HT2
APPLICATION OF COST-EFFECTIVENESS LOGIC TO US MANAGED CARE DRUG FORMULARIES: LONG TERM OUTCOMES OF A VALUE-BASED FORMULARY
Yeung K[1], Basu A[2], Hansen RN[1], Watkins J[3], Sullivan SD[2]
[1]School of Pharmacy, University of Washington, Seattle, WA, USA, [2]University of Washington, Seattle, WA, USA, [3]Premera Blue Cross, Mountlake Terrace, WA, USA
OBJECTIVES: Cost-effectiveness analysis (CEA) is explicitly used for informing drug coverage decisions in many countries but not in the United States. Evidence suggests that failure to incorporate value considerations in coverage decisions may lead to reduced economic efficiency in the form of increased costs or worsened health outcomes. Yet the use of CEA in the context of binary coverage decisions (yes or no) may not be politically or socially feasible in the US. In 2010, Premera Blue Cross implemented a value-based formulary (VBF) that uses CEA to determine the copayment level-not binary coverage-for each drug in the formulary; drugs with lower incremental cost-effectiveness ratios (ICERs) are assigned lower copayments, drugs with higher ICERs are assigned higher copayments. The objective of this study is to assess the impact of Premera's VBF on healthcare costs and outcomes. **METHODS:** We utilize an interrupted time series design with concurrent control group in order to examine the impact of the VBF on both pharmacy and medical costs for enrollees and the health plan separately and to examine the impact of the VBF on both emergency department visits and acute hospitalizations. In order to accomplish these aims, we utilize segmented regression models with two-part generalized estimating equations for analysis. **RESULTS:** Preliminary descriptive analysis suggests that over the 4 years of observation, comparing the period before VBF implementation to the period after VBF implementation, both medical and pharmacy costs increased more in the control group ($38.37 and $4.79 per member per month (PMPM)) than in the VBF group ($3.16 and -$0.54 PMPM). The number of emergency department visits and acute hospitalizations did not change in either group. **CONCLUSIONS:** Preliminary analyses suggest that the use of cost-effectiveness principles in the US context may lead to greater economic efficiency. Subsequent analyses utilizing greater control for confounding will establish more valid estimates of outcomes and costs.

HT3
ANALYSIS OF NICE DRUG TECHNOLOGY APPRAISALS (2001-SEPTEMBER 2014)
Briceno V, Seoane-Vazquez E
MCPHS University, Boston, MA, USA
OBJECTIVES: The UK National Institute for Health and Care Excellence (NICE) provides guidance and advice to improve health care in the UK. This study assessed the NICE Drug Technology Appraisals published in the period 2001-September 2014. **METHODS:** The list of NICE guidance, including published guidance, in development and consultations was extracted from NICE webpage. Descriptive statistics and chi-square were used in the analysis. **RESULTS:** In September 2014, NICE listed 994 guidance documents, including 246 technology appraisals (TA), of which 207 were drug TA. The drug TA assessed 158 different drugs, combinations, or drug classes. 75.8% of the drug TA evaluated was recommended by NICE in the National Health Service (NHS), however 17.0% of them were not recommended. NICE was unable to recommend them because no evidence submission was received from the drug sponsor in 7.2% of the TA. In 46.2% of the 91 TA published in 2010-Sep 2014 and recommended by NICE, the sponsor agreed a patient access scheme with the Department of Health to provide a confidential discount. The percentage of TA resulting in drugs with indications recommended by NICE decreased over time from 89.5% (n=49) in 2001-2004, to 71.7% (n=91) in 2005-2009, and 75.8% in 2010-September 2014 ($p<0.001$). There were six therapeutic classes with 10 or more TA: Cancer (68 TA, 57.4% recommended by NICE), blood and immune system conditions (27, 77.8%), cardiovascular conditions (20, 95.0%), musculoskeletal conditions (16, 81.3%), infections (11, 90.9%), and neurological conditions (10, 90.0%). **CONCLUSIONS:** Most of the TA resulted in a positive recommendation by NICE for using the drug in the NHS. Oncology and blood and immune system condition had the lowest percentage of TA resulting in a positive evaluation from NICE. Over 45% of the TA published after 2010 resulted in a confidential discount provided by the sponsor company to the NHS.

