patient-reported outcome (PRO) instruments in clinical trials, as recommended in the FDA draft guidance for industry. A systematic review of PROs assures that the best available instrument is used to measure the preferred endpoint. As systematic reviews are scientific exercises, they require the same rigour as other aspects of research, yet current methods used to conduct systematic reviews remain variable, meaning that the quality and comparability of such reviews is not assured. Our aim was to explore the comprehensiveness, understandability, and adaptability of two widely used methodologies in conducting and modifying a standard search. **METHODS:** We compared the most common systematic review method (syntax search) and the Cochrane-collaboration recommended "Population-Intervention-Comparator-Outcome" (PICO) strategy. SCOPUS was searched using terms devised to answer the research question "which PROs have been used to date in islet cell transplantation?" The output resulting from each strategy was independently evaluated by two researchers and the methods critiqued. **RESULTS:** Both methods returned 6486 abstracts for review. Researchers were asked to identify ways in which to combine search terms to present a more manageable number for abstract screening. Both researchers agreed that PICO allowed for greater adaptability and targeted reviewing without compromising quality. Combining a priori search terms systematically according to [P and (I or C) and O], resulted in 359 abstracts. **CONCLUSIONS:** The quality of a review depends on the extent to which scientific review methods are used to minimise the risk of error and bias, but also the extent to which the search strategy is replicable and flexible. The PICO method is comparable to the standard syntax search, but offers the added benefits of being easy to implement, and sufficiently versatile to allow further targeting according to subtle changes in the research question as desired.

PMC55

THE C-STATISTIC AND THE EFFICIENCY OF THE PROPENSITY SCORES MODEL: EVIDENCE FROM SIMULATED DATA

Kiri VA[1], Feudjo-Tepie M[2]
[1]PAREXEL International, London, UK, [2]GlaxoSmithKline R&D, London, UK

OBJECTIVES: Confounding is a common source of bias in outcome studies involving observational non-randomized data. The propensity scores methodology has been suggested as a good analytical approach for handling this problem without any indication on whether a threshold exits on its predictive ability. We investigate the usefulness of the C-statistic in this regard using simulated data. **METHODS:** In each simulation, we generated 100 sets of 10,000 patients; each patient being assigned probabilities of being treated and of experiencing the outcome of interest. The process involved two logistic models, one that related treatment to a set of 10 independent covariates and the other relating outcome to treatment and the same 10 covariates, using Bernoulli distributions that assumed an odd ratio (OR) for treatment between 0.14 to 1.00 for each dataset. Propensity scores from each dataset were estimated and propensity scores-matched analysis conducted using conditional logistic regression to estimate the OR and from the 100 sets, we obtained the mean, median and bias in the estimate. Bias was defined as the difference between actual and estimated ORs as a proportion of actual. **RESULTS:** We found evidence of correlation between the levels of bias in the OR estimates and the C-statistics, with level often exceeding 300% when the C-statistic was less than 80%. **CONCLUSIONS:** Where as an elevated value of the C-Statistic may not guarantee effective correction of confounding by the resultant propensity scores derived from a given data, our study indicates that a lower value does indicate a poor capability. We suggest the C-Statistic can be adopted as a simple reporting tool on the propensity scores model in respect of its efficiency.

PMC56

THE TRANSFERABILITY OF ECONOMIC EVALUATIONS. VALIDATING THE MODEL OF WELTE

Knies S[1], Ament AJHA[1], Evers SMAA[1], Severens JL[2]
[1]Maastricht University, Maastricht, The Netherlands, [2]University Hospital Maastricht, Maastricht, The Netherlands

OBJECTIVES: One of the existing methods to assess the transferability of economic evaluations is the model of Welte, which is a decision chart method that includes general and specific knock-out criteria and a transferability checklist. This study aims to validate Welte's model with the help of a case study. **METHODS:** In this study, foreign studies were transferred to The Netherlands and then compared with a Dutch reference study. Using the case study on non-specific subacute and chronic low back pain, the cost-effectiveness of physiotherapy was compared with a multidisciplinary treatment. With the help of a systematic search several foreign studies could be identified. Based on these foreign studies two different predictions regarding costs, effects and cost-effectiveness were produced for The Netherlands. In the "all studies predictions" all foreign studies were used. In the "Welte's model predictions" only the foreign studies were used which passed the general and specific knock-out criteria. Both predictions were compared with the Dutch reference case. **RESULTS:** A total of fourteen non-Dutch studies were identified. Seven studies did not pass the general knock-out criteria and one study did not pass the specific knock-out criteria. The decision if a study was transferable was based on double retriever scores. As a result fourteen studies were included in the "all studies prediction" and six studies in the "Welte's model prediction". The predictions yielded different results and the "Welte's model prediction" proved better on costs than the "all studies prediction". The effectiveness predictions were least accurate for the "Welte's model prediction". Because of the small effectiveness difference no ICER predictions were calculated. **DISCUSSION:** Application of the model of Welte does influence cost and effects estimates when transferring economic data between countries. However, more cases should be subjected to the Welte transferability model before a final conclusion can be drawn.

PMC57

DISCONTINUOUS COST MEASUREMENT: HOW TO FILL IN THE GAPS?

Hendriks MRC[1], Evers SMAA[2], Bleijlevens MHC[2], van Haastregt JCM[2], Crebolder HFJM[2], van Eijk JTM[2]
[1]University Hospital Maastricht, Maastricht, The Netherlands, [2]Maastricht University, Maastricht, The Netherlands

OBJECTIVES: Economic evaluations are preferably performed from a societal perspective. This implies that participants report their health care utilization continuously during follow-up. Because this is a burden to participants, often resulting in missing values or withdrawal, researchers advocate collecting data discontinuously (i.e. in at least three months a year). This study aimed to compare costs of discontinuous measurement of health care utilization with those of continuous measurement, using several discontinuous measurement patterns and three imputation techniques: Individual Mean (IM), Last Observation Carried Forward (LOCF) and Next Observation Carried Backward (NOCB). **METHODS:** We used continuous health care utilization data from a trial with twelve months' follow-up and simulated several discontinuous measurement patterns combined with different imputation methods, to calculate simulated annual