30 April 2025
Data Analytics and Methods Taskforce EMA/467789/2024
European Medicines Agency

Standard operating procedure

Title: **Validation, Publication, and Maintenance of real-world data sources and studies in the HMA-EMA RWD Catalogues**		
Status: **Public**		Document no.: **SOP/TDA/3543**
Lead author:	Approver:	Effective date: 02/05/2025
Name: Sophie Groeneveld	Name: Paolo Alcini	Review date: 02/05/2028
Signature: Groeneveld Sophie Digitally signed by Groeneveld Sophie Date: 2025.04.30 14:47:55 +02'00'	Signature: Alcini Paolo Digitally signed by Alcini Paolo Date: 2025.04.30 14:56:08 +02'00'	Supersedes: N/A
Date:	Date:	DREAM no.: EMA/467789/2024

1. Purpose

The purpose of this SOP is to describe the steps and procedures concerning the validation, publication, and maintenance of data sources and studies within the HMA-EMA Catalogues of real-world data (RWD) sources and studies. The principles and activities described will apply specifically to EMA staff responsible for the validation of the Catalogues' content. The same principles will apply to the Catalogues' users when submitting entries.

2. Scope

This SOP applies to TDA-HCD and to TDA-RWE staff (hereafter called the EMA Validator) specifically tasked with the validation and maintenance of the Catalogues' records once they have been submitted for publication by the user (hereafter called Editor).

3. Responsibilities

It is the responsibility of the heads of TDA-HCD and TDA-RWE to ensure that this procedure is adhered to. The responsibility for the execution of a particular part of this procedure is identified in the right-hand column of section 9.

4. Changes since last revision

New SOP.

5. Documents needed for this SOP

- Good Practice Guide for the use of the HMA-EMA Catalogues of real-world data sources and studies
- List of metadata for the HMA-EMA Catalogues of Real-World data sources and studies

6. Related documents

N/A

7. Definitions and Acronyms

7.1 Definitions

Catalogue: A collection of dataset descriptions (i.e. metadata), which is arranged in a systematic manner and consistsof a user-oriented public part, where information concerning individual dataset parameters is accessible by electronic means through an online portal. In the context of this document, this refers to the Catalogue of RWD sources and the Catalogue of RWD studies.

Data holder: Any natural or legal person, which is an entity or a body in the health or care sector, or performing research in relation to these sectors, as well as Union institutions, bodies, offices and agencies who has the right or obligation, in accordance with the Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space [1], applicable Union law or national legislation implementing Union law, or in the case of non-personal data, through control of the technical design of a product and related services, the ability to make available, including to register, provide, restrict access or exchange certain data.

Data source: Dataset sustained by a specified actor, which is the data holder. The data source is characterised by the underlying population that can potentially contribute records to it, the trigger that leads to the creation of a record in the data source, and the data model used in the data source.

Editor: Registered user of the Catalogues, responsible for the submission and maintenance of a record in the Catalogues of RWD sources and studies (i.e., the data holder for RWD sources and the study owner for RWD studies). A record can have one or multiple editors (also called co-authors).

EMA Validator: EMA staff responsible for the verification of the submitted data in a record prior to its publication.

Institution: An organisation connected to one or more data sources - such as a data holder, or a research organisation running a study.

Metadata: Metadata are defined as "data about data" providing context about their purpose and generation. It's a set of data that describes and gives information on other data providing context about their purpose, location, key-variables, generation, format, and ownership of a dataset. Metadata are often published in data catalogues, which have the purpose of allowing data to be discoverable and checked for fitness for purpose, without revealing the data themselves.

Network: A virtual structure defined by a formal agreement between individuals, organisations and/or structures sharing and collaborating towards the same objectives and quality standards.

