variability in CERs for same drugs for different indications, in some cases also varying by biomarkers. Primary care drugs had lower and less variable CERs than specialty drugs. Variations also exist in methodology used by different groups in modeling cost effectiveness, especially for time horizon and comparator. Majority of primary care drugs were modeled for a time horizon of 35–40 years or lifetime to demonstrate cost effectiveness. Among the top 10 drugs, quetiapine and erythropoietin had the highest variability across different studies, and atorvatstatin, salmetrol/fluticasone and clopidogrel had the most consistent ICER values across studies. **CONCLUSIONS:** This analysis shows the range, variability and methods used for calculation of ICER values for these high budget impact drugs and provides lessons for executives and policy makers.

CONCEPTUAL PAPERS & RESEARCH ON METHODS – Patient-Reported Outcomes Studies

PMC17

A COMPARISON OF THE DISCRIMINTATIVE AND EVALUATIVE PROPERTIES OF THE SF-36 AND THE SF-6D INDEX

Mutebi A[1], Brazier J[2], Walters S[2]
[1]University of Arizona, Tucson, AZ, USA, [2]University of Sheffield, Sheffield, UK

OBJECTIVES: To examine whether the move from the SF-36 to the SF-6D entails a loss in discriminative and evaluative strengths, the magnitude of that loss and whether it matters. **METHODS:** The study used relative validity (RV); a ratio of two F statistics, and standardized response means (SRM) to evaluate sensitivity and responsiveness of the SF-36 scales and SF-6D index. An RV of 1 reflected the most sensitive/responsive scale and the smaller the RV the less sensitive the measure would be. Cohen's criterion for interpreting effect sizes was used to interpret the SRMs. The data used were initially collected for prior studies in seven diseases/conditions: chronic obstructive pulmonary disease, leg ulcers, the elderly in exercise, osteoarthritis, irritable bowel syndrome, migraine and obesity. Identified discriminative and evaluative variables were used to compare RVs and SRMs of the SF-36 scales and the SF-6D index. The mean RV differences and mean SRMs differences between the SF-36 scales and the SF-6D index represented the loss or gain in sensitivity. **RESULTS:** Data were available from a total of 10,089 subjects. No single SF-36 scale consistently had the largest RV or SRM, and there was no largest RV or SRM observed for the SF-6D index in any condition studied. Comparisons showed the SF-6D index was more discriminative with a mean RV difference of 0.09, (95% CI; 0.07 to 0.12) and more responsive with a mean SRM difference of 0.08, (95% CI; 0 to 0.16) than the SF-36 scales. However, based on longitudinal RVs the index was less responsive with a mean RV difference of 0.07, (95% CI; 0.01 to 0.15) than the SF-36 scales. **CONCLUSIONS:** Moving from the SF-36 to the SF-6D index entails a loss in evaluative strength and a gain in discriminative strength, a loss/gain too small to matter given the merits of either instrument.

PMC18

ELECTRONIC PRO VERSUS PAPER PRO: WHAT DO THE PATIENTS THINK?

Ross J, Marcovitz M
Almac Clinical Technologies, Yardley, PA, USA

OBJECTIVES: To examine patients' preferences and satisfaction on completing Patient Reported Outcome (PRO) assessments in studies that compared paper-administered to electronic versions. To identify which data collection method patients prefer. To explore aspects that makes the PRO experience more positive or negative for patients. **METHODS:** A large literature search was conducted to gather articles that utilized ePRO. From that, articles were identified and reviewed that compared paper to ePRO and assessed for patient satisfaction/preferences. **RESULTS:** 119 articles were identified that utilized ePRO; 26 (21.8%) compared paper to ePRO. Of the 26, 17 (65.4%) reported on patient satisfaction/preferences. Electronic modalities consisted of handheld devices (70.6%), interactive voice response system (IVRS) (phone) (17.6 %), electronic data capture system (5.9%) and both IVRS and handheld (5.9%). Patient satisfaction/preference was assessed through either interviews (41.2%) or questionnaires (58.8%). Patients reported preferring ePRO over paper in 88.2% of the articles. Positives aspects of paper included: familiarity, not dependent on technology that may malfunction and ease of reading. Negative aspects of paper included: forgetting to complete and burden. Positive aspects of ePRO included: liked the diary's appearance, convenient, ease of data entry, fast/efficient, saves trees, reminders, overall survey experience, more fun/novel, easier on eye, more up-to-date, and comfort in handling. Negative aspects of ePRO included: system problems/failures, difficulty to read, difficult to use, instructions could have been simpler, and inability to change reminder time or enter data late. **CONCLUSIONS:** As PRO are measures that come directly from the patients, it is important to identify their preferences and aspects of what makes their experiences more positive. These findings suggest that patients overall preferred ePRO and identified more positive aspects for ePRO. Both positive and negative aspects reported are equally valuable in identifying how PRO data collection can be improved to provide patients with the most positive experiences.