HT4
IQWIG EARLY BENEFIT ASSESSMENTS OF TYPE 2 DIABETES THERAPIES
Oehrlein EM, Hanna ML, Perfetto EM
University of Maryland, School of Pharmacy, Baltimore, MD, USA
OBJECTIVES: Since 2011, the Institute for Quality and Efficiency in Health Care (IQWiG) has awarded very few anti-diabetic therapies an "added benefit" over current standard therapies status. IQWiG dossier evaluations were examined to determine if new therapies do not demonstrate "added benefit" or if results may be due to a lack of comparative evidence. **METHODS:** Results from IQWiG dossier assessments for all anti-diabetic drug therapies requiring early benefit assessment were examined. IQWiG's website was searched for "diabetes," identifying Albiglutide, Canagliflozin, Dapagliflozin, Empagliflozine, Linagliptin, Lixisenatide, Saxagliptin, Sitagliptin, and Vildagliptin. Dossier assessment results were reviewed for all sub-indications to determine if added benefit had been evaluated. **RESULTS:** In total, 14 dossier assessments (9 monotherapies; 5 combination therapies) were reviewed representing a total of 48 sub-indications. Of the 48 sub-indications, 2 were designated as having "added benefit" or "hint of added benefit," the remaining 46 did not. The most common reasons for no evidence of "added benefit" were: no evidence submitted (n=13) or evidence deemed irrelevant (n=33). The most common reasons for supporting studies being deemed irrelevant included: differing "therapeutic strategies" in study arms (n=4); deviation from G-BA recommended comparator (n=6); lack of power to demonstrate added benefit (n=1); non-compliance with approved dosage (n=12); non-compliance with approved population (n=12); insulin therapy not tailored to patient (n=6); study duration too short (n=5). **CONCLUSIONS:** In most cases, the evidence needed to perform an IQWiG early benefit assessment for drug therapies for Type II Diabetes did not exist. Moreover, studies with the potential to provide the appropriate evidence lacked use of approved dosage, indicated populations, or comparable study arms preventing a complete assessment of whether sub-indications of an anti-diabetic drug provided an "added benefit" to patients. Future clinical development plans should include well-designed comparative studies to improve likelihood of reimbursement and patient access.

PATIENT PREFERENCE STUDIES

PP1
PATIENTS' AND PHYSICIANS' TIME TRADE-OFF PREFERENCES FOR ADVERSE OUTCOMES ASSOCIATED WITH METASTATIC COLORECTAL CANCER TREATMENTS
Morlock R[1], Gonzalez JM[2], Ogale S[1], Sommer N[1], Posner J[2], Grothey A[3]
[1]Genentech, South San Francisco, CA, USA, [2]RTI Health Solutions, Research Triangle Park, NC, USA, [3]Mayo, Rochester, MN, USA
OBJECTIVES: To estimate health-state utilities for adverse outcomes associated with metastatic colorectal cancer. **METHODS:** Patients and physicians completed time trade-off (TTO) questions. Health states were drafted and refined based on literature review, and patient and clinician interviews. Four adverse conditions were described: severe papulopustular rash (rash), serious bleeding, severe heart attack, and gastrointestinal perforations. Respondents evaluated the risk of serious bleeding, heart attack, and gastrointestinal perforation. Three event risk levels were randomized across events and respondents. Rash was presented as a deterministic outcome, so respondents evaluated the impact of experiencing the rash, not as a 'risk' of developing rash. Patients and physicians evaluated the health states in TTO questions that provided a range of time in the adverse health that would leave respondents indifferent between the adverse health states and shorter life spans with perfect health. TTO data were analyzed using an interval regression model to estimate the health-state utility for each side effect. Results were used to infer the health-state utility of the outcomes' clinically relevant levels corresponding with the most commonly used targeted treatments for mCRC, VEGFi and EGFRi (20% chance of rash, 5% chance of serious hemorrhage, and a 2% chance of gastrointestinal perforations and cardiopulmonary arrest). **RESULTS:** A total of 127 patients and 150 physicians completed the TTO questions. Among clinically-relevant levels of the health states for patients, cardiopulmonary arrest had the lowest utility (0.68), with serious hemorrhage (0.74), GI perforation (0.79) and rash (0.91) having higher levels of utility. Utility scores for physicians followed a similar pattern: cardiopulmonary arrest (0.75), serious hemorrhage (0.76), GI perforation (0.82) and rash (0.92). **CONCLUSIONS:** Results add to previously published literature regarding utilities for adverse outcomes from patients' and physicians' perspectives. Results show that patient and physician ratings of health states were largely consistent, suggesting agreement in the perceived impact of these adverse events.