Study owner: The person or team responsible for the submission and maintenance of the study record

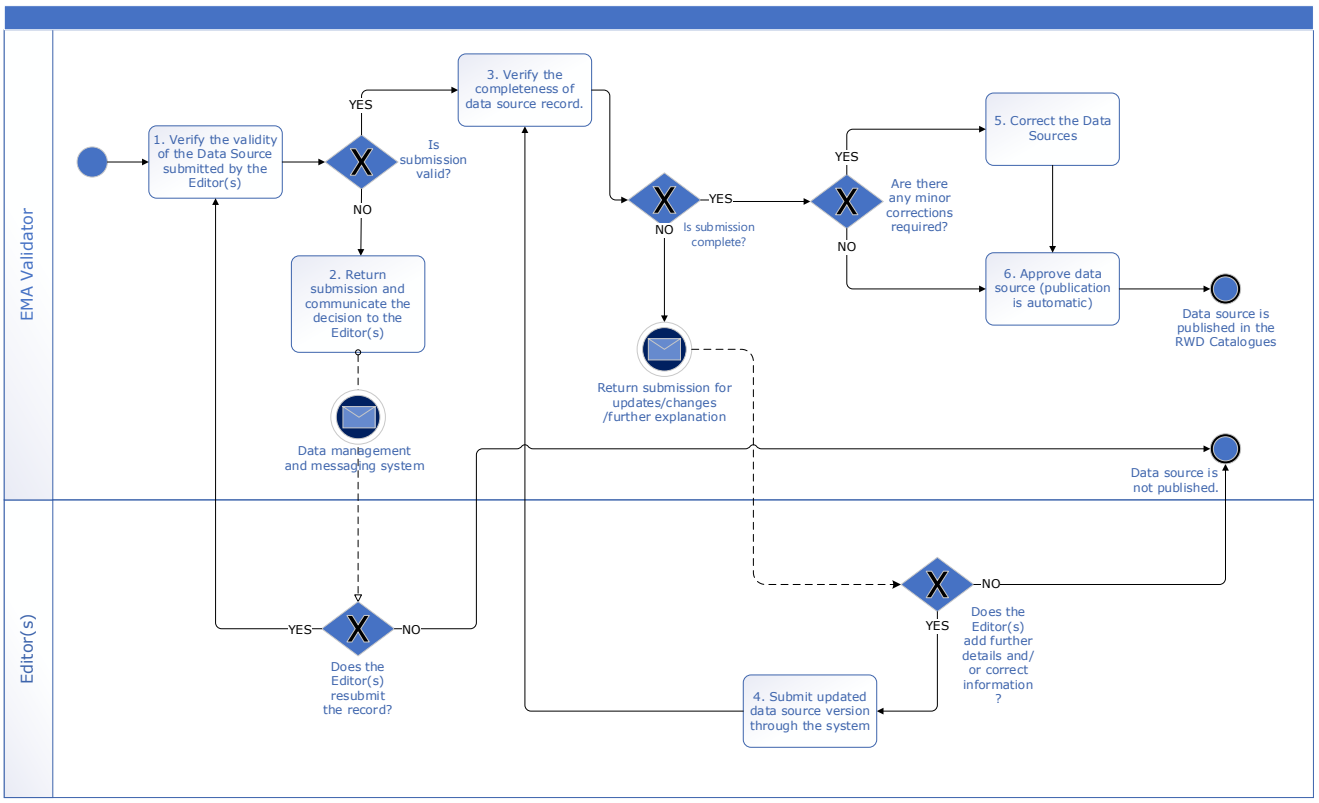
in the Catalogue of the RWD studies. This may be the primary investigator or any other delegated person or team, as deemed suitable.

