

PMc8
SYSTEMATIC REVIEW RELIABILITY: EFFICACY OF TWO INDEPENDENT REVIEWERS AND TWO-STEP REVIEWS

Pietri G, Muston D

Heron Evidence Development Ltd, Stopsley, Luton, UK

OBJECTIVES: In a systematic review of literature, reviewing papers twice by two reviewers improves the likelihood of correctly including/excluding citations. Also, reviewing first the abstract then the full text allows the reviewer to have had more exposure to the literature, and thus to be more accurate in his decision. A statistical model was fitted and paired t-tests were performed to determine between-reviewer and between-review reliability and variation. **METHODS:** Inclusion/exclusion decisions made by two reviewers at the abstract and the full-text reviews of six recently conducted systematic reviews (one economic and five clinical) were analysed and compared to the final inclusion/exclusion decision. For both reviewers, sensitivity (the proportion of correctly included citations) and specificity (the proportion of correctly excluded citations) were modelled using bayesian Poisson regression. Paired t-tests were performed to evaluate the improvements of sensitivity and specificity between the abstract and the full-text reviews. **RESULTS:** The sensitivity of reviewer two was significantly higher in one out of six systematic reviews at abstract review, and in another one at full-text review. Reviewer two's specificity was also significantly higher in three systematic reviews at abstract review, and in one systematic review at full-text review. Reviewers were on average 89.6% sensitive and 95.7% specific at including or excluding publications at abstract review. At full-text review, the average sensitivity increased significantly (97.1%, $p = 0.006$) whereas the specificity remained similar (95.9%, $p = 0.81$). **CONCLUSIONS:** While first and second reviewers tend to have similar sensitivity in including citations at abstract review, reviewer two tends to be more accurate at excluding citations, thus increasing the likelihood of correctly excluding citations. At full-text review, sensitivity and specificity tend to be similar between the two reviewers, but sensitivity increases significantly compared to abstract review. This shows that the full-text review provides further confirmation that the included citations are indeed relevant.

PMc9
SIGNIFICANCE OF INCORPORATING COMMUNITY-BASED DATA OF A TARGET POPULATION INTO PHARMACOECONOMIC MODELS

Farahani P

Berkshire Medical Center, University of Massachusetts Medical School, Pittsfield, MA, USA

Community-based aspects of therapeutic can influence the outcomes of pharmacoeconomic evaluations. **OBJECTIVES:** To evaluate the effect of incorporating data from RCTs versus data from community clinical practices on the results of an economic model of statins. **METHODS:** The benefit of reducing LDL-C was incorporated into a model to calculate reduction in cardiovascular events and resulted economic outcomes. Data for LDL-C reduction from a head-to-head RCT [Am Heart J 2002;144:1044-51] were obtained for rosuvastatin (starting 5 mg) versus atorvastatin (starting 10 mg) with up-titration doses. A distribution of cardiovascular risk for users [N = 100,000, duration 5 years] in Canadian population [Clin Invest Med 2007;30:E63-E69] was assumed. Then, to illustrate the significance of the population level data, the data from the Canadian community-based clinical practice settings was removed from the model and the original RCT probability distribution for cardiovascular risk strata was applied into the model. **RESULTS:** Using community-based data modelling rosuvastatin and atorvastatin can prevent 9505 and 8702 cardiovascular events (non-fatal MI and stroke). Reduction in non-fatal MI and stroke can be translated to \$252,300,392 (CDN), and \$230,980,624 direct cost savings, respectively. Incorporating the RCT cardiovascular risk distribution, rosuvastatin and atorvastatin can prevent 7129 and 6712 cardiovascular events. This could lead to \$180,214,565 and \$178,152,982 direct cost savings for the Canadian health care system (adherence to therapy was assumed to be at the level of RCT). **CONCLUSIONS:** The distribution of cardiovascular risk was dissimilar between the RCT and the Canadian community-based data. The proportion of low risk patients enrolled in the RCT was significantly higher in comparison with the proportion of low risk patients on statin therapy in the Canadian community. Therefore, in this case the magnitude of cost savings would considerably be reduced if the RCT data were incorporated into the model instead of the community-based data.

CONCEPTUAL PAPERS & RESEARCH ON METHODS – Cost Methods

PMc10
COMPARISON BETWEEN PROPENSITY SCORE AND TRADITIONAL COVARIATE ADJUSTMENT WITH LOGISTIC REGRESSION MODELS IN ESTIMATING THE AVERAGE TREATMENT EFFECTS (ATEs): RESULTS FROM MONTE CARLO SIMULATIONS

Le QA, Hay JW

University of Southern California, Los Angeles, CA, USA

OBJECTIVES: Several recent systematic reviews of published studies in clinical research using both traditional covariate adjustment regression and propensity score (PS) methodology to control for confounding reported that both methods produced similar estimating results. In this study, we used Monte Carlo simulations to evaluate performance of different models derived from the two methods. **METHODS:** Sixteen dichotomous variables (X1 – X16) with different levels of association between treatment and outcome were randomly and simultaneously generated from the independent

