECISION AID FOR PATIENTS WITH LOCALIZED PROSTATE CANCER USING THE MULTI-CRITERIA DECISION ANALYSIS METHOD

Decision Psychology and Shared Decision Making (DEC)

Isabel de Angst, MD[1], Paul J.M. Kil, MD, PhD[1], **Marieke G.M. Weernink, PhD**[2], Eric Cornel, MD, PhD[3], Janine A. van Til, PhD[2] and Johanna Takkenberg, MD, PhD[4], (1)Elisabeth-Tweesteden Hospital, Tilburg, Netherlands, (2)University of Twente, Enschede, Netherlands, (3)Ziekenhuis Groep Twente, Hengelo, Netherlands, (4)Department of Cardiothoracic Surgery, Erasmus University Medical Center, Rotterdam, Netherlands

Purpose:

Previously, a web-based decision aid (DA) with values clarification method (VCM) for patients with localized prostate cancer was designed and evaluated. Although the DA was positively evaluated, about 34% of the patients did not indicate a preferred treatment option after DA use. The VCM, consisting of a series of rating scales, was not designed to quantitatively show patients which treatment best matched their stated preferences. Our hypothesis is that receiving a value driven treatment preference might be beneficial for indecisive patients. Therefore, we aimed to develop an explicit VCM using the analytic hierarchy process (AHP), a multi-criteria decision analysis tool.

Method(s):

To identify the alternative treatment options and attributes, literature and the existing treatment DA were studied. Active surveillance (AS), radical prostatectomy (RP), external beam radiotherapy (EBRT), and brachytherapy (BT) were considered as treatment options. Internal project meetings, and a focus group with seven treated patients with localized prostate cancer was conducted to select the most important differences between treatments (attributes). Subsequently, the prototype VCM was designed and tested among participants of the focus group.

Result(s):

The final VCM consists of two separate exercises which can be used consecutively: VCM-1 (AS vs. curative treatment) and VCM-2 (RP vs. EBRT vs. BT). Patients have to answer a series of pairwise comparisons to estimate attribute importance (Figure 1a). In addition, for the subjective attributes, pairwise comparisons are used to elicit patient specific performance valuations for the treatments (Figure 1b). In total, patients have to answer eight pairwise comparisons for VCM-1 and fourteen for VCM-2. By combining attribute importance and alternative performance an overall value can be calculated for each treatment option, and subsequently, a most preferred treatment can be identified. A table and graphs are used to display patients' preferences. Usability test of the prototype VCM showed that 80% of the participants would recommend this VCM to other patients. The tool was well accepted and only demanded minor modifications.

Conclusion(s):

In this study, we extensively describe the design and development of an explicit AHP-based VCM, which is well accepted by patients. The VCM is also able to calculate the most preferred treatment and can therefore, give an actual treatment advice upon patients' request. The VCM is currently being evaluated in clinical practice.

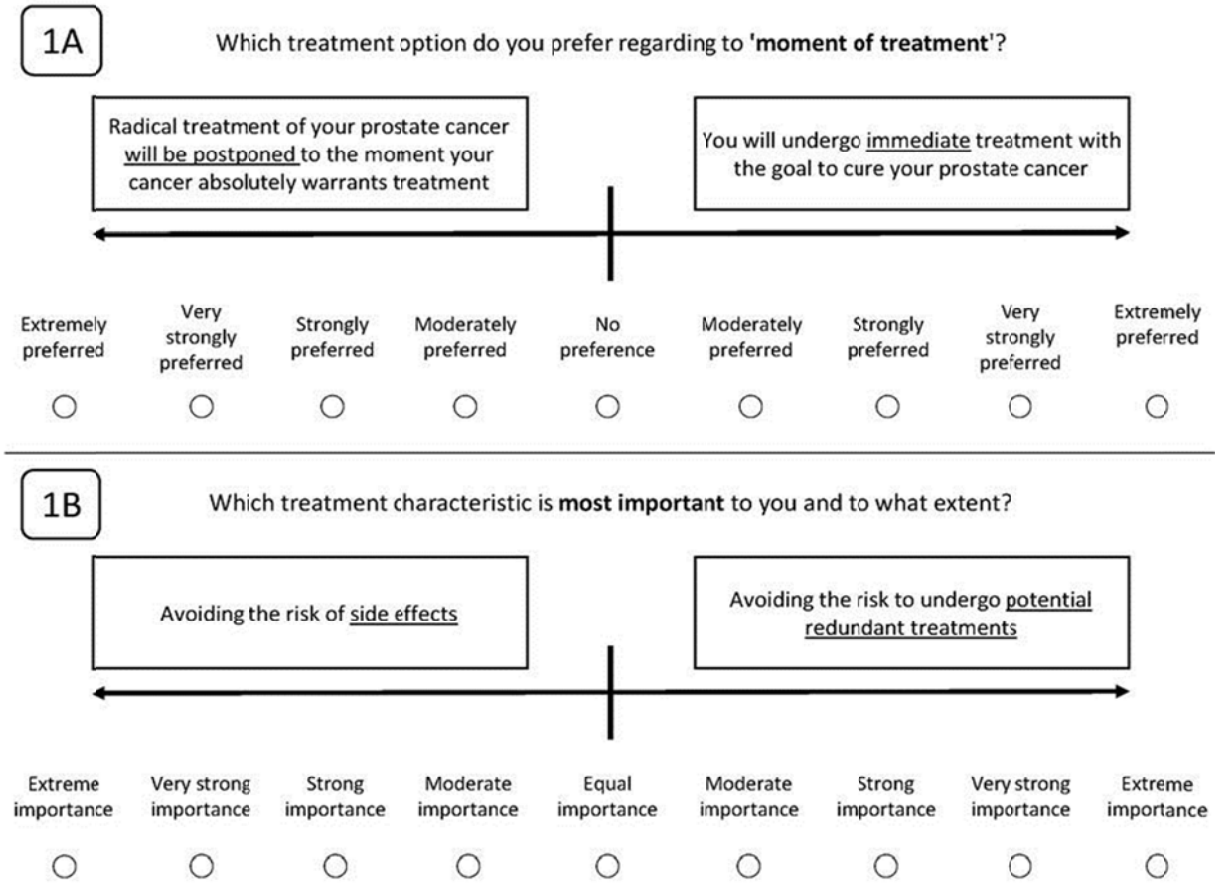


Figure 1.

1A-2 THE EFFECTS OF USING POPULATION DIAGRAMS AND 10-YEAR TIME FRAMES IN COMMUNICATING BREAST CANCER RISKS

Decision Psychology and Shared Decision Making (DEC)

Danielle Timmermans, PhD[1], Janet Geipel, PhD[1], Liesbeth Claassen, PhD[1], Christi van Asperen, PhD[2], Jan Oosterwijk, PhD[3], Fred Menko, PhD[4] and Lidewij Henneman, PhD[5], (1)Department of Public and Occupational Health, Amsterdam Public Health research institute, VU University Medical Center, Amsterdam, Netherlands, (2)Department of Clinical Genetics, Leiden University Medical Center, Leiden, Netherlands, (3)Department of Genetics,University Medical Center Groningen, Groningen, Netherlands, (4)Department of Clinical Genetics, VU University Medical Center, current affiliation Family Cancer Clinic, Antoni van Leeuwenhoek Hospital, Amsterdam, Netherlands, (5)Department of Clinical Genetics, Amsterdam Public Health research institute, VU University Medical Center, Amsterdam, Netherlands

Purpose: Women tend to inaccurately estimate their risk of breast cancer, which may lead to inappropriate surveillance behavior. We tested whether the inclusion of visual aids and narrower time frames, such as 10-year risk as compared to lifetime risk, can improve health risk understanding and perception.

Method(s): Current standard counseling in the Netherlands is based on lifetime risks. In a clinical trial, we tested four risk format interventions. After receiving standard genetic counseling, unaffected women at increased familial breast cancer risk (n = 288) were assigned to one of four intervention conditions: (1) lifetime risk in frequency format, (2) lifetime risk in both frequency and visual formats, (3) lifetime risk and 10-year risk in frequency format, and (4) lifetime risk and 10-year risk in both frequency and visual formats. Baseline, 2-week, and 6-month follow-up risk accuracy, risk perception, and surveillance intentions were collected via questionnaires.

Result(s): For 10-year age-related risk accuracy, we found that the inclusion of 10-year risk information improved the participants' risk understanding of developing breast cancer in the next 10 years at 2-week and 6-month follow-up. For risk perception at 6-month follow-up, it was found that women who received the 10-year risk information, the inclusion of population icons *decreased* their perception of lifetime personal breast cancer risk. However, for the group of women who were not presented with the 10-year risk information, the inclusion of population icons had *no effect* on their perception of lifetime personal breast cancer risk. In terms of preventive intentions at 6-month follow-up, it was found that the inclusion of population icons decreased intentions of breast surveillance for women with either low or moderate breast cancer risk, while it increased intentions for women with high risks of breast cancer.

Conclusion(s): The inclusion of visual formats decreased risk perception for the intervention condition that also included 10-year risk, and, more importantly, resulted in screening intentions more in line with medical recommendation. Based on our findings, we recommend the inclusion of 10 years risk frames with visual formats for genetic counseling of women at increased familial breast cancer risk.

1A-3 FAMILY DYNAMICS IN A MULTI-ETHNIC ASIAN SOCIETY: COMPARISON OF ELDERLY CKD PATIENTS AND THEIR FAMILY CAREGIVERS WITH MEDICAL DECISION MAKING FOR MANAGING END STAGE KIDNEY DISEASE

Decision Psychology and Shared Decision Making (DEC)

Semra Ozdemir, PhD[1], Tazeen Jafar, MD, MPH[1], Lina Choong Hui Lin, MBBS[2] and Eric A. Finkelstein, PhD, MHA[1], (1)Duke-NUS Medical School, Singapore, Singapore, (2)Singapore General Hospital, Singapore, Singapore

Purpose: Elderly end stage kidney patients face a decision concerning whether or not to initiate dialysis. In Asia, this decision is often highly influenced by family caregivers. The objective of this paper was to understand patients' experience with and preferences for family involvement in making treatment decisions, and via a series of hypothetical vignettes, to identify whether there was discordance in treatment preferences between patients and their caregivers, and, if so, how any potential conflicts were reconciled.

Method(s): We conducted a survey with 151 elderly (aged ≥ 65) chronic kidney disease patients and their caregivers at outpatient renal clinics. The survey asked, when making treatment decisions, whom they wish makes the final decisions (i.e., preference) and who usually makes the final decisions (i.e., experience). The survey also presented a series of choice vignettes for managing patient's condition and asked respondents to choose between two hypothetical treatment profiles in each vignette. Patients and caregivers were first interviewed separately in tandem. After the completion of the individual interviews, patient-caregiver dyads were brought together and they were asked to choose a treatment jointly for vignettes where the initial treatment choice differed within the dyad. We used logistic regressions to investigate the predictors of discordance and reconciliation.

Result(s): We found that most (51% (95% Confidence Interval (CI 43-59)) patients preferred and experienced (64% (CI 55-71) significant involvement from caregivers. However, 38% (CI 31-47) of patients preferred to make final decisions alone but only 27% (CI 20-34) of patients did. In the hypothetical vignettes, caregivers chose the more intensive option (i.e., dialysis) more than patients did (26% vs 19%; p value < 0.01). Overall, 44% (CI 36-53) of the dyads had discordance in at least 3 vignettes, and the odds of discordance within patient-caregiver dyads was higher when caregivers choose dialysis ($p < 0.01$). In only 50% (CI 45-56) of the cases, discordance resolved in the patients' favor. Discordance was more likely to reconcile in patient's favor if the patient had financial independence ($p = 0.03$).

Conclusion(s): Our results highlight the important role of caregivers in decision-making but also the potential for them to overstep. Clinicians should be aware of this challenge and identify strategies that minimize the chances that patients may receive treatments not consistent with their preferences.

1A-4 CLINICIAN-PATIENT COMMUNICATION ON THE DECISION TO PROCEED WITH DIAGNOSTIC TESTING FOR (ALZHEIMER'S) DEMENTIA: THE ABIDE-CLINICAL ENCOUNTER STUDY

Decision Psychology and Shared Decision Making (DEC)

Leonie N.C. Visser, PhD[1], Marleen Kunneman, PhD[2], Laxsini Murugesu, MSc[1], Femke Bouwman, MD, PhD[3], Wiesje van der Flier, PhD[3] and Ellen MA Smets, PhD[4], (1)Academic Medical Center, Amsterdam, Netherlands, (2)Leiden University Medical Center, Leiden, Netherlands, (3)Alzheimer center/VU University Medical Center, Amsterdam, Netherlands, (4)Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands

Purpose:

The development of novel diagnostic tests for Alzheimer's disease and dementia has increased the number of available diagnostic options. This study aimed to gain insight into the decision making process regarding diagnostic testing, by means of examining patient-clinician communication as observed during clinical encounters prior to diagnostic testing.

Method(s):

We performed an observational study in the routine diagnostic workup of dementia. In an unselected sample of 125 new patients from eight Dutch memory clinics, we audiotaped the encounters with their clinicians (N=36) prior to diagnostic testing. A study-specific coding scheme was used to code the audiotaped encounters and categorize communication behavior (e.g., '*Statements are made in which diagnostic testing or tests are explained/discussed: yes/no*'). In addition, the OPTION12-scale was used to assess the extent to which the clinicians involved patients in decision making regarding testing (range= 0 [no involvement] to 100 [high involvement]).

Result(s):

In 112/125 encounters, patients (mean age=70, SD=10; mean MMSE-score=25, SD=5) were accompanied by a caregiver. Mean duration of the pre-diagnostic testing encounter was 48 minutes (range 10-101). A reason for the encounter was voiced in 23/125 (18%) of the encounters, but this reason was never related to diagnostic testing. Most often the reason was 'talking about complaints' (21/125). In 74/125 (59%) encounters, testing was addressed to some extent. However, clinicians' behavior to involve patients in decision making regarding testing was low (mean OPTION12-score=16.6, SD=12.8). If a decision was made during the encounter (70/125), diagnostic testing was often presented by the clinician as the obvious next step (e.g., *"I am proposing..."*), and/or the decision related to testing in general (yes/no), rather than to decisions about specific tests.

Conclusion(s):

Results suggest that informing patients on (the available options for) diagnostic testing for Alzheimer's disease and dementia is not routine in pre-diagnostic conversations. The decision to proceed with diagnostic testing (or not) often seemed made prior to the encounter, potentially by the GP or fixed care path. If a decision was made during the encounter, this decision was often implicit, not related to specific tests, and with low levels of patient involvement. More research is needed to explore ways to improve patient involvement in this patient population, and assess their values and preferences related to diagnostic testing.

1A-5 SHARED DECISION-MAKING IN PROSTATE CANCER CARE: ENCOURAGING EVERY PATIENT TO BE ACTIVELY INVOLVED IN DECISION-MAKING, OR ENSURING PATIENTS' PREFERRED LEVEL OF INVOLVEMENT?

Decision Psychology and Shared Decision Making (DEC)

Marie-Anne van Stam, MSc[1], Arwen H. Pieterse, PhD[2], H.G. Van der Poel, MD, PhD[1], J.L.H.R. Bosch, MD, PhD[3], Corinne Tillier, MANP[1], S. Horenblas, MD, PhD[1] and N.K. Aaronson, PhD[1], (1)The Netherlands Cancer Institute, Amsterdam, Netherlands, (2)Medical Decision Making, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands, (3)University Medical Center Utrecht, Utrecht, Netherlands

Purpose:

The aims of this study were: (1) to describe preferred and experienced roles in treatment decision-making among patients with localized prostate cancer (PC); (2) to identify how often patients' experienced roles matched their preferred roles; and (3) to determine whether active involvement in decision-making regardless of role preferences, or concordance between preferred and experienced role is the strongest predictor of more favourable patient-reported outcomes.

Method(s):

In this prospective, multicenter, observational study we obtained serial questionnaire data from newly-diagnosed localized PC patients (cT1-cT2 or Gleason≤7, PSA≤20) (N=454). Questionnaires were completed prior to treatment, and at three, six, and twelve months post-treatment follow-up. Clinical data were obtained from patients' medical records. Active involvement and role concordance were operationalized using the Control Preferences Scale. Analysis of variance and effect sizes (Cohen's d; 0.2=small, 0.5=medium) were used to compare patients' knowledge of prostate cancer, decisional conflict, decisional regret, and overall health-related quality of life (HRQoL).

Result(s):

Most patients (87%, n=292) reported having been actively involved in treatment decision-making. However, 17% (n=78) indicated having had either less or more involvement than preferred. Active involvement was significantly associated with more PC knowledge (d=0.30), less decisional conflict (d=0.52), and less decisional regret (d=0.34). Role concordance was also, but less strongly, associated with less decisional conflict (d=0.41).

Conclusion(s):

Our findings support a policy of encouraging all localized PC patients, regardless of their stated role preferences, to be actively involved in the decision about their treatment.

1A-6 LIMITED INVOLVEMENT IN DECISION-MAKING BY THEIR ONCOLOGIST DOES NOT REDUCE BREAST CANCER PATIENTS' TRUST

Decision Psychology and Shared Decision Making (DEC)

Irini Sorial, MSc.[1], **Ellen Engelhardt, PhD**[2], Arwen H. Pieterse, PhD[3], J. (Hanneke) CJM de Haes, PhD[4], Ellen Smets[5], Anne M. Stiggelbout, PhD[3] and Marij Hillen, PhD[6], (1)Academic Medical Center, Dept. of Medical Psychology, Amsterdam, Netherlands, (2)VU University Medical Center, Dept. Epidemiology and Biostatistics and Amsterdam Public Health Research Institute, Amsterdam, Netherlands, (3)Medical Decision Making, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands, (4)Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands, (5)Department of Medical Psychology, Amsterdam Public Health research institute, Academic Medical Centre Amsterdam, Amsterdam, Netherlands, (6)Academic Medical Center, Amsterdam, Netherlands

Purpose: Adjuvant systemic treatment for early-stage breast cancer significantly reduces mortality, but is associated with many side-effects reducing patients' quality of life in the short- and long-term. Decisions about adjuvant treatment are preference-sensitive, i.e., need to be guided by the patient's informed preferences, and are ideally suited to a Shared Decision Making (SDM) approach. Implementing SDM might strengthen patients' trust in their oncologist, by enhancing openness and autonomy. Conversely, it could reduce trust by inducing uncertainty. Literature on the potential relationship between SDM and patient's trust in their oncologist is lacking. We investigated whether SDM in the consultation was associated with patients' trust in their oncologist, and explored if other patient and oncologist characteristics were associated with patients' trust in their oncologist.

Method(s): In total, 101 decision consultations between breast cancer patients and their medical oncologist were audio-recorded and transcribed verbatim. Observed level of SDM was scored from the transcripts by researchers (double-coded until an inter-rater kappa ≥ 0.7 was achieved) using the 12-item Observing Patient Involvement In Decision-making rating scale (OPTION-12). Patient's perceived involvement in decision-making was ascertained in a telephone interview (*Who made the final treatment decision in your opinion?*) and categorized as patient-driven, shared or oncologist-driven. Patients' trust in their oncologist was measured using the 18-item Trust in Oncologist Scale (TiOS). Multivariate linear regression was used to identify characteristics associated with patients' trust in their oncologist.

Result(s): The average TiOS score was 4.05 (*SD*=0.56) out of a maximum score of 5, whereas the average OPTION-12 score was 15.5 (*SD*=11.6; range scores: 2.08 and 56.25) on a scale from 0-100. OPTION-12-scores (Pearson's R =.021; p =.837) nor the patient's perception of involvement in decision-making (endocrine therapy: F(1)=.257; p=.613); chemotherapy (F(1)=0.079; p=.780)) were associated with TiOS-score. A larger tumor size (p=.035; β=.210) and the use of Adjuvant! (a prediction model) during the consultation to communicate survival probabilities (p =.031; β =.215) were associated with higher TiOS scores.

Conclusion(s): Patients' reported trust in their oncologist was strong despite the lack of SDM in virtually all consultations. In our study population, trust in the oncologist appeared not associated with patients' role in the decision-making process, but with disease severity (i.e., having a larger tumor) and oncologists' open communication about patients' prospects (i.e., disclosure of survival probabilities).

1B-1 VALUE OF INFORMATION CHOICES THAT INFLUENCE ESTIMATES: A SYSTEMATIC REVIEW AND DEVELOPMENT OF THE VOICE CHECKLIST

Applied Health Economics (AHE)

Hendrik Koffijberg, PhD, University of Twente, Enschede, Netherlands, Claire Rothery, PhD, University of York, York, United Kingdom, Kalipso Chalkidou, PhD, Global Health and Development Group, Institute for Global Health Innovation, Imperial College London, UK, London, United Kingdom and Janneke Grutters, PhD, Radboud university medical center, Nijmegen, Netherlands

Purpose:

Value of Information (VOI) analysis can be used to quantify the expected value of eliminating uncertainty as described in the ISPOR-SMDM Modeling Good Research Practices Task Force Report 6 (Briggs et al, 2012). Although VOI analyses are increasingly advocated and used for reimbursement decisions and for research prioritization, the interpretation and usefulness of VOI outcomes depends critically on the underlying choices and assumptions used in the analysis. To increase awareness and transparency of these choices and assumptions underpinning VOI outcomes we develop a reporting checklist.

Method(s):

A systematic literature review was performed in PubMed to identify all aspects of VOI analyses that were previously reported to potentially influence VOI outcomes. All records in which VOI-related terms were present in the publication title, abstract, or full text were included. No restrictions regarding type of records/publications, language, date of publication, study design, study outcomes, or funding, were applied. The identified aspects were grouped, to develop a checklist for standardizing the reporting of VOI outcomes. Explanations were defined, and examples were provided, for all checklist items.

Result(s):

We retrieved 687 unique papers, of which 71 original papers and 8 reviews were included. In the full text of these 79 papers a total of 16 aspects were found that may influence VOI outcomes. These aspects related to the underlying evidence (bias, synthesis, heterogeneity, correlation), uncertainty (structural, future pricing), model (relevance, approach, population), choices in VOI calculation (estimation technique, implementation level, population size, perspective) and to assessing the value of future study designs (reversal costs, efficient estimator). These aspects were aggregated into 7 items to form a reporting checklist, in which each item can be described and justified. The checklist items were 1) Health economic model; 2) VOI perspective; 3) VOI estimation; 4) Implementation; 5) Population benefitting from research; 6) Expected relevant future changes; and 7) Proposed research portfolio (optional), together covering the 16 aspects in subitems.

Conclusion(s):

The developed reporting checklist should increase awareness of key choices underlying VOI analysis, and facilitate structured reporting of such choices and interpretation of the ensuing VOI outcomes by researchers and policy makers. Use of this checklist should improve prioritization and reimbursement decisions, in particular when combined with good practice guidelines describing adequate use and justification of methodologies in VOI analysis.

1B-2 VALUE OF INFORMATION OF EARLY INTERVENTION FOR TREATING SUBTHRESHOLD PANIC DISORDER IN THE NETHERLANDS: HEALTHCARE VERSUS SOCIETAL PERSPECTIVE

Applied Health Economics (AHE)

Robbin H. Ophuis, MSc[1], Joran Lokkerbol, PhD[2], Juanita A. Haagsma, PhD[1], Mickael Hiligsmann, PhD[3], Silvia Evers, PhD, LLM[3] and Suzanne Polinder, PhD[1], (1)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (2)Centre of Economic Evaluation, Trimbos Institute, Netherlands Institute for Mental Health and Addiction, Utrecht, Netherlands, (3)Department of Health Services Research, Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, Netherlands

Purpose: Value of information (VOI) informs policy makers about the expected costs of uncertainty, which facilitates decision making regarding potential reimbursement of treatment options in healthcare. The aims of this study were to apply VOI analyses to a previously developed model and to examine the effect of perspective on VOI. The previously developed model assessed the cost-effectiveness of adding a cognitive behavioural therapy (CBT) based early intervention for adults with subthreshold panic disorder to the existing health care for panic disorder in the Netherlands.

Method(s): A four-state Markov model was developed that assessed the 5-year population cost-effectiveness using healthcare and societal costs per quality adjusted life-year (QALY) gained as outcome. The model showed that adding a CBT-based early intervention to the usual care for panic disorder was more effective at higher costs from healthcare perspective, with an average incremental cost-effectiveness ratio of €17,092 per QALY gained. From societal perspective, the early intervention was cost-saving on average. The early intervention was therefore the optimal choice based on current information when a willingness-to-pay (WTP) threshold of €20,000 per QALY was applied. The expected value of perfect information (EVPI) was calculated as the difference between the expected monetary benefit with perfect information and the expected monetary with current information.

Result(s): Given a WTP of €20,000 per QALY, the population EVPI equalled €756 thousand from healthcare perspective and €47.4 million from societal perspective. These values thus represent the estimated costs of eliminating all parameter uncertainty. The EVPI was valued at different WTP thresholds (figure 1). The probability that the early intervention is cost-effective increased for larger WTP thresholds. The EVPI thus decreased as the decision is less likely to change by performing additional research. Because the early intervention is cost-saving from societal perspective, the probability of a correct decision always increases at larger WTP thresholds. The societal EVPI therefore declines.

Conclusion(s): From healthcare perspective, additional research is potentially worthwhile to reduce decision uncertainty. When a societal perspective is applied, additional research is not likely to be worthwhile as additional information would not change the decision. Our results therefore underline that the VOI should be adjusted to the perspective that is most relevant for the decisions makers in order to avoid erroneous choices of research priorities.

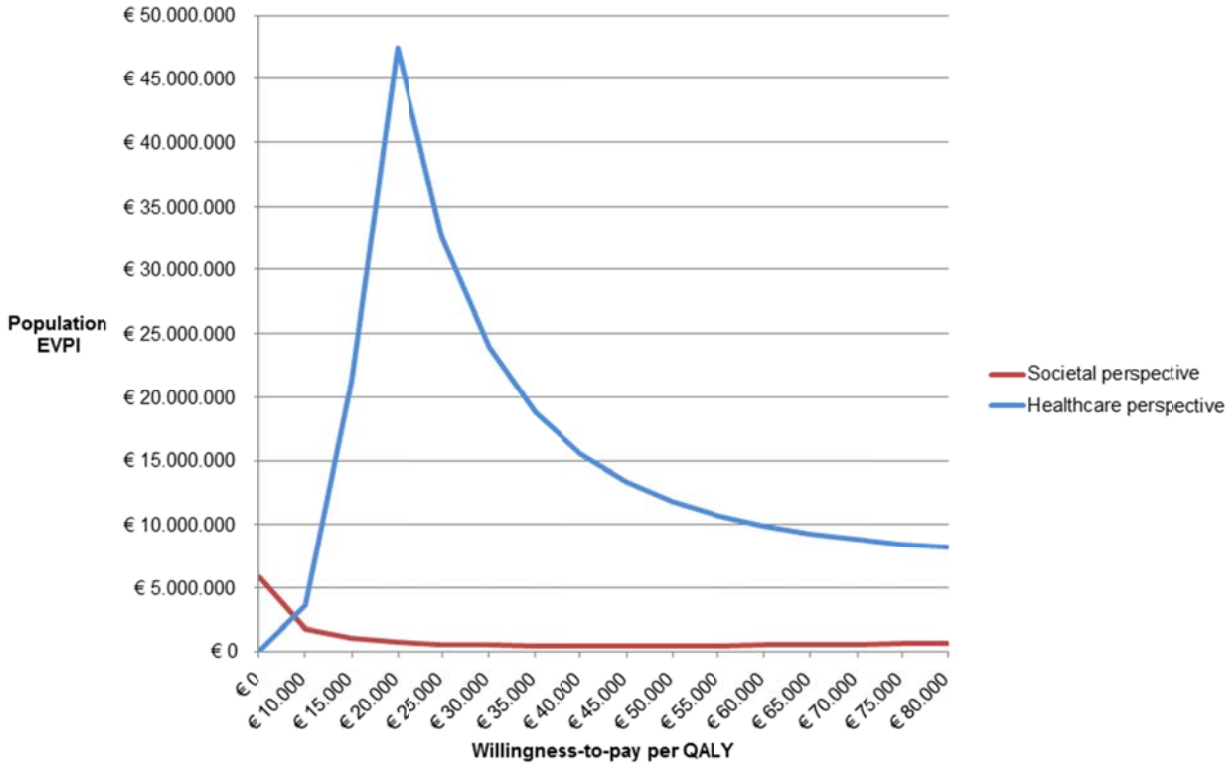


Figure 1: The population EVPI for different willingness-to-pay threshold values

1B-3 COST-EFFECTIVENESS OF TYPHOID VACCINATION STRATEGIES: COST-EFFECTIVENESS ACCEPTABILITY CURVES, COST-EFFECTIVENESS ACCEPTABILITY FRONTIER AND EXPECTED VALUE OF PARTIAL PERFECT INFORMATION FOR 54 COUNTRIES

Applied Health Economics (AHE)

Joke Bilcke[1], Marina Antillon[1], Zoe Pieters[1], Elise Kuylen[1] and Virginia Pitzer[2], (1)University of Antwerp, Wilrijk (Antwerp), Belgium, (2)Yale School Public Health, New Haven, CT

Purpose: To identify for each of the 54 GAVI-eligible countries the preferred (in terms of cost-effectiveness) vaccination strategy against typhoid fever: either no vaccination, routine immunisation alone or routine immunisation complemented with a catch-up campaign up to 5 or 15 years of age.

Method(s): Output from a transmission dynamic model was combined with information on typhoid treatment costs, vaccine-related costs and disability-adjusted life-years. Country-specific data was used if available. All uncertainty was quantified in a probabilistic way (probabilistic sensitivity analysis). The results for each country are presented on a website (including cost-effectiveness acceptability curves, cost-effectiveness frontier, and expected value of perfect information (EVPPI) for each uncertain input parameter for a range of willingness-to-pay values).

