Paper AS06

Statistical Training for Health Technology Assessment Statistical Programmers

Anastasia Kokoreva, MSD, Belgium
Dave Gelb, MSD, Switzerland
William Malbecq, MSD, Belgium
Kristel Vandormael, MSD, Belgium

ABSTRACT

Statistical programmers (SPs) are key contributors to the preparation of the statistical results that support Health Technology Assessments (HTA) submissions. In order to collaborate optimally with statisticians, programmers should have a good understanding of the statistical methods and approaches used in HTA submissions. These methods represent a mix of standard techniques taught in applied statistical courses and more innovative/recent techniques. At MSD in the Biostatistics and Research Decision Sciences (BARDS) Statistical Programming HTA team, statistical programmers have diverse educational backgrounds and work experience and consequently different levels of statistical knowledge.

The Statistical Training for SP in HTA initiative has been ongoing since 2020. Its main objective is to increase the level of statistical knowledge within the statistical programming team and bring it to a common level that optimizes analysis and reporting activities, both in efficiency and quality. It also increases the motivation of the statistical programmers.

During the COVID-19 pandemic a series of virtual classroom trainings have been organized. Statistical concepts are illustrated using examples from HTA dossiers and accompanied by practical exercises. Trainings on traditional statistical methodologies are given by HTA SP team members. For more HTA specific and innovative methodologies the sessions are delivered by colleagues from the BARDS HTA Statistics group.

INTRODUCTION

Health Technology Assessment (HTA) is a formal, systematic research process designed to synthesize and evaluate the existing evidence for a medical treatment or health delivery innovation (Guide to Understanding Health Technology Assessment 2018). For sponsors, the importance of HTA is constantly increasing as more countries request the submission of such a dossier and, in a parallel, the complexity and volume of this dossier increases. A typical HTA dossier may include statistical analyses varying from simple summaries of clinical trial results to implementation of more complex statistical methods.

Statistical Programmers are responsible for delivering solutions to analysis and reporting needs specific to HTA statistics reports. The HTA SP team is formed by professional statistical programmers. Understanding of scientific methodologies and businesses, collaborative work on specifications and ability to produce high quality deliverables are essential SP skills (Coppin et al. 2020). For SP, it is important to build a good statistics culture to foster optimal collaboration with statisticians. In addition, improving the HTA specific statistical understanding adds to the motivation and job satisfaction of statistical programmers. To address those needs the Statistical Training for Statistical Programmers in HTA initiative has been launched. The main purpose of the training is to improve the understanding of general statistical concepts and HTA specific methodologies by HTA SP team members.

MSD Biostatistics and Research Decision Sciences (BARDS) SP HTA is a fast-growing team. The number of employees grew substantially in recent years to address the growing HTA business needs. Programmers joining the team have diverse educational and working experiences, with different levels of statistical understanding. Typical HTA analyses use a very broad range of statistical methods from standard approaches to more specialized and innovative techniques. The training is not aimed to replace academic statistical courses but to present statistical approaches specific to HTA submissions, focusing on the concepts and their practical implementation. By organizing a series of different level statistical concepts trainings, we aim to provide a common statistical practical knowledge to SP's.

This paper will describe the content of the Statistical Training for Statistical Programmers in HTA initiative and its organization. At the end, we discuss the influence of the trainings on the collaboration between HTA Statisticians and Statistical Programmers and plans for future trainings development.

COLLABORATION BETWEEN STATISTICS AND STATISTICAL PROGRAMING IN HTA

The HTA SP provides analysis and reporting support to HTA projects in partnership with the BARDS HTA Statisticians, and global, regional, and local staff working in outcomes research, health economics, and market access (Figure 1); with the goal to provide fair accessibility for patients to our company's medicines and vaccines, making a real difference in patient's lives. To support MSD strategic goals, we provide comprehensive analytical and programming expertise. Our team offices are located in London, Brussels and Zürich.

