HT2 APPLICATION OF COST-EFFECTIVENESS LOGIC TO US MANAGED CARE DRUG FORMULARIES: LONG TERM OUTCOMES OF A VALUE-BASED FORMULARY
Young E1, Basu A2, Hansen RN2, Watkins J1, Sullivan SD1
1School of Pharmacy, University of Washington, Seattle, WA, USA, 2University of Washington, Seattle, 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 relative value considerations in drug 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, Fремera Blue Cross implemented a value-based formulary (VBF) that uses CEA to determine a coverage decision at the level of a single drug and did not include a binary coverage decision. The objective of this study was to assess the impact of Fремera’s VBF on healthcare costs and outcomes. METHODS: We utilized 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 payment 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 utilized 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 ($21.65 and $0.29 per member per month (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 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 V1, Socano-Vazquez E1
1MCPHS University, Boston, MA, USA
OBJECTIVES: The 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 594 guidance documents, including 246 technology appraisals (TAs), of which 29 were drug TA assessed 158 different drugs, combinations, or drug classes. 75.8% of the drug TA evaluated was recommended by NICE in the NICE Drug Technology Appraisals published in the period 2001-September 2014.
CONCLUSIONS: Preliminary analyses suggest that the use of cost-effectiveness principles in the US context may lead to greater efficiency. Subsequent analyses using more validated control for confounding will establish more valid estimates of outcomes and costs.

PP1 PATIENTS’ AND PHYSICIANS’ TIME TRADE-OFF PREFERENCES FOR ADVERSE OUTCOMES ASSOCIATED WITH METASTATIC COLORECTAL CANCER TREATMENTS
Molino R1, Gonzalez JM2, Ogale S1, Somer N1, Posner J, Grothey A1
1Gastroenterology, South San Francisco, CA, USA, 2RTI Health Solutions, Research Triangle Park, NC, 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 examined: 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 state that would leave respondents indifferent between the adverse health state and death. Life expectancy and health state were defined 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, which are clinically relevant levels corresponding to the most commonly used targeted treatments for mCRC, VEGFi and EGFi (20% chance of rash, 5% chance of serious hemorrhage, and a 2% chance of gastrointestinal perforations and cardiology cardiac arrhythmia). RESULTS: A total of 127 patients and 150 physicians completed the TTO questions. Among clinically-relevant levels of the health states for patients, cardiopulmonary arrhythmia had the lowest utility (0.68), with serious hemorrhage (0.74), GI perforation (0.79) and rash (0.91) having higher utilities for utility. Utilization of this tool can inform clinicians and patients of benefits and risks. Seven attributes described hypothetical treatments randomly

PP2 PATIENT PREFERENCES FOR FIRST-LINE MAINTENANCE TREATMENTS FOR OVARIAN CANCER
Mansfield CA1, Poulos C1, Hackshaw MD1, Pugh A1, Johansson MH1
1RTI Health Solutions, Research Triangle Park, NC, USA, 2GlaxoSmithKline, 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: In the United States with self-reported physician diagnosed cases of ovarian cancer and eligible for maintenance therapy completed an online discrete-choice experiment (DCE) survey. The survey presented nine choice questions describing five different treatments with characteristics including tolerability, efficacy, and side effects. Each choice question included: (1) a hypothetical treatment with six attributes, (2) an additive benefit, (3) a profile 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 M1, Gagne J1, Choudbury NK1, Polinkijn J1, Avorn J1, Schneeweiss S1
1Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA
OBJECTIVES: For 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 preferences for benefits and harms of anticoagulant treatments 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 eliciting 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), myocardial 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 non-inferior 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
Campioneschi GF, Bridges JF, Danoff SK, Richardson B.
Johns Hopkins University, Baltimore, MD, USA, 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 instrument to evaluate treatment preferences.
METHODS: Using patient and stakeholder engagement, researchers developed a survey instrument for measuring the treatment preferences of IPF patients. This led to the IPF 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, and impact on health. They developed a survey instrument using 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 minus the number of times it was chosen as worst and divided by the total number of times the level was chosen. 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 (21%) and impact on persistent cough (11%). Patients estimated the levels were most likely to be of the same level of importance (25%) 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
Johnson S1, Sheehan J2, Kiehle E1, Cappelli KA1, Varker H3, Aranzale D4, Ghanam A2, Kalsekar A2.
1Tuven Health Analytics, Bethesda, MD, USA, 2AstraZeneca, Fort Washington, PA, USA, 3Tuven Health Analytics, Ann Arbor, MI, USA, 4AstraZeneca, 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 m² 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 staging criteria as the dependent variable to develop a renal impairment algorithm (eGFR <15, 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 model, the contribution of each cohort from the classification algorithm 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 of 0.875 (sensitivity, 0.875; specificity, 0.875) was 75.9% for eGFR ≥90, 64.5% for eGFR <90, and 67.9% for eGFR ≥60 and <90. 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 IDENTIFICATION usING ADMINISTRATIVE CLAIMS-BASED ALGORITHMS, MODIFIED ROME III DIAGNOSTIC CRITERIA, AND PATIENT-REPORTED PHYSICIAN DIAGNOSES
Buono J.L1, Faghihian JF2, Slovik WM2, Faghihian BM2, Tan H1, Carson RT1, Doshi JA3.
1Forest Research Institute, a subsidiary of Actavis plc, Jersey City, NJ, USA, 2HealthCore Inc., Wilmington, DE, USA, 3Tonwood Pharmaceuticals, Inc., Cambridge, MA, USA, 4University 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, algorithmic and 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 IBS claim) and CC 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 III 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 B1, Ballinger BS2, Doil HA3, Bush EN4.
1UCON, San Francisco, CA, USA, 2UCON, Doha, UK, 3Elly Lilly, Indianapolis, IN, USA.
OBJECTIVES: The development of three Perfos (Timed Up-and-Go (TUG), 4-Step Stair Climb (ASC) 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 15 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 item concepts (IPC), and Perfos were administered. Perfos were scored as time (seconds), or to count correct test. Perfro 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 obtained from 75 patients with a mean age of 68.63 years, 68.0% female) at baseline, from 68 and 66 at visits 2 and 3. Inter-rater (CCs: 0.87 to 0.97) and test-retest (CCs: 0.91-0.95) reliability was excellent across the whole known groups validity. Those without an assistance device had equal 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 = 0.009). Convergent/divergent validity: the TUG, RCS-B, and 4SC demonstrated moderate correlations with SF-12 PCS (r ranged -0.227 to -0.449), and stronger correlation with the individual physical dimensions than the mental component (MC3) and dimension scores. Ability to detect change: patients demonstrated significant changes in Perfos from baseline to Visit 2 for RCS-B (p < 0.003) and 4SC (p = 0.034). MIDs ranging from 1.5s (4SC) to 6s (TUG) were found. Based on Best Cut Points (BCP) of one-point change in clinician GCIs and values of minimal detectable change (MDC300), 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 EXTRAPOLATING BEYOND THE DATA
Al-Mujaini B, Rittenhouse B.
MCPRU, University of 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 Fharar & GlaxoSmithKline (2003) suggesting 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 simulated the cohort from the original data and compared the impact of the intervention on survival over lifetimes, this increased to 2.35 in the extrapolation, clearly differently differ-