ical trials, observational studies, systematic reviews and economic evaluations. Studies that fulfilled the inclusion criteria were reviewed and data-extracted. The synthesis was narrative. **RESULTS:** After discarding duplicates, 1528 citations were screened, and 833 full papers were retrieved for consideration. Eight out of 145 studies were included, which focused in either diabetes or heart disease. From these, 21 utility values were identified and included in the database. Most of the excluded studies used SF-36 or EQ-5D, but only for descriptive purposes, and did not calculate the utility index. **CONCLUSIONS:** Few studies were available providing utilities elicited from the Spanish population for prevalent diseases such as diabetes or heart disease. As next steps, other diseases will be explored, searches will be conducted in other databases, authors and experts will be contacted to identify additional, relevant information, and included studies will be quality-assessed. This ongoing review and the database of Spanish utilities will be useful for researchers developing economic evaluations or requiring information on quality of life derived from the Spanish population. The project will also allow us to identify data gaps for which further research is needed.

PRM34
EXAMINING VARIATIONS IN ITEM STRUCTURE AND CONTENT IN PRO INSTRUMENTS, OR, THERE MUST BE 50 WAYS TO EXPRESS YOUR DISTRESS

Erickson P[1], Willke RJ[2]
[1]O.L.G.A., State College, PA, USA, [2]Pfizer, Inc., New York, NY, USA

OBJECTIVES: The content validity of an instrument depends not only on the concepts embodied in the items but also on how the items are structured to elicit responses from patients. This research explores the "grammar" of individual items and how it varies across a number of instruments in selected disease areas. The goal is to understand how such considerations affect consistency of content and to classify items by concept. **METHODS:** The structure of each item is characterized as an item stem with a core concept, with an implicit or explicit context (e.g., a disease), event (e.g. "felt frustrated or impatient"), and stimulus (e.g., "about your symptom"), as well as the recall period and response options. Concepts are classified using the WHO International Classification of Functioning (ICF). Similarities and differences across instruments within disease areas are analyzed. **RESULTS:** We decomposed over 600 items in at least 23 instruments across 5 disease areas and several generic instruments, capturing and classifying each aspect of the structure of each item. Most physical function items could be matched with specific 3-4 digit ICF codes; most emotional function items could not be matched as specifically. There was considerable variation across instruments regarding the explicit statement of context as well as the presence of a stimulus. We observed at least 8 different recall periods ranging from an implied present to "in the past year" to "in 10 years" with distinct patterns by disease area. We observed at least 9 types of response options, but the majority of items used 5-point scales. **CONCLUSIONS:** There are some commonalities but little standardization in how items are structured, within or across diseases. Classification and comparison of items and evaluating comparative content validity is complicated by the variation in most aspects of how the items are phrased.

PRM35
AGREEMENT BETWEEN PATIENTS' SELF-REPORT AND PHYSICIANS' PRESCRIPTIONS ON DRUGS AND VACCINE EXPOSURE: THE INTERNATIONAL PGRX DATABASE EXPERIENCE

Grimaldi-Bensouda L[1], Rossignol M[2], Aubrun E[1], Benichou J[3], Abenhaim L[4]
[1]LA-SER, Paris, Paris, France, [2]LA-SER Centre for Risk Research Inc., Montreal, Montreal, Canada, [3]INSERM U═, Rouen, Rouen, France, [4]LA-SER Europe Ltd, London, London, UK

OBJECTIVES: Patients' self-reporting of drug exposure is subject to memory errors and varying degrees of bias. Utilisation of prescription records is often impaired by non-compliance and use of over-the-counter (OTC) drugs. Our study compared patient self-reports (PSR) to physician's prescription reports (PPR) for: cardiovascular drugs (CVDs), allegedly used on a daily basis to treat chronic conditions; drugs for musculoskeletal disorders (MSDs) used intermittently; and vaccines. **METHODS:** The reference pool from the PGRx database consists of several networks of general practitioners. For every referent included, data was obtained for all drugs used within the two years preceding the consultation date via: 1) a structured telephone interview assisted by a guide listing pathologies and packaged visual display of drugs; 2) physician's prescriptions reports. Both PSR and PPR measurements were obtained independently and blindly by investigators. Comparisons were made on exposure to CVDs, MSDs and vaccines, for different time-windows up to 24 months prior to the index date. **RESULTS:** The concordance between physician and patient reports was assessed on 2702 and 4152 patient-physician pairs for CVDs and MSDs, respectively. Overall, agreement between PSR and PPR for all classes of CVDs was excellent (kappa = 0.83 [95% confidence interval (CI): 0.81 – 0.85]). Agreement was substantial for drugs for osteoarthritis (kappa = 0.62 [0.55 – 0.68]), fair for non-aspirin NSAIDs (kappa = 0.31 [0.26 – 0.36]) and low for muscle relaxants and non-narcotic analgesics. Use of OTC drugs was associated with greater disagreement (Odds ratio = 2.2 [95% CI: 1.1 – 1.4]), but not age. **CONCLUSIONS:** The PGRx standardised and systematic collection of drug exposure directly from patients provided similar data to physician prescription records for chronic drug exposure. Differences between PSR and PPR in estimating prevalence for drugs used in MSDs varied by type of drug and time elapsed up to the index date.