PMC19

DIMENSIONS CHARACTERIZING GOOD HEALTH BY CHINESE IN CHINA

Li M[1], Bao Z[2], Zhou J[2], Luo N[3]
[1]China Pharmaceutical University, Nanjing, Jiangsu, China, [2]The First Affiliated Hospital of College of Medicine, Zhejiang University, Hangzhou, Zhejiang, China, [3]National University of Singapore, Singapore, Singapore

OBJECTIVES: Health-related quality of life (HRQoL) instruments used in China are mainly from Western countries. Such instruments may not cover all the important health dimensions relevant to Chinese people as health is a culture-specific concept. However, there is a paucity of empirical data on what good health is to Chinese people. The objective of the current study is to identify health dimensions with which Chinese people use to define health. **METHODS:** A convenience sample of 200 adult Chinese (healthy persons: 80; inpatients: 120) were interviewed face to face. Open questions were used to elicit characteristics and life domains of good health. **RESULTS:** Fourteen health dimensions were identified. The 5 most frequently alluded dimensions were: mood (35.5%), absence of disease (33.3%), mobility (25.1%), ability to work (22.4%), and eating (17.5%). Other dimensions included vitality, pain or discomfort, physical fitness, sleep, freedom, self-care, social relationship, enjoyment, and cognition. More proportion of healthy persons than patients quoted mood and self-care as dimensions of health while more patients emphasized ability to work. Males regarded eating as a health dimension more often than females while females quoted self-care and social relationship more frequently than males. With regard to age, older persons valued ability to work more than younger people while more younger people thought absence of pain or discomfort is a characteristic of good health. **CONCLUSIONS:** This study provides useful information for assessing the adequacy of HRQoL instruments developed in Western countries for the Chinese population in China.

PMC20

THE TRANSLATION AND LINGUISTIC VALIDATION OF THE EQ-5D VISUAL ANALOGUE SCALE (VAS)

Clough LA, Ashcroft-Jones AJ, Furtado T, Wild D
Oxford Outcomes Ltd, Oxford, Oxon, UK

OBJECTIVES: The EQ-5D has been translated into many languages. The Euroqol group have recently altered and clarified the VAS scale. The objective of this study was to produce translations that are conceptually equivalent to the original and to other language versions, ensuring the relevance of the translations within the target cultures. **METHODS**: A standard methodology was employed: 1 forward and 1 back translation, review and developer review; or an in-country review and developer review; linguistic validation interviews with 8 subjects, a mix of healthy people and patients, a second developer review and 2 proofreadings. **RESULTS**: The translation process highlighted numerous cultural and linguistic issues, including: 1) Cognitive interviews showed that there was no clear Dutch word for *scale*, so an explanation likening the scale to a thermometer as in the previous 3L VAS was necessary; 2) In some cultures 'mark an X on the scale' was difficult to render, and had to be amended by using alternative verb formations and formatting; 3) Though the new VAS mentions only 'health', in some languages, it was necessary to use "health **state**" to avoid confusion, e.g. in Czech "health" alone means "good health."; 4) In some languages the concepts of "health" and "health state" had different temporal associations. In Korean, "health" referred to a longer period of time, so "health today" had to be expressed by "health state today"; 5) Russian patients understood "health state" as the evaluation given by a doctor or test results, therefore "in your opinion" was added. **CONCLUSIONS**: The EQ-5D VAS has been translated and linguistically validated using a rigorous translation process. A number of cultural and linguistic issues became apparent and were resolved. The measure is now appropriate for use in multinational trials.

