January 2015, compared with an estimated 20-30 patients diagnosed-per-year (estimated total uptake of 120-180 from recommendation). However, in some instances the number of patients treated with an HST were very low: zero patients for Strimvelis® (HST7), ≤10 for eliglustat (HST5), and ≤10 for metreleptin (HST14); much lower than any reported estimates in their respective HST appraisals (Strimvelis®: 3 diagnosed/year; eliglustat: 50-100 eligible patients; metreleptin: <200 eligible patients). Relatively few therapies have been recommended under the NICE HST pathway since 2015 compared to the standard NICE process with STA/MTA. Further, these recommendations have been over 19-months from EC-approval. Even following a positive NICE recommendation, several of these have not achieved meaningful patient access, potentially due to the lack of awareness for rare diseases, and likelihood of underdiagnoses.

P39
BETTER OUTCOMES AND VALUE FOR MONEY WITH COST-EFFECTIVENESS MODELLING OF CASCADE SCREENING STRATEGIES FOR FAMILIAL HYPERCHOLESTEROLAEMIA

Faria R,[1] Cox E,[1] Saramago P,[2] Haralambos K,[3] Watson M,[4] Humphries SE,[5] Qureshi N,[6] Woods B[7]
[1]*University of York, York, UK,* [2]*University of York, Heslington, York, UK,* [3]*University Hospital of Wales, Cardiff, UK,* [4]*University Hospital Southampton NHS Foundation Trust, Southampton, UK,* [5]*University College London, London, UK,* [6]*University of Nottingham, Nottingham, UK,* [7]*University of York, York, YOR, UK*

Objectives: Cascade screening for familial hypercholesterolaemia (FH) refers to the systematic testing of relatives of people known to have FH (termed 'indexes'). Cascade screening is cost-effective compared to no screening, but alternative screening strategies have not been studied. Our objective was to identify the cost-effective strategy to select indexes to cascade, and to contacting and testing their relatives. ***Methods:*** We developed a new cost-effectiveness model informed with routinely collected UK data from indexes and their relatives. Our decision tree model takes the UK National Health Service perspective and calculates, per index assessed to the cascade, the relatives diagnosed, cascade costs, quality-adjusted life years (QALYs), healthcare costs and incremental cost-effectiveness ratios (ICERs). ***Results:*** We compared 1792 strategies. The cost-effectiveness frontier was mostly formed by strategies which contacted 1st and 2nd degree relatives of indexes with genetic FH simultaneously and directly. The cost-effective strategy diagnoses relatives according to whether they were on lipid lowering treatment, cholesterol, and age, with some having confirmatory genetic testing – it diagnoses 52% affected relatives, at a cascade cost of £536; ICER = £13,996/QALY. Sequential contact (i.e. contacting second degree relatives only when their first degree relative was diagnosed with FH), indirect contact via their index/relatives and genetically testing them diagnoses 36% of relatives, while direct and simultaneous contact with genetic testing diagnoses more relatives (56%); neither are in the cost-effectiveness frontier. If genetic testing is not available, cascade screening remains cost-effective, diagnosing 41% of relatives (ICER=£5,603 vs no cascade). Results are robust to alternative scenarios bar those affecting long-term benefits of diagnosis. ***Conclusions:*** Simultaneous and direct contact of relatives of indexes with genetic FH and a mixed approach to testing relatives is cost-effective and achieves better outcomes than sequential and indirect contact. Identifying this strategy required systematic comparison of multiple alternatives, which is only achievable with cost-effectiveness modelling.

P40
HTA AGENCIES PERSPECTIVE ON SURVIVAL MODELLING IN CELL OR GENE THERAPY APPRAISALS

Lissdaniels J,[1] Medin E[2]
[1]*Parexel International, Vällingby, Sweden,* [2]*Parexel International, Stockholm, Sweden*