Result(s): At a vaccine price of 1.5 USD per dose, on average routine vaccination with catch-up up to 15 years of age is the preferred strategy in most GAVI-eligible countries. 'Preferred' means that the strategy has the highest expected net benefit at a willingness-to-pay value lower than or equal to one time the GDP per capita. But for most countries the uncertainty around this is huge: if we consider the lowest willingness-to-pay value for which any vaccination strategy has at least a 50%/80% probability to be the most optimal strategy, vaccination is the preferred strategy in only 20/5 countries. This within-country uncertainty is mainly due to the uncertainty about the average annual number of typhoid cases and the typhoid case fatality rate. Vaccination is unlikely to offer good value for money in Kyrgyz Republic, Lesotho and Tajikistan, mainly because of very low estimated typhoid incidence (i.e. less than 30 symptomatic cases per year). Vaccination against typhoid is very likely to be cost-effective in Yemen, India, Solomon Islands and Papa New Guinea, mainly because of high typhoid incidence and relatively high treatment costs, and in Nigeria because of an exceptionally high case fatality rate measured in the country.

Conclusion(s): Our results can be used directly to inform countries and agencies on the cost-effectiveness of introducing typhoid vaccination. Future efforts should focus on registering incidence, in particular age at which typhoid occurs, and the probability to die from typhoid fever.

1B-4 DEVELOPING AND VALIDATING A METAMODEL OF THE CEPAC-PEDIATRIC MODEL FOR INFANT HIV TESTING AND SCREENING STRATEGIES
Applied Health Economics (AHE)

Djora Soeteman, PhD[1], Lorna Dunning, MBiochem, MPH[2], Taige Hou[3], Caitlin Dugdale, MD[2], Simone Frank, BA[2], Pamela Pei, PhD[2], Rochelle Walensky, MD, MPH[3], Martina Penazzato, MD, PhD[4], Stephen C. Resch, MPH, PHD[1], Kenneth Freedberg, MD[5] and Andrea Ciaranello, MD, MPH[2], (1)Center for Health Decision Science - Harvard T.H. Chan School of Public Health, Boston, MA, (2)Medical Practice Evaluation Center - Massachusetts General Hospital, Boston, MA, (3)Massachusetts General Hospital, Boston, MA, (4)World Health Organization, Geneva, Switzerland, (5)Massachusetts General Hospital, Harvard Medical School, Boston, MA

Purpose: Metamodels to simplify complex health policy models may promote reproducibility of cost-effectiveness analyses. We investigated the predictive validity of regression-based metamodels of a microsimulation model of infant HIV testing/screening strategies.

Method(s): We developed metamodels from the Cost-effectiveness of Preventing AIDS Complications (CEPAC)-Pediatric model of two infant HIV testing/screening programs in South Africa. We evaluated early infant diagnosis *(EID)* per national guidelines compared to immunization clinic-based screening of all infants for HIV exposure, with further testing for those identified as exposed (*Screen and test*). Outcomes were life expectancy (LE), lifetime HIV-related costs, and net-health benefit (NHB, assuming willingness-to-pay of $790/life-year saved). We fitted: 1) linear regression models (for LE and NHB, ordinary least squares regression [OLS]; for costs, generalized linear models [GLM] with Gaussian family and log link), and 2) generalized additive models (GAM). For each strategy, we assessed agreement between metamodel and CEPAC estimates for all three outcomes with Bland-Altman plots, defining limits of agreement as the range within which 95% of differences between methods fell. When comparing the two strategies, predictive validity was determined by the percentage of simulations in which the metamodel accurately predicted the strategy with the greatest CEPAC-projected NHB. We also calculated NHB losses from "wrong" decisions by the metamodel.

Result(s): With *EID*, the Bland-Altman plots showed good agreement for LE and NHB between CEPAC and both metamodels. Comparing CEPAC and GLM/OLS, mean difference was 0.0000003 (95% limits: -2.1–2.1) life-months in LE, 0.05 (95% limits: -2.8–2.9) life-months in NHB, and -3.4 (95% limits: -92.4–85.6) USD in costs. Comparing CEPAC and GAM, mean difference was -0.000001 (95% limits: -2.1–2.1) life-months in LE, -0.000001 (95% limits: -3.4–3.4) life-months in NHB, and -0.0000000009 (95% limits: -98.4–98.4) USD in costs. Results were similar for *Screen and test*. The optimal strategy (maximal NHB) was predicted accurately by GLM/OLS in 92.7% of simulations and by GAM in 93.4%. Mean NHB lost was 0.013 (range: 0.0027–0.75) life-months for GLM/OLS and 0.009 (range: 0.0027–0.60) life-months for GAM.

Conclusion(s): Regression-based metamodels can predict the behavior of a complex microsimulation model. GLM/OLS offers similar predictive validity as GAM, with lower computational requirements. These metamodels can be made publicly available and allow policymakers to customize simulation model-based analyses to local settings.

1B-5 ESTIMATION OF LIFETIME INDIVIDUAL-LEVEL COSTS AND BENEFITS OF DIABETES MELLITUS PREVENTION PROGRAMS

Applied Health Economics (AHE)

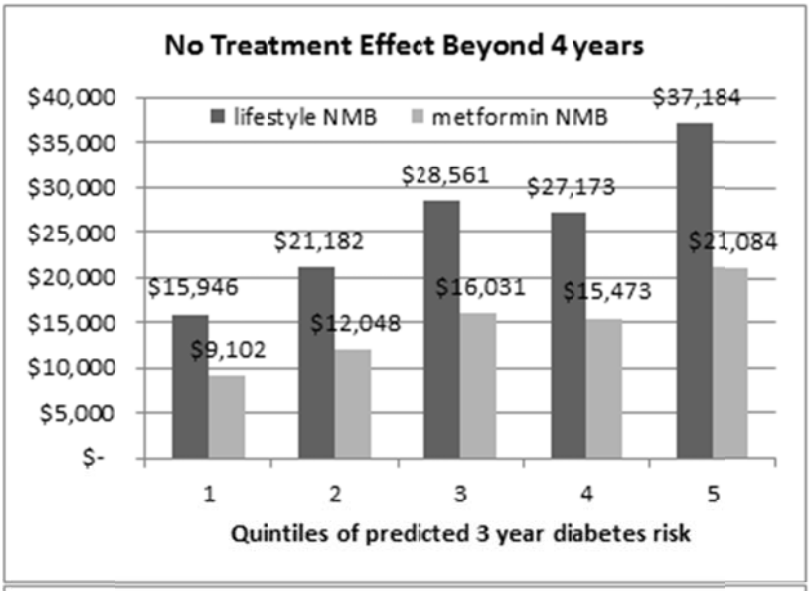
Natalia Olchanski, MS[1], David van Klaveren, PhD[2], Joshua T. Cohen, PhD[3], John B. Wong, MD[1] and David M Kent, MD, MS[1], (1)Tufts Medical Center, Boston, MA, (2)Predictive Analytics and Comparative Effectiveness (PACE) Center, Tufts Medical Center, Boston, MA, (3)Center for the Evaluation of Value and Risk in Health, Tufts Medical Center, Boston, MA

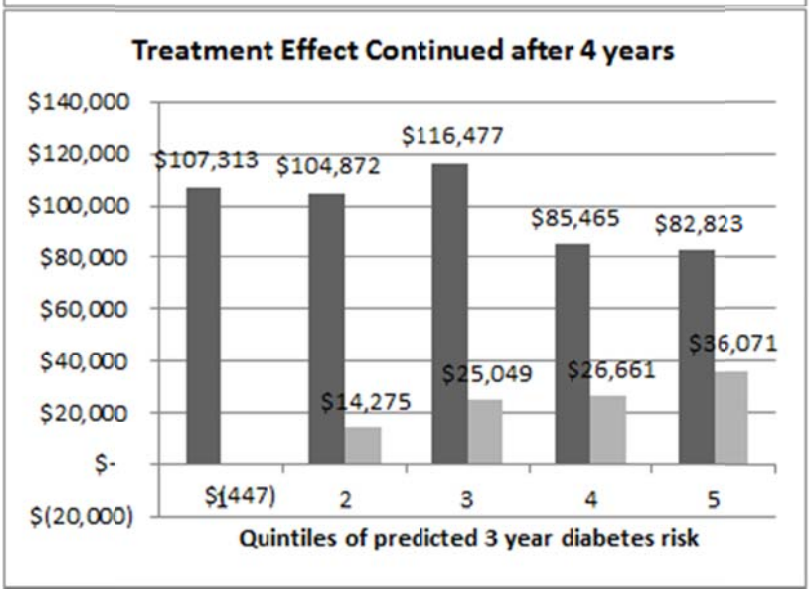
Purpose: To estimate individual-level cost-effectiveness of diabetes prevention programs in individuals with pre-diabetes over a lifetime horizon, and assess whether targeting treatment using individual information may increase population net monetary benefit (NMB).

Method(s): We estimated total costs, life expectancy (LE), and quality-adjusted life expectancy (QALE) using a decision-analytic micro-simulation model for three alternatives: (1) lifestyle intervention that involves an intensive program focused on maintaining healthy diet and exercise, (2) metformin administration, and (3) no intervention. The model combines several components. First we use a Cox proportional hazards model for predicting risk of onset of diabetes in pre-diabetics, given baseline characteristics and prevention strategy to derive a probability distribution for each individual. This risk model was derived from the Diabetes Prevention Program (DPP) clinical trial data and its follow-up study DPP-OS, and further extended using a parametric model beyond the first four years. The Michigan Diabetes Research Center Model for Diabetes was then used to estimate costs and outcomes for individuals after diabetes diagnosis, assuming standard of care diabetes treatment. We projected individual expected costs, LE, and QALE using age-specific probabilities and outcomes conditional on age of diabetes onset, and evaluated NMB of the two interventions. We conducted analyses varying assumptions for longterm duration of treatment effects.

Result(s): The DPP trial followed 3234 adults with impaired glucose tolerance for 3 years, and was continued in a 7 year open-label follow-up study with 2766 participants. In our models, for base case, assuming no additional treatment effect after 4 years, lifestyle intervention delayed onset of diabetes by 2.6 years on average, and metformin by 1.3 compared to no intervention. For the alternative scenario assuming continuing treatment effect, the onset of diabetes was delayed by 7 years with lifestyle intervention, and by 3.4 years with metformin. In both scenarios the interventions were cost-effective on average, and for most individuals NMB (Figure) were positive, though generally higher for the lifestyle intervention. The average NMB for metformin was negative in the lowest risk quintile.

Conclusion(s): The distribution of individual NMB suggests that while lifestyle intervention is favorable in all patients, some individuals derive greater value than others. Metformin therapy may be cost-effective in higher risk individuals, but not be worth the cost in individuals at lowest risk of developing diabetes.


No Treatment Effect Beyond 4 years
$40,000
$35,000
$30,000
$25,000
$20,000
$15,000
$10,000
$5,000
$-
lifestyle NMB
metformin NMB
$15,946
$9,102
$21,182
$12,048
$28,561
$16,031
$27,173
$15,473
$37,184
$21,084
1
2
3
4
5
Quintiles of predicted 3 year diabetes risk


Treatment Effect Continued after 4 years
$140,000
$120,000
$100,000
$80,000
$60,000
$40,000
$20,000
$-
$(20,000)
$107,313
$(447)
$104,872
$14,275
$116,477
$25,049
$85,465
$26,661
$82,823
$36,071
2
3
4
5
Quintiles of predicted 3 year diabetes risk

1B-6 COST-EFFECTIVENESS OF INITIATIVES TO EITHER INCREASE NUMBER OF PATIENTS ON HEPATITIS C TREATMENT OR REDUCE BURDEN OF DISEASE

Applied Health Economics (AHE)

Torbjorn Wisloff, PhD, MSc[1], Richard White[2], Olav Dalgard[3], Ellen Amundsen[2], Hinta Meijerink[2] and Hilde Klovstad[2], (1)University of Oslo, Oslo, Norway, (2)Norwegian Institute of Public Health, Oslo, Norway, (3)Akershus University Hospital and University of Oslo, Lørenskog, Norway

Purpose:

New drugs for treating hepatitis C have considerably increased the probability of being cured from the disease. Treatment uptake, however, is still low. The objective of this study was to analyse the impact of different initiatives that may increase the proportion of infected people on treatment and interventions aimed at reducing the incidence of new infection among people who inject drugs. The World Health Organization targets of reducing hepatitis C incidence with 90% within 2030 was set as target.

Method(s):

A tested and validated compartmental model for Norway was used to simulate hepatitis C and related complications such as cirrhosis and hepatocellular carcinoma. The model is implemented in R and simulates the complete population of people who are currently or have previously injected drugs. All uncertain input parameters are simulated in order to facilitate probabilistic sensitivity analysis. Cost estimates were based on official Norwegian sources.

Interventions aimed to increase the proportion of infected people on treatment were: screening for hepatitis C by general practitioners and screening at addiction treatment centres, analysed assuming either treatment only for those with cirrhosis or all infected cases. Interventions aiming at reducing the hepatitis C incidence analysed were opioid substitution therapy, a clean user equipment programme and a combination of both. Intervention effects was based on meta-analyses or randomised controlled trials for all four interventions.

Result(s):

No screening alternatives were cost-effective if only treating hepatitis C patients with established cirrhosis. The most cost-effective screening initiative when treating all infected, regardless of disease status, was screening conducted at general practitioners office. We estimated an incidence reduction of hepatitis C by 2030 of 63% compared to the current.
The two harm-reduction strategies both reduced the incidence of hepatitis C by more than 70%. Combining an increase of the current clean user equipment programme with opioid substitution therapy was clearly the most cost-effective harm reduction option. This strategy would reduce the incidence of hepatitis C by 84% compared to the current incidence by 2030.

Conclusion(s):

Interventions to reduce the burden and spread of hepatitis C are cost-effective. The WHO target of reducing hepatitis C incidence by 90% within 2030 may be difficult to reach without combining different initiatives.

2E-1 EVOLVING INDIVIDUAL PREFERENCES FOR VACCINATION POLICIES

Health Services, Outcomes and Policy Research (HSOP)

Nicolas W. Meier, M.Sc.[1], Robert Boehm, PhD[1], Cornelia Betsch, PhD[2] and Lars Korn, M.Sc.[2], (1)RWTH Aachen University, Aachen, Germany, (2)University of Erfurt, Erfurt, Germany

Purpose: Vaccination is one of the most successful preventive health measures. Despite its success in the past, vaccine hesitancy is an increasing threat to achieve high vaccination rates and to eliminate and eradicate infectious diseases. More and more countries discuss about introducing partial or full mandatory vaccination policies. For instance, in 2017, Italy and France introduced mandatory vaccination policies for standard childhood vaccines and the California State Senate passed one of the strictest vaccination bills (SB 277) in the United States. We investigate the development of individual preferences for mandatory and voluntary vaccination policies to understand the determinants of short-term and long-term support for these preferences.

Method(s): We conduct a controlled and incentivize-compatible laboratory experiment using an interactive vaccination game. In each round, participants could decide for either a voluntary or mandatory vaccination policy. This methodological approach has important advantages. First, choices are consequential, i.e., after opting for either a voluntary vs. mandatory vaccination policy, participants engage in an interactive vaccination game with decision-contingent monetary incentives. Second, participants face 20 decision rounds in which we implement a permanent competition between a voluntary and mandatory vaccination policy. Thus, we can identify changes in policy preferences over time under controlled conditions.

Result(s): Initially, the majority of participants (N = 168) opts for the voluntary vaccination policy (61%). Over all rounds, participants' preference for the voluntary vaccination policy even increases (average: 70%). Yet, we observe considerable switching in participants' preference for the voluntary vs. mandatory vaccination policy. Individual-level regression analyses reveal that the preference for the mandatory vaccination policy increases with decreasing vaccination rates under voluntary vaccination. However, experiencing vaccine-adverse events under mandatory vaccination increases the preference for voluntary vaccination. The latter effect is larger for individuals with a negative (vs. positive) attitude toward vaccination.

Conclusion(s): When individuals gain experience with both voluntary and mandatory vaccination policies, the preference for voluntary vaccination increases over time. This is due to an increased sensitivity for negative events under mandatory vaccination. The results provide important insights for policy makers: In the long run, such effects may spark opposition to mandatory vaccination policy.

2E-2 POLITICAL AND RELIGIOUS IDEOLOGICAL ASSOCIATIONS WITH SEASONAL INFLUENZA VACCINE UPTAKE IN EUROPE AND THE UNITED STATES

Decision Psychology and Shared Decision Making (DEC)

Aaron M. Scherer, PhD[1], Megan Knaus, MPH[2], Brian J. Zikmund-Fisher, PhD[2], Enny Das, PhD[3] and Angela Fagerlin, PhD[4], (1)University of Iowa, Iowa City, IA, (2)University of Michigan, Ann Arbor, MI, (3)Radboud University Nijmegen, Nijmgen, Netherlands, (4)University of Utah / Salt Lake City VA, Salt Lake City, UT

Purpose: Seasonal influenza vaccination rates among EU member states remain well below the targets established by the European Centre for Disease Control and Prevention (ECDC). The objectives of the current analyses were to examine a) whether political and religious ideologies are associated with seasonal influenza vaccine uptake and b) whether these relationships differ based on membership in a high risk group or across countries.

Method(s): We conducted an online survey in 10 European countries (Finland, Germany, Hungary, Italy, Netherlands, Norway, Poland, Spain, Sweden, and the United Kingdom) and the United States using stratified random sampling based on age and gender. Embedded in the survey were questions about whether the participant had received the influenza vaccine the previous year; measures of political ideology (political orientation; trust in the national ministry of health for their country) and religious ideology (religiosity; religious views); individual differences linked to medical decision making (numeracy; health literacy); and demographics (age; gender; education). Responses to the ideology items were on 7-point Likert scales. Data were analyzed in two ways: 1) regressing the ideology, individual difference, and demographic measures on self-reported influenza vaccination and 2) stratified analyses by age group (above or below recommended age range) and country.

Result(s): The survey was completed by 16,510 participants (81% completion rate). Controlling for numeracy, health literacy, age, gender, and education, each point increase in political conservatism (OR=1.07, 95% CI=1.04, 1.12), trust in the national ministry of health (OR=1.44, 95% CI=1.39, 1.48), progressive religious views (OR=1.03, 95% CI=1.01, 1.06), and religiosity (OR=1.08, 95% CI=1.05, 1.11) was associated with an increased likelihood of having received the influenza vaccine. These results were strongest among younger adults (i.e., under the national recommended age range)—for whom vaccination may be a values-sensitive decision—and varied significantly based on country (see Table 1).

Conclusion(s): The current study provides preliminary evidence that political and religious ideologies may contribute to decisions to receive the seasonal influenza vaccine. These results highlight the need for additional research examining how to communicate about the seasonal influenza vaccine uptake to appeal to individuals with more left-wing political ideologies, more traditional religious values, or who may distrust their national ministries of health, and how ideologies may shape medical decision making more broadly.

Table 1. Odds ratios for ideological associations with influenza vaccine uptake in total sample and by country.

	Political conservatism OR (95% CI)	**Trust in national ministry of health** OR (95% CI)	**Religiosity** OR (95% CI)	**Progressive religious views** OR (95% CI)
Total Sample	**1.07 (1.04, 1.12)**	**1.44 (1.39, 1.48)**	**1.08 (1.05, 1.11)**	**1.03 (1.01, 1.06)**
Finland (n=1050)	**1.26 (1.12, 1.43)**	**1.56 (1.38, 1.76)**	**1.22 (1.10, 1.35)**	1.05 (0.97,1.14)
Germany (n=878)	1.01 (0.86, 1.19)	**1.29 (1.15, 1.43)**	0.93 (0.84, 1.03)	0.91 (0.82, 1.01)
Hungary (n=600)	0.84 (0.69, 1.02)	**1.47 (1.25, 1.72)**	1.07 (0.92, 1.23)	0.96 (0.84, 1.10)
Italy (n=1160)	1.10 (0.98, 1.24)	**1.38 (1.24, 1.53)**	1.06 (0.96, 1.16)	1.04 (0.95, 1.13)
Netherlands (n=1296)	0.91 (0.82, 1.02)	**1.35 (1.22, 1.50)**	**1.09 (1.02, 1.17)**	1.03 (0.95, 1.12)
Norway (n=458)	0.97 (0.81, 1.16)	**1.34 (1.11, 1.61)**	1.12 (0.94, 1.33)	1.02 (0.87, 1.20)
Poland (n=1153)	0.94 (0.81, 1.09)	**1.46 (1.30, 1.66)**	1.00 (0.89, 1.12)	**1.18 (1.08, 1.29)**
Spain (n=1060)	0.99 (0.87, 1.12)	**1.37 (1.22, 1.53)**	**1.15 (1.04, 1.27)**	1.02 (0.93, 1.12)
Sweden (n=750)	1.08 (0.95, 1.23)	**1.29 (1.14, 1.47)**	1.01 (0.89, 1.11)	1.00 (0.89, 1.11)
United Kingdom (n=998)	1.09 (0.96, 1.23)	**1.24 (1.11, 1.38)**	1.05 (0.96, 1.14)	1.08 (1.00, 1.18)
United States (n=1356)	0.99 (0.91, 1.07)	**1.41 (1.30, 1.53)**	1.06 (0.99, 1.13)	1.00 (0.94, 1.06)

NOTE: Results controlling for numeracy, healthliteracy, age, gender, and education. Items in bold statistically significant at $p<.05$.

2E-3 COMPARISON OF INDIVIDUAL AND COLLECTIVE DECISION MAKING FOR HEALTH OUTCOMES IN TIME TRADE-OFF AND STANDARD GAMBLE

Applied Health Economics (AHE)

Stefan Lipman, M.Sc.[1], Arthur Attema, Ph.D.[1], Han Bleichrodt, Ph.D.[1] and Olivier L'Haridon, Ph.D.[2], (1)Erasmus University Rotterdam, Rotterdam, Netherlands, (2)Université de Rennes, Rennes, France

Purpose: Medical decision making is typically a collective process, where patients engage with physicians and significant others before reaching a decision. Nonetheless, the effects of such collective decision making on health outcome preferences are not well-documented within the health economic literature. Our focus is on Quality-Adjusted Life-Years (QALYs), the preferred outcome measure for cost-utility analyses (CUA). The health state weights relevant to this measure are typically derived from individual preferences over health episodes only, without consultation of others. This paper reports the first empirical investigation into the effects of collective decision making on QALY methodology, using both time trade-off (TTO) and standard gamble (SG) tasks.

Method(s): We investigated collective decision making in dyads, where subjects worked in pairs to reach a satisfactory decision. By means of a mixed between-within subjects design, we were able to distinguish between the effect of collective decision making and learning effects for both TTO and SG. Two experimental conditions were used: individual decision making (IDM) and collective decision making (CDM), for three mild health states, described by means of EQ5D. For subjects in both conditions (n = 163), a baseline measurement for both TTO and SG was obtained, afterwards either a filler task (IDM) or a group measurement (CDM) was completed, depending on their condition. Finally, all subjects completed another individual measurement to determine whether learning effects occurred.

Result(s): Our data suggested that collective decision-making has little to no effect on: 1) decision quality, and 2) decision outcomes. More specifically, no systematic discrepancies between CDM and IDM were observed, in terms of consistency and variance for both methods. Furthermore, TTO and SG utilities remained similar across conditions, and the typical difference in elicited utilities between these methods was unaffected by CDM. Learning effects were observed on decision quality and outcomes, collectively suggesting that repetition of TTO and SG measurement is beneficial.

Conclusion(s): These findings suggest that consulting with others has no effect on preferences with regard to health outcomes in TTO and SG, although learning effects may occur. This could be relevant for health state valuation studies, which increasingly utilize personal interview strategies. Additionally, our findings add to the literature of the de-biasing effect of collective decision-making, suggesting that no such effect occurs for TTO and SG.

2E-4 IS PATIENT CHOICE PREDICTABLE? THE IMPACT OF DISCRETE CHOICE EXPERIMENT DESIGNS AND MODELS

Patient and Stakeholder Preferences and Engagement (PSPE)

Esther de Bekker-Grob, PhD[1], Joffre Swait, PhD[2], Habtamu Tilahun Kassahun, PhD[2], Michiel Bliemer, PhD[3], Marcel Jonker, PhD[1], Jorien Veldwijk, PhD[1], Karen Cong[2], John Rose, PhD[4] and Bas Donkers, PhD[1], (1)Erasmus University Rotterdam, Rotterdam, Netherlands, (2)University of South Australia, Sydney, Australia, (3)University of Sydney, Sydney, Australia, (4)University of Technology Sydney, Sydney, Australia

Purpose: Increased use of discrete choice experiments (DCEs) in healthcare requires establishing whether stated preferences are predictive of observed healthcare utilization. This study aimed to determine whether the number of alternatives in a DCE choice task should reflect the actual decision context, and how complex the choice model needs to be to predict real-world choices correctly at an aggregate and individual level.

Method(s): Two randomized controlled trials (RCTs) involving choices for influenza vaccination and colorectal cancer screening were used. Each RCT had three study conditions: DCE choice tasks with (i) two alternatives, (ii) three alternatives, or (iii) both. Two samples of 1,200 respondents each were randomly assigned to one of the conditions. Each respondent answered 16 DCE choice tasks (for the derivation of the decision model) plus a choice task mimicking the real-world choice (to keep the decision context the same). The data was analysed in a systematic way using random-utility-maximization (RUM) and random-regret-minimization (RRM) choice processes with scale and/or preference heterogeneity (based on 19 patient characteristics) and/or random intercepts.

Result(s): Irrespective of the number of alternatives per choice task, the choice to opt for influenza vaccination or colorectal cancer screening was correctly predicted by DCE at an aggregate level, if scale and preference heterogeneity were taken into account. At an individual level, three alternatives per choice task and using heteroscedastic model plus preference heterogeneity seemed to be most promising, correctly predicting the real-world choice in 81.7% to 87.9% of the cases. No evidence was found that RRM outperformed RUM.

Conclusion(s): Our study shows that DCEs hold the potential of being externally valid if at least scale and preference heterogeneity are taken into account. Further research is needed to determine if this result remains in other contexts, and to optimise choice prediction at an individual level.

2E-5 DEVELOPMENT AND PILOTING OF A DISCRETE CHOICE EXPERIMENT FOR PREFERENCE ELICITATION FROM ADOLESCENT-PARENT DYADS

Patient and Stakeholder Preferences and Engagement (PSPE)

Sophy Barber, BDS, MSc, PG Cert in Health Research, Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom, David Meads, PhD, Leeds Institute of Health Sciences, Leeds, United Kingdom, Sue Pavitt, BSc, PhD, University of Leeds, Leeds, United Kingdom, Balvinder Khambay, BDS, PhD, University of Birmingham, Birmingham, United Kingdom and Hilary Bekker, PhD, Leeds Institute of Health Sciences School of Medicine - University of Leeds, Leeds, United Kingdom

Purpose:

To develop a Discrete Choice Experiment (DCE) to elicit adolescent and parent preferences for hypodontia treatment. Hypodontia is a developmental dental condition where one or more teeth fail to develop, requiring complex interdisciplinary dental care over time.

Method(s):

DCE development stages were:

- Attribute identification using mixed methods: a) systematic review of measures for evaluating hypodontia care (56 studies; 8 clinical audits); b) interview study with young people with hypodontia and parents (n=12); c) observational study of hypodontia clinical consultation (n=5); d) environmental scan of hypodontia patient information resources (n=30); d) systematic analysis of social-media posts (n=167).
- Attribute and levels selection from items identified during stage 1. Adolescents with hypodontia and parents (n=18) recruited from a Facebook group rated and ranked all items. Attribute scores were used to inform stakeholder consultation for final attribute selection.
- DCE survey included four sections: study information; demographic characteristics; DCE task; decision-making questionnaire.
- The pilot with adolescents (n=12) and parents (n=8) employed a Think Aloud method in interview setting.

Result(s):

Studies identified 30 attributes of hypodontia care focused broadly around service delivery and outcome factors. Adolescents and parents differed in their ratings; both sets of responses were used to select seven attributes with 2-3 levels for use in the DCE questionnaire. The pilot found young people differed from their parents in their understanding of the DCE and perceptions of utility. Young people found the choice tasks challenging to understand and the presentation of the DCE tasks required iterative refinement and testing. Parents expressed interest in their child's

choices and decision-making, suggesting the DCE may facilitate parent-child discussion about decision-making.

Conclusion(s):

Rigorous methodology ensured the attributes selected for the DCE were relevant to adolescent patients and parents. This stage is often under-reported in DCE development. Piloting identified adolescents were not thinking about their treatment in the same way as adults and health professionals. Preference elicitation from adolescent-parent dyads is novel and may help services deliver better patient-centred hypodontia care.

2E-6 HOW DOES OFFERING DISINCLINED PEOPLE CHOICE BETWEEN DIFFERENT SCREENING APPOINTMENTS AFFECT SCREENING INTENTIONS? EVIDENCE FROM AN ONLINE EXPERIMENT

Decision Psychology and Shared Decision Making (DEC)

Sandro Tiziano Stoffel, Alex Ghanouni, Yasemin Hirst, Jo Waller and Christian von Wagner, University College London, London, United Kingdom

Purpose:

Previous research has demonstrated how increasing access to cancer screening via single timed appointment can increase uptake. The current study investigated whether offering a range of different timed appointment slots would either further enhance this effect or lead to confusion and increased perceived difficulty to make the screening decision.

Method(s):

We recruited 9,127 men and women aged 35-54 without previous diagnosis of bowel cancer living in England for this online study. After being informed about bowel scope screening (BSS), 2,125 (23.3%) stated that they did not intend to participate. These respondents were randomised to receive either 1, 2, 4 or 6 available appointment alternatives. We used multivariable logistic regression adjusting for sociodemographic variables to predict intention by condition and report adjusted odds ratio (aOR) and 95% confidence intervals (CI).

