EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Doc.Ref. EMA/540136/2009

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

ENCePP Checklist for Study Protocols (Revision 4)

Adopted by the ENCePP Steering Group on 15/10/2018

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) welcomes innovative designs and new methods of research. This Checklist has been developed by ENCePP to stimulate consideration of important principles when designing and writing a pharmacoepidemiological or pharmacovigilance study protocol. The Checklist is intended to promote the quality of such studies, not their uniformity. The user is also referred to the ENCePP Guide on Methodological Standards in Pharmacoepidemiology, which reviews and gives direct electronic access to guidance for research in pharmacoepidemiology and pharmacovigilance.

For each question of the Checklist, the investigator should indicate whether or not it has been addressed in the study protocol. If the answer is "Yes", the section number of the protocol where this issue has been discussed should be specified. It is possible that some questions do not apply to a particular study (for example, in the case of an innovative study design). In this case, the answer 'N/A' (Not Applicable) can be checked and the "Comments" field included for each section should be used to explain why. The "Comments" field can also be used to elaborate on a "No" answer.

This Checklist should be included as an Annex by marketing authorisation holders when submitting the protocol of a non-interventional post-authorisation safety study (PASS) to a regulatory authority (see the Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies). The Checklist is a supporting document and does not replace the format of the protocol for PASS presented in the Guidance and Module VIII of the Good pharmacovigilance practices (GVP).

Study title: Use cases for development, optimisation and implementation of artificial intelligence methods for real world data analyses in regulatory decision-making and health technology assessment along the product lifecycle
Checklist for study protocol from Work Package 1, Use cases 1 & 2

EU PAS Register® number: EUPAS105544
Study reference number (if applicable):

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	16
1.1.2 End of data collection[2]	☐	☐	☒	
1.1.3 Progress report(s)	☒	☐	☐	16
1.1.4 Interim report(s)	☒	☐	☐	16
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	16
1.1.6 Final report of study results.	☒	☐	☐	16

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.
[2] Date from which the analytical dataset is completely available.

Comments:

1.1.2 presumably in 2024

Section 2: Research question		**Yes**	**No**	**N/A**	**Section Number**
2.1	Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	2
	2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	2.1
	2.1.2 The objective(s) of the study?	☒	☐	☐	2.2
	2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	2.3
	2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☒	☐	
	2.1.5 If applicable, that there is no *a priori* hypothesis?	☒	☐	☐	2.4

Comments:

Section 3: Study design		**Yes**	**No**	**N/A**	**Section Number**
3.1	Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	3
3.2	Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	3.1
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	3.2
3.4	Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☐	☒	☐	3.3
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☐	☒	☐	3.4

Comments:

3.4 and 3.5 are no study objectives

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	4.1
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	4.2, Tbl. A
	4.2.2 Age and sex	☒	☐	☐	4.3,T.C&D
	4.2.3 Country of origin	☒	☐	☐	4.1, T.A&B
	4.2.4 Disease/indication	☒	☐	☐	4.3, T.B
	4.2.5 Duration of follow-up	☒	☐	☐	4.2, TA&B

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	4.3

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☐	☐	☒	
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☐	☐	☒	
5.3 Is exposure categorised according to time windows?	☐	☐	☒	
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	☐	☐	☒	
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6 Is (are) (an) appropriate comparator(s) identified?	☐	☐	☒	

Comments:

5. This study describes study populations and the use of historical control arms and synthetic data. It does not have an exposure.