Objectives: The first cell- and gene therapies, so called advanced therapy medicinal products (ATMPs), have recently become available, making it possible to treat, and even potentially cure, very severe and sometimes previously untreatable conditions. These characteristics have led to a discussion among health economists about whether a specific methodological reference case is required for economic evaluation of gene therapies and the conclusion has been that a new methodological reference case is not required but that "the confluence of various characteristics can lead to specific methodological challenges...". Traditional survival modeling may underestimate outcomes by assuming the same mortality rate for all patients, in situations where the treatment could lead to the cure of a patients. Mixture cure models have been suggested as a supplementary analysis, alongside standard parametric models. The aim of this study was to review reimbursement appraisal reports of the 12 EMA approved ATMPs, to identify the differences in methods and assumptions in survival modelling of long-term treatment effects across different HTA agencies. ***Methods:*** Publicly available assessment reports were retrieved from each reimbursement agency's website in the Nordics, the Netherlands, England and Wales, Canada and Australia for the relevant drugs. ***Results:*** Across the HTA agencies different level of acceptance to non-standard survival modelling is seen. E.g. in appraisals of cell-therapies, mixture cure models have been accepted in the Nordic countries and in England and Wales but assumptions around the percentage of cured patients and the preferred source for the survival extrapolation post clinical trial follow-up differs. In appraisals of gene-therapies, the exploration of the impact of the main assumptions that drive model results have been recommended across the HTA agencies. ***Conclusions:*** There is yet not an established golden standard on how to apply survival modelling to ATPMs and the preferences on the methodology varies across HTA agencies.

Methodological Developments in Network Meta-Analysis and Comparative Effectiveness Research

P41
CROSNMA: A NEW R PACKAGE TO SYNTHESIZE CROSS-DESIGN EVIDENCE AND CROSS-FORMAT DATA

Hamza T, Salanti G
University of Bern, Bern, BE, Switzerland

Objectives: In network meta-analysis, we synthesize all relevant available evidence about health outcomes from competing treatments. That evidence might come from different study designs and in different formats: from non-randomized studies (NRS) or randomized controlled trials (RCT) as individual participant data (IPD) or as aggregate data (AD). To utilize all available evidence, we need a software that allows us to combine these different pieces of information accounting for their differences, e.g. RCTs have typically lower risk of bias than NRS. ***Methods:*** We integrate the three-level hierarchical model that combine IPD and AD with the following four models that incorporate both RCT and NRS evidence by (a) ignoring their differences in risk of bias (b) using NRS to construct discounted treatment effect priors (c,d) adjusting for the risk of bias in each study and controlling the contribution of high risk of bias information in two different ways. ***Results:*** We have implemented these models in a new R package, crosnma. This software allows us for conducting Bayesian network meta-analysis and meta-regression. Up to three study- or patient-level covariates can be also included, which may help explaining some of the heterogeneity and inconsistency across trials. The package runs a range of models with JAGS by generating the code automatically from user's input. ***Conclusions:*** crosnma is a new R package to conduct Bayesian network meta-analysis and meta-regression to synthesise cross-design evidence and cross-format data. We believe that this package will encourage the investigators to not discard any relevant evidence on their analysis. Authors are supported by the HTx-project. The HTx project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 825162

P42
COMPARISON OF ESTIMATION METHODS FOR SINGLE-ARM TRIALS IN RARE DISEASES WITH HISTORICAL CONTROL GROUPS

Barlev A,[1] Brookhart MA,[2] Xun P,[3] Thirumalai D,[3] Sadetsky N,[1] Suissa S[4]
[1]*Atara Biotherapeutics, South San Francisco, CA, USA,* [2]*Duke University, Durham, NC, USA,* [3]*Atara Biotherapeutics, Thousand Oaks, CA, USA,* [4]*McGill University and Jewish General Hospital, Montreal, QC, Canada*

Objectives: Randomized controlled trials are the gold standard for estimating treatment effects. However, in rare diseases with high unmet need, single-arm trials are used when randomization of patients to placebo or standard of care is infeasible or unethical. We evaluated various methods to control for confounding in estimating treatment effects in a small single-arm trial with a historical comparator group. ***Methods:*** We used simulation to evaluate different techniques to estimate the "true" treatment effect on overall survival (OS) and objective response rate (ORR) in a specific target population using an external comparator design. We varied effect size, sample size, confounders, and correlation between confounders. We assessed two broad categories of methods: i) requiring specification for treatment allocation [propensity score (PS)-based inverse probability of treatment weighting (IPTW), standardized mortality/morbidity ratio (SMR), stabilized IPTW (SIPTW), overlap weighting (OW), stratification, and matching], and ii) adding outcome information (g-computation). Their precisions and accuracies were evaluated by a combination of 95% confidence interval (CI) coverage, power, bias, and mean square error (MSE). ***Results:*** G-computation resulted in the most accurate and precise estimator of OS (95% CI coverage: 93.5%, power: 69.3%, bias: -0.001, MSE: 0.055) in a small sample size scenario of 30 treated subjects compared with 120 comparator subjects. Similar results were observed for ORR. In comparison, results for OS were: 95% CI coverage: 72.8%, 65.6%; power: 75.9%, 62.6%; bias: -0.026, 0.072; MSE: 0.114, 0.167 for SMR and IPTW, respectively. ***Conclusions:*** In our simulated example, the g-computation estimator performed best to control confounding in a small single-arm trial with an external comparator group. PS based methods (e.g., SMR & IPTW) may be suitable as an initial step in the creation of the comparator arm when researchers are blind to the outcome, while g-computation can be subsequently used to estimate the efficacy of treatment.