7.2 Acronyms

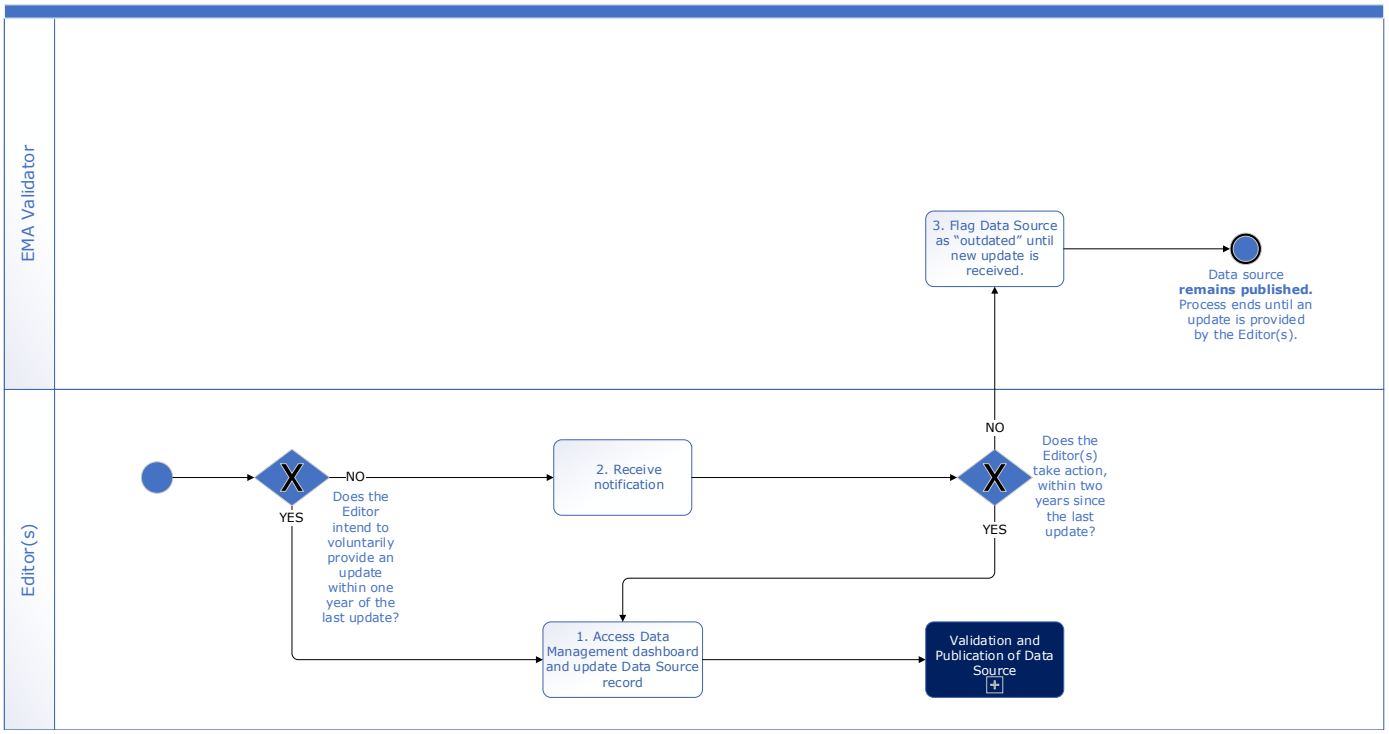
Acronym	Definition
EMA	European Medicine Agency
EU	European Union
GVP Module	Good Pharmacovigilance Practices Module
HMA	Heads of Medicines Agencies
MedDRA	Medical Dictionary for Regulatory Activities
PASS	Post-Authorisation Safety Studies
SOP	Standard Operating Procedure
TDA-HCD	Data Analytics and Methods – Healthcare Data workstream
TDA-RWE	Data Analytics and Methods – Real World Evidence workstream

