**HTA ESIG Parallel Session**

**A cacophony of comparisons? Multiplicity in future EU HTA Joint Clinical Assessments**

*Session Chair: Lara Wolfson, MSD*

The questions asked in Health Technology Assessment (HTA) are often different from those asked by regulators. In HTA, one is interested in the added clinical and economic benefit of a treatment within the context specific health care system. This can lead to requests for a sizeable number of post hoc comparative analyses in order to address endpoints, subgroups or subpopulations of particular interest in a local context. As the number of analyses increases, so does statisticians' concern of potential misinterpretation - because with enough looks at data, something will eventually stick out simply due to chance.

When pan European HTA (EU HTA) and the Joint Clinical Assessment (JCA) becomes a reality in 2025, the sheer scale of this multiplicity concern may reach new heights, as data statistical analyses must cover the needs of each of the 27 EU member states simultaneously. How should statisticians tackle this potential cacophony of comparisons in a way so that assessors and member states can navigate it in a scientifically sound way? And what about the high level of transparency of the JCA that will make multiplicity challenges very visible to the general public?

This session will discuss the potential extent of the multiplicity concern, alongside pros and cons of formal multiplicity adjustment in the context of EU HTA JCA.

**Presentation 1**

**One Size Fits All – The EU’s Joint Clinical Assessment System and its Implications for Statisticians**

*Martin Scott, Numerus*

In December 2021, the European Commission adopted the proposal for a new regulation on health technology assessment, which foresees the cooperation of EU member states on joint clinical assessments (JCA) of health technologies, including pharmaceuticals. One of its goals is to streamline the current process of having to submit evidence to each of the 27 member states. The regulation allows manufacturers to submit evidence as part of a joint clinical assessment just once at the EU-level.

Since January 2022, efforts have been made by HTA authorities of 12 EU member states to agree on guidance relating to the levels of evidence, types of endpoints and statistical methodology that should be considered for the JCA. As of January 2025 - only 2 years away - all cancer medicines and advanced therapy medicinal products (ATMPs) will fall under the new regulation. For all those working in the field of HTA this constitutes a significant logistical and methodological challenge.

This presentation will discuss the impact the regulation will have not just on statisticians employed within the HTA field but also those currently working on phase II-III clinical trials. It will highlight the areas of clinical trial analysis and design in which both regulatory and HTA statisticians will be required to work more closely. Lastly, it will demonstrate that the future success of bringing effective medicines to market lies squarely in the hands of the statisticians.

**Presentation 2**

**PICOs in EU HTA: How many, how varied?**

*James Ryan, AZ*

The PICO is one of the most critical components of the EU HTA, determining the scope, the evidence required and the types of analyses to be performed. Based on the process outlined in EUnetHTA 21’s PICO proposal, sets of PICOs were created for two common indications in NSCLC and multiple myeloma. The research showed that without change to the proposal, or increased transparency on the consolidation of additive Member State needs, we should expect between five and twenty PICOs per assessment, and potentially thousands of requested analyses. Of these, safety analyses could be a significant proportion, and indirect comparisons will be a necessary and dominant component of the assessment. The magnitude of such analyses raises concerns on multiplicity and misinterpretation of results, particularly considering that there is no proposed Member State guidance on selecting clinically appropriate populations and sub-group requests, and EUnetHTA 21’s proposal that clinical context and evidence availability have limited relevance in the formation of the PICO and assessment itself.

**Presentation 3**

**Can more really be less?**

*Stephen Senn*

It seems natural to many statisticians trained in frequentist statistics that you should adjust if you carry out many analyses in the same study and in drug regulation, elaborate efforts are made to control the type I error rate. But statisticians also know that rate figures require a clarifying ‘per’ statement. The usual choice is per study but why? The adjustment is not strict enough if we wish to control errors per development programme and it is too strict if we just need to control it per question and if we have a per test standard it is not needed at all.

Bayesians sometime claim that they can get way without making any adjustments. However, one defence of not making adjustments is that Bayesian results are in any case shrunk and the shrinkage deals with multiplicity. However, many Bayesian analyses these days seem to involve uninformative prior distributions and these do not shrink results.

I shall attempt to explain where how and why I think adjustment is necessary and where it is not.

**Presentation 4**

**Balanced Multiplicity Approach in Safety Evaluation**

*William Wang, MSD*

In the context of safety evaluation of pharmaceutical products, point estimates, their confidence intervals and sometime inferential statistics (aka p-values) are used to quantify the safety risk or risk elevation (ICH E9 Section 6.4). When these are applied across the multi-dimensional MeDRA dictionary with 27 body systems and thousands of preferred terms, the evaluation should take into account of the multiplicity issues arising from the numerous comparisons (ICH E9 Section 7.2.2).

In this presentation, we first examine the role of statistical uncertainty or play of chance in safety evaluation, we discuss a balanced approach that controls false discovery rates while maintaining the power of detecting potential safety events/terms of interest that require clinical evaluation Multi-disciplinary collaboration that integrates medical judgement and statistical quantification will be advocated.