P43
A CASE STUDY AND SIMULATION TO COMPARE DIFFERENT INDIRECT TREATMENT COMPARISON METHODS UNDER VARYING ACCESS TO INDIVIDUAL PATIENT DATA

Ayers D,[1] Cope S,[1] Phillippo DM,[2] Jansen JP,[3] **Park J,**[1] Yuan Y[4]
[1]*PRECISIONheor, Vancouver, BC, Canada,* [2]*University of Bristol, Bristol, BST, UK,* [3]*PRECISIONheor, San Anselmo, CA, USA,* [4]*Bristol-Myers Squibb, Princeton, NJ, USA*

Objectives: To compare alternative methods for unanchored indirect comparisons of two interventions for overall survival when individual patient data (IPD) is available for both studies (IPD-IPD analyses), and when IPD is only available for one study and aggregate data (AD) for the other study (IPD-AD analyses). ***Methods:*** A case study comparing nivolumab with standard of care (SoC) in 3L small cell lung cancer was performed using the CheckMate032 trial and the Flatiron Health database to compare unanchored comparisons without population adjustment, IPD-IPD adjustment analyses using inverse probability treatment weighting (IPTW), regression adjustment (RA), and doubly robust methods ($DRM_{IPD-IPD}$), and IPD-AD analyses using matching-adjusted indirect comparison (MAIC), simulated treatment comparison (STC), and a novel DRM_{IPD-AD}. Relative treatment effects were expressed with Cox hazard ratios (HRs) and differences in restricted mean survival time (DRMST). A simulation study was performed evaluating the performance of the alternative methods in different scenarios. ***Results:*** Conditional HRs from RA differed from the marginal HRs from IPTW due to non-collapsibility. Therefore, estimates of DRMST in months were used as the relative treatment effect measure for the case study, which provides a collapsible estimand for DR estimates and comparison between methods DRMST with nivolumab versus SoC was higher for RA (3.22) than IPTW (2.72) and $DRM_{IPD-IPD}$ (2.71). MAIC, STC, and DRM_{IPD-AD} estimates were similar (2.88; 2.92; 2.88) but with more uncertainty. The simulations found that when covariates differed substantially between studies, IPTW performed poorly in comparison to RA and DRM. ***Conclusions:*** Analysts should be aware of differences between marginal and conditional treatment effects when performing indirect comparisons of time-to-event outcomes; DRMST may be considered. Regression-based or DRM models may be preferable over other adjustment methods in cases with limited overlap in covariates between studies indirectly compared. Having IPD for both studies is preferable over having IPD for only one.

P44
NOVEL AND EXISTING FLEXIBLE SURVIVAL METHODS FOR NETWORK META-ANALYSES

Heeg B,[1] Garcia A,[2] van Beekhuizen S,[3] Verhoek A,[1] Roychoudhury S,[4] Cappelleri J,[5] Postma M,[6] Ouwens M[7]
[1]*Ingress-health, Rotterdam, ZH, Netherlands,* [2]*Ingress-Health, rotterdam, Netherlands,* [3]*Ingress-Health, Rotterdam, Netherlands,* [4]*Pfizer, Inc., Gladstone, NY, USA,* [5]*Pfizer Inc, Newington, CT, USA,* [6]*University of Groningen, Groningen, Netherlands,* [7]*AstraZeneca, Gothenburg, Sweden*