8. Process map(s)/ flow chart(s)

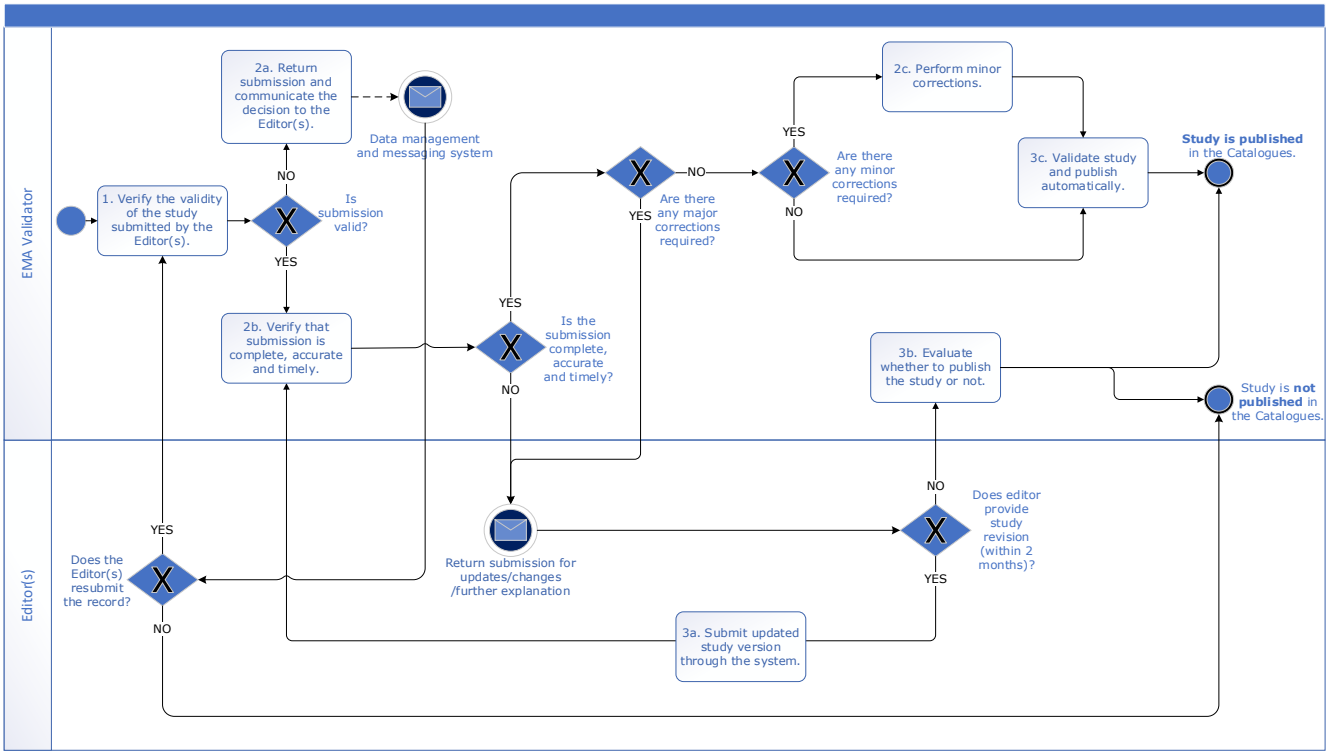
8.1 Validation and Publication of Data Sources