PMC21

PATIENT-REPORTED OUTCOMES IN PRODUCT DEVELOPMENT GUIDANCE

Mordin M[1], Lewis S[2], Gnanasakthy A[3], Demuro-Mercon C[2], Copley-Merriman K[1], Fehnel S[2]
[1]RTI Health Solutions, Ann Arbor, MI, USA, [2]RTI Health Solutions, Research Triangle Park, NC, USA, [3]Novartis Pharma, East Hanover, NJ, USA

OBJECTIVES: Patient-reported outcomes (PRO) have received increasing attention from regulatory agencies regarding intended use of the data for promotional labeling claims. However, some disease areas and/or regulatory bodies necessitate the use of PRO data to substantiate product efficacy for securing approval. Therefore, the research objective was to determine how many of the final product development guidance available from EMEA and FDA for clinical/medical research indicate PRO as a mandatory component of efficacy. **METHODS:** Final guidance documents from the EMEA and FDA were reviewed for mention of PRO. EMEA Guidance documents that fell under the following categories were excluded: Clinical Pharmacology and Pharmacokinetics, Blood and Blood Forming Organs, Blood products (including biotech alternatives), and Herbals. Included in FDA Guidance review were those listed under the "Clinical/Medical" heading. The following data were abstracted from each guidance: guidance number, name, issue date, disease area, body system classification, PRO requirement, PRO endpoint hierarchy, and a summary of the PRO language included in the guidance. PRO statements were then characterized within each of the following categories (yes/no): signs/symptoms, function/feeling, HRQOL, or patient global rating. **RESULTS:** Of the 134 final guidance documents reviewed (EMEA = 81, FDA = 53), 52 mention PRO (EMEA n = 39; FDA n = 13). Within EMEA, PRO is

indicated as primary (n = 5) or secondary (n = 22; of which 4 are secondary and/or exploratory) or both (n = 12).The majority of PRO statements are characterized as sign and symptom measures followed by HRQOL measures. Within FDA, 5 required PRO and 8 suggest use of PRO. The majority of PRO statements are characterized as sign and symptom measures, followed by measures of function/feeling. **CONCLUSIONS:** PRO data in many disease areas are viewed by regulatory agencies as supportive evidence of the primary endpoint. PRO data are essential in the support of product submissions to regulatory stakeholders, especially within EMEA.

PMC22

THE IMPACT OF A HOST COUNTRY'S CULTURE ON IMMIGRANT LANGUAGE

Hergibo P[1], Plowright R[2], Bradley C[2], Weatherley BD[3], Gekkieva M[4], Unemori E[5], Conway K[1]

[1]MAPI Institute, Lyon, France, [2]Royal Holloway University of London, Surrey, UK, [3]Momentum Research Inc, Durham, NC, USA, [4]Novartis Pharma AG, Basel, Switzerland, [5]Corthera, San Mateo, CA, USA

OBJECTIVES: To facilitate international comparison of data, PRO translations must be conceptually equivalent to the original and culturally relevant to the target country. To assess the relevance of conducting a multi-step process on a PRO translation with the aim of using it on an immigrant population speaking that language in a different country, we investigated the presence and nature of differences between the 2 language versions thus obtained. **METHODS:** Three translations were compared before and after adaptation to the context of a host country: 1) the Turkish and German Turkish version of the *Diabetes Treatment Satisfaction Questionnaire (DTSQ);* 2) the Indian Gujarati and UK Gujarati version of the *Subject Self Report on Symptoms Worksheet (SSRSW);* and 3) the Chinese Mandarin and US Mandarin version of the *National Eye Institute Visual Functioning Questionnaire (NEI-VFQ-25)*. **RESULTS:** Six of the eight items in the Turkish DTSQ were modified following cognitive debriefing with Turkish speakers in Germany. The Turkish population in Germany tends to use more old-fashioned wording which doesn't reflect the original language's recent evolution. All four items in the Gujarati SSRSW needed changing when adapting it to a UK context. Some initially translated wording was reverted back to English, or substituted with transliterated English terms. In the Mandarin NEI-VFQ-25, out of 29 items, 11 were modified when adapting it for the USA. The language used in the initial translation was considered too basic for the target population in the USA, which tends to have a higher level of education. **CONCLUSIONS:** Immigrant language is affected by the host country's culture and language, and/or by separation from the mother country, and is no longer fully comparable with the language in the country of origin. Adaptation and cognitive debriefing on immigrant populations in target countries is advisable to establish culturally relevant translations.