Currently we observe an increase in the number of requests for HTA dossiers as well as a greater complexity of the statistical methodology in these dossiers. Typically, HTA statistical reports contains hundreds of tables and figures from clinical trial data. Each report is the fruit of cross functional collaborations. Our team uses clinical data to document the efficacy and safety profile of our compounds. We create HTA dossiers using scientific concepts, statistical analysis and high-quality data. The goal of the training consists in giving all statistical programmers a good understanding of HTA-specific statistical methods. This will improve the collaboration between statisticians and statistical programmers and optimize the analysis and reporting activities for multiple reasons: (1) easier collaboration and understanding of the specifications provided by the statisticians to programmers, (2) easier initial quality control of the output of the deliverables by programmers and (3) job satisfaction of both functions in their respective roles.

Figure 1 HTA SP customers

TRAINING PREPARATION

The first step in preparing a training program in statistics for statistical programmers is to identify the list of statistical concepts regularly used in the HTA Analysis and Reporting (A&R) process. Several HTA projects have been evaluated to identify the statistical methodology used. Also, the team members can propose statistical subjects which they find relevant to the HTA process and where they feel a lack of knowledge.

The identified topics are organized in seven groups and divided in three modules of complexity. The list of topics is displayed in Table 1. Beginner modules contains topics covered in a basic statistics university course. Intermediate modules contain hypothesis testing and regression methods covered in more advanced statistical courses and survival analysis which is typically part of Master's program for biostatisticians. In the advanced module we grouped Indirect Treatment Comparison (ITC) relative effectiveness analysis and Rank-Preserving Structural Failure Time (RPSFT) models. These methods are HTA specific and usually not covered within standard university programs.

Most SP's in the BARDS HTA department are educated in Computer Science, Statistics, Applied Mathematics, Life Sciences, Engineering or related analytical fields and have work experience in the pharmaceutical industry. The level of statistical knowledge within the team varies from basic statistics university courses to Master's degrees in statistics. It is clear that the need for certain training courses is not equal. We launched a questionnaire across the team for self-evaluation of their current statistical knowledge. The survey results are displayed in Table 1. The need in basic module trainings is minor, junior programmers are interested in refreshing basic statistical concepts in relation with HTA analysis. Also, survival analysis training was identified as less of a need probably because this type of analysis is widely used in late stage development analysis and in HTA. Therefore, SP's gained required knowledge from everyday work. Most of the team were interested in advanced module HTA specific trainings.

Modules	Title	Trainer	Unmet need identified by SP staff
Beginner	Descriptive Statistics	HTA SP	20%
	Probability Distributions	HTA SP	20%
Intermediate	Hypothesis Testing	HTA SP	40%
	Regression Methods	HTA SP	60%
	Survival Analysis	HTA Statistician	20%
Advanced	Rank-Preserving Structural Failure Time model	HTA Statistician	60%-100%
	Matching Adjusted Indirect Comparisons	HTA Statistician	60%-100%

Table 1 Statistical concepts for SP's training plan

TRAINING IMPLEMENTATION

The Statistical Training for SP's in HTA initiative is a collaborative project between HTA statistics and HTA SP teams (Table 1). The Beginner and Intermediate module trainings are created and given by the HTA SP team with support from HTA Statisticians. These training sessions have to be delivered sequentially. HTA-specific Advanced module trainings based on innovative methodologies are created by HTA biostatisticians who also act as trainers. Advanced sessions can be delivered in any order. Examples from real projects are used during the training sessions, including tables and figures, helping to make trainings more useful and serving as an additional motivation factor for the audience.

The first series of trainings are organized in the following way. After each two hour training session, a set of practical exercises is proposed to the attendees. Within a two weeks' time a follow up session is organized. The purpose of the follow up session is to go through exercises and give additional explanations to the problems which were more difficult for the audience. Also, these sessions are dedicated to address the questions from the audience and to deliver additional topics. The whole set of trainings is organized across a nine-month time span and can be offered repeatedly as new staff members onboard.

The following section provides more information about the training content and the HTA specific statistical methods used during the training for illustrative purposes.

TRAINING CONTENT

The training content is divided in three modules according to complexity.

