volumes and costs. Correlations and paired T-tests were used to compare simulated annual costs with actual annual costs from the continuous measurements. RESULTS: Analyses confirmed that discontinuous measurements using cost diaries offer good estimates of annual health expenditures, but measurement patterns and imputation methods did influence the outcomes, as the correlations differed between methods. The best estimated annual costs were obtained by random cohort measurement, using three random cohorts, ensuring that at least a third of the participants were measuring costs each month, combined with IM imputation. Discontinuous measurement of health expenditures carries a small risk of missing infrequent expensive events, which may result in underestimation of annual costs. CONCLUSIONS: To reduce the burden on participants in future economic evaluation, we recommend calculating annual costs from discontinuous measurements in random cohorts, combined with IM imputation.

THE POTENTIAL PENALTY FOR NOT SAMPLING FROM THE RISK SET IN NESTED CASE-CONTROL DESIGNS: EVIDENCE FROM SIMULATED DATA
Kiri VA1, Feudjo-Tepie M1
OBJECTIVES: Appropriate design and efficient analytical strategy are generally considered as some of the prerequisites for a valid and reliable health outcomes research on non-randomized observational data. For rare outcomes, the case-control design is often presented as the most efficient, whereby, proper selection of controls is crucial. Using simulated data for the nested case-control design, we assess the relative efficiencies of two sampling strategies for the controls- the version where controls can never be cases (design 1) and the recommended approach in which controls are sampled from the risk sets such that some controls can be future cases (design 2). METHODS: In each simulation, we assumed an underlying hazard that follows a Weibull distribution with inputted values for the scale and shape parameters to generate 100 sets of cohorts of 4000 patients in treatment groups (i.e. treated and untreated). The process also involved an assumed hazard ratio for treatment and 3 factors that required adjustments. Designs 1 and 2 were then applied successively on each of the resulting datasets and then analysed to obtained for each design, the estimated odds ratio (OR- an approximate of the inputted hazard ratio) and its 1st and 3rd quartiles. RESULTS: We considered over 50 scenarios for hazard ratio that varied between 0.3 and 4.0. The absolute differences between the inputted hazard ratio and the estimated odd ratio ranged from 0.01–8.00 and from 0.01–0.50 for designs 1 and 2 respectively. The inputted hazard ratio was within the inter-quartile range of the OR in less than 5% of the runs with design 1 but more than 80% with design 2. CONCLUSIONS: Our study suggests that in nested case-control designed studies, if controls are not sampled from the appropriate risk sets, we can expect much larger bias in our estimates than with correct sampling.

CONSISTENTLY ESTIMATING RISK DIFFERENCE IN A JURISDICTION OF INTEREST: ODDS SOLUTION TO RELATIVE RISK FALLACIES
Eckermann S1, Coory M2, Willan AR3
1Flinders University, Adelaide, South Australia, Australia, 2University of Queensland, Brisbane, Queensland, Australia, 3SickKids Research Institute, Toronto, ON, Canada
OBJECTIVES: Economic analyses in health technology assessment often require estimation of absolute risk difference (ARD) for outcomes such as survival or progression, given base risk in the jurisdiction of interest and trial evidence of treatment effects. We demonstrate that odds ratios (OR) provide distinct advantages over relative risk (RR) in consistently estimating such ARD in direct and indirect comparisons. METHODS: Use of RR is shown to lead to inferential anomalies in estimating ARD, while consistently estimated using OR. These inferential anomalies and odds solution are illustrated for indirect comparison of Natiluzimab versus Interferon beta-1b for multiple sclerosis, as well as direct comparisons. RESULTS: Standard use of relative risk to calculate ARD in indirect comparison suggests Natiluzimab is more effective than Interferon for progression (RR = 0.70, ARD = 21% for a base risk of 70% progression) but less effective than Interferon for no progression (RR = 0.84, ARD = 4.8%).
inherent risk differences. For direct comparisons ARD is shown to be consistently estimated with OR but change with framing of effects using RR wherever epidemiological risk differs from trial risk in the comparator arm. CONCLUSIONS: Odds ratios allow consistent estimation of absolute risk differences regardless of framing of effects in direct and indirect comparisons. This overcomes inherent anomalies that arise with use of relative risk in such comparisons whenever base risk differs in the jurisdiction of interest from that in trials, or base risk in the common arms differs in indirect comparisons. Consequently, odds ratios avoid selection biases in framing of effects inherent with risk ratios and are suggested as the preferred metric in estimating such risk differences.