Result(s):

1,908 of 2,125 (89.8%) respondents successfully completed all comprehension and manipulation checks, most were female (57.8%), White-British (82.0%), married/cohabiting (65.5%) and working (76.2%). Compared to offering a single appointment option, we found that offering people the choice of 2, 4 or 6 appointments decreased their likelihood of confirming one of the given appointments (25.8%, 25.3% and 24.8% vs 34.5%). Giving people the choice between 2 appointments significantly decreased intention (aOR 0.64; CI 0.48-0.85). Similarly, offering 4 (aOR 0.63; CI 0.47-0.83) or all 6 alternatives (aOR 0.63; CI 0.47-0.85) decreased intentions to a similar extent. Interestingly, we found no evidence for confusion as individuals across the conditions perceived the invitation process as easy ($p=0.778$) and the screening decision as not difficult ($p=0.550$). Furthermore, individuals in the choice conditions perceived the offered appointments as more convenient ($p=0.003$).

Conclusion(s):

Although we could not find evidence for confusion, the present study supports the notion that offering more than one pre-set appointment may have a detrimental impact on uptake. Reminders for previous non-responders may therefore not benefit from being offered multiple pre-set appointments.

2F-1 ASSESSING THE IMPACT OF A MATCHING ADJUSTED INDIRECT COMPARISON IN A BAYESIAN NETWORK META ANALYSIS

Quantitative Methods and Theoretical Developments (QMTD)

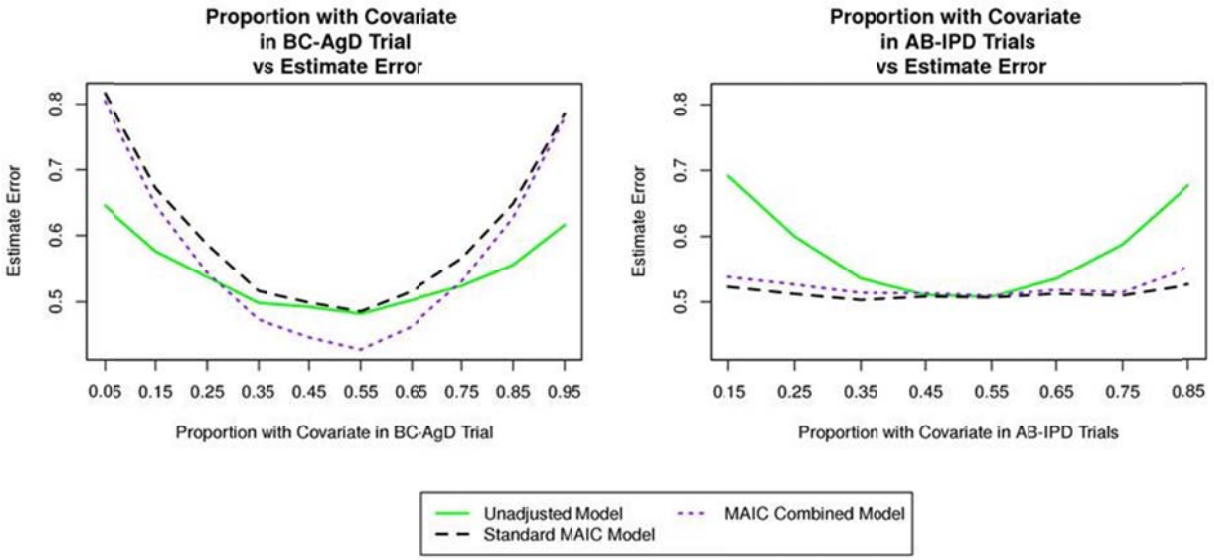
Joy Leahy, BSc, Trinity College Dublin, Dublin 2, Ireland and Cathal Walsh, PhD, University of Limerick, Limerick, Ireland

Purpose: When undertaking a Network Meta Analysis (NMA) a researcher may have Individual Patient Data (IPD) for trials concerning a particular treatment (for example from a sponsor), but none for other trials. Matching Adjusted Indirect Comparison (MAIC) is an increasingly popular method to use in this scenario, whereby the researcher re-weights the IPD to match covariates across trials.

Method(s): We carry out a simulation study to investigate this method in a Bayesian setting. We simulate 3 IPD trials comparing treatments A and B (AB-IPD trials), and one aggregate data trial comparing treatments B and C (BC-AgD trial). We investigate two options of weighting covariates: 1. all three studies are weighted separately to match the AgD trial (standard MAIC). 2. patients are weighted across all three IPD studies to match the AgD trial, but the NMA still considers each trial separately (MAIC combined). We apply these methods to a network of treatments for multiple myeloma.

Result(s): The left panel of the figure shows that when the BC-AgD trial has a particularly small or large proportion of covariates then the unadjusted models produce more accurate results than the MAIC, due to the difficulty in re-weighting trials to match the extreme values. However, when the proportion with the covariate in the BC-AgD trials is less extreme MAIC combined produces better results. In the right panel MAIC approaches are always at least as good as the unadjusted model. However, at the center of the graph when the proportion with the covariates are the same in both the IPD and the AgD trials there is not much difference between the methods. We can see also that the ranking of the standard MAIC and the MAIC combined is reversed between the two graphs. This emphasises the importance of the covariate make-up of the trials in choosing which method to use.

Conclusion(s): MAIC is beneficial as a sensitivity analysis to confirm results across patient populations. If there is a difference attributable to population imbalances, then it is useful to be able to quantify how big this difference is. Sometimes MAIC gives a more accurate estimate. However, this is not always the case. Given the increasing use of MAIC, it is important to further investigate the potential benefits and drawbacks of MAIC.


Proportion with Covariate
in BC-AgD Trial
vs Estimate Error
Estimate Error
0.5
0.6
0.7
0.8
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Proportion with Covariate in BC-AgD Trial
Proportion with Covariate
in AB-IPD Trials
vs Estimate Error
Estimate Error
0.5
0.6
0.7
0.8
0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85
Proportion with Covariate in AB-IPD Trials
Unadjusted Model
Standard MAIC Model
MAIC Combined Model

2F-2 MODEL-BASED NETWORK META-ANALYSIS FOR TIME-COURSE RELATIONSHIPS: A UNION OF TWO METHODOLOGIES

Quantitative Methods and Theoretical Developments (QMTD)

Hugo Pedder, MSc[1], Meg Bennetts, MSc[2], Sofia Dias, PhD[1], Martin Boucher, MSc[2] and Nicky J. Welton, PhD[1], (1)Population Health Sciences, University of Bristol, Bristol, United Kingdom, (2)Pfizer Plc., Sandwich, United Kingdom

Purpose: To develop a model-based meta-anaysis (MBNMA) framework that allows for non-linear modelling of multi-parameter time-course functions for comparative effectiveness, and can account for residual correlation between observations using a multivariate likelihood.

Method(s): Model-based meta-analysis (MBMA) is a technique increasingly used in drug development for synthesising results from multiple studies, allowing pooling of information on treatment, dose-response and time-course characteristics, which are often non-linear. Such analyses are used in drug development to inform future trial designs. Network meta-analysis (NMA) is used frequently in Health Technology Appraisals and by reimbursement agencies for simultaneously comparing effects of multiple treatments. Recently, a framework for dose-response MBNMA has been proposed that draws strengths from both MBMA and NMA.

We expand this framework for modelling of time-course functions that allows for the inclusion of observations from multiple study time points. This methodology preserves randomisation by aggregating within-study relative effects and, by modelling consistency equations on the time-course parameters, it allows for testing of inconsistency between direct and indirect evidence. We demonstrate our modelling framework using an illustrative dataset of 24 trials investigating treatments for pain in osteoarthritis.

Result(s): For our dataset, we report results from 10 different models. An E_{max} function allowed for the greatest degree of flexibility, both in the time-course shape and in the specification of time-course parameters (E_{max} and ET_{50}). Our final model had a posterior mean residual deviance of 291.4 (compared to 345 data points), indicating a good fit to the data. Some simplifying assumptions were needed to identify ET_{50}, as studies contained few observations at earlier follow-up times. Treatment estimates were robust to the choice of likelihood (univariate/multivariate), suggesting that accounting for residual correlation between time points may not be essential if the time-course function has been appropriately modelled and the parameters of interest are summary treatment estimates.

Conclusion(s): Time-course MBNMA combines strengths from MBMA and NMA to allow inclusion of multiple study time points into analyses whilst preserving randomisation and allowing for testing of inconsistency. This has the potential to be used both in helping design and predict studies in drug development, as well as for decision-making by reimbursement agencies, and can act as a bridge between early phase clinical research and Health Technology Appraisal.

2F-3 ESTIMATING THE EXPECTED VALUE OF SAMPLE INFORMATION ACROSS SAMPLE SIZE USING BAYESIAN NON-LINEAR REGRESSION

Quantitative Methods and Theoretical Developments (QMTD)

Anna Heath, Ioanna Manolopoulou and Gianluca Baio, University College London, London, United Kingdom

Background: The Expected Value of Sample Information (EVSI) determines the economic value of any future study with a specific design aimed at reducing uncertainty in a health economic model. This could be used as a tool for trial design, where the cost and value of different designs are compared to choose the trial with the greatest net benefit. However, despite the development of fast calculation methods, EVSI analysis can be slow especially when trying to optimise over a large number of different designs.

Purpose: The aim of this work is to develop an EVSI calculation method that allows researchers to investigate the optimal sample size for their trial with little to no extra computational cost.

Method(s): This has been achieved by extending an EVSI calculation method based on "moment matching" which approximates the EVSI by calculating the posterior variance for a small number of simulated datasets. The proposed extension uses Bayesian non-linear regression to estimate the EVSI based on one simulated dataset for each sample size.

Result(s): A health economic model developed to assess the cost-effectiveness of interventions for chronic pain demonstrates that this EVSI calculation method is fast and accurate for realistic models. This example also highlights how different trial designs can be compared using the EVSI.

Conclusion(s): The proposed estimation method is a quick, accurate method for calculating the EVSI across different sample sizes. This will allow researchers to realise the potential of using the EVSI to determine an economically optimal trial design for reducing uncertainty in health economic models.

2F-4 THE IMPACT OF SELECTION METHODS IN TWO-STAGE BOOTSTRAPPING TO ASSESS UNCERTAINTY IN HEALTH ECONOMIC OUTCOMES IN MULTI-CENTER RANDOMIZED CONTROLLED TRIALS

Applied Health Economics (AHE)

Michelle M.A. Kip, MSc.[1], Sofie Berghuis, MSc[1], Maarten W. N. Nijsten, MD, PhD[2], Prof. Maarten J. IJzerman, PhD[1] and Hendrik Koffijberg, PhD[1], (1)University of Twente, Enschede, Netherlands, (2)University of Groningen, University Medical Center Groningen, Groningen, Netherlands

Purpose:

Bootstrapping is often used to assess uncertainty in outcomes of randomized controlled trials (RCTs) due to sampling variation and limited sample sizes. Although guidance is available on two-stage bootstrapping for RCTs, specific guidance is lacking on sampling within multi-center RCTs to address the uncertainty in variation across centers. This study assesses the impact of using different selection approaches to sample patients in two-stage bootstrapping in a case study on procalcitonin-based antibiotic treatment in IC patients with sepsis.

Method(s):

The case study was a multi-center RCT including 16 hospitals (4 academic, 12 non-academic) with on average 48 patients per hospital in both groups (range n: 1-185). Five sampling approaches were investigated, based on random sampling of: 1) the intended number of patients, 2) 16 hospitals, 3) as method 2 while maintaining the total number of included patients, 4) 16 hospitals while maintaining the original ratio academic/non-academic hospitals, 5) as method 4 while maintaining the total number of patients. Additionally, a scenario analysis using half of the data was performed. Incremental cost differences and corresponding 95% CIs were determined based on 10,000 bootstrap samples.

Result(s):

Different approaches of bootstrapping resulted in variation in the mean incremental costs per patient. For bootstrap 1: €-24 (95% CI €-5,941-€6,088), bootstrap 2: €162 (95% CI €-5,699-6,566), bootstrap 3: €37 (95% CI €-5,745-€6,192), bootstrap 4: €176 (95% CI €-5,022-€6,241), bootstrap 5: € 183 (95% CI €5,124- €5,986). The mean incremental cost of the observed data was €16. Approach 5 deviated most from this observed mean, with a cost difference of €167. The 95% CIs varied in size (smallest 95% CI: €-5,123 - €5,986 [method 5], largest 95% CI: €-5,699 - €6,566 [method 2]). Differences in outcomes were more pronounced in the scenario analysis using half of the data, resulting in a mean incremental cost of €433 (95% CI €-6,980-€9,813).

Conclusion(s):

Using different approaches for sampling centers (i.e. hospitals) and individuals (i.e. patients) in two-stage bootstrapping may influence the mean outcomes and 95% CIs. The most appropriate sampling method based on outcomes and 95% CIs should be determined based on the approach for inclusion used in the real-world trial. When the inclusion strategy of the original study design is unknown, sensitivity analysis is recommended to assess uncertainty arising from this inclusion process.

2F-5 BAYESIAN INFERENCE ABOUT OPTIMAL TREATMENT REGIMENS USING OBSERVATIONAL DATA: A SIMULATION AND AN APPLICATION TO LARYNX CARCINOMA

Quantitative Methods and Theoretical Developments (QMTD)

Thomas Klausch, PhD[1], Peter van de Ven, PhD[1] and Johannes Berkhof[2], (1)VU University Medical Center Amsterdam, Amsterdam, Netherlands, (2)VU University Medical Center, Amsterdam, Netherlands

Purpose:

We present new methodology for learning of optimal treatment regimens from observational data, which allows (a) determining an optimal assignment rule for patients while penalizing sub-optimal decisions, (b) estimating the expected patient and population benefit under this optimal regimen, and (c) quantifying the degree of certainty about the correctness of treatment decisions and benefit estimates.

Method(s):

Using patient characteristics for personalizing treatment decisions is an increasingly important objective in evidence-based medicine and statistics. Based on the Rubin causal model and assuming observed confounding, we apply flexible Bayesian modelling strategies and machine learning techniques to approximate the functional relationship between potential outcomes and patient characteristics. Optimal treatment regimens are then determined using predictions of treatment outcomes and a loss function that penalizes sub-optimal decisions.

We evaluated the performance of a selection of modelling techniques -- Bayesian additive regression trees (BART), polynomial regression, and B-splines -- in nonlinear multivariate settings in a simulation study. We applied the methods to Larynx carcinoma data collected within the European BD2DECIDE consortium (n=419) and determined an optimal rule for assignment of patients to surgery or radiotherapy maximizing 2-year survival.

Result(s):

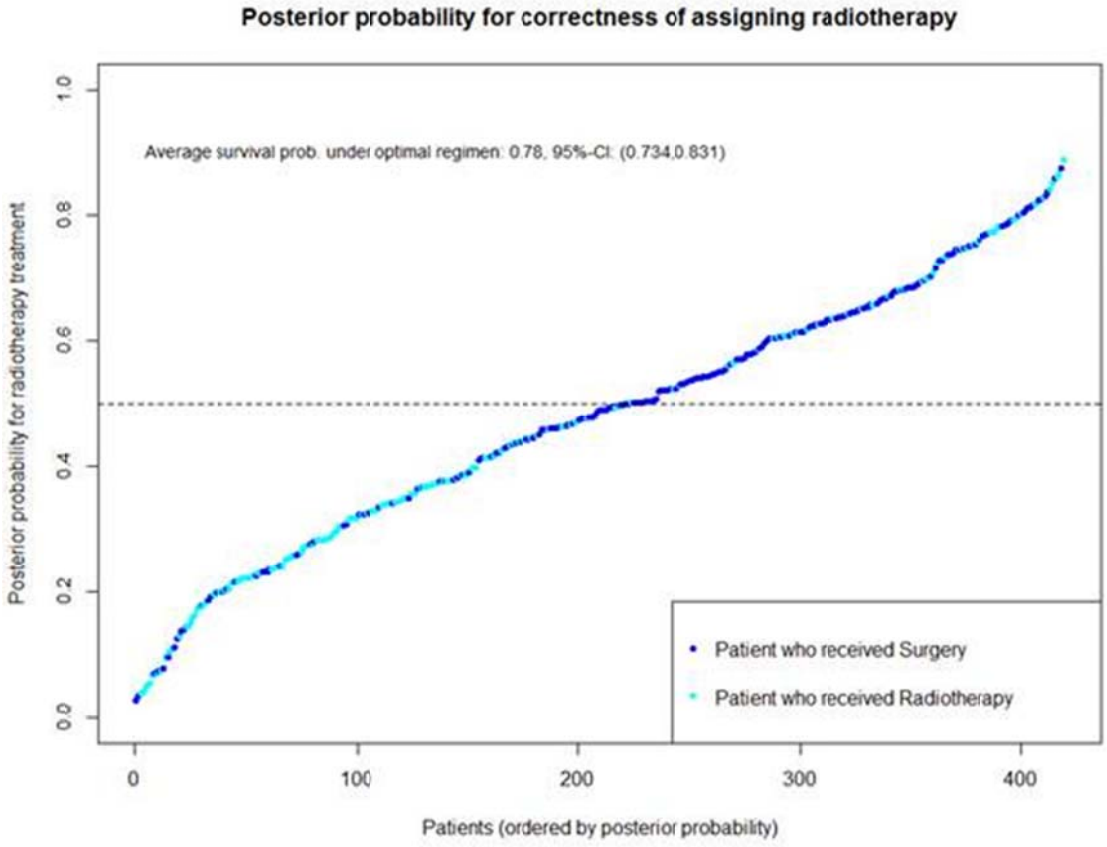
Across a series of simulation scenarios varying in terms of strength of selection bias and number of covariates (n=500), all techniques showed approximately 90-95% frequentist coverage of 95% credible intervals for the true optimal population benefit, small bias (<2%) and variance (sd=2-6%), and mostly (95%) correct treatment assignments.

Larynx 2-year survival probability was estimated at 0.716 which under the optimal assignment increased to 0.780 (95%-CI: 0.734,0.831). The attached figure shows patient-specific posterior probabilities for the correctness of radiotherapy assignment. In 34.8% cases optimal assignment was different from treatment received with proportion of patients receiving radiotherapy similar under optimal and actual assignment (44%).

We additionally considered another loss function giving preference to the less intense surgical treatment for patients with equal outcomes under both treatments. The proportion of patients receiving radiotherapy could be decreased substantially while keeping overall survival similar to that observed (0.716).

Conclusion(s):

We present novel methods that use state-of-the-art modelling techniques which can handle large numbers of predictors and unknown functional forms for estimating optimal treatment regimens. The Bayesian approach facilitates quantifying uncertainty concerning patient benefit with good coverage properties and high proportions of correct treatment assignment.



2F-6 ON THE EQUIVALENCE OF A COHORT MODEL AND THE MEAN EQUATION OF THE ASSOCIATED MASTER EQUATION

Quantitative Methods and Theoretical Developments (QMTD)

Rowan Iskandar, PhD, Center for Evidence Synthesis in Health, Brown University, Providence, RI

Purpose: Following its introduction over three decades ago, cohort model (CM) has been used extensively to model population trajectories overtime in decision modeling. However, the theoretical model underlying CMs has not been properly described. We derive the stochastic process underlying a CM by establishing the equivalence between the average of the process and CM.

Method(s): To show the equivalence between the average of a stochastic process and the CM, we derive the mathematical representation of each model and examine whether they are equal in expectation. First, we conceptualize a multi-state stochastic model that tracks the number of individuals in each state based on the following rules on the transition probabilities: (1) only one transition is allowed in an infinitesimal time-step (dt), and (2) the probability of transitioning from one state-configuration (i.e. a particular realization of the distribution of individuals across states) to another state-configuration depends on instantaneous transition rates and the number of individuals in each state and is linear in dt. The probability of a state-configuration at time t (a joint density on a lattice) is written as the sum of the probabilities of adjacent state-configurations at time t-dt multiplied by the probability of transitioning between two corresponding state-configurations. We then derive the difference equation of this joint density.

Secondly, given a matrix of one-step transition probabilities and an initial distribution of individuals across states, a CM generates the trajectory of a cohort and tracks the state-configuration at each time-step (Dt). In principle, the state-configuration at time t is calculated by multiplying the state-configuration at time t-Dt by the transition probability matrix. We then derive the difference equation of the state-configuration.

Result(s): By taking the limits of dt and Dt to 0, we recover the evolution equation of the probability of state-configuration (i.e., master equation) and the differential equation of the state-configuration (i.e., the continuous-time version of a CM), respectively. The equivalence is immediately established by taking the expectation of the master equation with respect to state-configurations, generating the deterministic version of a master equation, i.e. the differential form of CM.

Conclusion(s): The commonly-used CM represent the average of a continuous-time stochastic process on a multidimensional lattice governed by a master equation. Knowledge of the stochastic process underlying a CM provides a theoretical foundation for the modeling method.

3I-1 HARNESSING AN ITERATIVE SURVEY DESIGN WITH LARGE-SCALE OBSERVATIONAL DATA TO PREDICT INDIVIDUAL DECISION-MAKING BASED ON STATED AND REVEALED PREFERENCES: THE CASE OF BLOOD DONATION

Patient and Stakeholder Preferences and Engagement (PSPE)

Kaat De Corte, Sarah Willis, John Cairns and Richard Grieve, PhD, London School of Hygiene and Tropical Medicine, London, United Kingdom

Purpose:

Stated preference surveys can provide parameter estimates for decision models. However, a major concern is that they may not accurately predict actual decisions, i.e. revealed preferences. Few studies have contrasted stated with revealed preferences elicited from the same individuals. This research aims to illustrate how iterative survey design can be combined with a large observational dataset to minimize the discrepancy between stated and revealed preferences.

Method(s):

As part of the HEMO study (Health Economics MOdeling of blood donation), we undertook stated preference surveys to elicit blood donors' preferences for alternative changes to the blood service. The parameter of interest for decision modeling was the mean number of donations per year given a certain blood service configuration. Minimizing the discrepancy between stated and revealed preferences thus became a primary aim during the survey design process. The iterative design consisted of four stages: first, a round of consultation with the policy-maker (NHS Blood and Transplant) and blood donors, second, a pilot survey (5000 invitees), third, a review of the pilot results with NHSBT and donors to adjust the survey design before finally administering the main survey (100,000 invitees). Donors were invited to the surveys through selection from a large observational database – PULSE – of the routinely collected data for the 1.2 million blood donors in England. This selection process allowed for linkage of the survey responses to PULSE at an individual level. We defined 'discrepancy' as the difference between the mean number of donations per year predicted by the responses to the survey versus the observed annual donation frequency of donation recorded in PULSE. The discrepancy between stated and revealed preferences was calculated following both the pilot and main surveys, overall, and by subgroup for each gender.

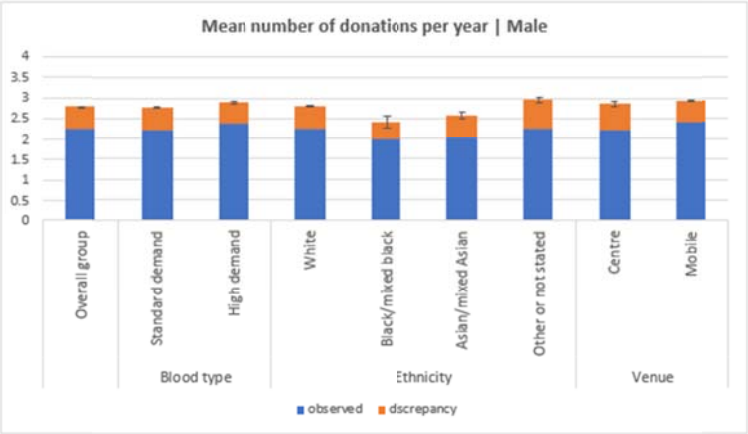
Result(s):

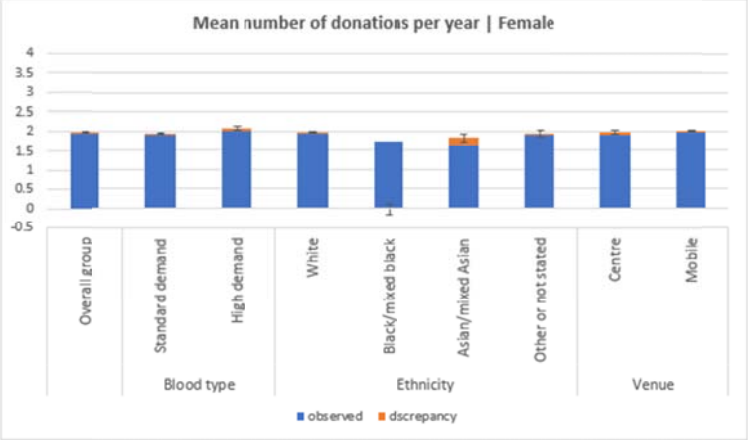
The figures show initial results for the main survey.

Conclusion(s):

This research finds that after combining an iterative survey design with large-scale observational data linked at the individual level, the discrepancy between stated preferences and revealed preferences was small. For each gender, the results were consistent across subgroups. However, the discrepancy is larger for men than women. This approach can be rolled out across other

contexts to help ensure that stated preference surveys are designed to provide accurate predictions for the parameters of interest required for decision-making.





3I-2 MAPPING THE GENERIC EORTC QLQ-C30 AND THE QLQ-H&N35 TO THE EQ-5D USING DATA FROM PATIENTS WITH HEAD AND NECK CANCER

Patient and Stakeholder Preferences and Engagement (PSPE)

Ann-Jean Beck, MD[1], Jacobien Kieffer, PhD[1], Valesca Retel, PhD[1], Lydia van Overveld[2], Robert Takes, MD PhD[2], Michiel van den Brekel, MD PhD[1], Wim van Harten, MD PhD[1] and Martijn Stuiver, PhD[1], (1)Netherlands Cancer Institute, Amsterdam, Netherlands, (2)Radboud University Medical Center, Nijmegen, Netherlands

Purpose:

To develop a mapping function to estimate utilities by translating QLQ-C30 outcomes to EQ-5D utilities through regression modeling, and to explore the value of adding disease-specific QLQ-H&N35 outcomes to this model.

Method(s):

Data of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC QLQ-C30), Head and Neck cancer module (QLQ-H&N35) and EuroQol five-dimensional questionnaire (EQ-5D) (361 observations), was obtained from the Dutch Head and Neck Audit database. Model development was conducted in two phases: 1) selection of possible predictor sets from the QLQ scales and 2) development of an optimal statistical model. Variable selection was done in a hierarchical fashion, using three predictor sets. Set1, which was not subject to selection, contained the QLQ-C30 scales that most closely correspond to EQ-5D-3L dimensions and its underlying construct; Set2 was selected from the remaining scales, based on their reported responsiveness; Set3 was constructed from QLQ-H&N35 scales that correlated >0.3 with the EQ-5D-3L utility score. Using Set1, four statistical approaches were compared: ordinary-least squares (OLS) regression, mixed-effects linear regression, Cox regression and beta regression, for best model fit. Subsequently, Set2 and Set3 were added to the best fitting model. Models were compared on goodness of fit by means of Akaikes Information Criterion, Bayesian Information Criterion and likelihood ratio test. Predictive performance of the final model was evaluated using R squared (R^2), mean absolute error (MAE), root-mean squared error (RMSE) and the Bland-Altman analysis. In addition, we calculated the estimated heuristic shrinkage factor of the final coefficients.

Result(s):

The beta regression model, consisting of the global health status, physical/role/emotional functioning and pain scale, showed best model fit. Adding Set2 and Set3 did not improve the base model. The final model showed reasonable performance of the model on a group level; R^2=0.39, MAE=0.095, RMSE=0.121, 95% limits of agreement were estimated at -0.24 to 0.23 (bias -0.01), with an error correlation of 0.32. The shrinkage factor was 0.90.

Conclusion(s):

Selected scales from the EORTC QLQ-C30 can be used to estimate utilities for head and neck cancer using a beta-regression model. Including QLQ-H&N35 scales does not provide additional value. Further research should assess the robustness and generalizability of the mapping function by validating it in an external cohort of head and neck cancer patients.

3I-3 MAPPING PEMB-QOL TO GENERIC EQ-5D-5L IN PATIENTS WITH PULMONARY EMBOLISM, A CROSS-WALK STUDY

Patient and Stakeholder Preferences and Engagement (PSPE)

Ling-Hsiang Chuang[1], **Sonja Kroep**[1], Pearl Gumbs[2], Jose M. Rodriguez[2], Alexander Cohen[3] and Ben Van Hout, Prof, PhD[4], (1)Pharmerit, Rotterdam, Netherlands, (2)Daiichi Sankyo Europe, Munich, Germany, (3)Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom, (4)University of Sheffield, Sheffield, United Kingdom

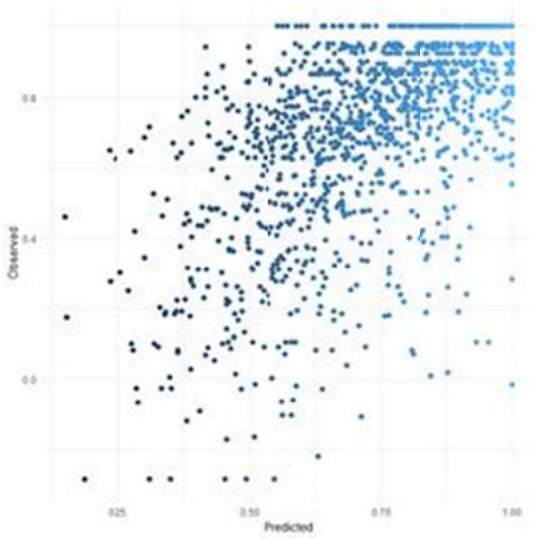
Purpose: Generic preference-based measurements are essential for economic evaluations of health interventions. However, disease-specific instruments are most frequently used, as disease-specific instruments are designed to inform and support clinicians in their decision-making when treating patients. For patients with Pulmonary Embolism (PE), the Pulmonary Embolism Quality of Life (PEmb-QoL) questionnaire is the most frequently used disease-specific instrument. This project aims to develop a mapping algorithm for estimating EuroQol five-dimensional- five-level- (EQ-5D-5L) index scores from the PEmb-QoL questionnaire.

Method(s): PREFER in VTE was a non-interventional disease registry conducted between 2013 and 2014 in primary and secondary care across seven European countries. In the current study, Pulmonary Embolism (PE) were included who completed both the EQ-5D-5L and Pemb-QoL questionnaire at diagnosis. Four modelling approaches (tobit model, beta-regression model, zero-inflated beta regression and the two-part model) were evaluated to estimate EQ-5D-5L index scores with each two different sets of explanatory variables. The simple set including only the six PEmb domain values and gender, and the extended set additionally including significant baseline characteristics as explanatory variables for the EQ-5D-5L index score in each model.