PP2
PATIENT PREFERENCES FOR FIRST-LINE MAINTENANCE TREATMENTS FOR OVARIAN CANCER
Mansfield CA[1], Poulos C[1], Hackshaw MD[2], Pugh A[1], Jobanputra MH[2]
[1]RTI Health Solutions, Research Triangle Park, NC, USA, [2]GlaxoSmithKline, Philadelphia, PA, USA
OBJECTIVES: Elicit medicine preferences of women eligible to receive first-line maintenance treatment for ovarian cancer and estimate benefit-risk trade-offs. **METHODS:** Women in the United States with self-reported physician diagnoses of ovarian cancer and eligible for maintenance therapy completed an online discrete-choice experiment (DCE) survey. The survey presented nine choice questions, each including a pair of hypothetical medicine profiles with varying efficacy, tolerability, and risks of side effects. Each profile was defined by the following attributes identified from the literature with clinical input and tested in patient interviews: progression-free survival (PFS), fatigue, diarrhea, nausea and vomiting, hypertension, and risk of gastrointestinal (GI) perforation. The profiles in the choice questions were based on an experimental design with known statistical properties. Random-parameters logit was used to estimate preferences. **RESULTS:** Two hundred women completed the survey; median age was 49 years, 26% were late stage (3/4), and 44% had been diagnosed within 2 years. Across the attributes, better outcomes were significantly preferred to worse outcomes, except that respondents did not distinguish between no nausea and mild nausea. Relative to the other attributes and levels, respondents placed the greatest weight on avoiding severe diarrhea, followed by reducing the risk of GI perforation, and increased PFS. Respondents were willing to give up 6.5 months of PFS to reduce diarrhea from severe to none. No statistical differences were found between the overall preferences of early versus late stage respondents, respondents above and below the median age in the sample, and respondents who had been diagnosed in the last 2 years compared to those diagnosed more than 2 years ago. **CONCLUSIONS:** Women with ovarian cancer were willing to trade-off efficacy (PFS) for improvements in side effects and risk. The lack of differences across subgroups suggest consistent preferences across the attributes within our sample. Funded by GSK.

PP3
PATIENT VERSUS GENERAL POPULATION PREFERENCES IN ANTICOAGULANT THERAPY
Najafzadeh M, Gagne JJ, Choudhry NK, Polinski J, Avorn JL, Schneeweiss S
Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
OBJECTIVES: Relative preference weights for outcomes of anticoagulant therapy can be used to inform quantitative benefit-risk analyses. Whether patients with cardiovascular diseases (CVD) and the general population have different preferences for benefits and risks of anticoagulant therapy is unknown. Using a Discrete Choice Experiment (DCE), we elicited and compared anticoagulant treatment outcomes preferences between patients and the general population. **METHODS:** A sample of patients with CVD and a general US population sample were selected from online panels. A DCE questionnaire was designed and administered to elicit preferences for benefits and risks. Seven attributes described hypothetical treatments randomly