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	6, Table G
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	6, Table G
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☒	☐	☐	6.3
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☒	☐	6.4

Comments:

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☒	☐	☐	7
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☒	☐	☐	7
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	7

Comments:

Section 8: Effect measure modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☐	☒	☐	8

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	9.1, Tbl. A
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9.1, Tbl. A
9.1.3 Covariates and other characteristics?	☒	☐	☐	9.1, Tbl. A
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☐	☐	☒	
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	Table G
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	Table E
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☐	☐	☒	
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	9.2
9.3.3 Covariates and other characteristics?	☒	☐	☐	9.2
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	9.3

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	10
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	10.2
10.3 Are descriptive analyses included?	☒	☐	☐	
10.4 Are stratified analyses included?	☐	☐	☒	
10.5 Does the plan describe methods for analytic control of confounding?	☐	☐	☒	10.3
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☐	☒	10.3
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	2.2, 10.1
10.8 Are relevant sensitivity analyses described?	☐	☐	☒	

Comments:

10.4. all analyses are stratified by country; further stratifications might be applied in an explorative manner on a country-specific level (Section 14)

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	11
11.2 Are methods of quality assurance described?	☒	☐	☐	11.4
11.3 Is there a system in place for independent review of study results?	☒	☐	☐	11.2

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☐	☐	☒	
12.1.2 Information bias?	☐	☐	☒	
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☐	☐	☒	
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☐	☐	☒	

Comments:

12.1. Since there are no comparative analyses in WP1, most biases are not applicable; however, selection bias and information bias may become applicable depending on the quality of algorithms.

12.2. This study investigates the value of real world data in population description and the application of historical control arms and synthetic data. Thus, one of the results of the study is a description of the feasibility of the study.

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	13
13.2 Has any outcome of an ethical review procedure been addressed?	☒	☐	☐	13
13.3 Have data protection requirements been described?	☒	☐	☐	13

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	14

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	15
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	15

Comments:

Name of the main author of the protocol: Elvira Bräuner, Vera Ehrenstein

Date: 11/April/2024

Signature: Britta Häußler

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Doc.Ref. EMA/540136/2009

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

ENCePP Checklist for Study Protocols (Revision 4)

Adopted by the ENCePP Steering Group on 15/10/2018

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) welcomes innovative designs and new methods of research. This Checklist has been developed by ENCePP to stimulate consideration of important principles when designing and writing a pharmacoepidemiological or pharmacovigilance study protocol. The Checklist is intended to promote the quality of such studies, not their uniformity. The user is also referred to the ENCePP Guide on Methodological Standards in Pharmacoepidemiology, which reviews and gives direct electronic access to guidance for research in pharmacoepidemiology and pharmacovigilance.

For each question of the Checklist, the investigator should indicate whether or not it has been addressed in the study protocol. If the answer is "Yes", the section number of the protocol where this issue has been discussed should be specified. It is possible that some questions do not apply to a particular study (for example, in the case of an innovative study design). In this case, the answer 'N/A' (Not Applicable) can be checked and the "Comments" field included for each section should be used to explain why. The "Comments" field can also be used to elaborate on a "No" answer.

This Checklist should be included as an Annex by marketing authorisation holders when submitting the protocol of a non-interventional post-authorisation safety study (PASS) to a regulatory authority (see the Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies). The Checklist is a supporting document and does not replace the format of the protocol for PASS presented in the Guidance and Module VIII of the Good pharmacovigilance practices (GVP).

Study title: Use cases for development, optimisation and implementation of artificial intelligence methods for real world data analyses in regulatory decision-making and health technology assessment along the product lifecycle
Checklist for study protocol from Work Package 2, Use Case 3

EU PAS Register® number:
Study reference number (if applicable):

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	16
1.1.2 End of data collection[2]	☐	☐	☒	
1.1.3 Progress report(s)	☐	☒	☐	16
1.1.4 Interim report(s)	☐	☒	☐	16
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	16
1.1.6 Final report of study results.	☒	☐	☐	16

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.
[2] Date from which the analytical dataset is completely available.