DESCRIPTIVE STATISTICS

Standard HTA dossiers contain hundreds of tables summarizing clinical trial results. Descriptive statistics is broadly used for creation of summary tables. The main purpose of this session is refreshing statistical methods used to calculate descriptive statistics, applicability of these methods and interpretation of results. Examples from the real AMNOG (Arzneimittelmarkt-Neuordnungsgesetz, Pharmaceuticals Market Reorganisation Act) dossier (Nutzenbewertungsverfahren zum Wirkstoff Pembrolizumab 2019) are used as an example to describe the calculation and presentation of descriptive statistics.

PROBABILITY DISTRIBUTION

Typical HTA dossiers use data coming from clinical trials, patient questionnaires and other sources. These data follow various probability distributions. Several probability distributions, their characteristics and corresponding parameters are covered in the second Beginner module training. Data applications are taken from HTA dossiers.

HYPOTHESIS TESTING

"In postmenopausal women with osteoporosis, Romo 210mg QM followed by ALN significantly reduced new vertebral, clinical, nonvertebral, and hip fracture risk vs ALN alone …." (Saag et al. 2017). How was this result obtained? What were the methods and assumptions used? The first session of the intermediate module is devoted to hypothesis testing. Parametrical and non-parametrical tests for discrete and continuous data illustrated by real data case studies are covered in this training session.

REGRESSION METHODS

Simple and multiple linear regression, logistic regression and Poisson regression models and their implementations in the BARDS HTA Standards library are the main subjects of the second Intermediate module training.

SURVIVAL ANALYSIS

Events are very important in a public health decision framework because events matter to patients and typically are associated with high healthcare costs. In HTA evaluations, there is a strong focus on patient-oriented outcomes. That is why time-to event data is important in clinical development and especially in HTA dossiers, in particular in the oncology therapeutic area. The last training session at the Intermediate level is devoted to Survival Analysis methods and their applications.

Time-to-event data cannot be analyzed by using the standard statistical techniques for experimental designs (such as t-test or analysis of variance or logistic regressions) because of the specificity of this type of data. Time-to-event random variables take positive values and follow statistical distributions not approximatively normal (Gaussian – the bell shape curve). Also, time-to-event data are typically incomplete. We cannot wait until every participant has experienced an event before to analyze the study. For example, in an oncology study we cannot wait until every participant has experienced an event including death or congestive heart failure before analyzing the study. There is a censoring process. Ignoring the censoring and applying standard methods for complete data will provide biased estimates and incorrect statistical tests. Survival analysis provides a methodology to analyse time-to event data.

During the training key concepts of Survival Analysis and methods of estimations are discussed. Comparison of treatments and existing regression models are covered and illustrated by applications from HTA dossiers. The session finishes with an overview of the existing SAS procedures for analysis of time-to event data.

RANK PRESERVING STRUCTURAL FAILURE TIME MODEL

The first training at the advanced level is devoted to the Rank Preserving Structural Failure Time (RPSFT) models, which are commonly used for analysis of overall survival data (OS) adjusting for treatment switchover (Robins et al. 1991). The estimation of the treatment effect after adjustment for switching is important in HTA filings: for both economic modeling and to provide a correct, unbiased clinical estimate in the HTA dossier. Two training sessions are organized by a team of biostatisticians and statistical programmers. During the first session the RPSFT methodology is covered. A case study and SAS implementation are subjects of the second session.

In randomized clinical trials an efficacy comparison of an experimental drug vs. a control drug is made. Treatment switching where patients randomized to the control group of the clinical trial are allowed to switch to the study treatment after the progression of the disease (PD), commonly occurs in oncology trials, see Figure 2. There are several different types of switching. The direct switching corresponds to switching from the control onto the experimental treatment or switching from the experimental onto the control. Indirect switching is switching from control/experimental onto other treatments. Also, RPSFT models can be applied for both types of switching. We implement the RPSFT for switching from control arm to experimental arm only. Not all patients who progress in the control arm switch to experimental treatment - treatment switch is a selective process. Also, timing of treatment switch varies across the patients. Since some patients in the control group receive the experimental drug, outcomes in the control arm partially reflect the effect of the experimental drug. This can cause bias in the analysis of endpoints accrued after the treatment switch.