labeled "new drug," "old drug," or "no drug": non-fatal stroke, non-fatal myocardial infarction (MI), cardiovascular death, minor bleeding, major bleeding, bleeding death, and need for therapeutic monitoring. We estimated preference weights and maximum acceptable risks. **RESULTS:** A total of 341 patients and 352 individuals from the general population completed the questionnaire. On average, patients perceived a 1% increased risk of a fatal bleeding equivalent to a 2% increase in non-fatal MI, a 3% increase in non-fatal stroke, a 3% increase in cardiovascular death, a 6% increase in major bleeding, or a 16% increase in minor bleeding. As compared to the patients, the general population had similar preferences except that they perceived a 3% increase in non-fatal MI or a 13% increase in minor bleeding equivalent to a 1% increase in risk of bleeding death. Patients were less likely to choose "no drug" (odds ratio, 0.72; 95% confidence interval, 0.61- 0.84) or "old drug" (odds ratio, 0.86; 95% confidence interval, 0.81-0.93) than "new drug." The general population sample was indifferent to the drug labels. **CONCLUSIONS:** Patients and the general population had similar relative preferences for anticoagulant treatment outcomes but were more likely to choose "new drug," irrespective of its relative benefits and risks.

PP4

MEASURING TREATMENT PREFERENCES OF PATIENTS DIAGNOSED WITH IDIOPATHIC PULMONARY FIBROSIS USING BEST-WORST SCALING

Camponeschi GF[1], Bridges JF[2], Danoff SK[1], Richardson B[1]
[1]*Johns Hopkins University, Baltimore, MD, USA,* [2]*John Hopkins Bloomberg School of Public Health, Baltimore, MD, USA*

OBJECTIVES: Idiopathic pulmonary fibrosis (IPF) is a rare, degenerative disease. While recently approved therapies provide hope, research is needed to assess the value of treatment benefits and risks. This study sought to develop and test a patient-centered survey instrument to value the benefits and risks of IPF therapies. **METHODS:** Using patient and stakeholder engagement, researchers developed a survey instrument for measuring the treatment preferences of IPF patients. This led to developing a novel best-worst scaling instrument to assess six treatment attributes, each defined across three levels, including lung function, shortness of breath, persistent cough, gastrointestinal problems, skin problems and risk of liver toxicity. Surveys were completed in person or by mail. Patients were shown 18 treatment profiles, created through a main-effect orthogonal experimental design, and asked to identify the best and worst aspect of each treatment. Preference weights were estimated using a simple score consisting of the number of times a level was chosen as best minus the number of times it was chosen as worst and divided by the total number of times the level was shown. Conditional on the level chosen in the experiment, attribute importance was estimated by comparing the range of scores across each attribute, relative to all such deviations. **RESULTS:** Thirty-five participants completed the survey. The most important attribute preferred was effect on lung function (35%), followed by risk of gastrointestinal problems (23%), risk of liver toxicity (12%) and impact on persistent cough (11%). Patients estimated the least important attributes to be risk of skin problems (9%) and impact on shortness of breath (9%). **CONCLUSIONS:** This research demonstrates the merits of a community-centered approach to survey instrument development to measure preferences and illustrates the value in quantifying preferences. Further research is needed to assess the generalizability of these findings and the implications for decision making.

RESEARCH POSTER PRESENTATIONS - SESSION I

RESEARCH ON METHODS STUDIES
RESEARCH ON METHODS – Clinical Outcomes Methods

PRM1

DEVELOPMENT AND VALIDATION OF A U.S. ADMINISTRATIVE CLAIMS-BASED ALGORITHM TO CLASSIFY PATIENTS WITH TYPE 2 DIABETES MELLITUS INTO RENAL IMPAIRMENT STAGES

Johnston S[1], Sheehan J[2], Riehle E[1], Cappell KA[3], Varker H[1], Anzalone D[4], Ghannam A[2], Kalsekar I[4]
[1]*Truven Health Analytics, Bethesda, MD, USA,* [2]*AstraZeneca, Fort Washington, PA, USA,* [3]*Truven Health Analytics, Ann Arbor, MI, USA,* [4]*AstraZeneca, Wilmington, DE, USA*

