naired to measure the severity of the disease were assessed to assess whether there was a difference between the treatment groups in terms of change in severity over the 12-month period. METHODS: Late transition Analysis is used to explain the responses to the questionnaire by grouping patients into categories (severity groups) based on their responses. There are three parameters that can be estimated using LTA: Membership probabilities (probability of belonging to a particular severity category), transition probabilities (probability of moving to a particular severity category) and item response probabilities. These parameters are compared between the two treatment groups to determine if there is a difference between them. Two covariates were included in the model to investigate their effect on the outcome. RESULTS: As shown in Table 1, there was no significant difference between the treatment groups in terms of Membership probabilities or Transition probabilities. One of these covariates was found to have a significant effect on the responses. The effect of the covariate was different for the two treatment groups and had a significant impact on the placebo effect. CONCLUSION: The objective of this contribution was to describe both approaches and compare them with respect to their suitability as a method for aggregating multiple patient-relevant endpoints within IQWiG’s efficiency frontier concept. METHODS: A catalogue with criteria has been established to assess both approaches with regard to their suitability for aggregating multiple patient-relevant endpoints. The catalogue comprises nine relevant legal and methodological aspects: two criteria were identified based on legal requirements; three criteria were included considering IQWiG method requirements; lastly, four general methodological requirements were considered. RESULTS: Both methods were assessed based on these criteria. Two criteria were identified that could be met by both CA and AHP. The remaining seven criteria could be met by either CA or AHP. None of IQWiG’s proposed approaches for prioritizing and weighting multiple patient-relevant endpoints is likely to fully address the methodological issues that have been assessed with regard to legal and methodological requirements. CONCLUSIONS: With the presence of multiple patient-relevant endpoints the implementation of the efficiency frontier concept remains unclear due to lack of methodological guidance. Further (empirical) research in methods for aggregating multiple patient-relevant endpoints is needed.

PRM153
IDENTIFYING MULTIPLE PATIENT-RELEVANT ENDPOINTS - A METHODOLOGICAL COMPARISON OF CONJUNCTIVE ANALYSIS AND ANALYTIC HIERARCHY PROCESS CONSIDERING IQWiG’S EFFICIENCY FRONTIER CONCEPT
Neidhardt K1, Wasmuth T1, Schmid A2
Mundipharma GmbH, Nürtingen, Germany, 1University of Bayreuth, Bayreuth, Germany
OBJECTIVES: The Institute for Quality and Efficiency in Health Care (IQWiG) in Germany evaluates benefits/harms and economic implications of medical interventions. For the purpose of cost-benefit analysis, IQWiG has developed the efficiency frontier concept to determine the maximum reimbursable price for pharmaceuticals. Within this concept benefits/harms are evaluated for each patient-relevant endpoint. Methodological problems arise with the presence of multiple patient-relevant endpoints because recommendations for the maximum reimbursable price will likely be imprecise. Conjoint analysis (CA) and analytic hierarchy process (AHP) are being discussed as potential approaches to aggregate multiple patient-relevant endpoints. The objective of this contribution was to describe both approaches and compare them with respect to their suitability as a method for aggregating multiple patient-relevant endpoints within IQWiG’s efficiency frontier concept. METHODS: A catalogue with criteria has been established to assess both approaches with regard to their suitability for aggregating multiple patient-relevant endpoints. The catalogue comprises nine relevant legal and methodological aspects: two criteria were identified based on legal requirements; three criteria were included considering IQWiG method requirements; lastly, four general methodological requirements were considered. RESULTS: Both methods were assessed based on these criteria. Two criteria were identified that could be met by both CA and AHP. The remaining seven criteria could be met by either CA or AHP. None of IQWiG’s proposed approaches for prioritizing and weighting multiple patient-relevant endpoints is likely to fully address the methodological issues that have been assessed with regard to legal and methodological requirements. CONCLUSIONS: With the presence of multiple patient-relevant endpoints the implementation of the efficiency frontier concept remains unclear due to lack of methodological guidance. Further (empirical) research in methods for aggregating multiple patient-relevant endpoints is needed.

PRM154
SOURCE DATA VERIFICATION -- A SURVEY OF CURRENT PRACTICE
Veelhus EM1, Milka ES2, Richards MS3, Orr M2, Semrau F4
PPD, Karlsruhe, Germany, 1Erasmus University Rotterdam, Rotterdam, Netherlands, 2Erasmus University, Rotterdam - Erasmus University Rotterdam, Rotterdam, The Netherlands
OBJECTIVES: We examined 38 HTAs published from 2005-2011 covering 13 disease conditions from 6 agencies and had an opposite effect on the Placebo group compared the effect on the Active group. CONCLUSIONS: It has not yet been demonstrated whether non-collapsibility affects T. Censoring induces bias when it is associated with patient characteristics and no adjustment is undertaken. The objectives of this study were to disentangle the effects of non-collapsibility and censoring bias and assess their impact on estimates of T. METHODS: Survival, treatment and five normally distributed prognostic covariates were simulated in RCT-like datasets with and without censoring. We investigated scenarios with an increasing number of covariates where we can calculate the HR for treatment and T. RESULTS: For uncensored data, HRs decreased with the inclusion of additional covariates, while T remained constant. For the censored data, T increased sharply with the inclusion of additional covariates, while the HRs decreased. The estimates of the full model of both outcomes were close to the means from the dataset, although the model was estimated on censored data. CONCLUSIONS: Analysis of the synthesized data makes it possible to distinguish between the impact of non-collapsibility and censoring on HR and T. While the HR is non-collapsible, T is collapsible. It can be used in cost-effectiveness analysis and meta-analysis of clinical trials to aggregate data when the HRs are estimated using a non-collapsible model. The efficiency frontier concept remains unclear due to lack of methodological guidance. Further (empirical) research in methods for aggregating multiple patient-relevant endpoints is needed.

