indicator of the influence of the UK in global RW research. METHODS: All 1455 abstracts presented at ISPOR 14th Annual European Congress in Madrid, Spain, in 2011 were reviewed in THE ISPOR OUTCOMES RESEARCH DIGEST, available via the ISPOR website. Posters were also reviewed where available. Those reporting RW studies were classified according to: therapeutic area, type of study, setting, source of data and methodology, country undertaken, country of authors and involvement in funding. RESULTS: A total of 278 abstracts (19%) described RW studies. Data were derived from a database in 55.8%, health service/patient medical records in 24.8%, surveys/questionnaires in 15.8% and other sources in 3.6%. 12% were conducted in the UK, a further 8% included UK centres in an international study. 34% were conducted in the US, 53% in another country (not UK) and 38 countries, most commonly Spain (6%), Canada (5%), Germany (5%), France (5%), The Netherlands (3%) and Italy (3%) and 3% were international without a UK source. NOT USA – 38 countries, most commonly Spain (6%), Canada 5%, Germany 5%, France 5%, the UK was the source of RW data in twice as many studies as any other country, lending weight to the opinion that the UK provides an excellent environment for conducting RW studies.

PRM1: INCREASING PHYSICAL ACTIVITY IN PATIENTS WITH CHRONIC DISEASE: WHAT IS THE LITERATURE TELLING US?
Leidy NK1, Kimel M1, Ajagbe L1, Kim K1, Hamilton A2, Becker K2
1United BioSource Corporation, Bethesda, MD, USA, 2Boehringer Ingelheim (Canada) Ltd., Burlington, ON, Canada, United Kingdom. OBJECTIVES: It is widely recognized that regular exercise improves fitness, with increasing evidence that physical activity (PA; movement resulting in elevated energy expenditure beyond basal levels) can affect health, particularly in chronic disease. While pharmacologic therapy and exercise training have been shown to improve capacity, persistent increase in PA requires behavior change. This review examined studies testing the effectiveness of behavioral interventions to increase PA in adults with chronic disease. METHODS: Embase and PubMed searches of interventions published in English, 1990–2011. Inclusion criteria: studies of adults > 18 years; COPD, diabetes, heart failure, obesity; exercise or PA endpoint; behavioral intervention described in sufficient detail to permit interpretation. RESULTS: A total of 392 abstracts screened; 169 articles retrieved; 36 reviewed. Most were randomized trials (n = 30, 83%) with 2 intervention arms (n = 29, 81%), medium to high quality (n = 34, 94%). Subjects were recruited through clinical settings (n = 28, 78%), with disease severity a primary eligibility criterion (n = 23, 64%); 15 (42%) had sample sizes 40–100. Mean study duration = 9.6 months (range: 1–84). Exercise intervention: 30–50 minutes aerobic activity 3–5 times/week (n = 22, 61%); 64% included walking. Instruction was individual (n = 25, 69%), initially supervised (n = 24, 67%) followed by unsupervised home exercise (n = 15, 42%). Behavioral intervention: counseling (n = 19, 53%) with personal contact follow-up (n = 12, 33%). Control group: exercise without behavioral intervention (n = 14, 39%) or usual care (n = 15, 42%). Significant effects of the intervention were reported in 15 of 25 (60%) studies testing exercise capacity (6-minute walk, cycle or treadmill), 19 of 26 (73%) testing PA outcomes (pedometer, activity log, questionnaire), 11 of 22 (50%) measuring HRQoL, and 8 of 13 (62%) capturing behavioral endpoints. CONCLUSIONS: This review provided insight into the range of designs, interventions, and outcome measures used in studies testing methods to improve PA in chronic disease. Results identify promising interventions, with implications for improving research methods and outcomes.

PRM9: TREATMENT OF RHEumatoid Arthritis – COMPARATIVE EFFECTIVENESS OF BIOLOGICS
Schiffrin-Rohé L1, Leverkus P1, Behmer O2, Kerkman L3
1Pfizer Deutschland GmbH, Berlin, Germany, 2Pfizer Pharma GmbH, Berlin, Germany. OBJECTIVES: Guidelines for treatment of rheumatoid arthritis (RA) advise initial therapy with non-biological disease modifying antirheumatic drugs (DMARDs). In case patients do not respond adequately, treatment should be switched to biological DMARDs. Aim of this research is to compare results of public available systematic reviews (SRs) on comparative effectiveness (CER) and potential impact of differences in methodology. METHODS: We performed literature research for SR on CER of biologics for the treatment of RA. Search was limited to reviews published in 2009 or later. Methods of the reviews and results were extracted from the publications. Results are summarized in narrative way and differences in results are reflected focusing on methodological key issues. RESULTS: Eleven recent SRs were identified addressing the question of CER of biologics. Since there are no head-to-head comparisons available for all but one biologic, reviews had to use indirect comparisons to assess CER. Authors used the Simon-Bucher approach or Bayesian methods (mixed treatment comparisons). Clinical trial guidelines for RA gave the advice to use the ACR20 (American College of Rheumatology 20%) as primary parameter. Therefore, all trials assessed ACR20 (20% improvement), ACR50 and ACR70 and could be used for comparison. Further parameters were not assessed in a uniform manner (e.g. quality of life or not assessed). Conclusions were drawn in the text (e.g. disease activity score) and therefore could not be used for indirect comparison of treatments. CONCLUSIONS: Due to lack of head-to-head data for comparison of biologics, statistical methods for indirect comparison have to be used to answer the question of CER. These methods have restrictions and base on assumptions that might be heavily violated. Substances were tested over a time period of more than 10 years with effects on study population and variation in study designs. Nevertheless, the results seem to be fairly consistent.