Result(s): Values for both instruments were available for 1334 patients with PE at diagnosis. Models with the extended set of explanatory variables performed best. The approach with the lowest root mean squared error was the beta-regression model (0.213), while the tobit regression model had the lowest mean absolute error (0.157). The tobit regression model was the only approach able to participate on the highly skewed EQ-5D-5L index scores in combination with a high truncation rate (14% of the observations). Evaluating the lower deciles of observed EQ-5D-5L index scores, the two-part model was the only modelling approach capable to estimate values below zero.

Conclusion(s): The modeling approaches used to map disease-specific instrument PEmb-QoL to the generic EQ-5D-5L index scores showed the ability to approach the data taking the highly skewed nature of the data into consideration. The modelling approaches enable the estimation of EQ-5D-5L index scores for economic evaluations when only disease-specific instruments are available.

Figure 1: Scatterplot for observed vs. predicted EQ-5D-5L index scores for the Tobit regression model.



3I-4 THE ROLE OF TRAINING MATERIALS IN HEALTHCARE STATED PREFERENCE STUDIES: IMPROVED RESPONSE EFFICIENCY?

Patient and Stakeholder Preferences and Engagement (PSPE)

Caroline Vass, PhD, Niall Davison, MSc and Katherine Payne, PhD, The University of Manchester, Manchester, United Kingdom

Purpose: To understand if, and how, training materials affect stated preferences elicited using discrete choice experiments (DCE).

Method(s): An online DCE was designed and piloted to elicit public preferences (recruited via an internet panel provider) for a targeted approach to using biologics (algorithm-based 'biologic calculator') compared with conventional 'trial-and-error' prescribing. The DCE comprised five attributes: delay to starting treatment; positive predictive value; negative predictive value; risk of infection; and cost saving to the National Health Service. Respondents were randomised to receive information about rheumatoid arthritis (RA), treatments, conventional prescribing and the biologic calculator as either: (survey-A) text or (survey-B) an animated storyline. The unlabelled DCE was blocked into four surveys. Each survey contained six choice-sets with three alternatives for prescribing: two biologic calculators and a conventional approach (opt-out). The design, generated using Ngene, incorporated a test for monotonicity. Background questions included socio-demographics, and self-reported measures of difficulty and attribute non-attendance (ANA). DCE data were analysed using standard and heteroskedastic conditional logit models (HCLM) allowing the scale parameter to be a function of the individual's characteristics including the type of training materials received.

Result(s): Three-hundred members of the public completed the DCE receiving either survey-A (n=158) or survey-B (n=142). The results of the conditional logit models showed all attributes were statistically significant and in line with *a priori* expectations. Respondents preferred the new targeted approach to prescribing with a negative and significant alternative specific constant for the opt-out. The results also showed those who received the storyline (survey-B) had larger estimated coefficients, suggesting they were more sensitive to the attributes. However, the HCLM showed a statistically significant ($p<0.005$) scale term, indicating those who received the text (survey-A) version were more random in their choices. Further statistical tests, after accounting for differences in scale, suggested no differences in preferences between respondents completing survey-A or survey-B. Respondents who completed the text (survey-A) version had lower rates of self-reported ANA to all attributes, apart from cost. There was no difference in failure of the monotonicity test between the two survey versions.

Conclusion(s): Using engaging (animated) training materials improved observed choice consistency but did not appear to bias preferences. Improved consistency may allow researchers to use a smaller sample size or more choice sets, improving the efficiency of their studies.

3I-5 COMPENDIUM OF METHODS FOR MEASURING PATIENT PREFERENCES IN MEDICAL TREATMENT

Patient and Stakeholder Preferences and Engagement (PSPE)

Vikas Soekhai, MSc, LL.M.[1], Chiara Whichello, MSc, MA[2], Bennett Levitan, MD, PhD[3], Jorien Veldwijk, PhD[4], Tarek Hammad, MD, PhD, MSc, MS, FISPE[5], Ulrik Kihlbom, PhD[6], Eline van Overbeeke, MSc[7], Selena Russo, PhD, MPsych(H), MOrgPsych[8], Ateesha Mohamed, MA[9], Richard Hermann, MD, MPH[10], Isabelle Huys, MSc, PhD[7], Vaishali Patadia, RD, MPH[11], Juhaeri Juhaeri, PhD[11] and Esther de Bekker-Grob, PhD[12], (1)Erasmus University Medical Centre Rotterdam/Erasmus University Rotterdam, Rotterdam, Netherlands, (2)Erasmus University Rotterdam, Rotterdam, Netherlands, (3)Janssen, R&D , Titusville, NJ, (4)Erasmus Choice Modelling Centre, Erasmus University , Rotterdam, Netherlands, (5)Merck & Co, North Wales, PA, (6)Centre for research ethics and bioethics Uppsala University, Uppsala, Sweden, (7)KU Leuven, Leuven, Belgium, (8)European Institute of Oncology, Milan, Italy, (9)Bayer, Whippany, NJ, (10)AstraZeneca, Wilmington, DE, (11)Sanofi, Bridgewater, NJ, (12)Erasmus University Rotterdam, Rotterdam, Netherlands

Purpose: Patient preference studies are taking on an increasingly important role in the medical product lifecycle. While there are numerous industry, academic, regulatory and patient group efforts addressing standards, quality and proper application of preference studies, there is limited understanding of the range of methods to assess preferences and the trade-offs between them. To develop evidence-based recommendations to guide different stakeholders on how and when patient preference studies should be performed, we developed a comprehensive overview of patient preference exploration and elicitation methods.

Method(s): We used a three-step approach to identify existing preference exploration (qualitative) and elicitation (quantitative) methods: 1) listing methods identified in previous preference method reviews; 2) conducting a systematic literature review on 4,572 unique papers identified through multiple scientific databases, using English full-text papers published between 1980 and 2016; and 3) having discussions with international experts (n=14) in the field of health preferences and/or medical decision making to validate the methods found.

Result(s): We identified 32 unique preference methods: 10 exploration and 22 elicitation methods. Consensus was reached among the experts interviewed to cluster exploration methods in three main groups: “Individual techniques”, “Group techniques” and methods that were both “Individual and Group techniques”. Elicitation methods were clustered in four groups: “Discrete Choice Based related techniques”, “Threshold related techniques”, “Rating related techniques” and “Ranking related techniques”.

Conclusion(s): This study identified 32 unique methods for exploring and measuring patient preferences, and reached consensus in clustering the methods. This compendium is a resource for researchers in the patient preference field and also serves as the basis to conduct additional studies that appraise the methods and determine which methods are most appropriate for measuring patient preferences in which phase of the medical product lifecycle to support patient-centric decision making.

3I-6 SHAPING PUBLIC EXPECTATIONS FOR ANTIBIOTICS: A UTILITY-BASED SIGNAL DETECTION APPROACH

Decision Psychology and Shared Decision Making (DEC)

Miroslav Sirota, PhD, Alistair Thorpe, BSc and Marie Juanchich, PHD, University of Essex, Colchester, United Kingdom

Purpose: A utility-based signal detection theory predicts that people inappropriately expect antibiotics because they adopt a liberal criterion of what establishes a signal due to conceptual confusion, and/or due to being oblivious to the costs associated with inappropriate use of antibiotics (i.e., antibiotic resistance). In two pre-registered experiments, we tested this explanation by devising three theory-driven interventions – i) signal-noise conceptual clarification (viral vs. bacterial illness aetiology), (ii) stressing personal costs associated with inappropriate antibiotic use, (iii) stressing the costs associated with inappropriate antibiotic use for others – that we expected to reduce antibiotic expectations and, in turn, requests compared with a baseline condition.

Method(s): Participants were assigned to one of the three interventions or a baseline condition and then read a hypothetical scenario of having a viral ear infection (Exp. 1, $n = 894$) or a cold (Exp. 2, $n = 879$). They expressed their expectations for antibiotics, and requests of antibiotics, using multi-item measures. To measure adherence to prescribed antibiotics, participants expressed their decision to take (or not) a course of antibiotics prescribed by their doctor for bacterial pneumonia (Exp. 1).

Result(s): Overall, the interventions decreased expectations for, and requests of, antibiotics for a viral ear infection (Exp. 1) to a similar extent compared with the baseline condition, $F(3, 890) = 14.88$, $p < .001$; $F(3, 890) = 11.41$, $p < .001$, as well as for a cold (Exp. 2), $F(3, 875) = 4.58$, $p = .003$; $F(3, 875) = 3.45$, $p = .016$. However, in Exp. 2, the intervention stressing the costs associated with inappropriate use of antibiotics for others did not reduce it significantly compared with the baseline condition, $p = .387$; $p = .427$. Importantly, the interventions did not decrease adherence to a prescribed course of antibiotics for bacterial pneumonia compared with the baseline condition, OR = 0.13, 95% CI[0.02, 1.06]; OR = 1.06, 95% CI[0.38, 2.96]; OR = 1.34, 95% CI[0.50, 3.60].

Conclusion(s): A utility-based signal detection theory explains why people expect antibiotics even when they do not need them. While a lot of information campaigns have focused on differentiating viral and bacterial infections, similar attention of public campaigns should be devoted to fleshing out the costs of inappropriate use of antibiotics.

3J-1 FORECASTING INFLUENZA VACCINATION COVERAGE IN ISRAEL: INTEGRATING PERSONAL AND SOCIAL DRIVERS

Decision Psychology and Shared Decision Making (DEC)

Adir Shaham[1], Gabriel Chodick[2] and Dan Yamin[1], (1)Department of Industrial Engineering, Faculty of Engineering, Tel Aviv University, Tel Aviv, Israel, (2)Maccabi Institute Research and Innovation, Tel Aviv, Israel

Purpose: influenza vaccination coverage in most developed countries including Israel is suboptimal, and varies substantially within and between seasons. This reality results in occasions with waste of vaccination doses and occasions with vaccine shortage. We 1) explore the drivers and predictors of an individual to become vaccinated against influenza, and 2) forecast the population-level vaccination coverage throughout the season.

Method(s): we utilized data from electronic medical records of 250,000 members of the second largest health maintenance organization in Israel, Maccabi Healthcare Services, between 2007 and 2017. Random forest and recurrent neural networks (RNN) were used to model the vaccination decision of an individual. An ordinary differential equation model was integrated into the model to consider the "virality" of vaccination behavior, and to capture vaccination coverage throughout the season.

Result(s): we found that not only does the vaccination decision of an individual in the year prior serve as a strong predictor for future uptake (RR: 9.7, 95% CI: 9.64-9.73), but also the time in the season when the individual got vaccinated compared to the rest of the vaccinated individuals (r = 0.98, p-value < 0.001). Furthermore, individuals who did not get vaccinated and were infected with influenza-like-illness in a given season, were 1.464 times more likely to become vaccinated in the subsequent season, compared to those who did not get vaccinated and were not infected (CI: 1.46-1.469). An analysis of the information gained from patients' historical behavior showed a diminishing marginal contribution that was negligible when considering more than three seasons. Our model yielded a 78% sensitivity and a 91% specificity.

Conclusion(s): our study provides a data-driven perspective on personal and subjective elements that are considered to explain the decision making of an individual in the case of influenza vaccination. At the individual level, our findings can help to better shape targeted seasonal vaccination campaigns, based on the individuals' main drivers. At the population level, our approach can lead to a more accurate prediction of vaccination coverage, thereby leading to more efficient resource allocation of the HMO.

3J-2 MULTIVARIABLE OUTCOME PREDICTION AFTER ENDOVASCULAR TREATMENT FOR ACUTE ISCHEMIC STROKE: DEVELOPMENT OF A PROGNOSTIC MODEL IN DATA FROM SEVEN RANDOMIZED CONTROLLED TRIALS

Decision Psychology and Shared Decision Making (DEC)

Vicky Chalos, MD[1], Esmee Venema, MD[2], Maxim J.H.L. Mulder, MD, PhD[1], Bob Roozenbeek, MD, PhD[3], Scott Brown, PhD[4], Andrew M. Demchuk, MD, PhD[5], Keith W. Muir, MD, PhD[6], Antoni Davalos, MD, PhD[7], Peter J. Mitchell, MD, PhD[8], Serge Bracard, MD, PhD[9], Michael Hill, MD, PhD[5], Philip White, MD, PhD[10], Bruce C.V. Campbell, MD, PhD[11], Francis Guillemin, MD, PhD[12], Jeffrey L. Saver, MD, PhD[13], Tudor G. Jovin, MD, PhD[14], Mayank Goyal, MD, PhD[5], Ewout W. Steyerberg, PhD[15], Aad van der Lugt, MD, PhD[16], Diederik W.J. Dippel, MD, PhD[1] and Hester F. Lingsma, PhD[2], (1)Department of Neurology, Erasmus University Medical Center, Rotterdam, Netherlands, (2)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (3)Department of Radiology and Nuclear Medicine, Erasmus University Medical Center, Rotterdam, Netherlands, (4)Altair Biostatistics, St Louis Park, MN, (5)Departments of Clinical Neuroscience and Radiology, Hotchkiss Brain Institute, Cummings School of Medicine, University of Calgary, Calgary, AB, Canada, (6)Institute of Neuroscience & Psychology, University of Glasgow, Queen Elizabeth University Hospital, Glasgow, United Kingdom, (7)Department of Neuroscience, Hospital Germans Trias y Pujol, Barcelona, Spain, (8)Department of Radiology, Royal Melbourne Hospital, University of Melbourne, Melbourne, Australia, (9)Department of Diagnostic and Interventional Neuroradiology, University of Lorraine and University Hospital of Nancy, Nancy, France, (10)Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, United Kingdom, (11)Department of Medicine and Neurology, Melbourne Brain Centre Royal Melbourne Hospital, University of Melbourne, Melbourne, Australia, (12)Department of Clinical Epidemiology, Inserm, University of Lorraine and University Hospital, Nancy, France, (13)Department of Neurology and Comprehensive Stroke Center, David Geffen School of Medicine, University of Los Angeles, Los Angeles, CA, (14)Stroke Institute, Department of Neurology, University of Pittsburgh Medical Center Stroke Institute, Presbyterian University Hospital, Pittsburgh, PA, (15)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands; Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, Netherlands, (16)Deptartment of Radiology, Erasmus University Medical Center, Rotterdam, Netherlands

Purpose: Even when the revascularization and clinical status of a patient after endovascular treatment (EVT) for acute ischemic stroke is known, outcome is still highly variable and difficult to predict. We aimed to develop a prognostic model that can be applied within one day after EVT to predict functional outcome at three months.

Method(s): We used data from patients in the treatment arms of seven randomized controlled trials within the HERMES collaboration (MR CLEAN, ESCAPE, REVASCAT, SWIFT-PRIME, EXTEND-IA, THRACE, and PISTE). Primary outcome was the ordinal modified Rankin Scale (mRS) score three months after EVT. Fifteen pre- and post-procedural variables, assessed within one day after EVT, were analyzed with univariable ordinal logistic regression analysis and multivariable ordinal logistic regression analysis with stepwise backward selection ($p<0.157$),

and a fixed effect for trial. From this model, we derived predicted probabilities of functional independence (mRS 0-2) and survival (mRS 0-5). Model performance was quantified with the c-statistic. Internal validation with bootstrapping was performed to estimate the degree of optimism in the final model. To correct for this optimism we reduced the regression coefficients by using penalized regression.

Result(s): The final model, based on 781 patients, included nine variables and explained 62.4% of the variance in outcome. Pre-procedural variables included age, diabetes mellitus, pre-stroke mRS, collateral vessel grade, occlusion location, and time from stroke onset to groin puncture. Post-procedural variables included revascularization grade, stroke severity measured with the National Institutes of Health Stroke Scale (NIHSS), and symptomatic intracranial hemorrhage. The NIHSS was the strongest predictor with 53.9% explained variance. The internally validated c-statistic was 0.83 for the prediction of the ordinal mRS, 0.89 for functional independence, and 0.80 for survival, indicating good model performance.

Conclusion(s): This model, which can be applied within one day after EVT, accurately predicts functional outcome at three months. It may provide physicians, patients, and family members with improved outcome expectations and improve decision making by personalizing the patients' treatment and rehabilitation plan.

3J-3 SCORE RISK FUNCTION TO PREDICT CARDIOVASCULAR DISEASE MORTALITY IN A POPULATION OUTSIDE OF EUROPE: PERFORMANCE AND CHALLENGES

Quantitative Methods and Theoretical Developments (QMTD)

Noushin Fahimfar, MD, PhD candidate of Epidemiology[1], Akbar Fotouhi, MD, PhD[1], Mohammad Ali Mansournia, MD, PhD[1], Reza Malekzadeh, MD[2], Fereidoun Azizi, MD[3], Marjan Mansourian, PhD[4], Sadaf G.Sepanlou, MD, PhD[2], Hamidreza Roohafza, MD[4], Farzad Hadaegh, MD[5], Mohammad Hassan Emamian, MD, PhD[6], Hossein Poustchi, MD, PhD[2], Akram Pourshams, MD[2], Ewout W. Steyerberg, PhD[7] and Davood Khalili, MD PhD[5], (1)Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran (Islamic Republic of), (2)Digestive Diseases Research Center, Digestive Diseases Research Institute, Tehran University of Medical Sciences, Tehran, Iran (Islamic Republic of), (3)Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran (Islamic Republic of), (4)Isfahan Cardiovascular Research Center, Isfahan Cardiovascular Research Institute, Isfahan University of Medical Sciences, Isfahan, Iran (Islamic Republic of), (5)Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran (Islamic Republic of), (6)Ophthalmic Epidemiology Research Center, Shahroud University of Medical Sciences, Shahroud, Iran (Islamic Republic of), (7)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands; Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, Netherlands

Purpose: To use a model in a new community, model validation is among the main objectives of prevention programs. We aimed to assess the external validity of the SCORE-high function and recalibrate it to predict cardiovascular (CVD) mortality in four Iranian cohort studies.

Method(s): We included 18189 individuals (8165 men) aged 45 to 65 year. We assessed the performance of original model and recalibrated the baseline survival and the slope of linear predictor. Model discrimination was assessed using C-index. To compare the predicted probability to the observed risk, calibration plot was depicted. Sensitivity and specificity were calculated at different risk thresholds, following the instruction of SCORE paper and considering the multiplication of predicted risk by two and four in diabetic men and women, respectively. Net Benefit Fraction (NBF) was calculated and decision curve was plotted to show the clinical usefulness.

Result(s): After 109872 person-year of follow-up, 259 cardiovascular deaths (158 men) occurred. The 10-year observed risk were 0.036 (95%CI: 0.030-0.043) in men and 0.018 (0.015-0.023) in women. The original model showed underestimation in women; the average predicted risks in original model were 0.036 and 0.010 in men and women, respectively. Calibration slope for CHD-CVD outcome were 0.56 in men and 0.38 in women. The corresponding values for the non CHD-CVD event were 1.01 and 1.33. The C-index of 0.73 for men and 0.71 for women were detected. At the threshold of 5%, we detected a sensitivity of 59% and 41% in men and women, respectively. The recalibrated model showed good clinical usefulness in both genders. NBFs were 0.26 in men and 0.13 in women, showing the fraction of incidence rate that could be

prevented properly regarding the usefulness of treatment for true-positives penalized by negative weight for harms of treatment in false-positives.

Conclusion(s): We showed good discrimination of the SCORE model in a large population outside of Europe. In total, recalibrated model showed good performance to predict CVD mortality. There were some technical challenges in the recalibration: Firstly, CVD outcome is contributed by two components, CHD-CVD and non CHD-CVD, which may have different calibration slopes. Secondly, diabetes is not included in the model and may cause challenges according to the proportion of diabetic patients.

3J-4 DEVELOPMENT OF A PERSONALIZED MODEL FOR PREHOSPITAL DECISION-MAKING IN ACUTE ISCHEMIC STROKE

Health Services, Outcomes and Policy Research (HSOP)

Esmee Venema, MD[1], Hester F. Lingsma, PhD[1], Vicky Chalos, MD[2], Maxim J.H.L. Mulder, MD, PhD[3], Maarten Lahr, PhD[4], Aad van der Lugt, MD, PhD[5], Adriaan C.G.M. van Es, MD, PhD[6], Ewout W. Steyerberg, PhD[7], M.G. Myriam Hunink, MD, PhD[8], Diederik W.J. Dippel, MD, PhD[3] and Bob Roozenbeek, MD, PhD[2], (1)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands, (2)Department of Radiology and Nuclear Medicine, Erasmus University Medical Center, Rotterdam, Netherlands, (3)Department of Neurology, Erasmus University Medical Center, Rotterdam, Netherlands, (4)Department of Epidemiology, University Medical Centre Groningen, Groningen, Netherlands, (5)Deptartment of Radiology, Erasmus University Medical Center, Rotterdam, Netherlands, (6)Department of Radiology, Erasmus University Medical Center, Rotterdam, Netherlands, (7)Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands; Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, Netherlands, (8)Erasmus University Medical Center, Rotterdam, Netherlands

Purpose:

Prehospital triage of suspected stroke patients requires a trade-off between avoiding delay of intravenous thrombolysis (IVT) and reducing delay of endovascular therapy (EVT) by direct transportation of patients at high risk of intracranial large vessel occlusion (LVO) to an intervention center. Our aim was to determine the optimal transportation strategy based on individual patient characteristics and expected treatment delays.

Method(s):

We constructed a decision-analytic model to estimate quality-adjusted life years (QALYs) for stroke patients after transportation to the nearest primary stroke center versus an intervention center further away. The base case concerned a 68-year old man with suspected stroke symptoms since 1 hour, who lives 25 minutes away from a primary stroke center and 75 minutes away from an intervention center. Risk of LVO was assessed with the prehospital Rapid Arterial oCclusion Evaluation (RACE) scale. Treatment delays were based on estimated transportation times and in-hospital treatment times. Decisional uncertainty was assessed with probabilistic sensitivity analysis using 10,000 Monte Carlo simulation runs.

Result(s):

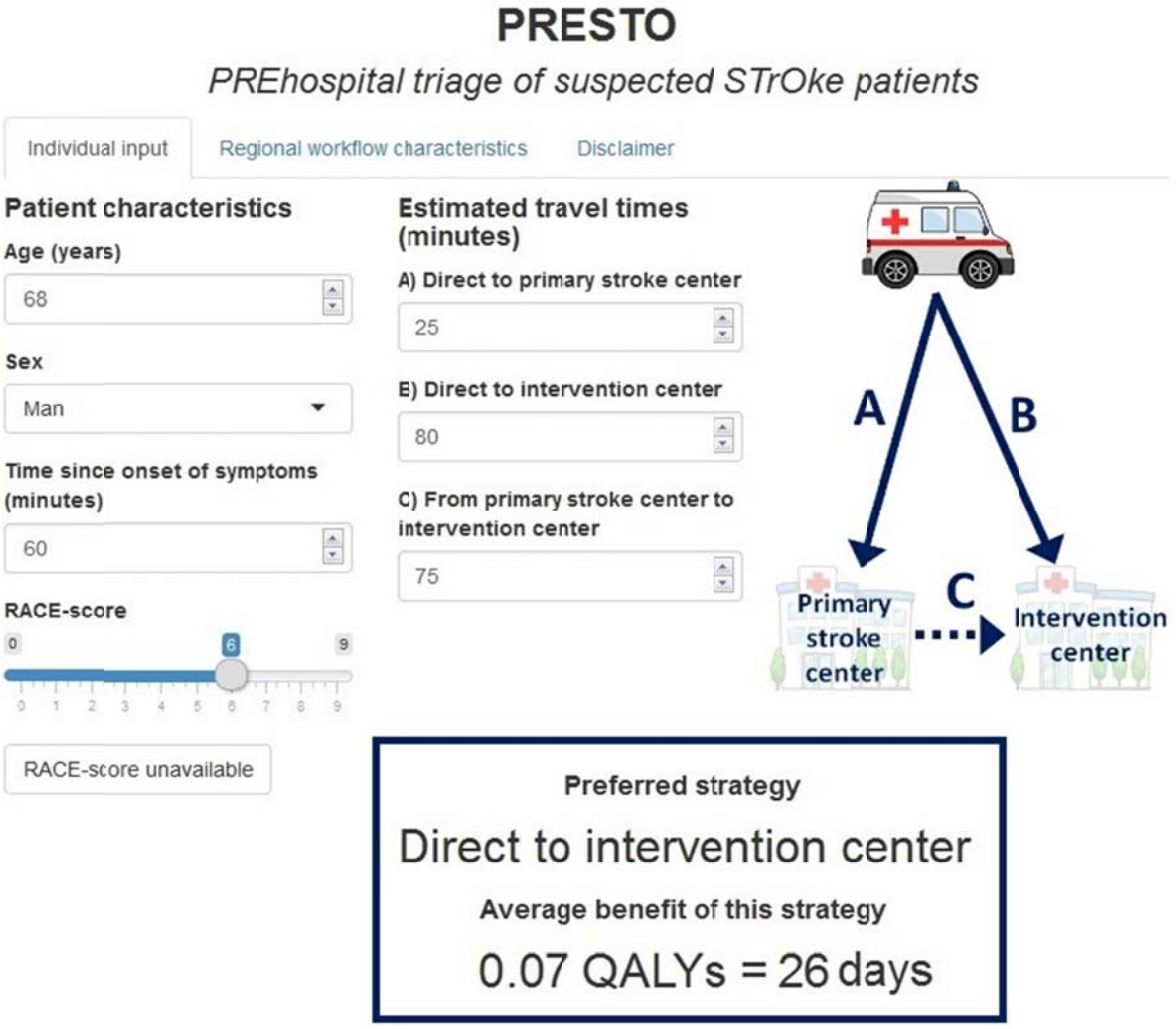
Direct transportation to an intervention center led to better outcomes in the base case analysis when the risk of LVO exceeded 43%. This threshold was lower in scenarios with increased transportation times to both centers. The preferred strategy was furthermore strongly influenced by local door-in-door-out time in the primary stroke center -in case of a transfer to an intervention center- and door-to-EVT time in the intervention center . The benefit of the previously suggested strategy of transporting all patients with a RACE score of 5 or more to the intervention center was 0.09 QALYs per patient (=33 days, interquartile range (IQR) -0.05–0.12)

compared with transportation of these patients to the nearest center. The effect of changes in the different input parameters are illustrated in an interactive table (Figure 1).

Conclusion(s):

Triage of suspected stroke patients depends strongly on the individual risk of LVO and local travel and workflow times. Direct transportation to an intervention center might be beneficial for patients with a high risk of LVO, defined as a RACE score of 5 or more, especially in rural areas with large between-hospital distances. Interactive tables facilitate the presentation of decision models and allows researchers or clinicians to vary input parameters to fit their own setting.

Figure 1. Screenshot of the interactive table



3J-5 DYNAMIC PREDICTION FOR PATIENTS WITH HIGH-GRADE EXTREMITY SOFT TISSUE SARCOMA

Decision Psychology and Shared Decision Making (DEC)

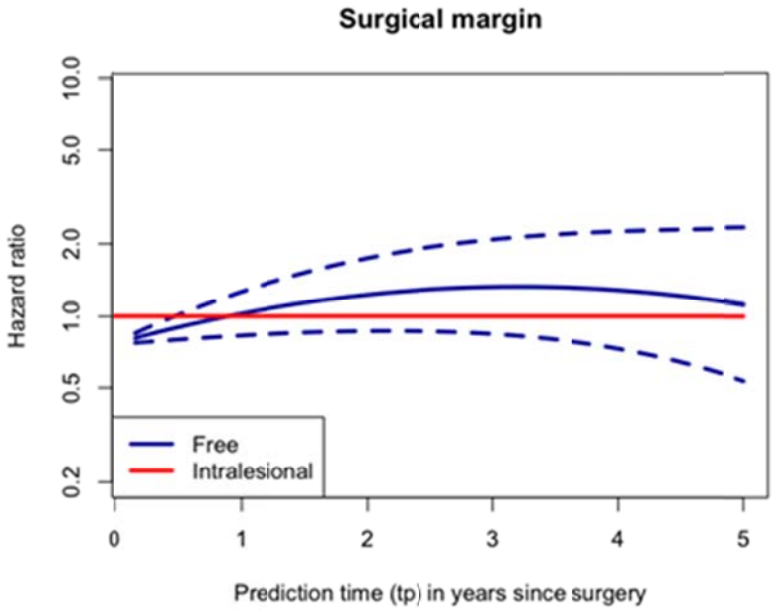
Anja Juana Rueten-Budde, Leiden University, Leiden, Netherlands and Marta Fiocco, Leiden University and LUMC, Leiden, Netherlands

Purpose: Increasing interest lies in personalised prediction of disease progression for soft tissue sarcoma patients. Currently, available prediction models are limited to predictions from time of diagnosis or surgery. However, updated patient information during follow-up may change a patient's prognosis, which is not accounted for in these models. The concept of dynamic prediction allows to include updated information as well as model time-varying covariate effects to make prediction of overall survival at different times during follow-up.

Method(s): Information from 2232 patients with high-grade extremity soft tissue sarcoma who underwent surgery at 14 specialised sarcoma centres, was used to develop a dynamic prediction model with primary endpoint overall survival. To estimate a patient's probability of surviving an additional 5 years from a particular prediction time point a proportional landmark supermodel was used. Landmark models are able to make predictions from a particular time, by using all (updated) information of patients still alive and in follow-up at that time. Several patient and tumour-specific risk factors were included into the model as well as time-dependent covariates, such as a patient's status of local recurrence and distant metastases. Additionally, time-varying effects of covariates on the outcome was investigated.

Result(s): Results show that surgical margin and tumour histology have a time-varying effect on overall survival. The effect of margin is strongest shortly after surgery and fades slightly over time (see Figure). The occurrence of local recurrence and distant metastasis during follow-up have a strong effect on overall survival and they must be accounted for to make updated predictions.

