intervals HPV-based screening was more effective than cytology alone, with a relative reduction in cervical cancer incidence of 49%-90% compared to 33%-80% with cytology alone (depending on screening intervals). In HPV- compared to cytology screening the incremental gain in effectiveness was higher with extended screening intervals and the increase in harms lower. Based on the BII, 12 of 17 screening strategies were more effective compared to conventional screening, both in Germany. Biennial HPV-screening was similarly effective as annual cytology and reduced unnecessary treatment. Moving from biennial HPV- with cytological triage to annual HPV-screening alone results in an incremental harm-benefit ratio of 15-53% for unnecessary treatments per additional prevented cervical cancer case (depending on screening adherence rate). CONCLUSIONS: The benefit-harm frontier is a useful tool to demonstrate the trade-off between expected gains and risks of different screening strategies. Based on the BII, HPV-based cervical screening is more cost-effective than cytology alone, but has a higher risk of overtreatment when used in annual screening. In the German health care context, depending on screening adherence rates, HPV screening in women every 3 or 5 years is only effective as annual cytology with significantly reduced unnecessary treatments.

PM2
EVALUATING WHETHER INCONSISTENCIES ARE PRESENT IN A MIXED TREATMENT COMPARISON OF TRough FORCED EXPIRATORY VOLUME IN 1 SECOND AT 12 WEEKS
Buckley F1, Baldwin M2, Kininger D3, Jansen J1, Coppe S4
1Mapitol, Boston, MA, USA, 2Novartis Pharmaceuticals UK Limited, Horsham, West Sussex, UK, 3Novartis Pharma AG, Basel, Switzerland, 4Mapitol, Toronto, ON, Canada
OBJECTIVES: To evaluate whether there are inconsistencies in the network of randomised controlled trials (RCTs) used for a network meta-analysis (NMA) comparing alternative long-acting bronchodilators among patients with moderate to severe chronic obstructive pulmonary disease (COPD) in terms of trough forced expiratory volume in 1 second at 12 weeks. METHODS: We evaluated the change from baseline in FEV1 as observed with placebo, tiotropium 18µg/5µg once daily (OD), salmeterol 50µg twice daily (BID), formoterol 12µg BID, aclidinium 400µg BID, glycopyrronium 50µg OD and indacaterol 150µg OD in RCTs identified with a systematic literature review were synthesized with a NMA. Where possible, treatment estimates from fixed effect (FE) and random effects (RE) NMA were used. RESULTS: Consistency across studies and independent means (iM) models (pooled direct evidence) were compared to assess whether any inconsistencies in the network were present. RESULTS: Thirty-two RCTs identified through a systematic literature review were included in the analysis. Direct evidence was available for the monotherapies versus placebo, the combination therapies versus tiotropium, for indacaterol+glycopyrronium versus placebo, and for tiotropium versus salmeterol. The largest differences between the estimated treatment effect estimates from the iM and the iM models were compared for the comparators between indacaterol 150µg versus tiotropium (FE difference=0.025 [95% Credible Intervals [95%CrI]: 0.002, 0.047]; RE difference=0.027 [95%CrI: 0.007, 0.046]; indacaterol+glycopyrronium versus placebo (FE difference=-0.022 [95%CrI: -0.053, 0.008]; RE difference=-0.018 [95%CrI: -0.059, 0.022]); and indacaterol+glycopyrronium versus tiotropium (FE difference=-0.011 [95%CrI: -0.014, 0.036]; RE difference=-0.015 [95%CrI: -0.024, 0.035]). CONCLUSIONS: Based on a comparison of the findings of a NMA and iM models, inconsistency across studies and independent means models more likely indicates an inconsistency in the network of RCTs identified that will be explored through additional sensitivity analyses.