RESEARCH ON METHODS - Study Design
PRM155
IS IT JUST SEMANTICS? THE USE OF “EFFICACY” AND “EFFECTIVENESS” IN COMPARATIVE EFFECTIVENESS REVIEWS AND HEALTH TECHNOLOGY ASSESSMENTS (HTAs)
Jaksa A, Rubinstein E, Ho YS, Daniel A
Context Matters, Inc., New York, NY, USA
OBJECTIVES: To explore and quantify the extent to which the terms “efficacy” and “effectiveness” are used consistently and correctly in and Health Technology Assessments (HTAs). Efficacy describes a drug’s effect in ideal and controlled circumstances (i.e. in clinical trials). Effectiveness describes the success of a drug in usual or “real world” practices in which all conditions cannot be controlled. Effectiveness is much more difficult to assess and is often measured by observational studies or regression modeling. While HRs are known to be non-collapsible, it has not yet been demonstrated whether non-collapsibility affects T. Censoring induces bias when it is associated with patient characteristics and no adjustment is undertaken. The objectives of this study were to disentangle the effects of non-collapsibility and censoring bias and assess their impact on estimates of T. METHODS: Survival, treatment and five normally distributed prognostic covariates were simulated in RCT-like datasets with and without censoring. We investigated scenarios with an increasing number of covariates where we can calculate the HR for treatment and T. RESULTS: For uncensored data, HRs decreased with the inclusion of additional covariates, while T remained constant. For the censored data, T increased sharply with the inclusion of additional covariates, while the HRs decreased. The estimates of the full model of both outcomes were close to the means from the dataset, although the model was estimated on censored data. CONCLUSIONS: Analysis of the synthesized data makes it possible to distinguish between the impact of non-collapsibility and censoring on HR and T. While the HR is non-collapsible, T is collapsible. It can be used in cost-effectiveness analysis and meta-analysis of clinical trials to aggregate data when the HRs are estimated using a non-collapsible model. The efficiency frontier concept remains unclear due to lack of methodological guidance. Further (empirical) research in methods for aggregating multiple patient-relevant endpoints is needed.

Czakonczy L, Solnica B
Polish Medical Journal, Cracow, Malopolska, Poland
OBJECTIVES: Acquisition of scientific data required for the rational decisions on health policy has become an important tool in determining the validity of the financing methods of treatment from public funds based on the health technology assessment (HTA). The primary source of scientific evidence for health technology assessment are randomized controlled trials (RCTs), because of their features (e.g. randomization or blinding) reducing methodological bias. These features may become less important when selecting the subset of trials for the evaluation of the results and conclusions to the everyday practice. In this situation an important role begin to play pragmatic randomized controlled trials (PRCTs), providing highly reliable information about the effectiveness in contrast to observational studies or registries. However, the impact of PRCTs on the efficiency frontier concept remains unclear due to lack of methodological guidance. Further (empirical) research in methods for aggregating multiple patient-relevant endpoints is needed.

CONCLUSIONS: The results of this survey and its planned follow-up survey will be helpful in evaluating optimal methods and levels of partial SDV in both clinical trials and observational studies.

PRM156
A PRAGMATIC RANDOMIZED CLINICAL TRIALS – DESIGN AND QUALITY ASSESSMENT OF THE SOURCE OF EFFECTIVENESS DATA
Kaczynski L, Solnica B
Polish Medical Journal, Cracow, Malopolska, Poland
OBJECTIVES: Acquisition of scientific data required for the rational decisions on health policy has become an important tool in determining the validity of the financing methods of treatment from public funds based on the health technology assessment (HTA). The primary source of scientific evidence for health technology assessment are randomized controlled trials (RCTs), because of their features (e.g. randomization or blinding) reducing methodological bias. These features may become less important when selecting the subset of trials for the evaluation of the results and conclusions to the everyday practice. In this situation an important role begin to play pragmatic randomized controlled trials (PRCTs), providing highly reliable information about the effectiveness in contrast to observational studies or registries. However, the impact of PRCTs on the efficiency frontier concept remains unclear due to lack of methodological guidance. Further (empirical) research in methods for aggregating multiple patient-relevant endpoints is needed.

CONCLUSIONS: The results of this survey and its planned follow-up survey will be helpful in evaluating optimal methods and levels of partial SDV in both clinical trials and observational studies.