PRM10: USE OF VA DATABASES FOR RETROSPECTIVE STUDIES IN ULTERATIVE COLITIS OUTCOMES RESEARCH
Kolva Y1, Shi L2, Abbas A2, Khan N4
1Tufts University / Southeast Louisiana Veterans Health Care System, New Orleans, LA, USA, 2Tufts University, New Orleans, LA, USA, 3Tufts University / Southeast Louisiana Veterans Health Care System, New Orleans, LA, USA, 4Memorial Healthcare System, New Orleans, LA, USA. OBJECTIVES: Veterans Administration Corporate Data Warehouse stores databases with standardized structure that could be used for automated data extraction, reviewer abstraction, and text mining to determine the association between health outcomes and disease-specific factors. This retrospective study provides assessment of VA administrative data used to examine the impact of pharmacological therapy on complications in ulcerative colitis (UC). METHODS: Previous studies investigating the effect of SASA on the risk for colorectal cancer (CRC) in UC patients have reported conflicting results. We obtained nationwide UC and CRC data from VA health care system for the period 2001-2011. Secondary relational databases were searched for clinical variables based on standardized criteria. ICD9 diagnoses, procedural and medication codes. Data extraction captured demographics, clinical information and pharmacy record for a cohort of 37,191 UC cases. We constructed a dataset of potential ulcerative colitis cases with CRC (n=1,087) defined by ICD9 codes 556.x and 153.x,154.x,159.0. A random sub-sample of 100 non-SASA users with CRC was compared to 100 controls without CRC. RESULTS: Diagnosis of ICD9 code for CRC had PPV 79% and NPV 100% in the random sample. Within the 1087 potential CRC cases, only 500 (46%) were found to have chart of both conditions at review with kappa agreement between automated and manual abstraction 0.73 (95% CI: 0.70-0.76) for CRC and significantly lower for UC: 0.60 (95% CI: 0.57-0.63). The initial overall prevalence of CRC in the UC cohort was 2.9% and decreased to 1.34% after human text search verification. CONCLUSIONS: Automated extracts have great potential for diseases surveillance but manual review yields more reliable data. Pre-defined diagnostic algorithms based on a combination of methods as well as further technology development like natural language processing and longitudinal patient record will improve accuracy of retrospective databases.

PRM11: HOW WORLD DATA TO CALCULATE COST-EFFECTIVENESS OF MONOCLONAL ANTIBODIES: PROBLEMS AND SOLUTIONS
van Rooijen EM, van der Linden N, van Gilis C, Oppe M, Uyl-de Groot C
1Memory for Medical Technology Assessment, Erasmus University, Rotterdam, The Netherlands. OBJECTIVES: Real-world data is considered to be the gold standard by decision makers to inform on cost-effectiveness of new drugs. Unfortunately real-world data are often lacking in important parameters needed to inform on cost-effectiveness, and RCT data can be used to address this problem. Illustrated by two cases this presentation will show that real-world data and RCT data combined can have a profound influence on the resulting ICER. METHODS: Two case studies in which real-world data on cetuximab for the indication of locally advanced head and neck cancer and panitumumab for the indication of chemo-refractory metastatic colorectal cancer were collected, respectively served as examples. RESULTS: The problem: In the case of cetuximab, patient selection in daily practice resulted in too much disparity in baseline characteristics between the treated and control group. The solution: survival data for both treatment groups from the pivotal RCT was used and corrected according to the results seen in the real-world data to better represent survival in daily practice. Using uncorrected RCT data resulted in a difference of approximately 5,000 euro/QALY in the
ICER. In the case of panitumumab no data were available on progression of disease in the control group. The solution: progression free survival and survival after progression of the control group were drawn from the pivotal RCT and adjusted according to the survival observed in the control group of the outcomes research. Unadjusted RCT data resulted in an ICER that was approximately 20,000 euro/QALY higher. CONCLUSIONS: RCT data are often necessary to supplement missing data that cannot be collected through outcomes research. However, if RCT data is used can have a profound effect on the resulting cost-effectiveness.