Comments:

1.1.2 presumably in 2024 1.1.3&1.1.4 no reports planned, but milestones for completion of specific analyses in 9/2025 and 2/2026

Section 2: Research question		**Yes**	**No**	**N/A**	**Section Number**
2.1	Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	2
	2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	2
	2.1.2 The objective(s) of the study?	☒	☐	☐	2
	2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	2
	2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☒	☐	2
	2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☒	☐	

Comments:

2.1.4 and 2.1.5: The objectives are stated in section 2

Section 3: Study design		**Yes**	**No**	**N/A**	**Section Number**
3.1	Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	3
3.2	Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	3
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	3
3.4	Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☒	☐	☐	3
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☒	☐	☐	5

Comments:

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	4
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	4
	4.2.2 Age and sex	☒	☐	☐	4
	4.2.3 Country of origin	☒	☐	☐	4
	4.2.4 Disease/indication	☒	☐	☐	4

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.2.5 Duration of follow-up	☒	☐	☐	4
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	4

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	5
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☒	☐	☐	4, 6
5.3 Is exposure categorised according to time windows?	☒	☐	☐	5
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	☐	☒	☐	
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6 Is (are) (an) appropriate comparator(s) identified?	☒	☐	☐	3, 4

Comments:

5.4. intensity not part of the study objective

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	5
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	5
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☒	☐	☐	5
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☒	☐	☐	9

Comments:

6.3. all-cause mortality and sudden cardiac death are presumably valid, but the validity of the most other outcomes is unknown

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☒	☐	☐	7
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☒	☐	☐	7
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	7

Comments:

Section 8: Effect measure modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☐	☒	☐	8

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	9
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9
9.1.3 Covariates and other characteristics?	☒	☐	☐	9
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	9
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	9
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	9
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	9
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	9
9.3.3 Covariates and other characteristics?	☒	☐	☐	9
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	9

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	10
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	10
10.3 Are descriptive analyses included?	☒	☐	☐	2, 3
10.4 Are stratified analyses included?	☐	☒	☐	
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	7
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☒	☐	
10.7 Does the plan describe methods for handling missing data?	☐	☒	☐	10.1
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	5, 10

Comments:

10.4. all analyses are stratified by country; further stratifications might be applied in an explorative manner on a country-specific level (Section 14); 10.6. we will investigate the possibilities for analytically controlling outcome misclassification after data inspection

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	11
11.2 Are methods of quality assurance described?	☒	☐	☐	6, 10.1, 11
11.3 Is there a system in place for independent review of study results?	☒	☐	☐	11

Comments:

11.2 each country will follow internal quality procedures

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☐	☐	☒	
12.1.2 Information bias?	☒	☐	☐	7
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☒	☐	☐	12
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	12

Comments:

12.1 The risk of selection bias is low since data is derived from national registries; information bias: misclassification is the same for both groups and thus no bias from this source is expected

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	13
13.2 Has any outcome of an ethical review procedure been addressed?	☒	☐	☐	13
13.3 Have data protection requirements been described?	☒	☐	☐	13

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	14

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	15
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	15

Comments:

Name of the main author of the protocol: Anna-Maija Tolppanen, Sirpa Hartikainen, Anne Paakinaho

Date: 11/April/2023

Signature: Britta Häuerle

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Doc.Ref. EMA/540136/2009

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

ENCePP Checklist for Study Protocols (Revision 4)

Adopted by the ENCePP Steering Group on 15/10/2018

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) welcomes innovative designs and new methods of research. This Checklist has been developed by ENCePP to stimulate consideration of important principles when designing and writing a pharmacoepidemiological or pharmacovigilance study protocol. The Checklist is intended to promote the quality of such studies, not their uniformity. The user is also referred to the ENCePP Guide on Methodological Standards in Pharmacoepidemiology, which reviews and gives direct electronic access to guidance for research in pharmacoepidemiology and pharmacovigilance.

For each question of the Checklist, the investigator should indicate whether or not it has been addressed in the study protocol. If the answer is "Yes", the section number of the protocol where this issue has been discussed should be specified. It is possible that some questions do not apply to a particular study (for example, in the case of an innovative study design). In this case, the answer 'N/A' (Not Applicable) can be checked and the "Comments" field included for each section should be used to explain why. The "Comments" field can also be used to elaborate on a "No" answer.

This Checklist should be included as an Annex by marketing authorisation holders when submitting the protocol of a non-interventional post-authorisation safety study (PASS) to a regulatory authority (see the Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies). The Checklist is a supporting document and does not replace the format of the protocol for PASS presented in the Guidance and Module VIII of the Good pharmacovigilance practices (GVP).