OBJECTIVES: The validity of diagnosis/procedure coding for determining the severity of renal impairment is unknown. This retrospective, observational study developed an administrative claims-based algorithm which classified patients with type 2 diabetes mellitus (T2DM) into renal impairment stages using estimated glomerular filtration rate/1.73 sq M by MDRD equation (eGFR) as the measure for renal function. **METHODS:** The data source was U.S. administrative claims collected from among a sample of 35,624 patients ≥18 years of age who, during the period from 1/1/2012-12/31/2012, had ≥1 laboratory result for eGFR, ≥2 medical claims with a diagnosis code for T2DM, continuous insurance enrollment, and no medical claims with a diagnosis/procedure code for type 1 diabetes, gestational diabetes, or pregnancy. The sample was divided into two equal random samples: a test set and validation set. Among the test set, four logistic regressions were fit modeling Kidney Disease Outcomes Quality Initiative-defined renal impairment stages (eGFR <15, <30, <60, and <90) as a function of age, sex, and 25 binary indicators for the presence of medical claims with renal impairment-related diagnosis/procedure codes. From each regression, a predicted probability was obtained for the validation set and performance of the algorithm was tested (e.g., by ROC analysis) at varying probability cutoff classification thresholds. **RESULTS:** In the validation set, the percentage of patients correctly classified by the test set algorithm using a standard probability cutoff=0.5 was 75.9% for eGFR <90, 82.1% for <60, 97.3% <30, and 99.3% for <15; in the test set, these same percentages deviated by less than 1 percentage point. Model C-statistics ranged from 0.79 for eGFR <90 to 0.89 for eGFR <15. Sensitivity/specificity varied considerably by selected probability cutoffs. **CONCLUSIONS:** This novel, replicable, administrative claims-based algorithm should prove useful to diabetes researchers who need to classify patients' renal impairment stage in the absence of detailed eGFR data.

PRM2

COMPARISON OF IRRITABLE BOWEL SYNDROME WITH CONSTIPATION AND CHRONIC CONSTIPATION PATIENT IDENTIFICATION UTILIZING ADMINISTRATIVE CLAIMS-BASED ALGORITHMS, MODIFIED ROME III DIAGNOSTIC CRITERIA, AND PATIENT-REPORTED PHYSICIAN DIAGNOSES

Buono JL[1], Stephenson JJ[2], Spalding WM[3], Cai Q[2], Tan H[2], Carson RT[1], Doshi JA[4]
[1]*Forest Research Institute, Inc., a subsidiary of Actavis plc, Jersey City, NJ, USA,* [2]*HealthCore Inc., Wilmington, DE, USA,* [3]*Ironwood Pharmaceuticals, Inc., Cambridge, MA, USA,* [4]*University of Pennsylvania, Philadelphia, PA, USA*

OBJECTIVES: Given the lack of specific ICD-9 codes, no definitive method exists for identifying irritable bowel syndrome with constipation (IBS-C) and chronic constipation (CC) patients in administrative claims. This study compared patients identified as having IBS-C and CC through claims-based algorithms with modified Rome III criteria and patient-reported physician diagnoses. **METHODS:** Consenting patients aged ≥18 years identified from the HealthCore Integrated Research Database as having IBS-C (≥1 IBS claim and ≥2 constipation claims or ≥1 constipation claim and ≥1 constipation-related pharmacy claim) or CC (≥2 constipation claims ≥90 days apart or ≥1 constipation claim and ≥1 constipation-related pharmacy claim ≥90 days apart and no IBS claims) completed a cross-sectional survey that included questions pertaining to IBS-C/CC symptoms based on modified Rome III criteria and patient self-report of IBS-C and CC physician diagnoses to confirm claims-based diagnoses. **RESULTS:** Among 236 claims-based IBS-C patients, 22% met Rome III IBS-C criteria and 43% reported being told by a physician they had IBS-C. In addition, 33% of claims-based IBS-C patients reported being told by a physician they had CC. Among 456 claims-based CC patients, 27% met Rome III CC criteria and 39% reported being told by a physician they had CC. However, 38% of claims-based CC patients met Rome III criteria for IBS-C and 18% reported being told they had IBS-C. Patients who did and did not meet Rome III criteria had similar demographic and clinical characteristics. **CONCLUSIONS:** A majority of patients identified as having IBS-C and CC via claims did not meet Rome III criteria. There was greater agreement between claims-based criteria and patient-reported physician diagnoses than Rome criteria. Our findings suggest that patients identified through claims may have been asymptomatic at the time of the survey, and those identified as CC patients may be IBS-C patients who never received an IBS claim.