PM12 JOINT ESTIMATION OF PROGRESSION FREE SURVIVAL AND OVERALL SURVIVAL
Ouwens MJ1, Bergman G2
1MAPF Consultancy, Houten, The Netherlands, The Netherlands, 2MAPF Consultancy, Houten, The Netherlands
OBJECTIVES: In cancer, treatments often aim to extend time to progression. The implications on overall survival are often inconclusive, as trials are too short and the majority of patients are still alive at the end of the trial. However, for decision making, it is important to estimate both the treatment effect on Progression Free Survival as well as the treatment effect on Overall Survival. This poster shows how the estimation of Overall Survival benefit can be improved by the use of Progression Free Survival data.

METHODS: The developed Network Meta-Analysis model uses the tested hypothesis that treatments provided until progression in general do not change the length of the post-progression period. This hypothesis is tested in detail based on systematic literature reviews concerning 4 different types of cancer. A test for equal lengths of the post-progression periods is described too.

RESULTS: A network meta-analysis model is described, which can be used to obtain estimates for OS from PFS data for treatments for which no OS data or insufficient OS data are available. In this situation, the assumption that the post-progression period is the same for all treatments is adopted. The methodology is applied to indirect comparisons of Erlotinib, Pemetrexed and Docetaxel. The results show that the assumption of equal post-progression periods among treatments, which is required for the equal period assumption, is often not valid.

CONCLUSIONS: Based on systematic literature reviews, a method is developed to use PFS as surrogate outcome for OS. In addition, a test is developed to justify the assumption of equal post-progression periods among treatments, which can be used to assess whether the translation of PFS time differences in OS time differences is appropriate.

PM13 METHODS FOR ESTIMATING SURVIVAL BENEFITS IN THE PRESENCE OF TREATMENT CROSSOVER: A SIMULATION STUDY
Laumond J2, Lambriksen C1, Andersen M1, Ahram A4, Morden P3
1University of Sheffield, Sheffield, UK, 2University of Leicester, Leicester, UK, 3The Institute of Cancer Research, Sutton, Surrey, UK

OBJECTIVES: We aimed to assess statistical methods for adjusting survival estimates in the presence of treatment crossover in order to identify which are the most appropriate in a range of scenarios. Treatment crossover is a common issue in clinical trials of cancer treatments. Crossover occurs when patients in the control group switch onto the experimental treatment at some point during follow-up. In such situations, the decision to treat (ITT analysis) does not address the decision problem faced by health technology assessment bodies, and will result in biased estimates of the overall survival advantage – and therefore the cost-effectiveness – associated with the experimental treatment.

METHODS: We conducted a simulation study to assess the performance of crossover-adjustment methods in a range of scenarios. We purposefully ran scenarios that did not satisfy the specific assumptions made by the methods, in order to assess their sensitivities.

RESULTS: Randomisation-based methods (eg Rank Preserving Structural Failure Time Models (RPSFTM) and Iterative Parameter Estimation (IPE)) were unbiased only when the treatment effect was not time-dependent. Observational-based methods (eg Inverse Probability of Censoring Weights (IPCW)) and Structural Nested Models (SNMs) with g-estimation) coped better with time-dependent treatment effects but are heavily data reliant, are sensitive to model misspecification and often produced high levels of bias in our simulations. Observational-based methods are particularly sensitive to the proportion of control group patients that crossover whereas randomisation-based methods are not.

CONCLUSIONS: Currently available randomisation-based methods may be preferred in health technology assessment compared to observational-based methods. However, in most circumstances they are likely to lead to lower bias than an ITT analysis, given the decision problem faced in an economic evaluation. Analysts should consider the treatment crossover mechanism, rather than just the crossover proportion, the treatment effect associated with different patient groups, and data availability when deciding which method to use to address treatment crossover.

PM14 EUROPEAN ASSESSMENT OF THE VALIDITY OF THE QALY OUTCOME MEASURE: RESULTS FROM THE EXPERIMENT CONDUCTED BY THE ECHOUCOMET PROJECT
Berserini A1, Auray F2, Duras G2, Medina-Lara A3, Tarricone R7, Samburok S7, Torbica A2, Lamoure M4, Esteva N2, Rizzi G3
1Data Science Consulting International, Switzerland, 2Cycligad Group, Ruelle la Page, France, 3Cycligad Group, Ruelle la Page, France, 4Boston Healthcare Associates International GmbH, Berlin, Germany, 5Boston Healthcare Associates, Inc., Boston, MA, USA

OBJECTIVES: Some European health authorities such as the National Institute of Clinical Excellence in the UK have published Health Technology Assessment (HTA) guidelines, which recommend the use of Quality Adjusted Life Years (QALYs) as the outcome measure as the reference case. The ECHOUCOMET project is an interdisciplinary European research platform funded by the seventh Framework Program of the European Commission which objectives are to assess the validity of the QALY calculation, its potential use in cost-effectiveness analyses in national HTA and proposing new European guidelines for conducting CEA studies.