Objectives: The NICE Decision Support Unit has recently published technical support document 21, focusing on flexible survival methods and considering excess mortality. In this study, we research these flexible survival methods by assessing the effect of treatment on the different parameters of the distributions and the implementation of these flexible excess survival models in network meta-analyses (NMA). ***Methods:*** Standard parametric, mixture- (MCM) and non-mixture-cure (nMCM), piecewise, and splines are used for comparison. For the base case model, selection treatment coefficients were assigned to all parameters. The lowest leave-one-out information criterion (LOOIC) defined the base case parametric distribution per flexible model. The best-fitting distribution per model was rerun with treatment coefficients only for those larger than their standard deviation, and the one with the lowest LOOIC was considered the base case. Models were compared based on incremental mean survival and uncertainty. We used a network of four trials in previously treated advanced non-small cell lung cancer, comparing nivolumab vs. docetaxel, pembrolizumab vs. docetaxel, and two trials comparing atezolizumab vs. docetaxel. ***Results:*** The base case parametric distributions were log-logistic for standard parametric distribution, log-logistic for MCM, lognormal for nMCM, Weibull for mixture, log-logistic for piecewise, and log-logistic for spline models, with corresponding LOOICs 13418, 13417, 13420, 13417, 13418, and 13419. Regarding treatment effect parametrization, we applied no treatment effect on cure, for instance, but treatment effects on the scale for MCM and nMCMs. For piecewise, treatment effects were only applied on the second piece. After reducing the number of treatment effects, the mixture Weibull had the lowest LOOIC with 13407. The mean incremental survival with docetaxel as reference for mixture Weibull was 3.58 [95%CI 1.2,5.26], 0.45[0.19,0.93] and 2.67[0.2,9.48] for nivolumab, pembrolizumab and atezolizumab, respectively. For standard log-logistic these were 1.67 [0.94,2.55],1.01[0.55,1.61], and 0.43[0.17,0.77]. ***Conclusions:*** Treatment effect specification is important as outcomes and uncertainty differ over the tested models.

Methodological Developments in Survival Analytic Methods: to Inform Cost-Effectiveness Models

P45
THE USE OF HISTORICAL CLINICAL TRIAL DATA TO INFORM SURVIVAL EXTRAPOLATION

Pham HA,[1] Smalbrugge D,[2] Kroi F,[2] Heeg B,[1] Ouwens M[3]
[1]*Ingress-health, Rotterdam, ZH, Netherlands,* [2]*Ingress-health, Rotterdam, Netherlands,* [3]*AstraZeneca, Mölndal, O, Sweden*

Objectives: Standard parametric distributions are commonly used for the extrapolation of survival data in cost-effectiveness analyses. However, survival data is often immature and uncertainty remains around the survival extrapolations. Mature historical data can be used to better predict survival beyond trial data. This study assessed two methods to incorporate historical data in the extrapolation of immature survival data. ***Methods:*** Immature data of a breast cancer trial comparing pertuzumab+trastuzumab+docetaxel versus trastuzumab+docetaxel (follow-up time 38 months; data-cut 2015) was extrapolated and mature survival data (follow-up time 120 months; data-cut 2020) from the same trial was used to validate the extrapolations. The historical data was from a previous breast cancer trial including mature survival data of trastuzumab+docetaxel (follow-up time 50 months; data-cut 2005). Two methods to quantitatively inform the extrapolation of immature survival data with historical data were compared to standard parametric distributions: 1) historical shape parameter as informative prior for the shape of the immature data; 2) historical data as a third arm. Predictions were assessed with delta area under the curve (AUC) values based on the mature survival data. ***Results:*** Without priors, the delta AUC was 7.59, 1.62, 13.15, 8.32, 25.15, and with the historical arm the delta AUCs were 9.65, 4.38, 6.79, 8.26, 21.81, for Weibull, loglogistic, lognormal, exponential, and Gompertz, respectively. With priors, the delta AUC were 8.43, 3.37, 9.11, 23.68, for Weibull, loglogistic, lognormal, and Gompertz, respectively (as for exponential there is no shape parameter). The loglogistic distribution without priors predicted the immature data the best. For three out of five distributions, the extrapolations with a historical arm resulted in better predictions compared to the extrapolations without prior. ***Conclusions:*** The impact of external data on clinically plausible survival extrapolations can further be improved by using historical data with longer follow-up with treatment patterns similar to the current standard of care.