PRM3

PSYCHOMETRIC VALIDATION OF PERFORMANCE OUTCOMES (PERFOS) FOR USE WITH HIP FRACTURE (HF) POPULATIONS

Gentile B[1], Ballinger RS[2], Doll HA[2], Bush EN[3]
[1]*ICON, San Francisco, CA, USA,* [2]*ICON, Oxford, UK,* [3]*Eli Lilly, Indianapolis, IN, USA*

OBJECTIVES: The measurement properties of three PerfOs [Timed Up-and-Go (TUG), 4-Step Stair Climb (4SC) and Repeated Chair Stands (RCS, in two versions with arms folded, RCS-A, or arm rests, RCS-B)] were evaluated in hip fracture (HF) surgery patients. **METHODS:** Patients were recruited from 11 clinical sites across 6 US states. Participants visited sites at designated time points after HF surgery when patient reported outcome (PRO) measures, patient- and clinician-reported global concept items (GCI), and PerfOs were administered. PerfOs were scored as time (seconds, s) to complete each test. PerfO measurement properties evaluated included: reliability (inter-rater, test-retest), construct validity (known-groups, convergent/divergent), ability to detect change, minimal important difference (MID) and responder definitions. **RESULTS:** Data were recruited from 75 patients (mean age 79.64, SD 6.83 years; 68.0% female) at baseline; from 68 and 66 at visits 2 and 3. Inter-rater (ICCs: 0.87 to 0.97) and test-retest (ICCs: 0.91-0.95) reliability was excellent across the PerfOs. Known-groups validity: Those without an assistive device had quicker mean completion times for all PerfOs but the RCS-A. In addition, TUG times were shorter for patients with high versus low SF-12 physical component summary (PCS) scores ($p = .009$). Convergent/divergent validity: the TUG, RCS-B, and 4SC demonstrated moderate correlations with the SF-12 PCS (rs ranged -0.227 to -0.449), and stronger correlations with the individual physical dimensions than the mental component (MCS) and dimension scores. Ability to detect change: patients demonstrated significant changes in PerfOs from baseline to Visit 2 for the RCS-B ($p = 0.030$) and 4SC ($p = 0.034$). MIDs ranging from 1.5s (4SC) to 6.0s (TUG) were found. Based on Best Cut Points (BCP) of one-point change in clinician GCIs and values of minimal detectable change (MDC90), responder definitions between 2.0s (4SC) and 3.5s (TUG) are recommended. **CONCLUSIONS:** Overall, the three PerfOs demonstrated adequate psychometric properties.

PRM4

GETTING THE FULL PICTURE: THE IMPORTANCE OF EXRAPOLATING BEYOND THE DATA

Al-Mutairi R, Rittenhouse B
MCPHS University, Boston, MA, USA

OBJECTIVES: Clinical assessments with a limited time horizon for data collection are typical. One example is a (now dated) 10-year study of statin therapy by Pharaoh & Hollingsworth, 1996 (P & H). Gray (2011) suggests that this study likely underestimates the effect of the intervention by limiting its time horizon. We explore this suggestion as a methodological point. **METHODS:** Using life table methods we analyzed one cohort from the P & H publication for which data were reported sufficiently (50-year old men with pre-existing CHD) and compared their limited horizon analysis with one which extrapolates until the cohort lives out its life expectancy under each treatment alternative. Because of the lack of longer term data, assumptions need to be employed in order to extend the results (e.g. the treatment effect persists, fully, partially, at different levels, or not at all). **RESULTS:** We replicated the P & H 10-year horizon that indicated a treatment/no treatment mean life expectancy difference of 0.071 life years (LYs). Assuming a fully persistent treatment effect over lifetimes, this increased to 2.35 LYs in the extrapolation, clearly sufficiently differ-