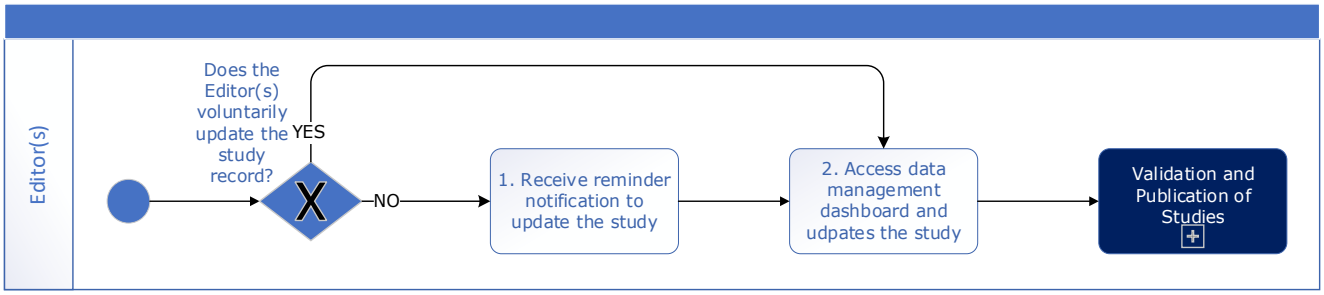
8.2 Maintenance of Data Sources



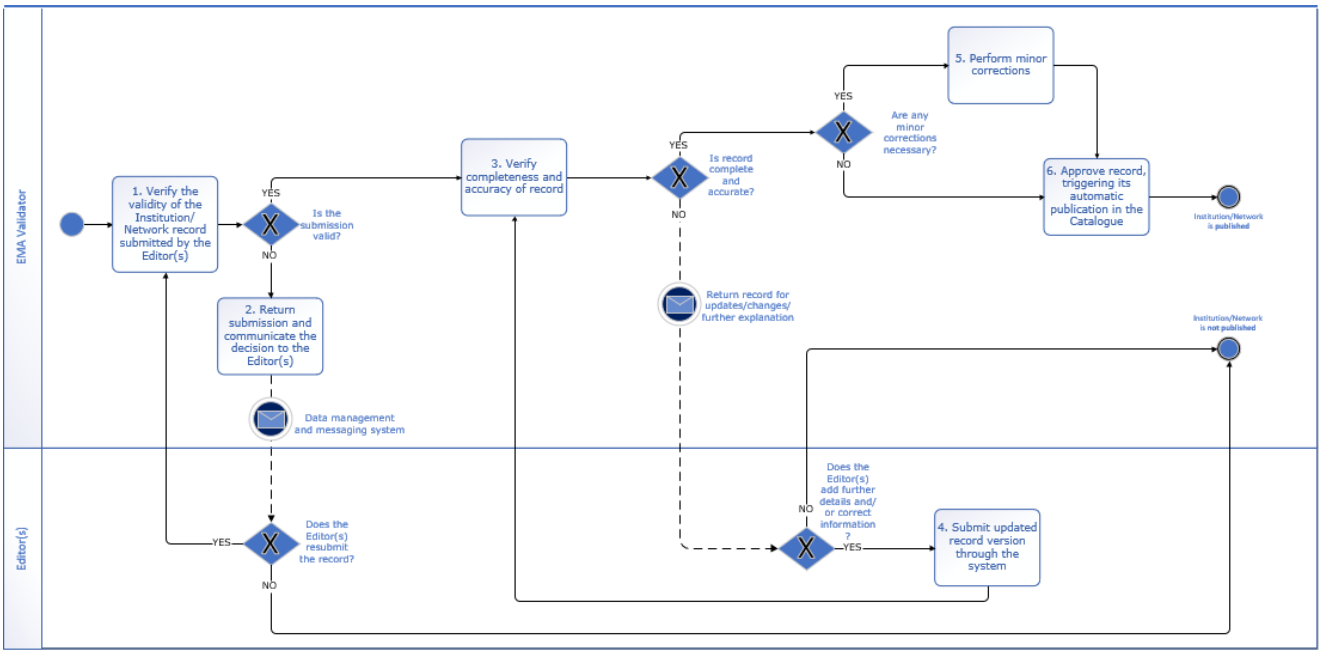
8.3 Validation and Publication of Studies