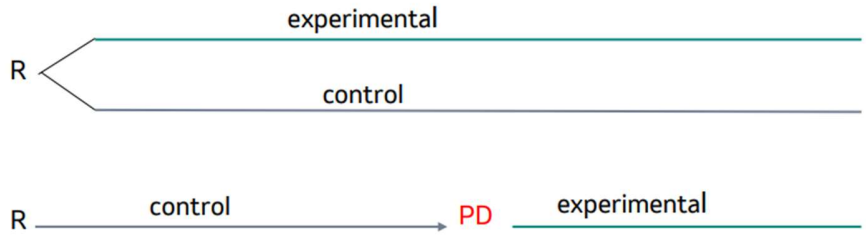

Figure 2 Treatment switching, R - randomization, PD - progression of disease

The goal of the RPSFT method is an estimation of the treatment effect of experimental treatment vs. control treatment if switching would not have occurred. Each patient proceeds through the disease process towards death at their own speed. An experimental therapy slows this speed down by a certain multiplying factor. This factor is assumed to be the same for the patients regardless of the treatment arm they were initially randomized to, and regardless of whether experimental therapy is given from the time of randomization (R) or from the time of treatment switching. It is known as the 'common treatment effect' assumption, which underlies the method. The RPSFT model applies this slow-down factor to patients who switched, to estimate event times that would have been observed had these patients not switched to the experimental treatment ('counter-factuals').

The method is illustrated by an example from the KEYNOTE-024 study (Reck et al. 2019). This study was in participants with previously untreated advanced non–small-cell lung cancer (NSCLC). Patients were randomized to either the Pembrolizumab or Standard of Care (SOC) arm. Participants assigned to the SOC arm were allowed to switch to Pembrolizumab after documented disease progression, potentially leading to a biased estimate of the treatment effect for overall survival. The RPSFT methodology was therefore used to estimate the treatment effect that would have been observed had switching from SOC to pembrolizumab not occurred.

At the end of the training, the %hta0a0rpsft macro is presented. The %hta0a0rpsft macro (developed by Paul Wicks and team in 2020 in collaboration with HTA Statisticians) is a SAS tool to create RPSFT statistics, datasets, tables and plots. It is designed to be used in SAS 9.4. The macro is robust and covers different analysis situations. Several scenarios covered by the macro are presented during the training. The macro flexibly controls the output tables and plot appearance. The macro user manual is presented during the training session. It covers the following topics: macro overview, basic features, macro syntax, statistical method used for calculations, detailed examples, references and frequently asked questions. The macro logic and sub macros used are discussed. Outputs and data sets produced by the macro are demonstrated on the real study data. Interpretation of macro outputs is an essential part of quality control at the development stage of HTA deliverables. The trainee goes through the standard macro outputs and interprets obtained results. At the end of the presentation common mistakes possibly arising during the macro application are covered.

MATCHING ADJUSTED INDIRECT COMPARISONS

The indirect treatment comparison methodology is covered during the last training session organized by BARDS HTA Statisticians.

Results of randomized clinical trials are the standard foundation for evidence-based clinical guidelines and reimbursement decisions. In the presence of multiple approved treatment options, comparison of all relevant treatments may be required for clinical and reimbursement decision making. Randomized clinical trials directly comparing all available treatments head-to-head would in some ways be an ideal solution but such evidence is unlikely to exist. Most often a new treatment is compared to placebo or the standard of care in the clinical trial and rarely with a contemporaneous, competitive treatment, as such treatments may not have had regulatory approval at the time of the trial initiation. Sometimes only data from single-arm studies is available.