PMC23

A UNIVERSAL SCORING SYSTEM FOR EQ-5D : A VASTLY SIMPLER SOLUTION

Kind P

University of York, York, UK

OBJECTIVES: Country-specific social preference sets have been estimated to support the use of EQ-5D in computing QALYs for cost-utility analysis. However these value sets have limited applicability in non-economic applications since a) they incorporate the state "dead" which is irrelevant in many therapeutic settings, and(b) they are based on hypothetical preferences from 3rd parties who may not have any experience of specific EQ-5D health states. This paper reports on the construction of a scoring system for EQ-5D based on self-rated VAS values generated by individuals with current experience of those states. **METHODS:** EQ-5D data from several different UK sources were pooled yielding a total of 23,679 useable observations. The health state defined by each respondent's self-rated problem level on the 5 EQ-5D dimensions was determined, yielding a total of 139 unique EQ-5D health states. The mean VAS rating was computed for each of these states. 0/1 dummy variables were defined for each of the EQ-5D dimensions and an OLS regression analysis was performed with the mean self-rated VAS rating as the dependent variable. **RESULTS:** The model fitted the mean VAS ratings data very well ($r^2 = 0.985$) when forced through the origin. All decrements within dimension were monotonic and internally consistent. Residuals were 5 points or lower when observed and estimated values were compared. Estimated values for all EQ-5D health states were computed so that full health (11111) has a value of 100 and worst possible health (33333) has a value of 0. **CONCLUSIONS:** This methodology contrasts markedly with the more complex requirements of utility estimation and produces a weighting system that can be used to meaningfully report health status. It has applicability as a performance measurement tool with real-world interpretability. If corresponding data from other countries were included then a single global scoring system for EQ-5D could be established.

PMC24

THE EFFECT OF FRAMING ON PREFERENCES FOR MAXIMIZING QALYS

Turpcu AH[1], Le QA[2], Bleichrodt H[3], Doctor J[2]

[1]University of Southern California, Los Angeles, CA, USA, [2]USC School of Pharmacy, Los Angeles, CA, USA, [3]Erasmus University, Rotterdam, The Netherlands

OBJECTIVES: To test whether social preferences for allocating health resources are affected by the framing of questions **METHODS:** 162 students from the University of Southern California were asked four questions. Each question asked participants to select one of two possible treatments with each treatment resulting in a different distribution of outcomes for the treated population. After treatment, patients could have one of three outcomes: "Good Health", "Poor Health", or "Death." The first medication listed always had 19 fewer people in "Good Health", 23 more people in "Poor Health" and 4 fewer people in the "Death" state relative to the second medication listed. The only aspect that varied between questions was the number of patients unaffected by treatment choice. Two questions had a "standard frame", indicative of commonly asked questions in the equity literature. The remaining two questions had a "sure thing" frame, in which common outcomes between the two treatments were made apparent. Frame order was randomized for each of the participants. A key qualitative principle behind QALY maximization is that those individuals unaffected by a policy choice should not influence the policy choice. Violations of this principle were measured for each of the frames. **RESULTS:** The proportion of violations of QALY maximization (indicated by switched preference) in the "standard" frame was 0.31 (56/183); while in the "sure-thing" frame, the proportion was 0.08 (15/183). The difference between groups was statistically significant ($p < 0.001$) **CONCLUSIONS:** The most common way of asking for preferences for equality tends to foster aversion to inequality, which does not support QALY maximization. In contrast, a frame that separates common outcomes between choices may occasion preferences that maximize QALYs. These results have implications for measurement techniques such as the person tradeoff which assumes framing has no effect on preferences for health allocation.