PM3
TESTING THE EUUNETHA INTERNAL VALIDITY OF RANDOMIZED CONTROLLED TRIALS GUIDELINE AND TOOL IN HUNGARY
He Review 1, Cope S2, Kasper S3, Crombet T4
1National Institute for Quality- and Organizational Development in Healthcare and Medicines, Budapest, Hungary
OBJECTIVES: The reliability of the results of a randomized trial depends on the extent to which potential selection bias has been avoided. We tested the EUUNETHA Internal Validity guideline so that to harmonize our risk of bias assessments with the European standard and finally to improve the reliability of relative efficacy and cost effectiveness assessments for decision makers in Hungary. METHODS: We translated the risk of bias standardized assessment questions of the EUUNETHA Internal Validity guideline so that to harmonize our risk of bias assessments with the European standard and finally to improve the reliability of relative efficacy and cost effectiveness assessments for decision makers in Hungary. RESULTS: We found adequate randomization sequence generation in seven studies and we marked it unclear in one trial where the non-inferiority results in the per-protocol subscale in non-pharmacological trials (50%). However, even within the common outcomes, the specified level of reduction to define a relevant response varied. Among trials reporting PANS total, five different levels of reduction were defined (≥20%, ≥25%, ≥30%, ≥40%, ≥50%). Common outcomes were also measured differently in terms of improvement from baseline and proportion achievement/ remission, with little consensus on clinical meaningfulness. CONCLUSIONS: The RCTs included in this review reported a broad range of outcomes, making comparison of different therapies a complex task. The disparity in outcomes between pharmacological and non-pharmacological outcomes scales highlights the challenges in designing trials to demonstrate clinical benefit.

PM5
MULTI-DIMENSIONAL CAPTURE OF PATIENT-RELEVANT ENDPOINTS IN REGULATORY TRIALS AND HEALTH TECHNOLOGY ASSESSMENTS IN ONCOLOGY TWO YEARS AFTER INTRODUCTION OF THE GERMAN AMNOG HEALTH CARE REFORM
Dintsios CM1, Knoerzer D2, Duuenae AA3, Schwartz FW4, Roei5
1German Association of Research-based Pharmaceutical Companies (vfa), Berlin, Germany, 2Roche Grenzach-Wyhlen, Germany, 3Hannover Medical School, Hannover, Germany, 4Novartis Pharma AG, Basel, Switzerland, 5Medical University of Vienna, Vienna, Austria
OBJECTIVES: To study whether there was an improvement in patient-reported outcome (PRO) endpoints biennial or triennial HPV-screening for women of reproductive age and 30+ years is similarly effective and safe. We used the German HTA tool to determine the acceptability of mortality and QoL endpoints typically evaluated in oncology. Considerable variability was observed in the acceptance of PROs in morbidity. Evaluating additional benefit based only on morbidity and QoL endpoints underestimates the potential value of new drugs. Multiple endpoints, which capture all three dimensions, should be evaluated in regulatory trials and accepted by IQWiG and G-BA to confirm patient-relevant additional benefit.

PM6
THRESHOLD SELECTION IN BIOMARKERS USING COX REGRESSION: AN APPLICATION TO NON-SMALL-CELL LUNG CANCER
Lelles P1, Sanchez L2, Ballesteros J3, Garcia B4, Rodrigues FC5, Viada C1, Popa X7, Crombet T7, Lage A1
1Center of Molecular Immunology, Havana, Cuba, 2University of the Basque Country UPV/EHU, Leioa, Spain
OBJECTIVES: To select thresholds for predictive biomarkers using Cox regression. METHODS: We used data from a Cuban trial designed to assess the efficacy of immunotherapy against the epidermal growth factor (EGF) to test our approach. The trial included 122 patients diagnosed with non-small-cell lung cancer (NSCLC) who had basal EGF concentration available. The EGF concentration was analysed as a predictor of immunotherapy success over the range of all possible values of the biomarker (0-515 ± 115 nmol/l). For each candidate threshold, we developed a Cox model adjusted to assess survival. We then identified the w*, with significant treatment results to find (a) the lowest biomarker threshold where the effect of treatment was significant and (b) also to find (b) the highest biomarker threshold where the highest difference between treatment results. RESULTS: For NSCLC we observed that EGF concentration thresholds range from 870 pg/ml to 2000 pg/ml were significant. At the lowest threshold (870 pg/ml) the immunotherapy group showed a 6-month difference for the survival endpoint (p = 0.014). A further 10% increase in the threshold (up to 1750 pg/ml) the immunotherapy group showed a 10-month difference for the median survival (p = 0.004). CONCLUSIONS: The evaluation of the p-values of the effect of treatment for each w*, allows the selection of the thresholds where the treatment result is significant. Whereas the