Conclusion(s): The presence of time-varying effects for some prognostic factors as well as the effect of the time-dependent variables local recurrence and distant metastasis on survival suggest the inadequacy of baseline models for predictions during follow-up. A model designed for dynamic prediction, which updates survival probabilities is developed. Such model can be used for better individualized treatment depending on a dynamic assessment of the prognosis of the patient. To the best of authors' knowledge this is the first dynamic prediction model in this field. The model will be made freely available through the Personalised Sarcoma Care (PERSARC) mobile application – *after surgery part two*.


Surgical margin
Hazard ratio
10.0
5.0
2.0
1.0
0.5
0.2
Free
Intralesional
0
1
2
3
4
5
Prediction time (tp) in years since surgery

3J-6 VALIDATION OF MICROSIMULATION MODEL PREDICTIONS AGAINST LONG-TERM COLORECTAL CANCER INCIDENCE AND MORTALITY OUTCOMES FROM RANDOMISED-CONTROLLED TRIALS OF COLORECTAL CANCER SCREENING USING GFOBT

Quantitative Methods and Theoretical Developments (QMTD)

Jie-Bin Lew[1], **Marjolein JE Greuter**[2], Michael Caruana[1], Emily He[1], Joachim Worthington[1], Eleonora Feletto[1], D. James St. John[3], Finlay A Macrae[3], Paul Grogan[4], Veerle MH Coupe[2] and Karen Canfell[1], (1)Cancer Council NSW, Sydney, Australia, (2)VU University Medical Center, Amsterdam, Netherlands, (3)The Royal Melbourne Hospital, The University of Melbourne, Melbourne, Australia, (4)Cancer Council Australia, Sydney, Australia

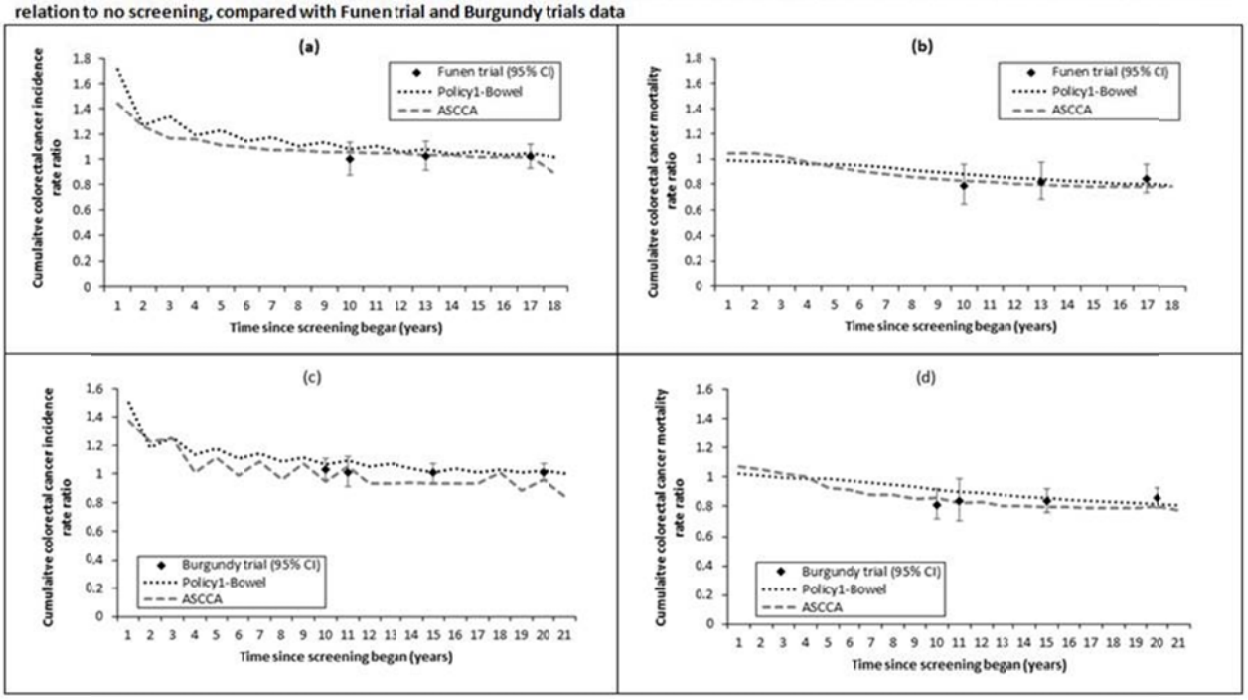
Purpose: Predictive models of colorectal cancer (CRC) natural history and screening inform screening policy by predicting health outcomes and cost-effectiveness of alternate strategies. Confidence in these tools is greatly increased when they can be successfully validated against major trials reporting long term cancer and mortality outcomes. Two comprehensive microsimulation models*, Policy1-Bowel* and *ASCCA*, which simulate both the adenoma-carcinoma pathway and the serrated pathway in CRC development and CRC screening have been developed. The models have been previously used to evaluate CRC screening in Australia and the Netherlands. This study aimed to compare each model's predictions of long-term CRC incidence and mortality reductions in screened vs. unscreened arms, to the reported data from randomised-controlled trials (RCTs) evaluating gFOBT screening.

Method(s): A systematic review of CRC screening RCTs was used as the basis for this study. For each RCT, we explicitly modelled the age distribution of participants, the timing and number of screening rounds, age- and sex-specific screening participation and compliance to diagnostic test referral. The model-estimated relative reduction in long-term CRC incidence (RR_{inc}) and mortality (RR_{mort}) in screened vs. unscreened were then compared with observational data. Example results for the Funen trial (9 biennial screening rounds) and the Burgundy trial (6 biennial screening rounds) are presented here. Data of the RCTs modelled in this study were not used to inform the development of the colorectal cancer natural history component of *Policy1-Bowel* and *ASCCA* model.

Result(s): Both models reproduced the observed relative reductions in incidence and mortality at different follow-up times (Figure 1). The Funen trial reported RR_{inc} 1.02 (95%CI: 0.93-1.12) at 17 years after screening began; model estimates of RR_{inc} at 17 years were 1.04 (*Policy1-Bowel*) and 1.01 (*ASCCA)*. Reported RR_{mort} was 0.84 (95%CI: 0.73-0.96); model estimates were 0.80 and 0.78, respectively. The Burgundy trial reported RR_{inc} at 20 years of 1.01 (95%CI :0.96-1.07); model estimates of RR_{inc} were 1.02 and 0.96 respectively. Reported RR_{mort} was 0.86 (95%CI: 0.79-0.93); model estimates were 0.82 and 0.80, respectively.

Conclusion(s): Carefully calibrated microsimulation models can successfully reproduce observed clinical trial outcomes up to 20 years of follow-up. These findings increase confidence in the predictive capacity of the two model platforms which will continue to be harnessed to consider a range of important policy questions in future.

Figure 1 Estimated relative reductions (a) (c) in cumulative colorectal cancer incidence rates and (b) (d) in cumulative colorectal cancer mortality rates in relation to no screening, compared with Funen trial and Burgundy trials data



4M-1 PERCEIVED AWARENESS OF CHOICE: A PILOT STUDY

Decision Psychology and Shared Decision Making (DEC)

Marleen Kunneman, PhD[1], Corrie A.M. Marijnen, MD PhD[2], Kim Brandes[3], Anne M. Stiggelbout, PhD[3] and **Arwen H. Pieterse, PhD**[3], (1)Knowledge and Evaluation Research (KER) Unit, Mayo Clinic; and Medical Decision Making, Dept of Biomedical Data Sciences, Leiden University Medical Center, Rochester, MN, (2)Dept of Radiotherapy, Leiden University Medical Center, Leiden, Netherlands, (3)Medical Decision Making, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands

Purpose:

Fostering choice awareness (acknowledging that there is more than one way forward, a decision needs to be made and the patient's views matter) may be a prerequisite for successful shared decision making (SDM). The aim of this study was to assess whether clinicians' behavior to foster choice awareness leads to more awareness of choice of patients.

Method(s):

We extracted fragments from previously recorded encounters between radiation oncologists and patients with rectal cancer facing a preference-sensitive decision about short-course radiotherapy treatment. To protect privacy, the fragments were re-recorded, with comparable pace and tone. In an online experiment, participants were randomly offered one of two fragments in which the radiation oncologist indicated that the reason for the encounter was: 1) to explain the treatment, or 2) to decide about treatment. Participants were asked to rephrase the reason for the encounter in their own words, and additionally to indicate on a six-point Likert-type scale the extent to which the decision to receive radiotherapy seemed already made (1 "Not definitive at all", 6 "As definitive as can be"). We qualitatively analyzed participants' written responses and used Chi-square tests to assess the differences in mean scores on the Likert-scale.

Result(s):

In total, n=28 lay people, n=6 treated rectal cancer patients, and n=10 oncologists participated in the study. In fragment 1, participants seemed to agree that the reason for encounter was 'providing patient information about the next steps', 'convincing the patient that he needs to undergo radiotherapy treatment', or 'informing the patient why he needs radiotherapy'. None of the participants used phrases that implied choice. In fragment 2, we found more variation in responses, including examples mentioned above, but also 'deciding about the treatment', 'making a choice about radiotherapy or not', and 'patient and clinician can decide together what to do'.

Participants allocated to fragment 2 scored significantly lower on the extent to which the treatment decision had already been made compared to those allocated to fragment 1 ($M = 4.3$ vs $M = 5.3$, $\chi^2=10.72$, $p = 0.03$).

Conclusion(s):

Our results suggest that phrases used by oncologists to indicate the reason for the encounter can influence awareness of choice. We are currently performing a vignette-study to confirm and elaborate on these results.

4M-2 DO WOMEN BEING TREATED FOR BREAST CANCER TALK WITH THEIR ONCOLOGIST ABOUT THE POORLY CONTROLLED SYMPTOMS THEY EXPERIENCED IN THE WEEK PRIOR TO THEIR CLINIC VISIT?

Decision Psychology and Shared Decision Making (DEC)

Kathi Mooney, PhD, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT and Meagan Whisenant, PhD, University of Texas MD Anderson Cancer Center, Houston, TX

Purpose:

The purpose of this study was to determine the extent to which women being treated for breast cancer discussed with their oncology team the symptoms that bothered them during the week prior to a clinic visit and whether the time spent in patient-provider symptom talk changed over subsequent visits.

Method(s):

Utilizing an observational design, women reported the severity of 11 common symptoms (0-10 scale) on a daily basis during the interim between clinic visits using an automated telephone system. These reports were not shared with the oncology team. Clinic visits were audio-recorded then transcribed and symptom encounters were coded using a pre-defined codebook. Utilizing descriptive statistics, symptoms reported at moderate-to-severe levels in the week prior to the visit were compared to the symptoms actually discussed at the visit and the length and focus of symptom discussions were compared across serial visits.

Result(s):

Twenty-six clinic visits of 10 women receiving chemotherapy for breast cancer were recorded. Participants mean age was 51.6 years; half had stage II disease and half had stage III or IV disease. In the week prior to their visits, participants provided 183 reports of moderate-to-severe level symptoms. Most common were fatigue, disturbed sleep, and pain. Reported symptoms were only discussed at 49.5% of visits and only 36% of symptom discussions were initiated by the patient. Symptom discussions were more likely to occur with younger women (60% of visits for women 40-49, 52.3% of visits for women 50-59, and 18.8% of visits for women 60 or older). The discussions averaged 27.8% of visit time across all visits but, despite continued moderate-to-severe symptom reports, symptom talk decreased over subsequent visits (37.8% baseline visit, 29.5% visit 2 and 3, 19.5% visit 4).

Conclusion(s):

Optimal symptom management requires effective communication and engagement between patient and provider, yet only half of the time moderate-to-severe symptoms are discussed. Most discussions are provider initiated, thus women are less likely to self-advocate for improved symptom care, particularly older women. While symptoms continue over chemotherapy cycles, their discussion and attempts to improve their management decreases, impacting quality of life and functioning. Strategies to overcome this barrier, such as routine use of patient-reported outcomes (PROs) at visits and guideline-based symptom decision support systems are needed to add value and facilitate personalized care.

4M-3 IMPROVING THE DECISION QUALITY FOR MEN WITH LOWER URINARY TRACT SYMPTOMS DUE TO BENIGN PROSTATIC HYPERPLASIA BY USING A DECISION AID

Decision Psychology and Shared Decision Making (DEC)

Isabel de Angst, MD[1], Fieke van der Wijden, MD[1], Maarten Cuypers, MSc[2], Marieke de Vries, PhD[3], Romy R.E.D. Lamers, MD[1], Harm van Melick, MD, PhD[4], Jeltje de Beij, MD[5], Dennis Oerlemans, MD, PhD[6], Cees van de Beek, MD[7] and Paul J.M. Kil, MD, PhD[1], (1)Elisabeth-Tweesteden Hospital, Tilburg, Netherlands, (2)Radboud University Medical Center, Nijmegen, Netherlands, (3)Radboud University, Nijmegen, Netherlands, (4)St. Antonius Hospital, Nieuwegein, Netherlands, (5)Zuyderland Medical Center, Heerlen, Netherlands, (6)Laurentius Hospital, Roermond, Netherlands, (7)Maastricht University Medical Center, Maastricht, Netherlands

Purpose: For men with lower urinary tract symptoms caused by benign prostatic hyperplasia (LUTS/BPH), guidelines recommend integrating patient preferences in treatment decisions. Therefore, a newly developed web-based decision aid (DA) was used to provide patients with information about treatments and to help patients clarify their preferences using values clarification exercises (VCE). We evaluated the effectiveness of the DA on improving well-informed and value congruent decision-making in patients with LUTS/BPH.

Method(s): Between July 2016 and January 2017 all new patients who consulted the urologist because of LUTS/BPH, were invited to use the DA and to participate in this questionnaire study. Results were compared with a historical control group of patients who had consulted the urologist between December 2015 and February 2016. The primary outcome 'well-informed choice' was measured using a knowledge questionnaire and 'value congruence' was measured by the relationship between responses on value statements and received treatment. Secondary outcomes were decisional conflict, shared decision, decisional regret, and treatment choice.

Result(s): A total of 109 DA users and 108 control patients were included. There were no significant differences in baseline characteristics between both groups, except for age (DA: 68.4 vs. control: 71.5 years, $p=0.003$) and higher education levels in the DA group ($p=0.047$). DA users made a 'well-informed and value congruent choice' more often than control patients (43% vs 21% $p=0.012$). Decision conflict was lower in the DA group (33.2 vs. 46.6, $p<0.001$), in particular the informed (37.3 versus 57.9, $p<0.001$) and value clarity (36.9 versus 58.2, $p<0.001$) subscale. DA users had a more active role in decision-making ($p=0.019$) and reported less process regret (2.4 versus 2.8, $p=0.003$). Furthermore, DA users who did not use prior medication chose conservative treatments more often than control patients ($p=0.001$).

Conclusion(s): A LUTS/BPH DA improves the decision quality by supporting patients in making well-informed and value congruent treatment decisions, and its use is feasible in the clinical setting. Patients with LUTS/BPH who use the DA feel better informed and are clearer about their values when making treatment decisions. Furthermore, they take a more active role in the decision-making, have less regret about the process leading to the decision, and choose more conservative treatments if they do not use prior medication.

4M-4 IMPLEMENTATION OF AN INFORMATION PORTAL FOR PATIENTS WITH CONGENITAL AORTIC AND PULMONARY VALVE DISEASE: A STEPPED-WEDGE CLUSTER RANDOMIZED TRIAL

Decision Psychology and Shared Decision Making (DEC)

Jonathan R.G. Etnel[1], Lidia Bons[2], Frederiek De Heer[3], Daniëlle Robbers-Visser[3], Ingrid M. Van Beynum[2], Bart Straver[3], Monique R.M. Jongbloed[4], Philippine Kies[4], Martijn G. Slieker[5], Arie P.J. Van Dijk[5], Jolanda Kluin[3], Robin A. Bertels[4], Elisabeth M.W.J. Utens[2], Regina The, Ir.[6], Eugene Van Galen[7], Barbara J.M. Mulder[3], Nico A. Blom[4], Mark G. Hazekamp[4], Jolien W. Roos-Hesselink[2], Willem A. Helbing[2], Ad J.J.C. Bogers[1] and Johanna J.M. Takkenberg[1], (1)Department of Cardiothoracic Surgery, Erasmus University Medical Center, Rotterdam, Netherlands, (2)Erasmus University Medical Center, Rotterdam, Netherlands, (3)Academic Medical Center, Amsterdam, Netherlands, (4)Leiden University Medical Center, Leiden, Netherlands, (5)Radboud University Medical Center, Nijmegen, Netherlands, (6)ZorgKeuzeLab, Delft, Netherlands, (7)Patient Association 'Patiëntenvereniging Aangeboren Hartafwijkingen', Maarssen, Netherlands

Purpose: In response to an increased need for patient information in congenital heart disease, we previously developed an online, evidence-based information portal for patients with congenital aortic and pulmonary valve disease. To assess the effectiveness of this information portal, we conducted a stepped-wedge cluster randomized trial.

Method(s): Adult patients and caregivers of pediatric patients with congenital aortic and/or pulmonary valve disease and/or Tetralogy of Fallot who visited the outpatient clinic at any time during follow-up (eg. at diagnosis, perioperative or routine check-up) at one of the four participating centers between 1/2/2016 and 31/6/2017 were prospectively included. The intervention was implemented according to a stepped-wedge randomized design (control group: standard care; intervention group: standard care + information portal). One month after their outpatient clinic visit, each subject completed a questionnaire on disease-specific knowledge, anxiety and depression, mental quality of life, involvement and autonomy and views on patient information and decision-making. All analyses were adjusted for center and time effects.

Result(s): 327 subjects were included (213 control, 114 intervention). Respondent diagnosis (p=0.968), education level (p=0.099) and gender (p=0.923) were comparable between the two groups. Outcomes are listed in Table 1 and were comparable between groups in the intention-to-treat analyses. However, only 54.4% of subjects in the intervention group (n=62) reported actually visiting the portal. In these subjects (as-treated), disease-specific knowledge was significantly better (p=0.019) than in control subjects, while education level and other baseline characteristics and outcomes were comparable. In the effect of the implementation of the portal on knowledge, there were significant interactions with diagnosis (p=0.019) and age group (parents of children vs. adults, p=0.013), but not with education level (p=0.665) or gender (p=0.187).

Conclusion(s): At one month after outpatient clinic visit, implementation of the information portal did not lead to improved outcomes overall. However, only approximately half of the subjects in the intervention group actually visited the portal and knowledge was significantly

better in these subjects. This demonstrates the potential effectiveness of an online evidence-based patient information portal in improving knowledge in patients with congenital heart disease. Active use of the portal and careful integration in the care path are crucial in optimizing its effectiveness and further investigation of why a large proportion of subjects did not visit the portal and efforts aimed at improving this are warranted.

Table 1. Outcome in the control (no access to information portal) vs. the intervention (access to information portal) groups.

	Control	Intervention		P-value	
	(n=213)	Intention-to-treat (n=114)	As-treated (n=62)	Intention-to-treat	As-treated
Disease-specific knowledge	5.04 ± 1.28	5.11 ± 1.27	5.33 ± 1.22	0.635	**0.019**
HADS	7.22 ± 5.66	7.33 ± 5.92	7.65 ± 6.12	0.872	0.556
-Anxiety	4.76 ± 3.26	4.70 ± 3.55	5.05 ± 3.97	0.888	0.421
-Depression	2.44 ± 2.87	2.60 ± 2.93	2.57 ± 2.59	0.628	0.806
SF-36 MCS*	75.66 ± 15.84	76.10 ± 16.23	75.77 ± 15.82	0.821	0.984
-Vitality	65.37 ± 18.20	67.58 ± 19.56	67.01 ± 17.57	0.320	0.681
-Social functioning	84.51 ± 20.29	83.41 ± 19.64	83.47 ± 19.04	0.646	0.770
-Role-emotional	78.51 ± 22.97	77.08 ± 22.73	74.04 ± 24.44	0.600	0.131
-Mental health	78.37 ± 16.11	79.81 ± 15.38	80.57 ± 15.36	0.449	0.348
Autonomy preference index	77.46 ± 8.24	78.01 ± 7.67	78.13 ± 7.94	0.567	0.604
-Information seeking	88.64 ± 8.91	89.60 ± 7.81	90.53 ± 7.64	0.353	0.110
-Decision-making	62.42 ± 13.79	62.93 ± 14.18	62.15 ± 14.85	0.759	0.780
Control preferences scale (who should make the final decision for treatment?)				0.661	0.742
-Physician	0.5% (1)	0.0% (0)	0.0% (0)		
-Physician, after considering patient opinion	13.4% (28)	11.2% (12)	14.5% (9)		
-Physician and patient together	82.8% (173)	82.2% (88)	79.0% (49)		
-Patient, after considering physician opinion	2.9% (6)	5.6% (6)	4.8% (3)		
-Patient	0.5% (1)	0.9% (1)	1.6% (1)		

Data are presented as "mean ± standard deviation" or "percentage (count)". *scores transformed to a 0-100 scale, non-normative. HADS=hospital anxiety and depression scale. SF-36 MCS = Short form-36 Mental Component Scale.

4M-5 ASSESSING THE EFFICACY OF AN EDUCATIONAL APP FOR SMARTPHONE OR TABLET, ACTIVELY PROVIDING SUBDIVIDED AND INTERACTIVE CONTENT TO INCREASE PATIENTS' ILLNESS KNOWLEDGE: A RANDOMIZED CONTROLLED TRIAL

Decision Psychology and Shared Decision Making (DEC)

Thomas Timmers, MSc, Interactive Studios, Rosmalen, Netherlands

Purpose:
Modern healthcare focuses on Shared Decision Making (SDM) because of its positive effects on patient satisfaction, therapy compliance, and outcomes. Patients' knowledge about their illness and available treatment options, gained through medical education, is one of the key drivers for SDM. Current patient education relies heavily on medical consultation and is known to be ineffective. The objective of this study was to determine whether providing patients with information in a subdivided, categorized, and interactive manner via an educational app for smartphone or tablet may increase the knowledge of their illness.

Method(s):
A surgeon blinded randomized controlled trial was conducted with 213 patients who were referred to one of the six Dutch hospitals by their general practitioner owing to knee complaints that were indicative of knee osteoarthritis (OA). Patients were offered an interactive app that, in addition to standard care, actively sends informative and pertinent content to patients about their illness on a daily basis by means of push notifications in the week prior to their consultation.

Result(s):
The primary outcome was the level of perceived and actual knowledge that patients had about their knee complaints and the relevant treatment options after the intervention. After the intervention, the level of actual knowledge (measured on a 0-36 scale) was 52% higher in the app group (26.4 versus 17.4, $P < 0.001$). Moreover, within the app group, the level of perceived knowledge (measured on a 0-25 scale) increased by 22% during the week within the app group (from 13.5 to 16.5, $P < 0.001$), compared to no gain in the control group.

Conclusion(s):
Actively offering patients information in a subdivided (per day), categorized (per theme), and interactive (video and quiz questions) manner increases their levels of knowledge (both perceived and actual), compared to standard care educational practices.

4M-6 MAPPING DECISION-MAKING PROCESSES AND IDENTIFYING DECISION POINTS WITH THE POTENTIAL TO INCLUDE PATIENT PREFERENCE INFORMATION THROUGHOUT THE MEDICAL PRODUCT LIFECYCLE AMONG ITS MAIN STAKEHOLDERS

Patient and Stakeholder Preferences and Engagement (PSPE)

Chiara Whichello, MSc, MA[1], **Karin Schoelin Bywall**[2], Jonathan Mauer[3], Stephen Watt[3], Irina Cleemput[4], Cathy Anne Pinto, PhD[5], Eline van Overbeeke, MSc[6], Rosanne Janssens[7], Isabelle Huys, MSc, PhD[6], Richard Hermann, MD, MPH[8] and Jorien Veldwijk, PhD[9], (1)Erasmus University Rotterdam, Rotterdam, Netherlands, (2)Uppsala University, Centre for research ethics and bioethics, Uppsala, Sweden, (3)Pfizer, New York, NY, (4)Belgian Health Care Knowledge Centre, Brussels, Belgium, (5)Merck Research Laboratories, Center for Observational and Realworld Evidence, Rahway, NJ, (6)KU Leuven, Leuven, Belgium, (7)Department of Pharmaceutical and Pharmacological Sciences, University of Leuven, Leuven, Belgium, (8)AstraZeneca, Wilmington, DE, (9)Erasmus Choice Modelling Centre, Erasmus University , Rotterdam, Netherlands

Purpose: The aims of this study were: 1) Identify the decision-making processes and decision points throughout the Medical Product Lifecycle (MPLC) for pharmaceutical industry, regulatory authorities, and health technology assessment (HTA) bodies and payers, and 2) Determine which decision points have potential to include Patient Preference Information (PPI). These aims serve to improve patient-centric decision-making throughout the MPLC.

Method(s): A 3- step approach was conducted, including a scoping literature review identifying relevant white and grey literature, validation meetings with stakeholders to confirm the decision-making process, and semi-structured interviews with representatives of 3 stakeholder groups (pharmaceutical industry n=24, regulatory authorities n=22, HTA bodies and payers n=24). The literature review was conducted using databases: Guidelines International Network, Embase, PubMed (including Cochrane Central and Medline), PsycINFO and EconLit). Validation meetings were conducted with European Union (EU) and United States (US) industry representatives, HTA experts from EU (associated with EUneHTA) and US, and regulatory representatives from US, EU, and EU-notifying bodies. The interviews were conducted within seven different European countries (Sweden, Romania, Italy, UK the Netherlands, Germany, France) and the US.

Result(s): Six decision points were identified for the *industry decision-making process*: informing the prioritization of treatment targets, studies, assets, or post approval opportunities; and decisions to advance product development, submission, or launch. There were four decisions points identified to the *regulatory decision-making process*: submission and validation, scientific opinion, orphan designation and commission decision. Six decision points were identified for the *HTA decision-making:* filtration for HTA, prioritization for HTA, and appraisal for reimbursement decision-making. These three decision points are repeated if a reassessment needs to be performed. Patient representatives are often involved in the decision-making process within industry, regulatory authorities and HTA bodies. PPI is currently not considered as required information to be submitted for decision-making, but would be of great benefit if it were incorporated in the future, according to all stakeholder groups.

Conclusion(s): A total of 16 decision points were identified throughout the MPLC. Each of the decision points requires different information and/or data to be submitted to the decision-makers in order for them to assess the information based on pre-set decision criteria. PPI within all stakeholder groups is considered an important component to inform future decision-making across the MPLC, but it is not currently included in a structured way.

4N-1 THE VALUE OF AIR POLLUTION INTERVENTIONS IN WEST YORKSHIRE: AN ECONOMIC EVALUATION CONSIDERING MULTIPLE OUTCOMES, INEQUALITY IMPACTS AND COSTS FALLING ON DIFFERENT DECISION MAKERS

Applied Health Economics (AHE)

Simon Walker, MSc[1], James Lomas, PhD[1], Susan Griffin, PhD[2] and Mark Sculpher, PhD[2], (1)Centre for Health Economics, University of York, York, United Kingdom, (2)University of York, York, United Kingdom

Purpose:

40,000 deaths are attributable to exposure to outdoor air pollution each year in the UK alone. Different policy options exist to tackle air pollution, and the costs of these policies often fall on decision makers whose mandates extend beyond health. Cost-effectiveness analysis assuming a single budget and an objective of maximising health is not suitable for the analysis of such policies. We apply a new framework to estimate the cost-effectiveness of different air pollution interventions in the West Yorkshire Low Emission Zone (WYLEZ).

Method(s):

The new framework is based on an 'impact inventory' which considers the impact of interventions on multiple outcomes and budgets whilst allowing for competing objectives of different decision makers (including alternative notions of equity). We compare two policies for reducing air pollution based on a previously published economic model. The outcomes include the impact on health, inequality in health, consumption and wider social benefits, and costs potentially falling on the local authority, central government, NHS budgets or the private sector. We compared scenarios for policy implementation costs allocated to different budgets, and alternative decision makers' preferences for different outcomes and different levels of inequality aversion.

Result(s):

The policies resulted in improved health and reduced health inequality. The impact on consumption and wider social benefit was mixed. The scale of the gains depended on which budgets the costs of the intervention fell and the level of opportunity costs associated with those budgets. The results allow for an informed discussion between the different decision makers on who should bear the costs of the intervention, and allow for consideration of potential compensation payments.

Conclusion(s):

This study shows how analysts can inform the assessment of policies which involve a range of public sector decision makers with heterogeneous and potentially conflicting priorities. Through appropriate assessments and by clearly distinguishing how value judgements feed into the evaluation process, assessments of policies with multiple outcomes and costs falling across budgets can be reliably informed.

4N-2 USING LARGE DATA TO PRESENT UNCERTAINTY FOR RISK PREDICTION IN THE ERA OF PRECISION MEDICINE: THE RESPECT ALGORITHM FOR PREDICTING DEATH AT END-OF-LIFE

Quantitative Methods and Theoretical Developments (QMTD)

Douglas Manuel, MD, MSc, Amy Hsu, PhD, Sarah Spruin, MSc, Carol Bennett, MSc and Peter Tanuseputro, MD, MHSc, Ottawa Hospital Research Institute, Ottawa, ON, Canada

Purpose: To predict and communicate risk of death for community dwelling older people using a precision medicine approach. A risk prediction algorithm was developed: Risk Evaluation for Support: Predicting Elder-life in the Community Tool – End-of-Life (RESPECT-EOL). Algorithm development and application considered personalized medicine by including patient-oriented risk reporting and use of adaptive questionnaires (varying the number of questions and predictors for respondents depending on their risk level).

Method(s): The study base was all community-dwelling Ontarians who received home care from 2007 to 2015. Algorithm development was pre-specified and published (Trial registration NCT02779309). There were 488,636 participants with 836,012 assessments and 298,657 deaths in the combined derivation and calibration cohort. The primary outcome was median survival time with 25th to 75th survival percentiles. Survival time was generated in a two-step process by rank ordering participants into 61 groups based on six-month probability of death (from a Cox-proportional hazard model) and generating Kaplan-Meier five-year survival curves for each group.