PMC25

PATIENT PREFERENCES FOR ADHERENCE TOOLS ACROSS 10 MEDICAL CONDITIONS

Cascade E[1], Rosen SH[2], Plummer R[3], Bharmal M[4]

[1]iGuard, Falls Church, VA, USA, [2]Quintiles, Parsippany, NJ, USA, [3]iGuard, Inc, Princeton, NJ, USA, [4]Quintiles, Falls Church, VA, USA

OBJECTIVES: Recognizing that medication non-adherence is a significant cause of suboptimal health outcomes, the objective of this study was to obtain patient feedback on communication with providers and preferences for various adherence tools. **METHODS:** Online cross-sectional survey of patients with: asthma/COPD, allergies, bipolar disease, cardiovascular disease, depression, diabetes, multiple sclerosis, osteoporosis, pain syndromes, and rheumatoid arthritis. Patients completed close-ended questions about the amount of information they received from health care providers; the usefulness of various adherence tools in managing their condition; and the impact of additional disease/product information. Patient ratings of each adherence tool were scored on a scale ranging from 0 ("not at all valuable") to 3 ("very valuable"). Paired t-test was used to compare the preference for explicit adherence reminders (medication reminders via email, telephone and SMS text) versus each of the other adherence tools. The association of patient preferences for each tool with age was evaluated using Pearson correlation coefficients and using one-way analysis of variance (ANOVA) for associations with medical conditions, gender, education, family income and health insurance source. **RESULTS:** A total of 642 patients completed the survey. Forty percent reported receiving inadequate information from their physician (range: 22% for rheumatoid arthritis to 54% for cardiovascular disease). Across medical conditions groups, patients preferred adherence tools that conveyed information about medication dosing, safety, and drug interactions. Explicit adherence reminders were uniformly deemed least valuable compared to other adherence tools (all p-values < 0.0001 based on paired t-tests). There were some differences observed in preferences for adherence tools across condition, gender, and age; no significant associations were found between patient preferences and education level, family income, or source of health insurance. **CONCLUSIONS:** Patients often receive inadequate information about their medications and conditions. Medication adherence tools that educate patients may simultaneously address their desire for more information and reinforce adherence.

PMC26

EVALUATING TREATMENT SATISFACTION ENDPOINT EVIDENCE FOR EMEA REGULATORY APPROVALS

Nixon A, Quadri N, Gallop K

Oxford Outcomes Ltd, Oxford, Oxon, UK

OBJECTIVES: To document the extent to which treatment satisfaction evidence is provided in support of EMEA regulatory approvals and to evaluate the quality of evidence provided in support of treatment satisfaction claims.**METHODS:** A review of EMEA published reports for all drugs approved since a centralised process was established in 1995 was undertaken: specifically the Scientific Discussion/Public Assessment Reports were reviewed for evaluations of patient-reported treatment satisfaction. The wording and types of PROs contained within approved product labels were examined in order to establish the nature and extent of previous successful claims for treatment satisfaction.**RESULTS:** A total of 508 currently authorised medicinal product approvals were reviewed, 26 made reference to 'satisfaction' or 'satisfied' but 9 were excluded for not focusing on patient-reported treatment satisfaction thus 17 medicinal products were identified as having a direct reference to evaluating patient-reported treatment satisfaction. These 17 approvals ranged from July 1998 to July 2008, and were distributed across a broad range of pharmaco-therapeutic groups with a cluster of approvals for 'insulin analogues for injection, long lasting' (n = 4): 10/17 approvals provided limited reference to the way in which treatment satisfaction was evaluated e.g. reference to a total satisfaction score without any further details, 2/17 measured treatment satisfaction using a VAS; 5/17 referenced a specific treatment satisfaction measure. 5/17 provided treatment satisfaction of results, yet only two of these gave any details on the way in which treatment satisfaction was measured.