METHODS: Over a period of 3 months, a total of 1,200 students from Belgium, France, Italy and the UK answered hypothetical health states in which the health state of patients in a given health state were varied according to Neumann-Morgenstern assumptions, mutual independence in utility, and the relevance of the multi-linear multi-attribute utility theory, which are the basis for the QALY calculation, as currently performed in the HTA literature.

RESULTS: The preliminary results provided of the experiment showed that students obtained by varying the health states and the duration of a given health state fail to comply with the theoretical basis of the QALY. CONCLUSIONS: The results suggest that the underlying assumptions of the QALY calculation model are not in line with behavior from a real life population implying that the QALY outcome measure might not be a valid measure for supporting health decision making in Europe. The findings of this first European experimental survey testing the validity of the QALY outcome measure should be considered by European member states before recommending such approach in HTA guidelines.

PM15 ASSESSING THE BROADER IMPACT OF VACCINATIONS: A GOVERNMENT PERSPECTIVE
Beresniak A1, Auray J2, Duru G3, Medina-Lara A4, Tarricone R5, Sambuc R6, Torbica A7, Ouwens MJ1, Bergman G2
1University of Sheffield, Sheffield, UK, 2University of Leicester, Leicester, UK, 3The Institute of Cancer Research, Sutton, Surrey, UK

OBJECTIVES: In clinical trials of cancer treatments. Crossover occurs when patients in the control group switch onto the experimental treatment at some point during follow-up. In such situations, the decision to treat (ITT analysis) does not address the decision problem faced by health technology assessment bodies, and will result in biased estimates of the overall survival advantage – and therefore the cost-effectiveness – associated with the experimental treatment.

METHODS: We conducted a simulation study to assess the performance of crossover-adjustment methods in a range of scenarios. We purposefully ran scenarios that did not satisfy the specific assumptions made by the methods, in order to assess their sensitivities.

RESULTS: Randomisation-based methods (eg Rank Preserving Structural Failure Time Models (RPSFTM) and Iterative Parameter Estimation (IPE)) were unbiased only when the treatment effect was not time-dependent. Observational-based methods (eg Inverse Probability of Censoring Weights (IPCW)) and Structural Nested Models (SNMs) with g-estimation) coped better with time-dependent treatment effects but are heavily data reliant, are sensitive to model misspecification and often produced high levels of bias in our simulations. Observational-based methods are particularly sensitive to the proportion of control group patients that crossover whereas randomisation-based methods are not.

CONCLUSIONS: Currently available randomisation-based methods may be preferred in health technology assessment compared to observational-based methods. However, in most circumstances they are likely to lead to lower bias than an ITT analysis, given the decision problem faced in an economic evaluation. Analysts should consider the treatment crossover mechanism, rather than just the crossover proportion, the treatment effect associated with different patient groups, and data availability when deciding which method to use to address treatment crossover.

PM16 ACCESS TO COST DATA CAPTURE USING PUBLIC DATABASE, WEBSITE AND LITERATURE IN GERMANY, FRANCE, SPAIN AND USA
Alvarez-Ossorio L1, Ezat V2, Catiou Y4, Tarab A2

OBJECTIVES: There is a great heterogeneity across health economic studies with regard to claim of cost inputs making comparison of costs complicated. Objective was to analyze data availability, corresponding limitations and improvement approaches.

METHODS: We evaluate the availability of cost data capture for Germany, Spain, France and USA. In Germany data from the Hospital Remuneration, the German Hospital Society (DKG) and the two outpatient tariffs were considered as well as the Federal Health Care and HCUP databases for USA. Database of Ministry of Health and national/regional official bulletin for Spain and lastly data from SNIRAM (Social Security information system) and FMSI (Programme de médicalisation des systèmes d'information) for France.

RESULTS: DRG database used in Germany and USA reflect the reimbursement level more than real cost per indication. Fragmentation of these costs is not possible. The DKG normal tariff (DKG-N) is listing detailed procedures and services used for example for reimbursement between two hospitals. The Spanish inpatient tariffs are difficult to collect due to the prospective hospital global budget; health authorities publish annually DRG and outpatient procedure tariffs as reference of their own resources cost. In France SNIRAM data is limited to Social Security own needs. In contrast publicly available FMSI data allows inpatient information tracking related to specific medical procedure use regarding Outpatient setting, the physician fee schedule is based on the Uniform Evaluation Scale (EBM) and the medical fee schedule (GOA) for SH and