Result(s): The median predicted six-month probability of death was 0.1095% (0.1093 to 0.1097, 95% CI). Risk varied among the 61 groups from 0.0158 (0.0158-0.0159) to 0.9820 (0.9810-0.9830). Median observed survival time varied from 27 days (10 to 81 days, 25th and 75%ile) in the highest risk group to 10 years (3655 days (2111 to > 3655 days)) in the lowest risk group. Discrimination and calibration were satisfactory the validation (temporal split sample), with C statistic of 0.77 and discrimination plot intercept 0.000, slope 1.000 in validation data. A web-application was used to generate six different approaches for visual display of risk and uncertainty. In each, the uncertainty definition can be modified (e.g., percentile range for median survival (from 1st to 99th percentile) and probability of death for different time cut-offs (from six months to five years)). The web application updates risk calculations after each question, providing user live feedback regarding influence of specific predictors.

Conclusion(s): Algorithms generated using a large number of participants allow for a variety of approaches to report patient-comprehensible prognosis and uncertainty. Development using population-based data allows a range of approaches to re-calibrate the algorithm in different populations. Community-based patient and caregiver studies are required to assess different approaches to communicate risk of death at end-of-life (ongoing).

4N-3 IMPLEMENTING INTERVENTIONS WITH VARYING MARGINAL COSTS AND BENEFITS: AN APPLICATION IN PRECISION MEDICINE

Quantitative Methods and Theoretical Developments (QMTD)

Stuart Wright, BSc, MSc, Manchester Centre for Health Economics, The University of Manchester, Manchester, United Kingdom, Mike Paulden, PhD, University of Alberta, Edmonton, AB, Canada and Katherine Payne, PhD, The University of Manchester, Manchester, United Kingdom

Purpose:

Economic evaluations of healthcare interventions are used to provide evidence of their cost-effectiveness in order to inform funding decisions. However, little attention is given to the economic issues involved in implementing cost-effective interventions. A range of barriers may constrain effective implementation and if the costs and benefits of an intervention vary at different levels of implementation, this can have consequences for the cost-effectiveness of the technology. The aim of this study is to illustrate the consequences of varying marginal costs and benefits for the cost-effective implementation of healthcare interventions.

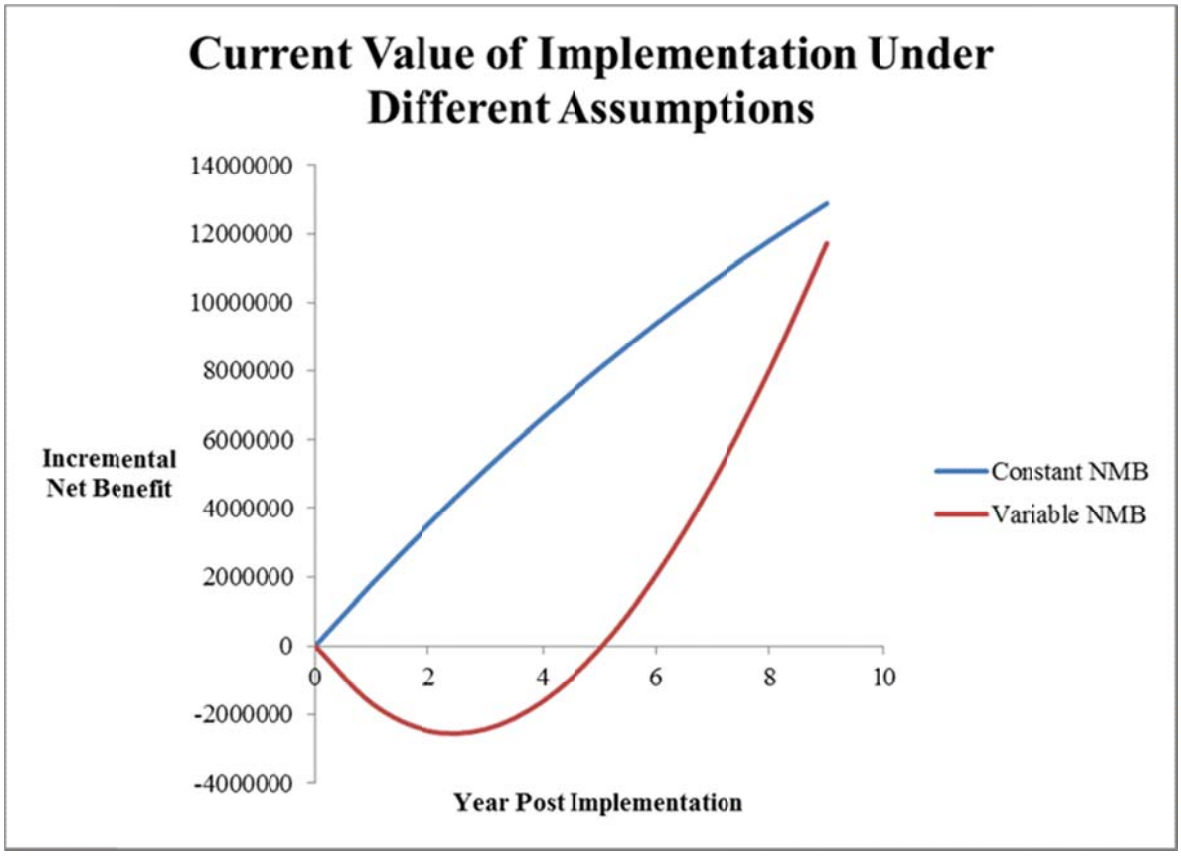
Method(s):

This study adapted the existing value of implementation methodology to allow costs and benefits to vary with differing levels of implementation. The expanded value of implementation framework is applied to a published case study of the evaluation of a 70-gene recurrence score for breast cancer. In this study, the marginal costs of the test decreased with increased implementation while the marginal benefits increased. The consequences of allowing for varying costs and benefits for the value of the intervention and of implementation strategies are illustrated graphically and numerically in both static and dynamic forms. Graphical representations of the consequences of other forms of non-linear costs and benefits are also presented.

Result(s):

The increasing returns to scale exhibited by the technology mean that while it appears cost-effective for high levels of implementation, it is not an effective use of resources at lower levels of implementation. The expanded value of implementation formulae presented can be solved to show that a minimum of 51% of patients must be treated for the intervention to be cost-effective. The non-constant marginal net benefit also means that the value of implementation strategies is dependent on the initial and ending levels of implementation as opposed to just the magnitude of the increase in patients receiving the intervention. In dynamic models, incremental losses caused by low implementation can accrue over time if implementation is slow (figure 1).

Figure 1: Current Value of Implementation for a 70-gene recurrence score assay



Conclusion(s):

Poor implementation of apparently cost-effective interventions can have a significant economic impact on health systems, especially if they exhibit increasing returns to scale. More attention should be given to the cost-effective implementation of healthcare interventions after funding decision are made.

4N-4 COMBINING MULTIPLE IMPUTATION AND BOOTSTRAP IN THE STATISTICAL ANALYSIS OF COST-EFFECTIVENESS TRIAL DATA

Quantitative Methods and Theoretical Developments (QMTD)

Wilbert van den Hout, PhD[1], Jaap Brand[1], Saskia le Cessie[2] and Stef van Buuren[3], (1)Medical Decision Making, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands, (2)Departments of Epidemiology and Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands, (3)Department of Methodology & Statistics, University of Utrecht, Utrecht, Netherlands

Purpose: In cost-effectiveness trial data, probability distributions are typically skewed and missing data are frequent. Bootstrap and multiple imputation are well-established resampling methods for handling skewed and missing data. However, it is not clear how these resampling techniques should be combined.

Method(s): We compared ten candidate methods to combine multiple imputation and bootstrap, in order to estimate 95% confidence intervals for the mean difference in outcome between two treatment groups. Statistical validity was assessed using simulation, generating 1000 incomplete data sets from 30 different data simulation models. The data simulation models varied in correlation and skewness of the data, percentage zero costs, sample size, percentage missing, and missing (completely) at random mechanisms. A method was considered statistically valid if it was unbiased with proper coverage for the confidence interval.

Result(s): Single imputation nested in the bootstrap percentile method emerged as the method with the best statistical properties (valid for 29 out of 30 data simulation models). Second best was multiple imputation nested in the bootstrap percentile method (valid for 23 out of 30). Third best was standard parametric multiple imputation (valid for 19 out of 30), for which the confidence intervals had a coverage of at least 89% for all 30 data simulation models.

Conclusion(s): Single imputation nested in the bootstrap percentile method (with added noise to reflect the uncertainty of the imputation) showed the best statistical validity. However, this method can require extensive computation times and the lack of standard software makes this method not accessible for a larger group of researchers. Using a standard unpaired t-test with standard multiple imputation without bootstrap appears to be a robust alternative with acceptable statistical performance for which standard multiple imputation software is available.

4N-5 SENSITIVITY ANALYSIS FOR NOT-AT-RANDOM MISSING DATA IN TRIAL-BASED COST-EFFECTIVENESS ANALYSIS USING MULTIPLE IMPUTATION

Applied Health Economics (AHE)

Baptiste Leurent[1], Manuel Gomes, PhD[2], Rita Faria, MSc[3], Steve Morris[4], Richard Grieve, PhD[2] and James Carpenter, DPhil[2], (1)London School of Hygiene & Tropical Medicine, London, United Kingdom, (2)London School of Hygiene and Tropical Medicine, London, United Kingdom, (3)University of York, York, United Kingdom, (4)University College London, London, United Kingdom

Purpose:

Multiple imputation is being increasingly used to address missing data in cost-effectiveness analysis (CEA) of randomised trials, assuming data are missing-at-random (MAR). However, in many CEA settings the missing data is related to unobserved values, i.e. data are missing-not-at-random (MNAR). For example, if patients in poorer health are less likely to complete EQ-5D questionnaires, given the observed data. Guidelines recommend sensitivity analyses under plausible MNAR assumptions, but this is rarely done in practice. We aim to illustrate an accessible framework to conduct sensitivity analysis for departure from MAR in trial-based CEA, using the Ten Top Tips (10TT) trial evaluating a weight management intervention.

Method(s):

We illustrate the implementation of pattern-mixture models using multiple imputation, modifying the imputed data to reflect a plausible departure from the MAR assumption. Sensitivity analyses are conducted under a range of plausible values (deviations from MAR), to assess the robustness of the study conclusions.

We applied this framework to the 10TT CEA, to address the concern that participants who dropped out of the study (about 42%) could be those less successful at losing weight (i.e. data were MNAR).

Result(s):

Under the base-case MAR assumption using multiple imputation, 10TT resulted in 0.004 fewer quality-adjusted life years (95%CI -0.074 to 0.066) and ≤35 lower costs (-504 to 434) compared to the control arm, with a 48% probability of being cost-effective at ≤20,000 per QALY. The sensitivity analysis illustrated that the study conclusions were sensitive to small departures from MAR (e.g. assuming missing quality-of-life was 10% lower than observed), with the probability of 10TT being cost-effective ranging from 16% to 75% (Figure 1).

Conclusion(s):

MNAR is an important concern in trial-based CEA and sensitivity analyses are recommended to assess whether the study conclusions are robust to departures from the missing-at-random

assumption. We illustrated an accessible framework to conduct and report these sensitivity analyses.

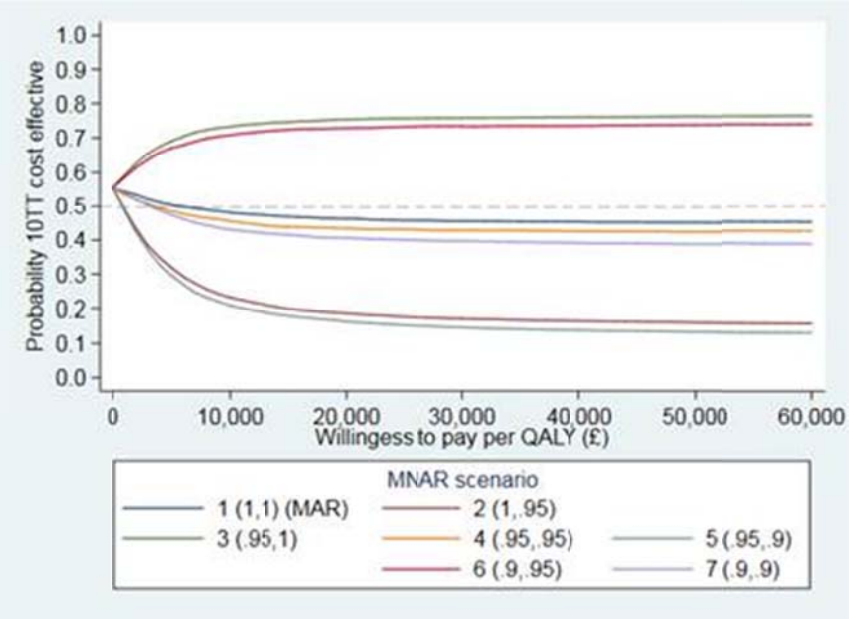


Figure 1. Cost-effectiveness acceptability curves under different MNAR assumptions. Numbers in brackets indicate the MNAR rescaling parameters in the control and 10TT arm. For example, scenario 7 (0.9,0.9), missing quality-of-life are assumed to be 10% lower than under MAR in both arms.

4N-6 CASE-MIX ADJUSTMENT OF CLINICAL AND PATIENT REPORTED OUTCOMES FOR VALUE BASED HEALTH CARE IN ACUTE STROKE CARE
Health Services, Outcomes and Policy Research (HSOP)

Arvind Oemrawsingh, MSc, Erasmus Medisch Centrum (EMC), Rotterdam, Netherlands, Nikki van Leeuwen, MSc, Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, Netherlands and Hester F. Lingsma, PhD, Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands

*On behalf of the CVAB 3.0 study group

Center for Medical Decision Making, Department of Public Health, Erasmus Medical Center, PO Box 2040, 3000 CA, Rotterdam, The Netherlands

Purpose:

Value based health care is built on comparing Patient Reported Outcome Measures (PROMs) between centers and relate these to potential difference in quality of care. Such comparisons requires case-mix adjustment to account for differences in patient characteristics between centers, and assessment of statistical uncertainty. The aim of this study was to develop a case-mix model for a PROM for acute ischemic stroke and assess statistical uncertainty, in comparison with clinical outcome measures.

Method(s):

1022 patients with an acute ischemic stroke from four stroke centers in the Netherlands, included in a quality registry between 2014 and 2016 were analysed. Case-mix models for mortality, the modified Rankin Scale (mRS) and EuroQol-5D (EQ-5D) utility score at three months were developed with respectively binary logistic, proportional odds and linear regression models with stepwise backward selection. The predictive abilities of the models were determined by the Area Under the Curve (AUC). Random effect regression models were used to quantify the role of statistical uncertainty.

Result(s):

Age, gender and National Institute of Health Stroke Scale (NIHSS) score were the strongest predictors for 3-month mortality. For the mRS score at 3 months the strongest predictors were age, ethnicity and NIHSS score. The strongest predictors for the EQ-5D utility score at 3 months were age, sex, ethnicity and socio-economic status. Preliminary analyses showed significant hospital effects for all three outcome measures.

Conclusion(s):

The case-mix adjustment models were different for the PROM compared to clinical outcome measures. Detecting differences in quality of care based on PROMS, as propagated in the Value Based Health Care approach, thus requires specific case-mix models. Because of its continuous nature, the EQ-5D might be less sensitive for statistical uncertainty than dichotomous or ordinal outcome measures.

5Q-1 THE OPTION-OLD: MEASURING THE EXTENT TO WHICH CLINICIANS INVOLVE OLDER PATIENTS AND INFORMAL CAREGIVERS IN TRIADIC DECISION-MAKING

Decision Psychology and Shared Decision Making (DEC)

Ruth Pel-Littel, Msc[1], Bianca Buurman, PHD[2], Marjolein van de Pol, PHD[3], Nida Gizem Yilmaz, Msc[4], Linda Tulner, MD[5], Mirella Minkman, PHD[1], Wilma Scholte op Reimer, PHD[6], Glyn Elwyn, MD, MSc, PhD[7] and Julia Van Weert, PhD[8], (1)Vilans, Utrecht, Netherlands, (2)Academic Medical Centre, University of Amsterdam,, Amsterdam, Netherlands, (3)Radboud University, Nijmegen, Netherlands, (4)Academic Medical Centre, University of Amsterdam, Amsterdam, Netherlands, (5)MC Slotervaart, Amsterdam, Netherlands, (6)Amsterdam University of Applied Sciences, Amsterdam, Netherlands, (7)The Dartmouth Institute for Health Policy and Clinical Practice, Lebanon, NH, (8)University of Amsterdam, Amsterdam, Netherlands

Purpose:

In shared decision making (SDM) with older people a shift from disease-oriented to a goal-oriented-approach is needed because of multiple chronic conditions (MCC) and involvement of informal caregivers (IC). The 'Dynamic Model for SDM in Frail Older Patients' adds steps to the regular SDM process to address these matters. The aim of this research is to develop a valid and reliable tool to measure SDM in older populations with multiple chronic conditions (MCC) and their informal caregivers: the OPTION-OLD, based on the OPTION-5 and OPTION-12

Method(s):

The study is part of the decision making in older persons (DICO) study and includes video consultations of 108 older patients and their IC that visited the geriatric outpatient department in two Dutch hospitals. The OPTION-OLD scale is a revised version of the OPTION-5 and OPTION-12 scale with further adaptations based on the 'Dynamic Model for SDM in Frail Older Patients'. Consultations were coded with the OPTION-OLD by three researchers with a health science background. First the raters used a sample of ten videotaped consultations as training-sample. They used the OPTION-5 manual and additional instructions concerning the rating of the added steps. Final interrater- reliability was established for another sample of ten videotaped consultations rated by the three raters independently. The remaining 88 consultations were rated independently, whilst every tenth consultation was discussed in order to ensure a stable interrater-reliability.

Result(s):

The mean age was 78 years, and 55% was female, polypharmacy was present in 69% of the patients. In 63% of the consultations an IC was present. The mean duration of consultations was 40.3 minutes. Most frequently discussed decisions were about additional diagnostics, medication, follow up, referral to primary care and lifestyle. In 88% of the decisions, more options were available, 43% of those options were considered equal. Comorbidities were often discussed (87%) and considered in relation to the decision (73%). Interrater-reliability was good with an

ICC of 0.95. The mean OPTION-OLD score was 42.5 (range 0-100), items about 'team talk' and 'evaluation' scored the lowest (resp. 31 & 36.5).

Conclusion(s):

The video observations showed that SDM was applied moderately by clinicians. Based on preliminary data the OPTION-OLD shows a good interrater-reliability, although further psychometric properties and participation of patients and IC are currently being analyzed.

5Q-2 REFERRAL DECISION MAKING OF GENERAL PRACTITIONERS: A SIGNAL DETECTION STUDY

Decision Psychology and Shared Decision Making (DEC)

Olga Kostopoulou, PhD[1], Martine Nurek, PhD[1], Grace Okoli, PhD[2], Francesca Fiorentino, PhD[1] and Brendan Delaney, MD[1], (1)Imperial College London, London, United Kingdom, (2)King's College London, London, United Kingdom

Purpose: Signal Detection Theory (SDT) describes how respondents categorise ambiguous stimuli over repeated trials. It measures separately 'discrimination' (sensitivity to the strength of evidence) and 'criterion' (inclination to respond 'yes' vs. 'no'). This is important because respondents may produce the same accuracy rate for different reasons. We employed SDT to measure the discrimination and criterion of General Practitioners (GPs) when making referral decisions about suspected lung cancer.

Method(s): Using published evidence, we constructed 44 vignettes describing patients presenting to the GP with symptoms suggestive of lung cancer. The 1-year cancer risk varied from 0.03% to 14.23% across vignettes. Under current UK risk-based guidance, half of the vignettes required urgent referral. We recruited 216 GPs from 151 practices across England. Practices represented a range of referral performance, measured by the positive predictive value (PPV) of their referrals (chance of referrals identifying cancer) and the sensitivity (chance of cancer patients being picked up via referral from their practice) – publicly available data. The GPs saw the vignettes online, and indicated whether they would refer the patient urgently or not. We calculated each GP's discrimination (d') and criterion (c) from their 'hit' and 'false alarm' rates, i.e., referral decisions for vignettes that should and should not be referred. We regressed d' and c on practice PPV and sensitivity, and on GP experience and gender.

Result(s): The sample's criterion ranged from -1.44 to 2.02 (median 0.43). Positive values of c indicate a conservative approach and bias towards 'no' responses, negative values indicate the opposite. Criterion was associated with practice PPV: as PPV increased, GPs' c also increased, indicating *lower* inclination to refer (b=0.06 [0.02-0.09] p=0.001). Female GPs were more inclined to refer than male GPs (t_{214}=2.08, p=0.04). Average discrimination was above chance (>0) but modest (d'=0.77), and highly variable (range -0.28 to 1.91). Discrimination was not associated with practice characteristics. The lowest discrimination was measured in the most experienced group (>18 years in practice).

Conclusion(s): Practices with high referral PPV achieve this by referring fewer patients (avoiding false positives), not by discriminating better. Rather than bluntly mandating more or fewer referrals, it is necessary to improve discrimination by providing GPs with better evidence**.** This is the first study to associate practitioner decision making with organisational performance in cancer referral decisions.

5Q-3 LOGISTIC AND HYBRID LENS MODELS APPLIED TO PHYSICIAN DECISION WHETHER TO REFER FOR POSSIBLE CANCER

Decision Psychology and Shared Decision Making (DEC)

Robert M. Hamm, PhD[1], Olga Kostopoulou, PhD[2] and Martine Nurek, PhD[2], (1)University of Oklahoma Health Sciences Center, Oklahoma City, OK, (2)Imperial College London, London, United Kingdom

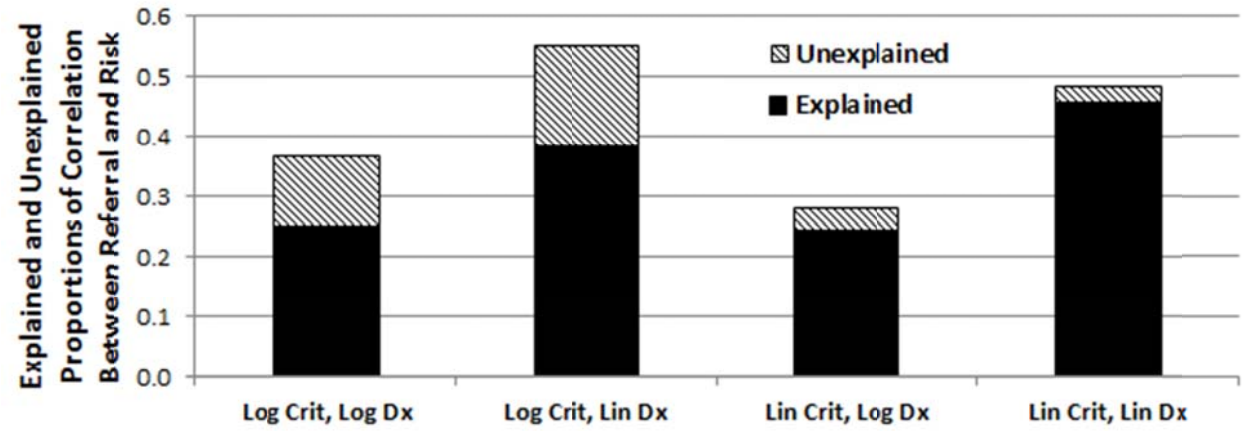
Purpose: Primary care physicians sometimes miss recognizing a patient might have cancer. To study patient features which may be neglected, we applied the lens model equation (LME) to analyze the accuracy of medical judgment or diagnosis. It statistically models the relation between disease states and patient features, and compares that to a descriptive model of the relation of physicians' diagnoses to the same features. Researchers may use continuously scaled measures or collapse them to binary variables. How might models differ?

Method(s): Over 200 UK general practitioners (GPs) each read over 40 hypothetical case descriptions of patients with symptoms suggesting cancer, and stated whether they would refer patient urgently for assessment. The criterion for each case could be the patient's continuous-scaled cancer risk from an established risk model (given age, smoking history, etc.), or it could be the dichotomous recommendation to refer for urgent assessment when risk exceeds a threshold. Thus for each GP it is possible to calculate a logistic lens model (BeBj: Binary Ecology, Binary Judgment) or a hybrid model (CeBj: Continuous Ecology, Binary Judgment).

Assuming GPs think alike, we can analyze group judgments, permitting both a continuous group judgment (vote count for referring patient) and a dichotomous group judgment (the majority decision). Four group lens models may be compared: Logistic (BeBj), Hybrid (BeCj and CeBj), and Linear (CeCj).

Result(s): Individual physicians referred various proportions of patients. When very few patients were referred, models explained less. Usually the proportion of an individual physician's accuracy explainable by the logistic LME was larger than by the linear LME.

Analyzing group decisions, the figure shows appropriate LMEs fit to all combinations of continuous (risk, or vote count) and binary (risk category, majority decision) measures. Both physician accuracy and proportion explained by the LME differed greatly according to nature of the measures of the risk definition and the group diagnosis.



Conclusion(s): The choice of binary or continuous criterion and judgment for analysis of individual or group diagnostic accuracy makes a large difference in the observed accuracy and in the proportion of the accuracy explainable with lens model equations. Researchers should consider the implications of the choice of measure, and possibly analyze several alternatives using the appropriate linear, logistic, or hybrid LMEs.

5Q-4 IMPACT OF PROBABILITY DISTORTION ON MEDICAL DECISIONS: A FIELD STUDY

Health Services, Outcomes and Policy Research (HSOP)

Marine Hainguerlot[1], Vincent Gajdos[2], Karen Milcent[3] and Jean- Christophe Vergnaud[1], (1)University Paris 1 Pantheon- Sorbonne, Paris, France, (2)INSERM, CESP Centre for Research in Epidemiology and Population Health, Villejuif, France, (3)Department of Pediatrics, Antoine Beclere University Hospital, Assistance Publique-Hopitaux de Paris, Clamart, France

Purpose: Estimating the probability that a patient has a disease is essential in the diagnostic process. We evaluated whether physicians' probability estimates are distorted compared to the probabilities generated by a statistical model. We investigated whether distortion in physicians' probability estimates impacts negatively the accuracy of medical decisions.

Method(s): Data come from a prospective, multicenter, cohort study in 15 French pediatric emergency departments. Physicians were asked to record the medical information they collected about febrile infants younger than 3 months (N=1848) from the admission to the discharge. Physicians were required to report their probability estimate that the infant had a bacterial infection (BI), on a scale from 0 to 100%. Then, they reported their decision to treat with antibiotics or not. The presence or absence of BI was categorized by the attending physician and reviewed by a committee of experts.

We used the Lens model approach to estimate: (i) how physicians weighted the medical information (linear judgment) compared to our statistical model (linear model); (ii) the physician's probability of BI predicted by the linear judgment and the probability of BI predicted by the linear model, for each patient. To quantify probability distortion, we compared the physician's probability of BI predicted by the linear judgment and the probability of BI predicted by the linear model in log odds form across all patients.

To assess the impact of probability distortion on antibiotic treatment, we categorized probability distortion into a high and a low probability distortion group. We considered that antibiotic treatment was necessary only if the infant had a BI. We compared across the two groups the sensitivity (i.e. the proportion of infants with BI who received antibiotic treatment) and the specificity (i.e. the proportion of infants without BI who did not receive antibiotic treatment).

Result(s): We found that physicians' probability estimates were distorted and followed an inverted S-shaped function. They over-estimated small probabilities and under-estimated large probabilities. Proportion tests showed that the specificity of antibiotic treatment was significantly lower in the high distortion group compared to the low distortion group (mean: 64.2% vs 77.6%, $p<0.001$) while the sensitivity did not differ across the two groups (mean: 98.7% vs 97.7%, $p=0.48$).

Conclusion(s): Our results suggest that probability distortion in clinical judgment might cause unnecessary health care.

5Q-5 DEVELOPMENT AND VALIDATION OF THE HEALTH-RISK ATTITUDE SCALE (HRAS)

Decision Psychology and Shared Decision Making (DEC)

Samare Huls, MSc[1], Sylvie van Osch, PhD[2], Werner Brouwer, Prof. Dr.[1], Job van Exel, Prof. Dr.[1] and Anne M. Stiggelbout, PhD[3], (1)Erasmus School of Health Policy & Management, Rotterdam, Netherlands, (2)Leiden University Medical Centre, Leiden, Netherlands, (3)Medical Decision Making, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands

Purpose: People differ in their attitude towards health risks, and this results in different health behaviours. The main aim of this study is to develop and assess the reliability and validity of a scale to measure how individuals value their health and manage health risks, the Health-Risk Attitude Scale (HRAS).

Method(s): Four studies were conducted to develop and validate the HRAS. The constructs of the scale were developed using interviews, critical review and extensive pilot study. In study 1, an 18-item pilot version of the HRAS was assessed (n=130). In study 2, the HRAS containing 13 items was assessed (n=132). In study 3, the HRAS was assessed in a representative sample (n=930). In study 4, a 6-item short form was assessed, the HRAS-SF. Reliability was assessed using Cronbach's alpha, factor analysis and test-retest. Validity was assessed using correlations with similar scales and subgroup analyses according to socio-demographic characteristics and health behaviour variables.

Result(s): In study 1, based on factor analysis and expert review, 5 items were excluded resulting in an increased Cronbach's alpha and a high test-retest reliability. The 13-item version showed a relatively good construct and discriminant validity. Study 2 confirmed the results of study 1 and showed that individuals involved in risky health behaviour had a higher HRAS score. Study 3 again showed similar results and, based on factor analysis, a short form 6-item scale (HRAS-SF) was defined. The correlation between the HRAS and HRAS-SF was 0.85. Respondents who engaged in risky health behaviour (i.e. smoking, poor nutrition, drinking alcohol and sedentary behaviour) had higher mean HRAS and HRAS-SF scores. Study 4 replicated the findings for the 6-item HRAS-SF in terms of descriptive properties, internal consistency, and associations with socio-demographic properties and health behaviours.