Bernoulli distribution with a probability of 0.5. Twelve variables (X1, X2, X3, X5, X6, X7, X9, X10, X11, X13, X14, and X15) are associated with treatment assignment; while (X1 – X12) are associated with the outcome; and X16 is neither associated with treatment assignment nor outcome. We investigated six models, of which two models using traditional covariate adjustment with logistic regression (LR), one model without adjustment for confounders, and three models using PS as followed: (1) model 1: adjusted for 16 variables; (2) model 2: adjusted for 12 variables associated with outcome; (3) model 3: no adjustment for confounders; (4) model 4: adjusted for the PS; (5) model 5: stratified on 5 quintiles of the PS; and (6) model 6: 4-digit match on the PS. **RESULTS:** The conventional covariate adjustment with LR (model 1 and 2) consistently and unbiasedly estimated the specified ATEs. Without covariate adjustment, model 3 produced biased estimations. Among all three PS models (models 4, 5, and 6), only model 6 resulted in unbiased estimation of the specified ATEs. However, compared with traditional covariate adjustment with LR, model 6, though produced unbiased estimates, had larger variances. **CONCLUSIONS:** For dichotomous confounders, adjustment and stratification on the quintiles of the PS, though widely used in clinical research, produce biased estimations; while conventional covariate adjustment with LR resulted in unbiased estimations with small variances.

PMc11
GROWTH, CHARACTERISTICS AND QUALITY OF THE COST-UTILITY ANALYSIS LITERATURE THROUGH 2006

Greenberg D¹, Fang C¹, Cohen JT¹, Eldar-Lissai A², Neumann PJ¹

¹Tufts-New England Medical Center, Boston, MA, USA, ²University of Rochester, Rochester, NY, USA

OBJECTIVES: To describe the growth over time and methodological quality of cost-utility analyses (CUAs) published in the peer-reviewed literature through 2006. **METHODS:** This study updates and expands our previous work, which examined CUAs through 2003 (n = 795). We systematically searched the English-language literature for original CUAs published through 2006 using Medline and other databases. Two trained readers independently audited each study and collected data on a wide variety of elements related to study origin, methods, and reporting of results. **RESULTS:** We identified 1431 original CUAs (currently included in the Tufts Medical Center Cost-Effectiveness Analysis Registry, www.cearegistry.org). The more than 200 CUAs published annually from 2004 to 2006 nearly doubled the size of the CUA literature. Most analyses pertain to the U.S. (51%), followed by the U.K. (16%), and Canada (7%). CUAs cover a wide range of disease areas, including cardiovascular diseases (15%), infectious diseases (14%), and cancer (13%), and examine interventions for tertiary care (62%), secondary prevention (23%), and primary prevention (15%). Although most studies adhere to guidelines for conducting and reporting CUA results, their average quality (4.2 ± 1.1 on a 1–7 Likert scale) did not change substantially over time. Study quality was higher for CUAs published in experienced journals (publishing a total of ≥ 10 CUAs); (n = 31 journals), vs. other (n = 390 journals); (4.5 ± 1.1 vs. 3.9 ± 1.1 , respectively; $p < 0.0001$), for CUAs evaluating pharmaceuticals vs. other interventions (4.3 ± 1.0 vs. 4.1 ± 1.2 , respectively; $p < 0.0001$), and for CUAs supported by a government organization vs. other sponsors (4.4 ± 1.0 vs. 4.1 ± 1.2 , respectively; $p < 0.0001$). **CONCLUSIONS:** The rapid growth in published CUAs demonstrates the increased role these analyses play in informing resource allocation decisions. Journals should conduct more rigorous reviews and assure adherence to guidelines for conducting and reporting CUA results.

PMc12
VALIDATION OF HEALTH CARE RESOURCE USE QUESTIONNAIRES IN MEXICO

Tenorio C¹, Vargas J², Ivanova S³, Martínez-Fonseca J³, Paladio A³, Mould-Quevedo J⁴

¹Instituto Nacional de Cancerología, Mexico, DF, Mexico, ²Econopharma Consulting SA de CV, Mexico, DF, Mexico, ³Econopharma Consulting SA de CV, Mexico City, Mexico, ⁴Pfizer Mexico, Mexico City, Mexico

OBJECTIVES: Current Mexican health care resource-use questionnaires are subjective and not validated through a scientific technique. This generates heterogeneity in costs obtained from several economic evaluations (EE), uneasy to compare methodologically and useless to include in future EE alongside clinical trials. The objective of this study was to develop and validate three instruments to assess resource utilization from the societal perspective. **METHODS:** Three specific instruments were constructed for patients with cardiovascular diseases (hypertension and acute myocardial infarction); oncological (breast, renal and lung cancer) and HIV/AIDS with the aid of a special Mexican consensus. Instruments were subdivided into two sections: recover medical data from hospital records and a survey with patients. The latter was applied to adult Mexican patients from the Social Security Mexican Institute (IMSS) between June and October 2008. Instruments include five dimensions: (diagnostic, treatment and medical follow-up, support services, clinical data and epidemiological data) and pretend to be helpful to obtain direct costs (data collected from hospital records); and indirect and out-of-pocket expenses (data collected from the patient). Sample size estimate used epidemiological Mexican data and statistical tests were performed to demonstrated internal consistency reliability (Cronbach's alpha), interclass correlation coefficient, component analysis (parametric bootstrapping) and construct validity was also evaluated. **RESULTS:** A total of 213 patients were included in the study (31.9% cardiovascular, 30.0% oncological; 38.1% HIV/AIDS). Mean age of participants was 54.9 ± 9.08 years; 49.3% were female. Internal consistency was found in all questionnaires: Cronbach's alpha reported range for cardiovascular patients from 0.73–0.78; cancer patients 0.67–0.80, and HIV/AIDS patients from 0.82–0.87. Higher internal consist-