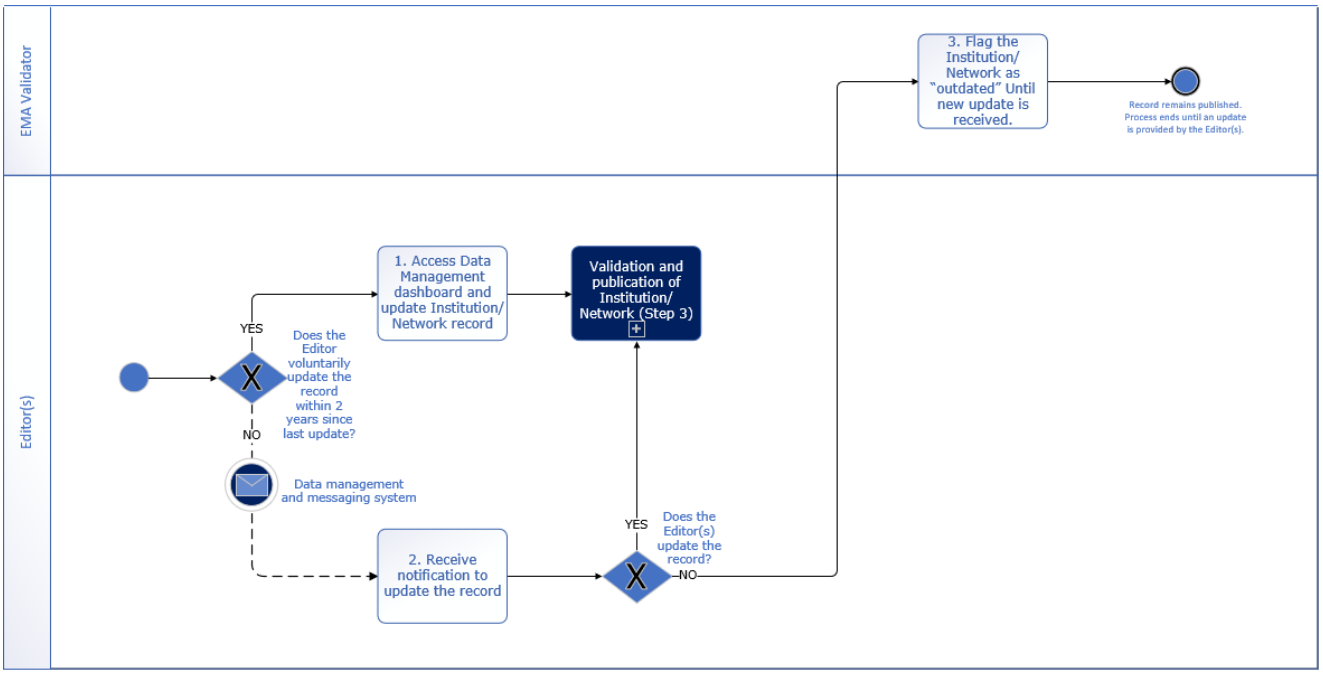
8.4 Maintenance of Studies



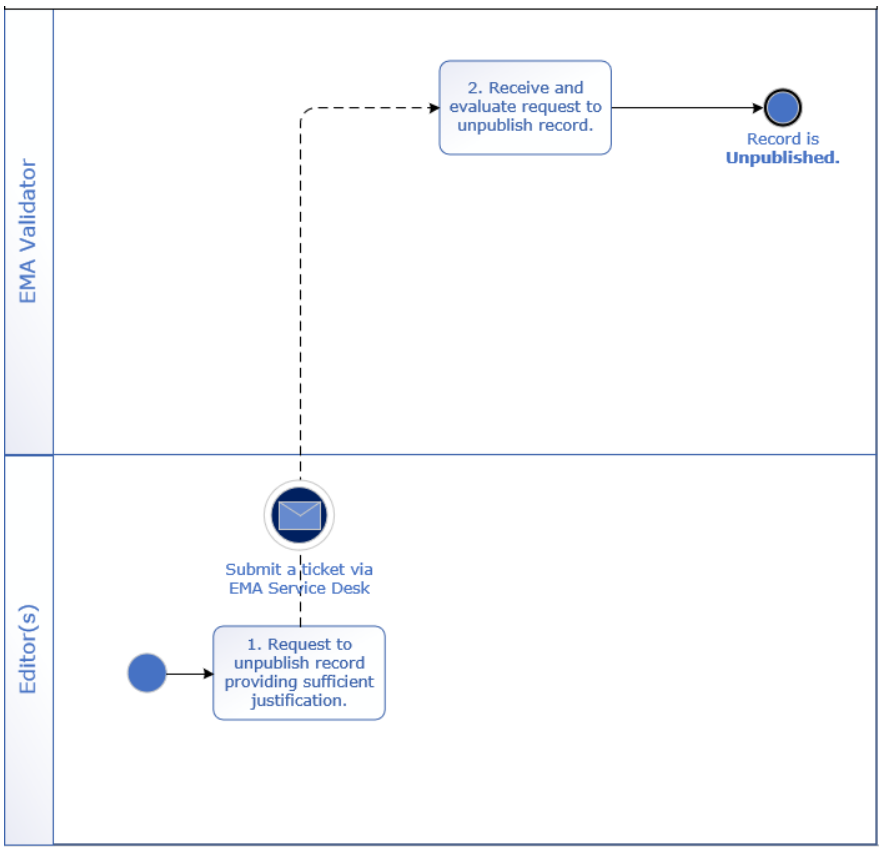
8.5 Validation and Publication of Institutions and Networks