Conclusion(s): The four studies show that the HRAS has high internal consistency, scores stable over time, correlates moderately to strongly with other measures of risk attitude and relates to respondents' characteristics and health behaviour as expected. The same holds for the HRAS-SF. These findings support the reasoning that, in general, a person who lives a healthy lifestyle is more risk-averse in the domain of health, and that the HRAS is a promising instrument for measuring health-risk attitude. Further research in decisions where health-risk plays a role is warranted.

5Q-6 SOCIAL NUDGING: THE EFFECT OF SOCIAL FEEDBACK INTERVENTIONS ON VACCINE UPTAKE

Decision Psychology and Shared Decision Making (DEC)

Lars Korn, M.Sc.[1], Cornelia Betsch, PhD[1], Robert Böehm, PhD[2] and Nicolas W. Meier, M.Sc.[2], (1)University of Erfurt, Erfurt, Germany, (2)RWTH Aachen University, Aachen, Germany

Purpose: Most vaccines provide indirect community protection by preventing the transmission of a disease. Psychologically, this effect can also motivate omission of vaccination because increasing vaccination rates reduce the risk of infection and, thus, the individual benefit of vaccination. Hence, vaccination becomes a social dilemma where individuals' interests conflict with group interests. The current study investigated two social nudge interventions aiming at increasing individuals' motivation to act in the group's interest.

Method(s): Previous research showed that goal-setting strategies enhance coordination and cooperation among individuals. It is, therefore, hypothesized that rewarding a group's goal-attainment (i.e., disease elimination) results in an increase in goal-directed behavior (i.e., vaccination). Secondly, the inter-group comparison -- intra-group cooperation hypothesis assumes that salient out-groups activate a comparative focus, which instigates individuals to increase their cooperation within their own group. Thus, we assume that comparison with another group increases cooperative vaccination within the own group.

In a laboratory experiment, the interactive vaccination (I-Vax) game was used to model the direct and indirect effects of vaccinations. The game was played by 288 participants over 20 rounds. The experimental setup varied the feedback information after each round to implement a 2 (*rewarding goal-attainment*: present vs. absent) × 2 (*inter-group comparison*: present vs. absent) between-subjects design. Participants in the *rewarding goal-attainment* condition obtained badges as an additional symbolic reward of the elimination history of the disease. In order to manipulate *inter-group comparison*, information on the performance of the out-group was either provided or not provided.

Result(s): Multi-level logistic regressions with mixed effects revealed the expected positive effect of *rewarding goal-attainment*. However, this effect lost its impact over the course of the game (Figure 1, Panel A). The effect of *inter-group comparisons* was also positive but weaker and less consistent (Figure 1, Panel B) than *rewarding goal-attainment*. Further, vaccination attitude as well as social value orientation were strong predictors of vaccination behavior - supporting the validity of the I-Vax game.

Conclusion(s): The current experiment shows that communicating and rewarding "small wins" obtained by a group may increase individuals' willingness to act in the group's interest. Inter-group processes deserve further attention and investigation as potential strategies for improving vaccine communication and advocacy.

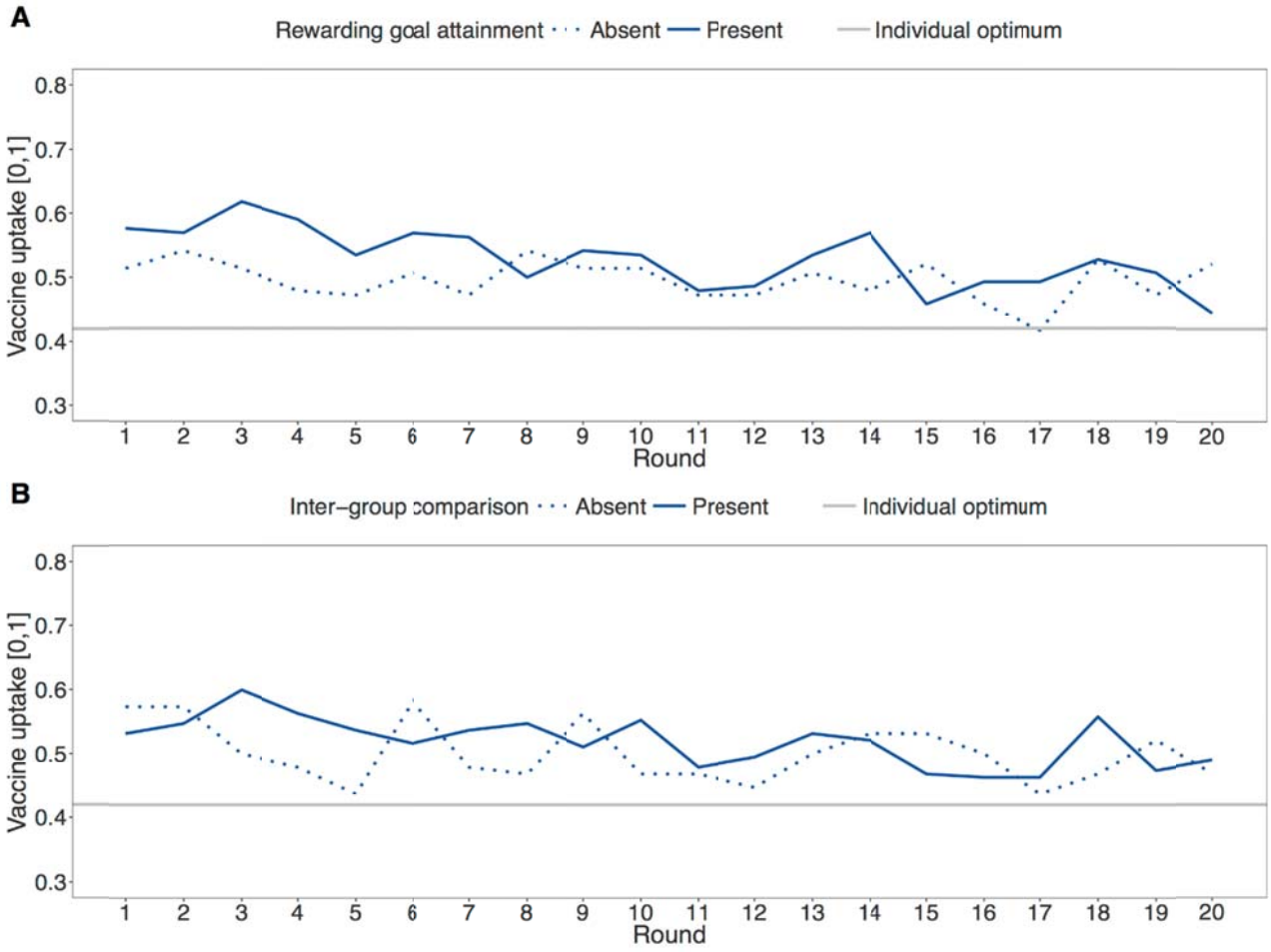


Figure 1. Vaccine uptake as a function of rewarding goal-attainment (A) and inter-group comparison (B).

5R-1 MISS RATE FOR COLORECTAL ADENOMA OF DAILY PRACTICE CONVENTIONAL COLONOSCOPY DETERMINED BY TANDEM STUDIES: A SYSTEMATIC REVIEW AND META-ANALYSIS

Health Services, Outcomes and Policy Research (HSOP)

Marvin Bundo, MD[1], Beate Jahn, PhD[1], Marjan Arvandi, MS[2], Gaby Sroczynski, MPH, Dr.PH[1] and Uwe Siebert, MD, MPH, MSc, ScD[3], (1)Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (2)Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, (3)UMIT, Dept. of Public Health, Health Services Research & HTA, Harvard Univ., Dept. of Health Policy & Management, Institute for Technology Assessment, Hall i.T./Boston, MA

Purpose: Accurate estimation of adenoma miss rate (AMR) of conventional colonoscopy (CC) is important as CC is one of the important tools for screening and surveillance for colorectal cancer (CRC). The aim of this study is to combine the existing evidence and to obtain a pooled estimate for AMR (p-AMR) for CC as determined by tandem studies, which are the best method for assessing AMRs.

Method(s): We conducted a systematic search in PubMed, Embase and Cochrane Library with no time restriction. We included randomized same-day tandem colonoscopy studies with cross-over design, comparing CC to novel colonoscopic technologies (NCTs) in patients undergoing daily practice colonoscopy. We performed meta-analyses using a generalized linear mixed model to estimate AMR with 95% confidence intervals (95%CI). We conducted subgroup analyses for size, histology and shape of adenomas. A likelihood ratio test (LRT), I^2 statistics and meta-regressions were used to address heterogeneity of the studies.

Result(s): Eleven studies with a total of 1,324 patients were included. In most of the studies, quality-adjusted colonoscopy was performed (i.e., optimal bowel preparation, experienced endoscopists, high quality of image). The overall p-AMR was 34% (95%CI: 28-40%). Subgroup analyses showed p-AMRs: 1) by size 37% (95%CI: 32-43%) for 1-5 mm, 28% (95%CI: 20-39%) for 6-9 mm and 12% (95%CI: 7-20%) for ≥10 mm adenomas (7 studies; 852 patients) 2) by histology : 41% (95%CI: 36-48%) for non-advanced and 19% (95%CI: 11-30%) for advanced adenomas (6 studies; 593 patients) 3) by shape: 51% (95%CI: 33-69%) for flat and 28% (95%CI: 21-39%) for polypoid (5 studies; 783 patients). Smaller, non-advanced and flat adenomas were significantly missed more by CC ($p< 0.05$). Significantly more adenomas were missed by CC in comparison to NCTs (OR=3.55; $p< 0.05$).There was significant (LRT: $p< 0.05$) and substantial heterogeneity ($I^2 = 79.5\%$) for the AMR but none of the tested covariates (prevalence, indication, NCT type) were significant.

Conclusion(s): Our study suggests that during daily practice CC one third of adenomas could be overlooked. A considerably high number of high risk adenomas (≥10 mm, advanced) is also missed by CC. The performance of CC depends on size, histology and shape of the adenoma. The study suggests also a worse performance of CC compared to NCTs, the implementation of which should be investigated.

5R-2 CALIBRATING MODELS OF CANCER TREATMENT TO OVERALL AND PROGRESSION-FREE SURVIVAL CURVES: CONSISTENTLY REFLECTING UNCERTAINTY AND HETEROGENEITY

Quantitative Methods and Theoretical Developments (QMTD)

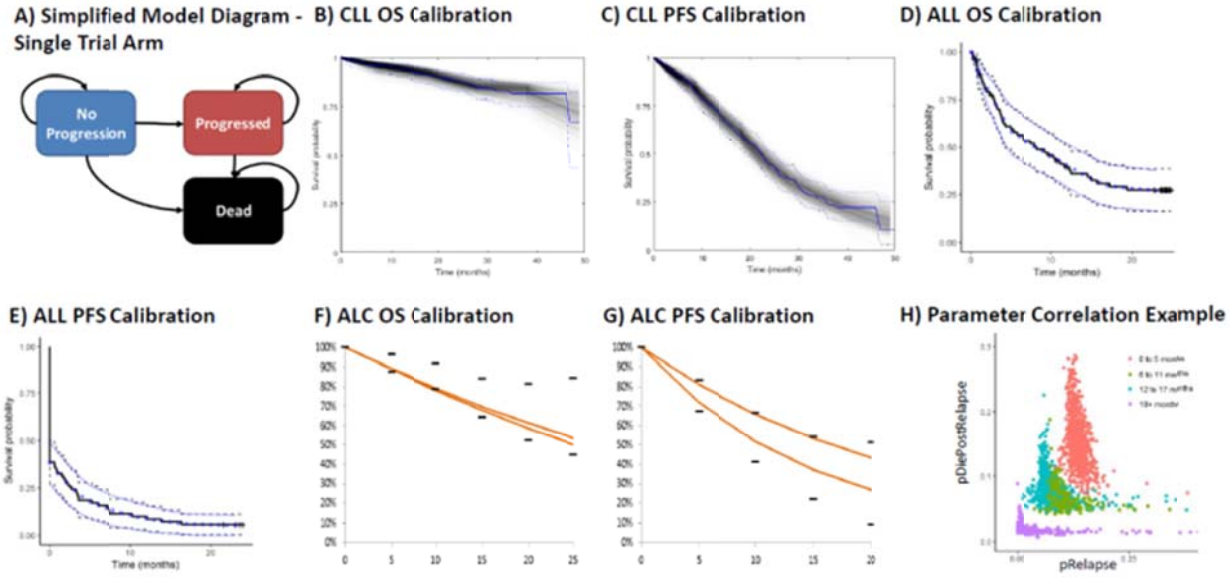
James Barnes, MD, MS, John Lin, MD and **Jeremy D. Goldhaber-Fiebert, PhD**, Stanford University, Stanford, CA

Purpose: New cancer therapies have improved efficacies but often very high prices. Assessing cost-effectiveness is critical, requiring simulation models because: 1) trials do not directly compare all treatments; 2) survival curves for new treatments must be extrapolated; 3) post-progression survival in older trials may depend on treatments that are no longer standard-of-care. When individual patient data are unavailable, models are difficult to parameterize because trials report Overall Survival (OS) and Progression-Free Survival (PFS), while the model requires risks of progression, pre-progression mortality, and post-progression mortality. Partitioned survival models (PSM) are commonly used, but PSMs require costs and utilities to be constant (per time period), which is a challenging assumption for many chemotherapies. One implication of this assumption is that PSMs cannot easily accommodate more than 2 therapy lines. We developed methods to address these issues, demonstrating feasibility and utility for treatments for Chronic Lymphocytic Leukemia (CLL), Relapsed/Refractory Pediatric B-Cell Acute Lymphoblastic Leukemia (ALL), and ALK-positive Lung Cancer (ALC).

Method(s): For each cancer and treatment, we extract OS and PFS survival curves using a validated digitizer. We derive exponential Greenwood confidence intervals, converting the digitized curves to failure times for trial participants and fitting a Kaplan-Meier curve using a validated method. We sample repeatedly from the confidence intervals to generate many sets of calibration targets (OS and PFS curves), ensuring that OS is greater than PFS using a published method. For each target set sampled, we use interior-point optimization algorithms to calibrate the model's uncertain parameters (i.e., risks of progression, pre-progression mortality, and post-progression mortality) (Figure 1, Panel A). For our example cancers, we extend this approach: 1) allowing time-varying risks; 2) reflecting patient heterogeneity (e.g., host or cancer genetic subtypes; history of stem cell transplantation); 3) accounting for post-progression treatments and treatment failures.

Result(s): Our method produces OS and PFS curves and uncertainty regions consistent with clinical trials for 3 cancers (e.g., Figure1, Panels B-G). These curves result from the model's calibrated joint posterior uncertainty distributions of inputs that can be used for probabilistic sensitivity analyses (e.g., Figure 1, Panel H).

Conclusion(s): Given cancer trial reporting practices, our method provides a feasible approach, enabling cost-effectiveness analyses of new expensive, effective cancer treatments and assessment of the uncertainty in such evaluations.


A) Simplified Model Diagram - Single Trial Arm
No Progression
Progressed
Dead
B) CLL OS Calibration
C) CLL PFS Calibration
D) ALL OS Calibration
E) ALL PFS Calibration
F) ALC OS Calibration
G) ALC PFS Calibration
H) Parameter Correlation Example
pDiePostRelapse
pRelapse

5R-3 RISK STRATIFICATION OF HPV-POSITIVE WOMEN IN CERVICAL CANCER SCREENING USING RESULTS FROM TWO CONSECUTIVE SCREENING ROUNDS

Health Services, Outcomes and Policy Research (HSOP)

Nicole Polman, Nienke Veldhuijzen, Danielle Heideman, Peter Snijders, Chris Meijer and **Johannes Berkhof**, VU University Medical Center, Amsterdam, Netherlands

Purpose: The replacement of cytology-based screening by human papillomavirus (HPV)-based screening for cervical cancer may lead to an increase in screening-related harms because many HPV infections have a transient nature. Adjunct testing of HPV-positive women is advisable to limit the number of colposcopy referrals and treatments. We studied whether the balance between benefits and harms of screening can be improved by managing HPV-positive women on the basis of results from two consecutive HPV-based screening rounds.

Method(s): We conducted a post-hoc analysis within a cohort of 21,996 Dutch women screened with cytology and HPV co-testing at baseline and at the second screen five years later (POBASCAM trial). We included only women who tested HPV-positive in the second screening round (n=366). Histology was collected up to nine years after baseline. Missing screening test results were imputed under the missing at random assumption. We evaluated the performance of sixteen adjunct testing strategies for detection of cervical precancer (CIN3+) in the second HPV-based screening round and examined whether risk stratification can be improved by including information from two consecutive screening rounds.

Result(s): The current Dutch policy for HPV-positive women, which consists of cytology and repeat cytology at six months, has a sensitivity of 79% for detecting CIN3+ and a colposcopy referral rate of 39%. Adding the HPV test result of the previous screen while omitting the six month repeat cytology test yields a CIN3+ sensitivity of 85% and a colposcopy referral rate of 39%. Adding repeat cytology at six months to the latter strategy yields a CIN3+ sensitivity of 90% but also a colposcopy referral rate of 47%.

Conclusion(s): The HPV test result of the previous screen is a useful marker for risk stratification of HPV-positive women that allows us to omit repeat cytology testing at six months without a decrease in CIN3+ detection. This illustrates that HPV-based cervical screening programmes can be further optimized by stratifying women on the basis of results from multiple screening rounds.

5R-4 COST-EFFECTIVENESS OF OVARIAN CANCER SCREENING IN AUSTRIA: A MODEL-BASED ECONOMIC EVALUATION

Applied Health Economics (AHE)

Artemisa Gogollari, Mag.sc.hum. Dr. med., Department of Public Health, Health Services Research and Health Technology Assessment, UMIT and ONCOTYROL - Center for Personalized Cancer Medicine, Division of Health Technology Assessment and Bioinformatics, Hall in Tirol, Austria, Gaby Sroczynski, MPH, Dr.PH, Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria, Wilhelm Oberaigner, Assoc.-Prof. Priv.-Doz. Dr., UMIT - University for Health Sciences, Medical Informatics and Technology, Institute of Public Health, Medical Decision Making and HTA, Department of Public Health, Health Services Research and HTA, Hall i.T., Austria and Uwe Siebert, MD, MPH, MSc, ScD, UMIT, Dept. of Public Health, Health Services Research & HTA / Harvard Univ., Dept. Health Policy & Management, Institute for Technology Assessment, Hall i.T. (Austria)/Boston (USA), Austria

Purpose: Our objective was to evaluate the long-term clinical and health-economic consequences of different ovarian cancer screening strategies in postmenopausal women for the Austrian health care context.

Method(s): A decision-analytic Markov cohort simulation model was developed to simulate the natural history of ovarian carcinogenesis and calibrated for the Austrian epidemiological and clinical context. We evaluated (1) annual multimodal screening with CA-125 and transvaginal ultrasound as a follow-up test (MMS), (2) annual ultrasound screening (USS), and (3) no screening (current practice) in asymptomatic postmenopausal women with screening starting at age 50 and ending with 85 years. We used Austrian epidemiological cancer data from the Tyrolean Cancer Registry, overall mortality from other causes from Austria's statistics bureau. Direct medical costs were derived mainly from Austrian sources. Cost data were transformed into 2016 Euros using gross domestic product purchasing power parity and the consumer price index. Predicted outcomes included the relative reduction in cancer mortality, number of false-positive tests and unnecessary oophorectomies, remaining life expectancy and quality-adjusted life expectancy, lifetime costs, and the discounted incremental cost-effectiveness ratio (ICER in Euro/LYG or Euro/QALY). The analyses were conducted for a lifelong time horizon applying a 5% annual discount rate for costs and effects and adopting the health care payer perspective. We considered a willingness-to-pay threshold of 50,000 Euro/LYG. Additionally, deterministic sensitivity analyses were performed to assess uncertainty.

Result(s): In the base-case analysis, MMS was the most effective strategy in terms of remaining life expectancy yielding an ICER of 46,000 Euro/LYG (51,000 Euro/QALYG) compared to no screening. USS was dominated and resulted in higher total costs and more false-positive cases compared to MMS. With USS screening, eleven women would need to undergo unnecessary oophorectomy in order to prevent one ovarian cancer death compared with three women with MMS screening. In sensitivity analyses, results were sensitive to discount rates above 5%, attendance rates, test costs, reduced test accuracy, and starting screening earlier.

Conclusion(s): Our decision analysis suggest, that annual multimodal ovarian cancer screening in Austrian women age 50-85 years has the potential to increase life expectancy and may be considered cost-effective compared to no screening. However, further research is needed to investigate the screening test related work-up strategies.

5R-5 A CONCEPTUAL MODEL FOR TECHNOLOGY ASSESSMENT OF NEXT GENERATION SEQUENCING IN PERSONALIZED ONCOLOGY (TANGO)

Health Services, Outcomes and Policy Research (HSOP)

Valesca Retel, PhD[1], Veerle MH Coupe[2], Talitha Feenstra, PhD[3], Prof. Maarten J. IJzerman, PhD[4], Hendrik Koffijberg, PhD[4], Manuela Joore, PhD[5], Carin Uyl-de Groot, PhD[6], Geert Frederix, PhD[7], Michiel van de Ven, Msc[4], Martijn Simons, Msc[8], Clemence Pasmans, Msc[7] and Wim van Harten, MD PhD[9], (1)Netherlands Cancer Institute, Amsterdam, Netherlands, (2)VU University Medical Center, Amsterdam, Netherlands, (3)RIVM, Bilthoven, Netherlands, (4)University of Twente, Enschede, Netherlands, (5)Department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Medical Center, Maastricht, Netherlands, (6)Erasmus University Rotterdam, Rotterdam, Netherlands, (7)University Medical Center Utrecht, Utrecht, Netherlands, (8)Maastricht University Medical Center, Maastricht, Netherlands, (9)University of Twente and the Netherlands Cancer Institute, Amsterdam, Netherlands

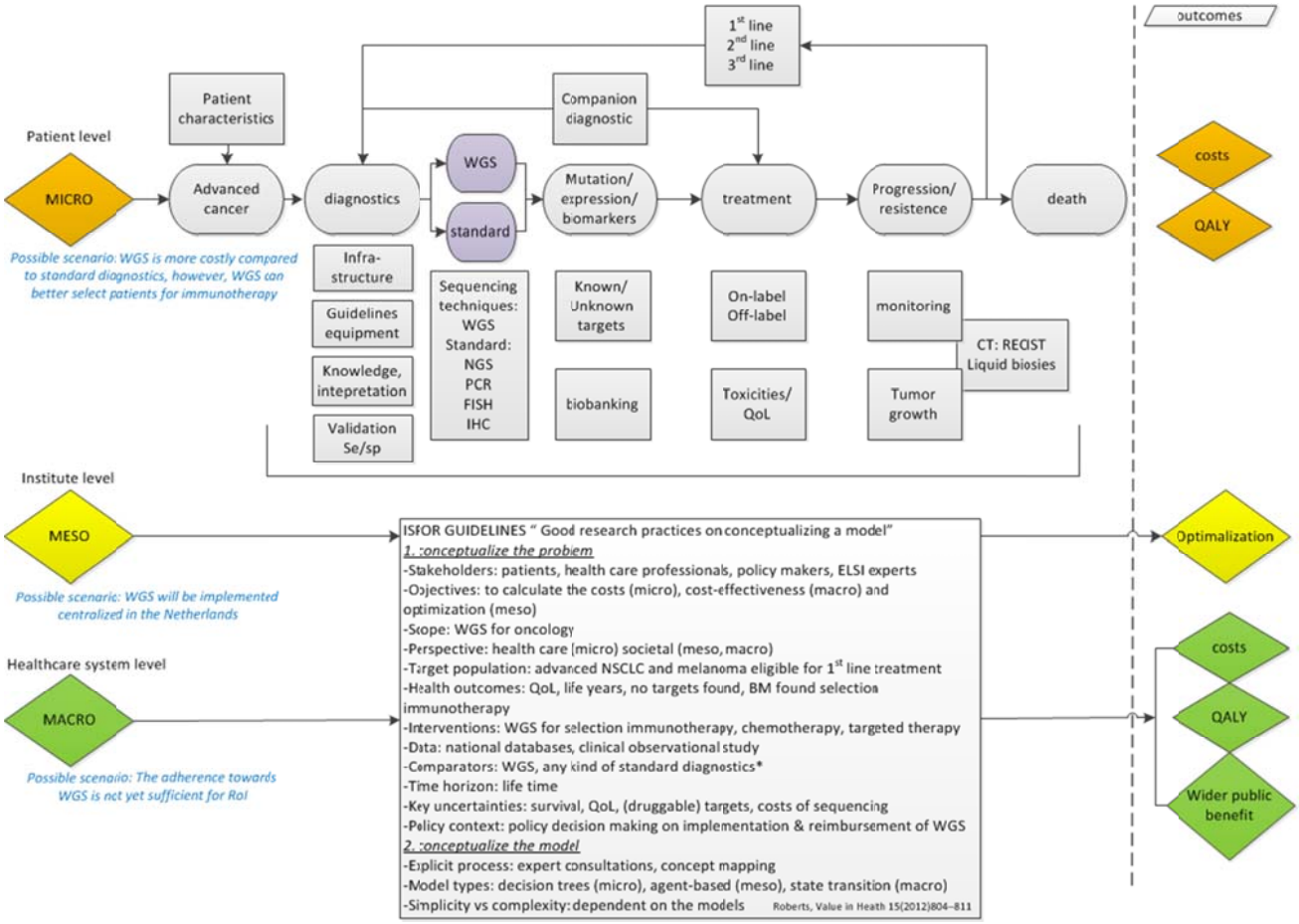
Purpose: To develop a conceptual model for Technology Assessment (TA) of Next Generation Sequencing (NGS) in Personalized Oncology (TANGO), to support optimal and cost-effective implementation of Whole Genome Sequencing (WGS) in the Netherlands.

Method(s): To decide upon implementing WGS in clinical practice, the ongoing TANGO project was initiated with collaboration of several consortia in fields of genetics, medicine, bio statistics, legal, ethical experts and health economists. The TANGO project consists of 6 Work Packages (WPs): 1)Diagnostic value of WGS, 2)Treatment decisions based on WGS, 3)Prediction of long-term health benefits and harms by micro-simulation, 4)Tumor-overarching early cost-effectiveness modelling, 5)Nation-wide organization of WGS, 6) Responsible implementation of WGS according to ethical, legal and societal implications (ELSI) principles. In order to align the TA WPs (1,3,4,5) and reach an aggregated conclusion on cost-effective implementation, a conceptual model was used as a starting point. We approached this model from three different angels; using the ISPOR conceptual model guideline for the content, scenario drafting for future perspective, and micro/meso/macro level for the different layers in policy making.

Result(s): Based on the ISPOR guidelines, the "base-case" prerequisites of the problem were identified (Figure 1). The base-case pathway for advanced non-small cell lung cancer (NSCLC) and melanoma was determined, including health outcomes, interventions, comparators. Subsequently, diversion from the base-case was set by means of the different levels (micro/meso/macro) and scenarios. On micro-level; WP1 & 3 used individual patient data to map the total pathway, to inform on test costs and optimal testing strategies based on tumor-growth. Possible scenario "WGS is more costly compared to standard diagnostics, however, can better select patients for immunotherapy" was identified. On meso-level; WP5 will use a system dynamic approach, to seek the most optimal infrastructure and logistics for implementing WGS. A possible scenario of a central and de-central perspective was identified. On macro-level; WP4 will perform a cost-effectiveness and budget impact analysis using a Markov modeling "tumor-overarching" approach. In this WP, different endpoints such as "wired public benefits" were identified. Possible scenario of adherence towards WGS testing was taken into account.

Conclusion(s): The initiative and practice of this project is innovative. This conceptual model can act as a basis for (future) Technology Assessments in decision making of NGS in oncology.

Figure 1:



5R-6 ARE PROPHYLACTIC SURGERIES AN EFFECTIVE AND COST-EFFECTIVE STRATEGY FOR BREAST AND OVARIAN CANCER PREVENTION IN GERMAN WOMEN WITH BRCA-1/2 MUTATION? - A DECISION-ANALYTIC EVALUATION

Applied Health Economics (AHE)

Lara R. Hallsson, MPH, Dr., Department of Public Health, Health Services Research and Health Technology Assessment, UMIT and ONCOTYROL - Center for Personalized Cancer Medicine, Division of Health Technology Assessment and Bioinformatics, Hall in Tirol, Austria, Gaby Sroczynski, MPH, Dr.PH, Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria, Jutta Engel, Prof., Dr.med., MPH, Ludwig-Maximilians-University, Munich, Germany, Martin Widschwendter, PhD, UCL - University College London, Department of Women's Cancer, London, United Kingdom and Uwe Siebert, MD, MPH, MSc, ScD, UMIT, Dept. of Public Health, Health Services Research & Health Technology Assessment / Harvard Univ., Dept. Health Policy & Management, Institute for Technology Assessment / ONCOTYROL, Division for HTA, Hall i. T. (Austria) / Boston (USA), Austria

Purpose: The aim of this study was to systematically evaluate the long-term effectiveness and cost-effectiveness of different strategies to prevent breast and ovarian cancer including prophylactic bilateral mastectomy (PBM) and prophylactic bilateral salpingo-oophorectomy (PBSO) compared to intensified surveillance (IS) in German BRCA-1/2 gene mutation carriers.

Method(s): A decision-analytic Markov model simulating breast and ovarian cancer development in BRCA-1/2 mutation carriers was developed and applied to evaluate the following different strategies: (1) IS, (2) PBM at age 30, (3) PBM at age 40, (4) PBSO at age 30, (5) PBSO at age 40, (6) PBM plus PBSO at age 30 or (7) PBM plus PBSO at age 40. The model was parameterized, calibrated and validated with German epidemiological data from German cancer registries and the Federal Statistical Office (Destatis) as well as with data from published literature. Assessed outcomes include relative reduction in cancer incidences and mortality (in %), remaining life years (LYs), quality-adjusted life years (QALYs), total costs (€), and discounted incremental cost-effectiveness ratios (ICER). The German health care system perspective was adopted with a 3% annual discount rate for costs and health effects. To assess robustness of the results, extensive deterministic sensitivity analyses were performed.