Study title: Use cases for development, optimisation and implementation of artificial intelligence methods for real world data analyses in regulatory decision-making and health technology assessment along the product lifecycle
Checklist for study protocol from Work Package 2, Use Case 4

EU PAS Register® number: EUPAS105544
Study reference number (if applicable):

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	16
1.1.2 End of data collection[2]	☐	☐	☒	
1.1.3 Progress report(s)	☐	☒	☐	
1.1.4 Interim report(s)	☐	☒	☐	
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	16
1.1.6 Final report of study results.	☒	☐	☐	16

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.
[2] Date from which the analytical dataset is completely available.

Comments:

1.1.2 presumably in 2024 1.1.3&1.1.4 no reports planned, but milestones for completion of specific analyses in 9/2025 and 2/2026

Section 2: Research question		**Yes**	**No**	**N/A**	**Section Number**
2.1	Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	2
	2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	2
	2.1.2 The objective(s) of the study?	☒	☐	☐	2
	2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	2
	2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☒	☐	2
	2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☒	☐	

Comments:

2.1.4 and 2.1.5 the objectives are stated in section 2

Section 3: Study design		**Yes**	**No**	**N/A**	**Section Number**
3.1	Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	3
3.2	Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	3
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	3
3.4	Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☒	☐	☐	3
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☐	☐	☒	

Comments:

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	4
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	4
	4.2.2 Age and sex	☒	☐	☐	4
	4.2.3 Country of origin	☒	☐	☐	4
	4.2.4 Disease/indication	☒	☐	☐	4

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.2.5 Duration of follow-up	☒	☐	☐	4
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	4, Table C

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	5
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☒	☐	☐	7
5.3 Is exposure categorised according to time windows?	☒	☐	☐	10
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	☐	☒	☐	
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6 Is (are) (an) appropriate comparator(s) identified?	☒	☐	☐	2

Comments:

5.4. intensity not part of the study objective

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	6
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	6
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☒	☐	☐	6, 7
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☒	☐	☐	10

Comments:

6.3. all-cause mortality is presumably valid; validity of death with heart failure recorded in death certificate is checked if the information is deficient, inconsistent, or difficult to classify

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☒	☐	☐	7
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☒	☐	☐	7
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	7

Comments:

Section 8: Effect measure modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☐	☒	☐	8

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	9
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9
9.1.3 Covariates and other characteristics?	☒	☐	☐	9
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	9
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	9
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	9
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	9
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	9
9.3.3 Covariates and other characteristics?	☒	☐	☐	9
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	9

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	10
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	10
10.3 Are descriptive analyses included?	☒	☐	☐	2.2
10.4 Are stratified analyses included?	☐	☒	☐	
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	7
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☒	☐	
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	11
10.8 Are relevant sensitivity analyses described?	☐	☒	☐	

Comments:

10.4. all analyses are stratified by country; further stratifications might be applied in an explorative manner on a country-specific level (Section 14) 10.6. we will investigate the possibilities for analytically controlling outcome misclassification after data inspection

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	11
11.2 Are methods of quality assurance described?	☒	☐	☐	11
11.3 Is there a system in place for independent review of study results?	☒	☐	☐	11

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☐	☐	☒	7
12.1.2 Information bias?	☒	☐	☐	7
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☒	☐	☐	7,12
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	12

Comments:

12.1. The risk of selection bias is low since data is derived from national registries; information bias: misclassification is the same for both groups and thus no bias from this source is expected

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	13
13.2 Has any outcome of an ethical review procedure been addressed?	☒	☐	☐	13
13.3 Have data protection requirements been described?	☒	☐	☐	13

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	14

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	15
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	15

Comments:

Name of the main author of the protocol: Anna-Maija Tolppanen, Sirpa Hartikainen, Anne Paakinaho

Date: 11/April/2024

Signature: Riitta Hämäläinen