8.6 Maintenance of Institutions and Networks



8.7 Unpublishing a record



9. Procedures

9.1 Validation and Publication of Data Sources

Step	Action	Responsibility
1	In an initial step, the EMA Validator verifies the validity of the Data Source record provided by the Editor(s) via the Catalogues website. The EMA Validator assesses the below criteria: • **Relevance to Health Context**, with a focus on data sources supporting medicines regulation. The data source should be considered relevant to the healthcare domain, aligning with EMA's regulatory focus and objectives. As one of the main objectives of the Catalogue of RWD sources is to promote data discoverability, this assessment is made in an inclusive manner, aiming to reflect the diversity of existing landscape of data sources. • **Uniqueness and Non-Duplication:** A check of the data source as a 'new' entity is performed in the context of the other data sources already registered in the Catalogue. An assessment of a potential record duplication is carried out and clarifications are sought from the Editor(s) when record duplication is suspected. Data source records created through linkage of other data sources, as well as data sources resulting from the transformation process connected to a common data model are considered independent (unique) data entries, while the connections between these data sources should be documented in the metadata elements. The criteria describing what constitutes a duplicate record are adjusted and further aligned as further information is received via communication with the Editor(s). • **Comprehensiveness and inclusion of mandatory elements:** The data source record should contain at a minimum the complete set of mandatory metadata elements – the provision of this information is enforced technically; the EMA Validator performs an initial content validation of the mandatory fields, ensuring that the content is meaningful and correct, to the extent that correctness can be assessed. If the submission meets all the validation criteria, go to step 3.	**EMA Validator**

2	If the submission does not meet all the validation criteria described, the EMA validator returns the submission, using the Catalogue's data management and notifications system, adding a message in the Revision Log containing the justification for the return. The Editor(s) receive(s) an e-mail notification regarding the validation outcome of their submission and can further refine the information provided and resubmit the record, as needed. If the Editor(s) resubmits the record, go to step 1.	**EMA Validator** **Editor(s)**
3	The EMA Validator verifies that the data source record submission is complete, meaning that metadata information expected on a given data source is submitted. This encompasses information in both mandatory and non-mandatory fields and the expected information may vary between each data source. For example, a disease registry would need the information on the 'Disease' field filled in. In the case of a prescription database the information on the type of vocabulary used would be expected to be filled in. The accuracy of the information provided is similarly verified in this step. The EMA Validator verifies the information submitted in its entirety, against publicly available information (e.g.: data source's website, article on the data source) and against the guidance provided in the User Guide. • If the submission is complete and accurate, and does not require minor corrections, go to step 6. • If the findings related to completeness or accuracy are minor and the EMA Validator can resolve them, go to step 5. • If the metadata submission is assessed as incomplete or inaccurate (major findings), the EMA validator returns the submission and requests the Editor(s) to make the necessary changes or to provide further explanations to the data source record. Go to step 4.	**EMA Validator**
4	The Editor(s) receive(s) the returned submission and is requested to add further details and/or to correct information. The Editor(s) now has/have the possibility to submit additional information. There are no prescribed timelines associated with this step, but it is recommended to submit the update within 2-3 weeks. If the Editor(s) submit(s) the updated data source record, go back to step 3. For data sources where no updates are received, the information will remain in a 'returned' state and will not be published in the Catalogue. The process is halted until the Editor(s) update(s) the record and resubmit(s) for publication.	**Editor(s)**