Result(s): The combination PBM plus PBSO for 30-year old German women with BRCA-1/2 mutation was the most effective strategy compared to PBM or PBSO alone or IS, followed by PBM plus PBSO at age 40 in the base-case analysis. The undiscounted remaining life expectancy increased by 6.2 LYs (8.6 QALYs) for PBM plus PBSO at age 30 and by 5.0 LYs (7.1 QALY) for PBM plus PBSO at age 40, compared to IS. In the economic analysis, PBM plus PBSO at age 30 dominated all other strategies saving discounted € 22,250 compared to IS. Multiple sensitivity analyses showed model results to be robust against variation in all relevant parameters.

Conclusion(s): Based on this decision-analysis, German women with BRCA-1/2 mutation benefit from PBM plus PBSO between age 30 and 40. PBM plus PBSO at age 30 is likely cost saving compared with other prevention strategies or intensified surveillance. However, individual preferences external to our analyses, such as a woman's family planning situation, must be considered in the final decision.

TRA 1-1 THE IMPACT OF DESCRIPTIVE NORMS ON MOTIVATION TO PARTICIPATE IN CANCER SCREENING

Decision Psychology and Shared Decision Making (DEC)

Sandro Tiziano Stoffel, Alex Ghanouni, Yasemin Hirst, Jo Waller and Christian von Wagner, University College London, London, United Kingdom

Purpose:

There is a strong relationship between descriptive norms and cancer screening participation. Thus informing people that uptake is high could motivate them to attend. The current study presented uptake information as part of a hypothetical vignette about Bowel Scope Screening (BSS). The aim of this study was to test whether manipulating normative beliefs increases intention to take part among a group of previously disinclined individuals.

Method(s):

We recruited 7,912 men and women aged 35-54 living in England from an online panel. After an explanation of BSS, 1,733 (21.8%) stated that they did not intend to participate. Of these 1,494 (86.2%) estimated uptake as less than 6 out of 10 ($x<6$) and were randomised to one of four social norm conditions: (1) *Echo and confirm* ('you guessed uptake is *x*, uptake is *x*'), (2) *Echo with proportional augmentation* ('you guessed *x*; uptake is *x*+3'); (3) *Echo with standard augmentation* (you guessed *x*; uptake is 8 out of 10); (4) *standard augmentation alone* ('uptake is 8 out of 10'). We used multivariable logistic regression adjusting for sociodemographic variables to predict intention by condition and report adjusted odds ratio (aOR) and 95% confidence intervals (CI).

Result(s):

1,432 of 1,494 (95.9%) respondents completed all comprehension and manipulation checks. Most were female (62.2%), White-British (80.5%), married or living with someone (59.0%) and working (70.4%). The three augmented messages led to a significantly greater proportion of BSS intenders (17.8%, 27.7% and 28.2% for conditions 2, 3 and 4 respectively) compared with 'Echo and Confirm, Condition1' (10.5%); (aOR 1.91; CI 1.23-2.96; aOR 3.42; 2.25-5.20; aOR 2.27; CI 2.27-5.31). Conditions 3 and 4 were also associated with a significantly greater proportion of intenders compared with Condition 2 (aORs 3.42 vs 1.91; $p<0.01$; aORs 3.47 vs 1.91; $p<0.01$).

Conclusion(s):

High uptake messages are positively associated with screening intention. Future randomised controlled trials should evaluate its impact on screening attendance.

TRA 1-2 CROSS-NATIONAL PSYCHOLOGICAL AND DEMOGRAPHIC ASSOCIATIONS WITH MEDICAL MAXIMIZING-MINIMIZING TENDENCIES

Decision Psychology and Shared Decision Making (DEC)

Aaron M. Scherer, PhD[1], Megan Knaus, MPH[2], Brian J. Zikmund-Fisher, PhD[2], Enny Das, PhD[3], Laura D. Scherer, PhD[4] and Angela Fagerlin, PhD[5], (1)University of Iowa, Iowa City, IA, (2)University of Michigan, Ann Arbor, MI, (3)Radboud University Nijmegen, Nijmgen, Netherlands, (4)University of Missouri, Columbia, MO, (5)University of Utah / Salt Lake City VA, Salt Lake City, UT

Purpose: Medical maximizing-minimizing is a recently developed construct that refers to a person's predisposition for seeking medical interventions regardless of symptom severity (medical maximizing) or avoiding medical interventions unless absolutely necessary (medical minimizing). The objectives of the current analyses were to examine a) the predictive validity of the Medical Maximizer-Minimizer Scale (MMS) in non-United States countries and b) psychological traits that may be associated with the MMS.

Method(s): We conducted an online survey in 10 European countries (Finland, Germany, Hungary, Italy, Netherlands, Norway, Poland, Spain, Sweden, and the United Kingdom) and the United States using stratified random sampling based on age and gender. Embedded in the survey were measures of medical maximizing-minimizing, medical ambiguity aversion, emotional reactivity to rare events, subjective numeracy, health literacy, and subjective health, along with demographics (age, gender, education, healthcare worker). Data were analyzed by regressing the reported measures and country on self-reported influenza vaccination and on MMS score.

Result(s): The survey was completed by 16,510 participants (81% completion rate). Consistent with predictions, each point increase on the MMS was associated with an increase in the odds of having received the seasonal influenza vaccine (OR=1.56, 95% CI=1.49, 1.62). Higher ambiguity aversion, emotional reactivity to rare events, numeracy, health literacy, older age, and being a healthcare worker were associated with higher medical maximizing tendencies, while being male, more educated, and having better subjective health were associated with increased medical minimizing (see right-hand column in Table 1). In comparison to the United States, residents of Italy, Poland, and Spain had higher medical maximizing tendencies, while residents of the Netherlands, Sweden, and the United Kingdom had lower medical maximizing tendencies (see right-hand column in Table 1).

Conclusion(s): The current study provides preliminary evidence that the MMS has predictive validity outside of the United States. Additionally, it provides suggestive evidence that a person's medical maximizing or minimizing tendency may be rooted in their aversion to medical ambiguity and emotional reactions to the potential of experiencing rare medical events. Future research should focus on how medical maximizing-minimizing influences medical decision making in other contexts and the antecedents to maximizing-minimizing tendencies.

Table 1. Odds ratios and beta values for associations with influenza vaccine uptake and medical maximizing-minimizing score.

	Influenza Vaccination OR (95% CI)	Maximizing-Minimizing b (p-value)
Maximizing-Minimizing	1.65 (1.58, 1.72)	*n/a*
Ambiguity Aversion	0.67 (0.62, 0.74)	0.05 (<.001)
Emotional Reactivity	0.92 (0.89, 0.95)	0.16 (<.001)
Numeracy	1.10 (1.05, 1.16)	0.07 (<.001)
Health Literacy	1.26 (1.20, 1.33)	0.16 (<.001)
Subjective Health	0.91 (0.87, 0.95)	-0.03 (.001)
Gender	0.84 (0.78, 0.91)	-0.03 (<.001)
Age	1.03 (1.03, 1.04)	0.12 (<.001)
Education	1.08 (1.05, 1.10)	-0.02 (.007)
Healthcare Worker	2.64 (2.32, 3.00)	0.03 (<.001)
Countries		
United States	*ref*	*ref*
Finland	0.47 (0.40, 0.56)	-0.00 (.776)
Germany	0.34 (0.29, 0.40)	-0.01 (.153)
Hungary	0.20 (0.16, 0.25)	0.00 (.641)
Italy	0.24 (0.20, 0.28)	0.03 (.005)
Netherlands	0.57 (0.49, 0.67)	-0.15 (<.001)
Norway	0.18 (0.14, 0.23)	-0.01 (.156)
Poland	0.16 (0.13, 0.19)	0.09 (<.001)
Spain	0.26 (0.22, 0.31)	0.10 (<.001)
Sweden	0.33 (0.28, 0.40)	-0.09 (<.001)
United Kingdom	0.76 (0.64, 0.90)	-0.09 (<.001)

TRA 1-3 COST-SENSITIVE MEASURES VS. YOUDEN`S INDEX TO DETERMINE AN OPTIMAL THRESHOLD FOR BIOMARKERS: A PRACTICAL APPROACH TO DEFINE PRE-DIABETES

Decision Psychology and Shared Decision Making (DEC)

Davood Khalili, MD PhD, Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran (Islamic Republic of) and Farideh Bagherzadeh-Khiabani, Msc, Research Institute for Endocrine Sciences, Tehran, Iran (Islamic Republic of)

Purpose: Cut-off selection for continuous biomarkers is still an attractive practice in clinical contexts. Cutoff points are commonly estimated using simple measures such as Youden`s index which do not consider clinical consequences of decisions. However, cost-sensitive measures could estimate appropriate cut-off points based on physician/patients' preferences i.e. based on a cost jointly agreed upon by physician and/or patient.

Method(s): We used Youden`s index and cost-sensitive measures including Generalized-Youden (GY), Misclassification-Cost-Term (MCT), and Net-Benefit-Fraction (NBF) to determine a threshold for fasting plasma glucose (FPG) and define pre-diabetes as a high-risk sate for diabetes. We used convertible terms such as cost, harm-to-benefit ratio and threshold probability for treatment to calculate these measures. A cohort of 1212 men and 1758 women aged 20-60 years, from the Tehran Lipid and Glucose Study, were included for this analysis. FPG was measured at the baseline and the incidence of type 2 diabetes, during 12 years of follow-up, was considered as the outcome.

Result(s): All cost-sensitive measures found the same optimal cut-off point and this consistency was proved mathematically. The area under the curve (AUC) showed a discrimination power of 0.77 (95% CI: 0.73-0.81) for FPG to predict incidence of diabetes in men; this value was 0.79 (0.76-0.82) in women. The Youden`s index resulted in the threshold of 94 for men (sensitivity=0.73, specificity=0.71) and 95 for women (sensitivity=0.62, specificity=0.84); however, cost-sensitive measures showed different thresholds for a range of threshold probability for treatment of 0-0.5, corresponding to harm-to-benefit ratios of 0-1 (odds of the thresholds). The preferred threshold probability of %10 for preventive treatments, resulted in FPG cutoffs of 94 in men and 89 in women. These cutoffs correspond to net benefit fraction of 0.58 in both genders; this is sensitivity penalized for false positive rate with a weight of 1/9.

Conclusion(s): Since harm-to-benefit of interventions may differ in different patients or subgroups, cost-sensitive measures are preferred to Youden`s Index for selection of an optimal cutoff. The probability threshold for treatment is a simple index which is defined as the specific probability of disease at which the clinician prefers to make intervention. The NBF as an tangible measure uses this index; it is defined as the fraction of the incidence rate that could be predicted and prevented appropriately regarding harm-to-benefit of treatment.

TRA 2-1 SUB-CATAGORISING EXPECTED VALUE OF PERFECT IMPLEMENTATION (EVPIM) TO IDENTIFY WHEN AND WHERE TO INVEST IN IMPLEMENTATION INITIATIVES

Quantitative Methods and Theoretical Developments (QMTD)

Kasper Johannesen, MSc, Magnus Janzon, MD and Martin Henriksson, PhD, Linkoping University, Linkoping, Sweden

Purpose:

The expected value of perfect implementation (EVPIM) is a framework for estimating the value of improving implementation of healthcare technologies. The purpose of this work was to sub-categorise EVPIM in order to identify where policy interventions is expected to have greatest value. The implementation of a new antiplatelet therapy for the treatment of myocardial infarction (MI) in Sweden was used as an illustrative example.

Method(s):

Using the established EVPIM framework, EVPIM was estimated for a new antiplatelet MI therapy, based on observed implementation patterns in Sweden between 2011 and 2015. In addition, EVPIM was sub-categorised to estimate the expected value of: A) eliminating regional variation in implementation; B) implementing in remaining patients treated with other antiplatelet treatment; C) eliminating implementation delay; and D) implementing in patients not receiving any treatment.

Published estimates of incremental treatment effect and cost of drug therapy, and a threshold of 250.000 SEK (≈≤22,500) per QALY, were used to estimate incremental net health benefits (INHB) and incremental net monetary benefits (INMB).

Result(s):

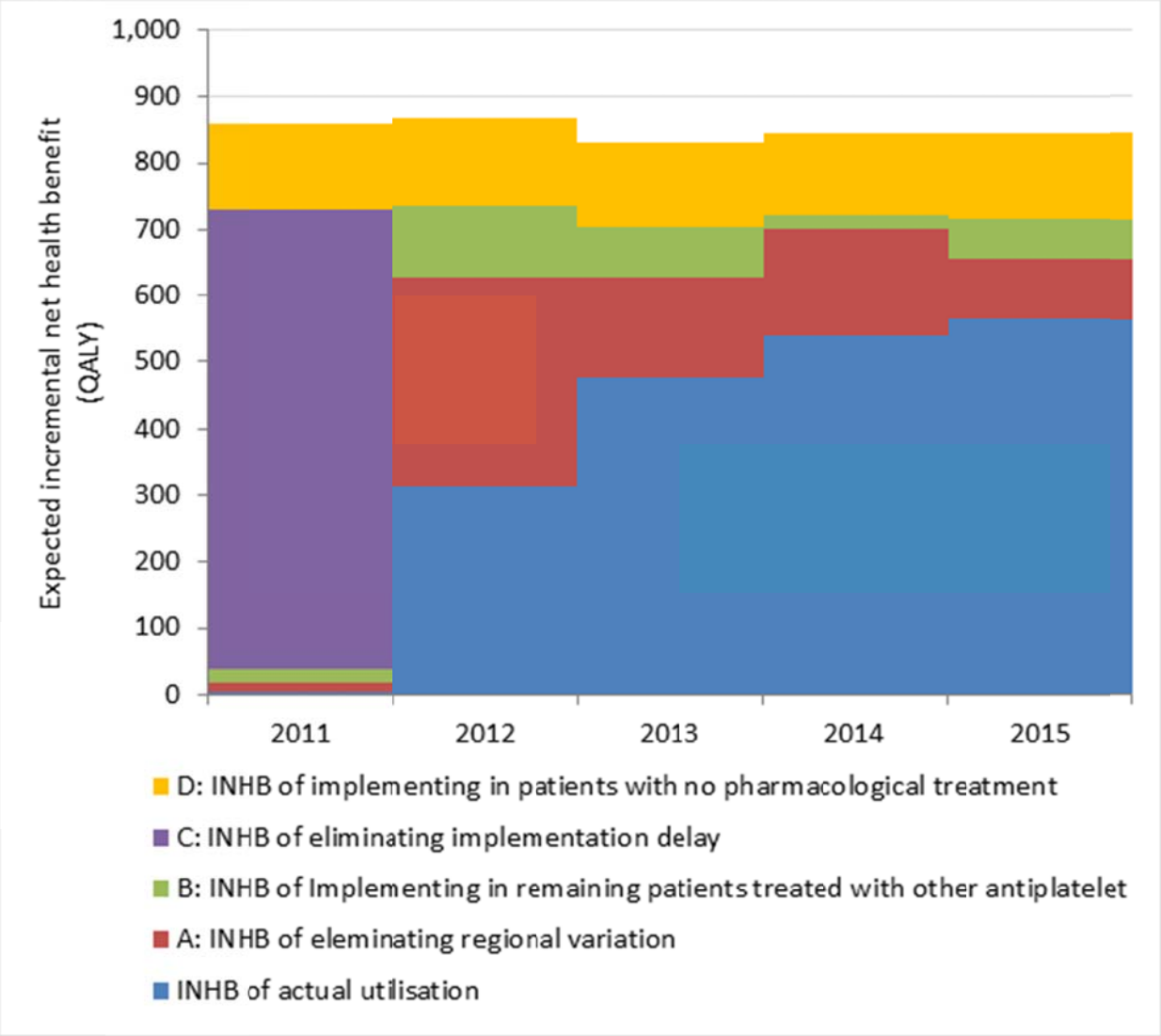
Out of 65,672 MI patients, 29.318 were treated with the new antiplatelet therapy. The total EVPIM was estimated to 4,254 QALYs or SEK 1.1 billion.

An estimated additional 11,303 patients would have been treated if all healthcare regions had implemented the antiplatelet therapy to the same extend as the regions with highest implementation in each year, representing an INHB of 1,322 QALYs and 31% of the total EVPIM (area A). Further 4,313 patients would have been treated if all healthcare regions had further increased implementation to all patients receiving any antiplatelet treatment, representing 505 QALYs and 12% of the total EVPIM (area B). Eliminating the delay in implementation would have resulted in 10,707 more patients treated and an INHB of 1,253 QALYs, 30% of total EVPIM (area C). Implementing the treatment in all patients with no current antiplatelet treatment would have resulted in 10,031 more patients treated, 1,174 QALYs, 28% of total EVPI (area D).

Conclusion(s):

Sub-categorizing EVPIM illustrates the value of eliminating different implementation challenges. This is crucial for determining when and where there are greatest returns of investing in implementation improvements.

Figure: Expected INHB of observed utilisation and different EVPIM sub-categories from 2015-2011



TRA 2-2 TWO-STEP REGRESSION APPROACH FOR INCLUDING COVARIATE ADJUSTED MEAN UTILITIES' UNCERTAINTY IN COST-EFFECTIVENESS MODELS

Applied Health Economics (AHE)

Arantzazu Arrospide, MSc, PhD, Integrated Health Organization Alto Deba, Osakidetza, REDISSEC., Arrasate, Spain, Juan Manuel Ramos-Goni, EuroQol Research Foundation, REDISSEC, Rotterdam, Netherlands, Petros Pechlivanoglou, Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, ON, Canada, Igor Larranaga, Integrated Health Organisation Alto Deba, Arrasate, Spain and Javier Mar, MD, PhD, Alto Deba Hospital, Mondragón, Spain

Purpose: To develop a statistical procedure to estimate covariate adjusted utilities taking into account both possible ceiling effect and negative values. This approach allows including utilities' second order uncertainty for correlated parameters in a probabilistic simulation model.

Method(s): We propose that mean utility values (u) can be estimated through a two part model: $u(x)=p(x)\cdot 1+(1-p(x))\cdot(1-w(x))$. First a logistic regression model is applied in order to estimate the probability (p) of utilities equal to 1, conditional on covariates (x) Subsequently, a generalized linear model can be applied on the disutility w=1-u for individuals not being in perfect health. This way estimated values would always be in the range $(-\infty,1]$.
To illustrate the method we used the Spanish Health Survey performed in 2012 (21,007 individuals, 53% women, mean age 51.3) including EuroQol-5L questionnaire. The two steps to obtain utilities described above were adjusted by age, sex, social class and body mass index category.

In order to include second order uncertainty in cost-effectiveness models, both coefficients (C_1,C_2) and Cholesky decomposition of variance-covariance matrix (T_1,T_2) were obtained from regression analyses. Using a vector of random values following a Gaussian distribution (Z_1,Z_2), random mean utility values can be obtained for each simulation: C_1*= C_1+(T_1*Z_1) and C_2*= C_2+(T_2*Z_2).

Result(s): Related to EuroQol 5D-5L, 63% of the sample were in perfect health and the mean utility value for the participants was 0.91. Based on the regression analyses (C_1,C_2, T_1,T_2) and random vectors (Z_1,Z_2), the coefficients for the first simulation in our model were:
C =(constant,age,age>70,female,middle social class,high social class,bmi 25-30,bmi>30,age*(age>70),age*female,middle social class*bmi 25-30,high social class*bmi 25-30,middle social class*bmi>30,high social class*bmi>30)
C_1* = (0.87,-0.04,0.88,-0.73,0.60,0.43,-0.01,-0.45,-0.02,-0.00,-0.21,0.03,-0.36,0.20)
C_2* = (-1.69,0.01,-0.50,0.10,-0.05,-0.29,-0.00,0.15,0.01,-0.00,-0.03,0.11,-0.04,0.25)
Therefore, the estimated mean utility value for, for example, men aged 60, high social class and BMI<25 would be 0.9559, calculated as:
$u= \exp(0.87 – 0.04 * (60-50)+ 0.43)/(1+\exp(0.87 – 0.04 * (60-50)+ 0.43))$
$+(1-\exp(0.87 – 0.04 * (60-50)+ 0.43)/(1+\exp(0.87 – 0.04 * (60-50)+ 0.43)))*(1-\exp(-1.69 + 0.01 (60-50)– 0.29))$
$=0.9559$

Conclusion(s): The proposed approach allowed estimating mean utility values based on individual characteristics and permitted managing also its associated uncertainty. In addition, it ensured proper range of utility values, i.e. $(-\infty,1]$, avoiding values higher than 1 as could occur using linear regression models.

TRA 2-3 FROM EVALUATION TO OPTIMIZATION: CAPACITY-CONSTRAINED OPTIMIZATION USING A META-MODEL TO MAXIMIZE THE BENEFITS OF COLORECTAL SCREENING

Quantitative Methods and Theoretical Developments (QMTD)

Hendrik Koffijberg, PhD[1], Veerle Coupe, PhD[2], Prof. Maarten J. IJzerman, PhD[1], Koen Degeling, MSc[1] and Marjolein JE Greuter[3], (1)University of Twente, Enschede, Netherlands, (2)Decision Modeling Center VUmc, Amsterdam, Netherlands, (3)VU University Medical Center, Amsterdam, Netherlands

Purpose:

Model-based analyses are typically useful for assessing the cost-effectiveness of a few but not a vast number of alternative health care strategies, because this is computationally infeasible, in particular when contraints apply. We illustrate the potential advantages of using a meta-model to identify the best screening strategy for colorectal cancer, while accounting for colonoscopy capacity constraints.

Method(s):

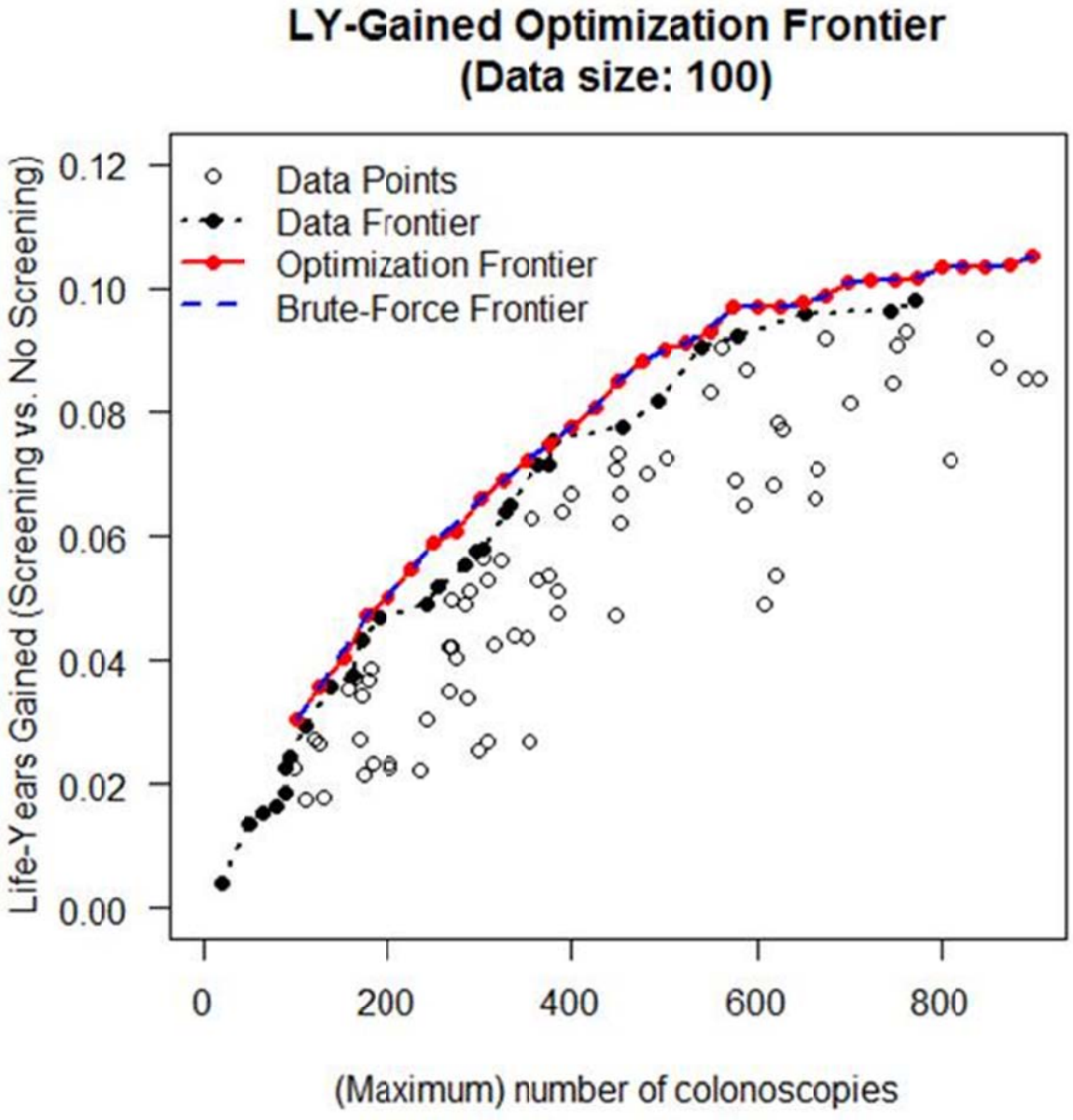
Screening strategies were defined by starting age, screening interval, number of screening rounds, and screening test positivity threshold (>70,000 unique strategies). We evaluated a limited sample of predefined strategies (n=200) with the validated ASCCA model and identified the best screening strategy therein, in terms of life-years gained (LYG), compared with no screening. In a first analysis, a Gaussian Process meta-model was fitted to this limited sample and then discrete evolutionary programming was applied to identify the best possible screening strategy (GP-DEP approach). The optimization was performed for different capacity constraints on the number of colonoscopies per 1,000 simulated individuals (n=100-900). In a second analysis, colonoscopy demand was restricted to 550/1,000 individuals, reflecting current colonoscopy capacity in the Netherlands, and the sample size of predefined strategies was varied (n=25-150). For each sample size, GP-DEP performance was assessed with bootstrapping (n=500), brute force exhaustive search, and comparison with ASCCA outcomes.

Result(s):

GP-DEP resulted in stable predicted best screening strategies when applied to a sample of >=100 strategies, identifying the exact same best strategy as exhaustive search in 96% of bootstrap samples. Compared with ASCCA, predicted colonoscopy demand, LYG and costs of the best strategies from GP-DEP were accurate, slightly too high and slightly too low, respectively. However, strategy ranking (in deciles) according to ASCCA and GP-DEP were similar. As presented in Figure 1, for sample size 100, GP-DEP resulted in better screening strategies (higher number of LYG) compared to just evaluating predefined strategies, for different capacity constraints. For sample size 100, average predicted benefit of the best strategy identified by GP-DEP compared to the best strategy identified by ASCCA equalled 0.028 LYG (95%CI 0.013-0.043) per individual.

Conclusion(s):

Creating a meta-model of the ASCCA model and applying GP-DEP enhances performance: the best screening strategy can be identified much faster, even when constraints apply, and will outperform the best screening strategy as typically identified from a limited sample of predefined strategies.



TRA 2-4 CAN IRELAND'S COLORECTAL SCREENING PROGRAMME SAVE MORE LIVES, SAVE MONEY AND LIVE WITHIN EXISTING COLONOSCOPY CAPACITY LIMITS? EVIDENCE FROM MICROSIMULATION MODELLING
Health Services, Outcomes and Policy Research (HSOP)

Ethna McFerran, BSc Hons Nursing / PGC Specialist Practice Oncology / PhD Candidate Health Economics[1], **James O'Mahony, PhD**[2], Duncan McVicar, PhD[3], Ann G. Zauber, PhD[4], Iris Lansdorp-Vogelaar, PhD[5] and Frank Kee, BSc, MSc, MD[3], (1)QUEENS UNIVERSITY BELFAST, BELFAST, United Kingdom, (2)Trinity College Dublin, Dublin, Ireland, (3)Queen's University Belfast, Belfast, United Kingdom, (4)Memorial Sloan-Kettering Cancer Center, New York, NY, (5)Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, Netherlands

Purpose: To determine if BowelScreen, Ireland's current population colorectal cancer screening programme, can be adjusted to make it more effective and less costly while not exceeding current colonoscopy capacity constraints.

Method(s): BowelScreen was established on the basis of a health technology assessment (HTA) conducted in 2009. BowelScreen employs biennial FIT screening. Ireland lacked sufficient colonoscopy capacity to provide the necessary number of follow-up tests required by the new programme. Accordingly, two programme modifications have been made to reduce the number of colonoscopies required: reducing the screening age range from 55-74 to 60-69; raising the FIT cut-off from 100 ng of haemoglobin /ml of buffer to 225. Prior literature indicates that reducing the screening age range and raising the FIT cut-off are not likely to be the most efficient ways of reducing colonoscopy requirements. We use the MISCAN microsimulation cancer screening model to expand the range of screening intervals, screening age ranges and FIT cut-offs beyond what was considered in the initial 2009 HTA. We simulate net costs, quality-adjusted life-years (QALYs), cancer deaths prevented and the number of colonoscopies required for 315 simulated strategies. The cost and effects of strategies that require less or equal colonoscopy capacity than the current strategy are compared to those of the status quo.

Result(s): Relative to the current strategy, a combination of reducing the FIT cut-off to 50ng Hb /ml and extending the screening interval to 4 years saves 15% more QALYs, prevents 13% more CRC deaths and yields a 6% reduction in colonoscopy requirements, which could be achieved with a modest 9% cost saving. Relaxing the constraint that a new policy has to cost less than the status quo permits the choice of a strategy that provides 37% more QALYs, prevents 29% more CRC deaths and still does not exceed the current colonoscopy capacity.

Conclusion(s): Simple changes to BowelScreen could save lives, reduce costs and relieve pressure on colonoscopy capacity. The extent of the potential improvements depends in part on the policy acceptability of lengthening the screening interval. These findings highlight the important methodological insight that a full range of policy alternatives should be considered when conducting cost-effectiveness analyses.