In the absence of head-to-head comparisons of relevant treatment regimens, using indirect treatment comparisons (ITCs) is a useful approach to evaluate the comparative effectiveness of two medications (Jansen et al. 2011). For example, Pembrolizumab + chemotherapy and nivolumab + ipilimumab are two approved therapeutic options in metastatic NSCLC. These two therapeutic options have never been compared against each other, but each combination has been compared to a placebo + chemotherapy. In this case an ITC can be used to evaluate relative effectiveness of two treatments, via a common comparator. Relative effectiveness analyses are increasingly becoming standard components of HTA submissions, as they provide decision-makers with evidence-based information required for making sound decisions. NICE (UK) and G-BA (Germany) HTA agencies documented their interest in obtaining results from such analyses in the absence of direct comparisons (Massaad et al. 2018).

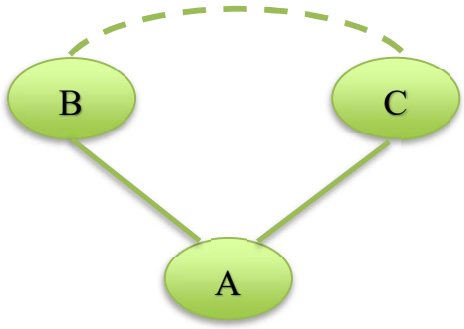

Figure 3 Indirect Treatment Comparing (‘anchored’)

An illustration of an ITC scheme is given in Figure 3. Each node corresponds to a treatment and the solid lines each represent a direct comparison of the two treatments. For each pair of treatments in a connected network, relative effectiveness can be estimated. The main objective of the exercise is to compare two approved treatments B and C (dotted line). No trial comparing treatment B and treatment C is available, but there are available trials comparing B and A and A and C. The indirect treatment comparing of treatment B versus treatment C is called ‘anchored’ (Jansen et al. 2011). A could represent an active treatment comparator or placebo.

If there is no common comparator for the trials B and C or only single arm studies are available, the indirect treatment comparison of treatment B versus treatment C is called 'unanchored', see Figure 4.

Figure 4 Indirect Treatment Comparing ('unanchored')

Within RCTs, results are not biased due to the randomization that determines treatment assignment. Compare the proportion of participants responding to treatment B in trial AB with the proportion of participants responding to treatment C in trial AC is prone to bias because it doesn't take in consideration the effects of treatment A in each of the two trials. Another source of bias is cross-trial differences in study design, conduct, populations and characteristics. Naïve ITCs which do not attempt to adjust for cross-study differences will to some extent be biased.

Matching Adjusted Indirect Comparisons (MAICs) were developed to address these issues (Signorovitch et al. 2012). Application of the MAIC methodology can limit potential bias. In this method, relevant, pre-defined baseline characteristics are selected and an algorithm is used to define unique weights for all participants in the sponsor trial, such that when a new baseline characteristics table is produced, the results match those of the corresponding published table for the comparator study. The method was developed with the knowledge that individual patient data (IPD) is rarely available for both trials. More often a researcher has access for IPD only for one study. For the comparator treatment, only published summary data is usually available. MAICs can combine IPD from the experimental study with published summary results for other treatments of interest. The MAIC approach and its limitations is discussed during the training and illustrated by case studies (Halmos et al. 2020, Halmos et al. 2021).

CONCLUSION

Internal department training in statistical methodology for statistical programmers is not unique in the pharmaceutical industry. According to our state of knowledge, training SPs in new, innovative and HTA specific methodologies covered in advanced module sessions is however novel in the industry. Preparation of presentations and other training materials is time-consuming, but efforts are being rewarded by increased levels of understanding of methods and approaches realized in HTA dossiers. From 30% to 70% of the team depending on the module were trained. SP are responsible for producing high quality deliverables for HTA dossiers. Improved statistical knowledge positively affects collaboration between statisticians and SP and motivated SP for further professional development.

BARDS HTA SP is a fast growing team, with new members onboarding frequently. This year we are planning to organize a second edition of online trainings. Instead of only Rank-Preserving Structural Failure Time model training we will propose a new training covering various methods used for overall survival analysis adjusted for switchover, including following approaches: Simplified 2-stage method, RPSFT method and Inverse probability of censoring weighting method. Also, we are planning to prepare training on patient reported outcome analysis widely used in clinical research and HTA submissions.