P46
EFFECTIVE USE OF RECONSTRUCTED SURVIVAL AND COMPARATIVE EFFECTIVENESS DATA: A CASE STUDY FROM ESTIMATING UNREPORTED SUBGROUP SURVIVAL IN ADVANCED STAGE GASTROINTESTINAL CANCERS

Alagoz O,[1] Xiao H,[2] Singh P,[3] Gricar J,[3] Dixon M,[3] Kim I,[3] Kurt M[2]
[1]*University of Wisconsin-Madison, Madison, WI, USA,* [2]*Bristol Myers Squibb, Princeton, NJ, USA,* [3]*Bristol Myers Squibb, Lawrenceville, NJ, USA*

Objectives: This study devises a systematic approach that can utilize aggregate level survival and comparative effectiveness data published from randomized controlled trials (RCT) to assist subgroup-specific health economic and meta-analyses. ***Methods:*** We developed a soft-constrained optimization model, which approximates the restricted mean survival time (RMST) for the overall population in each arm via weighted sum of the RMSTs of two subgroups of interest. Survivals of both subgroups in each arm were assumed to follow Weibull or log-logistic distribution. The constraint ensured that cumulative hazards between the arms were proportional for each subgroup at a sufficiently long pre-specified time point. Estimated subgroup-specific survival functions for the control arm were direct outputs of the model and were shifted by applying the reported hazard ratios from the forest plots to generate their counterparts for the intervention arm assuming proportional hazards between the arms. For validation, we tested our approach in a case study consisting of 10 distinct RCTs with reported subgroup-specific Kaplan-Meier (KM) curves from advanced stage gastrointestinal tumors. ***Results:*** Across all 48 subgroups, on average, loglogistic model performed equally or better than Weibull model in performance criteria comparing overall survival (OS) rates, median OS and RMSTs. Predicted survival curves laid within the 95% confidence intervals (CIs) of reported KM-curves in 75% and 81% of the time for Weibull and loglogistic models, respectively. Predicted median survivals were within the 95% CIs of the reported medians in 34 and 40 subgroups for Weibull and loglogistic models, respectively. Average relative gap between the predicted and reported RMSTs was 10% in both models. Predicted RMSTs were within the 95% CI of reported RMSTs in 34 and 37 subgroups for Weibull and loglogistic models, respectively. ***Conclusions:*** Our elicitation approach is effective and demonstrably reliable in deriving unreported subgroup survival with flexible time-varying hazard functions.

P47
A BAYESIAN HIERARCHICAL MIXTURE CURE MODELLING (MCM) FRAMEWORK FOR THE JOINT UTILIZATION OF PROGRESSION FREE SURVIVAL (PFS) AND OVERALL SURVIVAL (OS) IN ESTIMATING LONG-TERM SURVIVORSHIP RATES IN PREVIOUSLY UNTREATED METASTATIC MELANOMA: A CASE STUDY FROM CHECKMATE-067 TRIAL

Green N,[1] Paly V,[2] Youn JH,[3] Kurt M,[4] Moshyk A,[4] Gianluca Baio G[1]
[1]*University College London, London, LON, UK,* [2]*ICON plc, New York, NY, USA,* [3]*ICON plc, Marlow, Bucks, UK,* [4]*Bristol Myers Squibb, Princeton, NJ, USA*

Objectives: Differences in emergent survival plateaus between PFS and OS may imply clinically unintuitive dichotomy between the resulting proportions of long-term survivors (LTS) when they are analyzed separately via mixture cure models (MCM). We present a novel Bayesian hierarchical (BH) MCM framework assuming a dependency between PFS and OS to estimate LTS rates in CheckMate 067 and demonstrate its practical utility over frequentist MCMs in long-term QALY estimations. ***Methods:*** Frequentist and BH MCMs were fitted to PFS and OS data from the trial with minimum 60-months of follow-up. In the frequentist MCMs, PFS and OS were modelled separately whereas in BH MCMs both endpoints were modelled jointly with a shared LTS rate. In both approaches, background mortality rates were taken from World Health Organisation's age, gender and country-specific lifetables and time-to-event outcomes for the non-LTS were modeled using a range of standard parametric distributions. Estimated incremental QALYs gains for nivolumab containing therapies versus ipilimumab under both approaches were compared using