PRM36
COMPARING THE PERFORMANCE OF THE SF-6D AND EQ-5D ACROSS DISEASES

Ferreira L[1], Ferreira P[2], Pereira L[1]
[1]University of the Algarve-ESGHT, Faro, Portugal, Portugal, [2]University of Coimbra-Faculty of Economics, Coimbra, Portugal, Portugal

OBJECTIVES: Given that the number of preference-based instruments has been growing over the past years, it is important to compare their performance. This work seeks to compare the performance of the SF-6D and EQ-5D across four diseases: asthma, COPD, cataracts and rheumatoid arthritis. **METHODS:** The overall sample consists of 643 cases. The indexes are analysed by disease group to explore the instruments' ability to distinguish between socio-demographic groups. Ceiling and floor effects are calculated for both instruments. The level of agreement between the instruments is analyzed using correlation coefficients. Paired samples t-tests are used to identify differences between the indexes. Regression analyses are used to explore the relationship between the indexes. The discriminative properties of both indexes are also compared using ROC curves. **RESULTS:** Mean values were the same for both indexes (0.72). However in the analysis by disease the mean EQ-5D index was 0.05 higher than the mean SF-6D index for asthma and COPD. There was a strong correlation between both indexes (0.68). Similar results were found by disease group. The agreement level between both instruments was higher between similar dimensions. Both instruments showed a similar ability to distinguish between socio-demographic groups. There was a significant ceiling effect in the EQ-5D. The results of the regression models indicate that the relationship is not uniform between the two indexes. These results are supported by specific hypothesis tests. The analysis of the area under the curves showed that the SF-6D is more efficient in detecting differences between groups in almost all cases. **CONCLUSIONS:** The SF-6D generates higher values in disease groups. The SF-6D and the EQ-5D perform differently in each of the diseases studied. These results do not allow looking for a global adjustment between both measures regardless of the health state of the individual. These differences should be further investigated.

PRM37
ASSESSING THE HEALTH STATUS OF ROMA POPULATION BY USING SF-36 HEALTH SURVEY: EVIDENCE FROM GREECE

Pappa E[1], Rotsiamis N[1], Halkiopoulos G[1], Papadopoulos A[2], Niakas D[1]
[1]Hellenic Open University, Patras, Greece, [2]ATTIKON University Hospital, Athens, Greece

OBJECTIVES: Roma people compose a vulnerable minority with poor health which has been the subject of discrimination. The aim was to provide a valid estimate of the health status of Roma people in Greece, by using the validated instrument of the SF-36 questionnaire, which has been widely used in previous surveys of general and clinical populations, and also to determine whether SF-36 is a valid and reliable instrument in assessing self-assessed health status of Roma population **METHODS:** The study was carried out in 2009 in two geographically dispersed Roma settlements in Greece. A sample of 433 Roma people was face to face interviewed. The survey included the SF-36, questions on socio-demographic and health related characteristics, health service use and factors associated with material deprivation. Construct validity was investigated with "known group" validity testing and reliability with chronbach alpha coefficient. Statistical significance was accepted at the 5% level. All statistical analyses were undertaken using SPSS v.17. **RESULTS:** Roma responders are young with mean age of 33.5 years old. However they rate their health very low with highest score in PF (66.1%) and the lowest score in MH (41.5%). RP and RE scales had high ceiling and floor effects. Cronbach's alpha coefficient met the criterion (>0.70) for all eight scales with two exceptions. SF-36 scale scores distinguished well, and in the expected manner, between groups of respondents providing evidence of construct validity. Significant statistical differences in mean scores were observed in relation to demographic characteristics, socio-economic status, existence of chronic disease, health services utilization and variables related to material deprivation. **CONCLUSIONS:** The findings support the validity and reliability of the SF-36 when used in assessing Roma's health. On the other hand, Roma experience social exclusion and deprivation which profoundly affect their health. Tackling the poor health of Roma acquires certain public health interventions and health promotion programs.

PRM38
ESTIMATING SF-6D HEALTH STATE UTILITY VALUES FOR COMORID HEALTH CONDITIONS

Ara R, Brazier J
University of Sheffield, Sheffield, UK

OBJECTIVES: The objective of the current study was to compare the accuracy of different methods to estimate health state utility values (HSUVs) for comorbid health conditions. **METHODS:** Data collected during five rounds of the Welsh Health Survey (n=64,437) were used to generate mean SF-6D scores for cohorts with specific health conditions. These data were then used to estimate mean SF-6D scores for cohorts with comorbid health conditions. **RESULTS:** The mean SF-6D scores for the subgroups with comorbidities ranged from 0.465 to 0.607. The minimum and additive methods overestimated and underestimated the majority of actual SF-6D scores, had mean absolute errors (MAE) of 0.056 and 0.121 and just 15% and 3% of estimated values were within the mimimum important difference (MID) for the SF-6D (0.041) respectively. While the multiplicative method also tended to underestimate the actual SF-6D scores (MAE 0.075) it performed better when estimating scores below 0.50 and 47% of estimated values were within the MID. A linear model obtained by mapping the disutilities associated with the mean SF-6D scores for two subgroups with single conditions (plus the interaction between the two) onto the mean SF-6D scores for subgroups with comorbidities gave the most accurate results overall. The predicted SF-6D scores had a mean absolute error of 0.0191 and 88% of predicted SF-6D scores accurate to within the MID. **CONCLUSIONS:** While in our data the linear model gave the most accurate results, additional research is required to validate our results.