**PMC61**
The Implicit Value of Statistical Life: Estimates Derived from Public Interventions Implemented in The Netherlands

Goegebels AFG1, Ament AJHA1, Novák A2, Veraart CPWM1, Severens JL2

1Maastricht University, Maastricht, The Netherlands, 2Anovák-Services, Apeldoorn, The Netherlands, 3Roche Nederland BV, Woerden, The Netherlands, 4University Hospital Maastricht, Maastricht, The Netherlands

OBJECTIVES: The economic literature suggests the Value of Statistical Life (VSL) as a common measure of efficiency for life saving interventions throughout different societal sectors. Policy decisions in The Netherlands have not yet been explicitly based on this measure, however a trade off between wealth and mortality risk is made implicitly when deciding whether or not to implement a life saving intervention. This study aimed to gain insights into this trade off, referred to as Implicit Value of Statistical Life (IVSL), by means of a retrospective investment analysis of interventions implemented in The Netherlands.

METHODS: A literature search was conducted to find life saving intervention cases meeting the requirements for a uniform IVSL calculation and additional inclusion criteria. A sample of 10 cases was included in the study and concerned interventions implemented in the water control, consumer safety, transport and health care sector. RESULTS: IVSL estimates derived from the cases ranged from €1 to almost €1 million. Differences were most extreme when comparing IVSL estimates of interventions implemented in different societal sectors. However, estimates also varied greatly between interventions in the same sector and even within the same intervention, when critical assumptions were altered. CONCLUSIONS: Despite limited comparability of IVSL estimates, our findings suggest that there are great imbalances between societal investments for preventing a statistical death. This highlights the need to develop ways to increase transparency and efficiency of policy decisions by systematically taking the Value of Statistical Life into account. Given the conceptual problems inherent to the IVSL, future research should focus on the potential merit of explicit VSL measures for decision making. Since the consequences of life saving interventions are not restricted to mortality reduction, research should also address the question whether there is a need to incorporate broader health and other consequences of life saving interventions in the measure of efficiency.

**PMC62**
Universal Translation and Neuropsychological Comparisons of Patients from US-Mexico Border Region and Spain

Herzberg T, Arnold B

Evanston Northwestern Healthcare, Evanston, IL, USA

OBJECTIVES: There are two main approaches to questionnaire translation for languages spoken in multiple countries called “universal approach” and “country specific approach”. The universal approach postulates that one translation can be developed via participation and consensus between native speaking translators from various countries where that language is spoken. Demonstrating that the universal approach is possible for questionnaires has been the life work of neuropsychologist Dr. Lidia Artiola Fortuny. METHODS: This poster will share the results of a study published in the Journal of the International Neuropsychological Society where participants from the US-Mexico border region (N = 185) and Madrid, Spain (N = 205) were compared on 16 Spanish language neuropsychological measures, with special attention to avoid item content that was specific to one geographic group. Differences in socio-economic, education and health were considered. Samples were drawn from volunteers in each community between the ages of 18 and 76 with 0 to 20 or more years of formal education who claimed Spanish as their first language and demonstrated native fluency in the language. Participants were excluded from the study if they had past neurological, emotional, psychological issues or learning difficulties. RESULTS: Analyses of variance were performed to study place of birth effects on performance on each measure. Findings indicate that the populations from Spain and the Borderland obtained similar results for most of the measures. Participants did not report difficulties with the instructions or test items. CONCLUSIONS: Dr. Artiola Fortuny asserts that Spanish speaking populations do not differ any more than mainstream English speaking populations such as the United States, England or Australia, and that one should capitalize on the great amount of linguistic overlap across populations that share the Spanish language.

**PMC63**
Systematic Review Reliability: Sensitivity and Specificity of One vs. Two Reviews

Muston D, Alnwick K

Heron Evidence Development Ltd, Letchworth, Hertfordshire, UK

OBJECTIVES: In a systematic review of literature, reviewing abstracts twice by two reviewers improves the likelihood of correctly including relevant citations and excluding irrelevant ones. During the second abstract review, the reviewer has had more exposure to the literature, and thus may be more accurate than the first reviewer. A statistical model was fitted to determine between-reviewer and between-review reliability and variation.

METHODS: Inclusion/exclusion decisions made by two reviewers in the abstract review stage of six recently conducted clinical and economic systematic reviews were analysed in the context of the final inclusion/exclusion decision. For the first and second reviewers, sensitivity (the proportion of correctly included citations) and specificity (the proportion of correctly excluded citations) were modelled using bayesian poisson regression.

RESULTS: Across one economic and five clinical systematic reviews, the sensitivity of reviewer one ranged from 82% to 95%; the second reviewer’s sensitivity ranged from 80% to 98%. The specificity of reviewer one ranged from 94% to 98%; the second reviewer’s specificity ranged from 92% to 99%. The pattern of results varied substantially between reviews. In the breast cancer, hyperlipidemia, and anaesthesia reviews, the