REFERENCES

Halmos, B, Burke, T, Kalyvas, C, Insinga, R, Vandormael, K, Frederickson, A, Piperdi, B. Matching-Adjusted Indirect Comparison of Pembrolizumab + Chemotherapy vs. Nivolumab + Ipilimumab as First-Line Therapies in Patients with PD-L1 TPS ≥1% Metastatic NSCLC. *Cancers*, *12*(12), 3648. 2020.

Halmos B, Burke T, Kalyvas C, Vandormael K, Frederickson A, Piperdi B. Pembrolizumab+chemotherapy versus atezolizumab+chemotherapy+/-bevacizumab for the first-line treatment of non-squamous NSCLC: A matching-adjusted indirect comparison. Lung Cancer. 2021 May;155:175-182. doi: 10.1016/j.lungcan.2021.03.020. Epub 2021 Mar 30. PMID: 33839603.

Coppin, F, Kampouris, H, Cao, Y. (2020) Managing an International Statistical Programming Team in the Era of New Ways of Working. Proceeding of PhUSE EU Connect 2020.

Guide to Understanding Health Technology Assessment (HTA). Institute for clinical and economic review. 2018

Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, Lee K, Boersma C, Annemans L, Cappelleri JC. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the

ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. Value Health. 2011 Jun;14(4):417-28. doi: 10.1016/j.jval.2011.04.002. PMID: 21669366.

Massaad R, Malbecq W, Kampouris H, Gelb D. Matching-Adjusted Indirect Comparison (MAIC) in HTA and its application using SAS Procedures. PhUSE Single Day Events. 21st September. 2018.

Nutzenbewertungsverfahren zum Wirkstoff Pembrolizumab (neues Anwendungsgebiet: nicht-kleinzelliges Lungenkarzinom, Nicht-Plattenepithelhistologie, Erstlinie, Kombination mit Pemetrexed und Platin-Chemotherapie). 2019 https://www.g-ba.de/downloads/92-975-3022/2019-03-29_Modul4B_Pembrolizumab.pdf

Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, Gottfried M, Peled N, Tafreshi A, Cuffe S, O'Brien M, Rao S, Hotta K, Vandormael K, Riccio A, Yang J, Pietanza MC, Brahmer JR. Updated Analysis of KEYNOTE-024: Pembrolizumab Versus Platinum-Based Chemotherapy for Advanced Non-Small-Cell Lung Cancer With PD-L1 Tumor Proportion Score of 50% or Greater. J Clin Oncol. 37(7):537-546. Mar 1, 2019

Robins JM, Tsiatis AA: Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Commun Stat Theory Methods. 20:2609-2631, 1991

Saag K, Petersen J, Brandi ML, Karaplis A, Lorentzon M, Thomas T, Maddox J, Fan M, Meisner PD, Grauer A. A Randomized Alendronate-Controlled Trial of Romosozumab: Results of the Phase 3 Active-Controlled Fracture Study in Postmenopausal Women with Osteoporosis at High Risk. Arthritis Rheumatol. 2017; 69

Signorovitch et al, Matching-Adjusted Indirect Comparisons: A New Tool for Timely Comparative Effectiveness Research, Value in Health 15, 940-947. 2012

ACKNOWLEDGMENTS

The authors would like to thank our MSD BARDS HTA colleagues and Frederic Coppin for they input during the preparation of this paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Anastasia Kokoreva
MSD
Clos du Lynx 5, 1200 Brussels [Belgium]
+32 473 834691
anastasia.kokoreva@merck.com
Web: http://www.msd.com

William Malbecq
MSD
Clos du Lynx 5, 1200 Brussels [Belgium]
+32 2 776 6447
william_malbecq@merck.com
Web: http://www.msd.com

Dave Gelb
MSD
The Circle 66
8058 Zürich-Flughafen [Switzerland]
dave.gelb@merck.com
Web: http://www.msd.com

Kristel Vandormael
MSD
Clos du Lynx 5, 1200 Brussels [Belgium]
+32 (2) 7766480
kristel_vandormael@merck.com
Web: http://www.msd.com