5	If necessary, the EMA Validator performs minor corrections. Minor corrections refer to adjustments made by the EMA Validator ensuring that the structured data provided is following the rules set out, consistently, throughout the Catalogues. Minor corrections do not refer to changing the content provided by the Editor(s), but, matching existing information to the correct field where necessary. This step is performed to streamline the data submission process wherever possible. For example, if the Editor(s) enter(s) information in a free-text field (e.g., 'we use MedDRA dictionary to code indication'), while a structured value (e.g., MedDRA) is available in a predefined lookup, the EMA Validator harmonises the entry with the structured value entry. If there is a formatting error or a typo, the EMA Validator may also correct it. The EMA Validator may also complete with the remaining Age Groups when the Editor has selected "All". However, any other inaccuracies will be considered major and will be returned to the Editor for correction. Once corrections are performed, if any, go to step 6.	**EMA Validator**
6	EMA Validator approves the data source record, which triggers the automatic publication of the record in the Catalogue.	**EMA Validator**

9.2 Maintenance of Data Sources

Step	Action	Responsibility
1	The Editor(s) is expected to keep the information of their record up to date on a regular basis. It is encouraged to provide updates as often as needed, but at a minimum this is expected on a yearly basis. The Editor(s) will receive a notification when 1 year has passed since they last updated the data source record. To follow the update of the data source record process, go to step 3. The Editor(s) may voluntarily update the data source within one year of the last update. Any update performed by the Editor(s) initiates a return to step 3 of the Validation and publication of Data Source Process (Section 9.1).	**Editor(s)**
2	If one year has passed since the last update, the Editor(s) receive(s) notifications via the messaging system, requesting them to update the data source. • If the Editor(s) update(s) the data source record, go to step 3 of the Validation and publication of Data Source Process (Section 9.1). If the Editor(s) do(es) not provide an update for two years since the last update, go to step 3.	**Editor(s)**
3	The EMA Validator flags the data source as "outdated" and keeps the entry public to support the assessment of study data and data discoverability. Note: In case the data source has been marked as "outdated" due to lack of an update for two years, the Editor(s) may update the data source record at any point after the two-year period after which the "outdated" flag will then be automatically removed.	**EMA Validator**

9.3 Validation and Publication of Studies

Note: *For any new study requesting the ENCePP Seal, the ENCePP Secretariat performs the validation process of the study record following the EMA Validator steps as required. The ENCePP Secretariat grants the ENCePP Seal to the study if the necessary requirements are met.*

Step	Action	Responsibility
1	The EMA Validator verifies the validity of the study record submitted, by the user, by assessing against the criteria listed below: • **Non-Duplication:** The study record should not duplicate another existing record in the Catalogue. • **Exclusion of Clinical Trials outside the scope of the Catalogue:** In case of a clinical trial, the EMA Validator needs to verify that the study falls within the scope of the Catalogue. Clinical trials that are outside the scope of the Catalogue will be excluded. • **Relevance:** the study should include a medicinal product/device or, a class or a specific disease or disease area. • **Required documentation for PASS imposed studies**: protocols and abstract of results (as applicable for the study timeline) should be published.	**EMA Validator**
2a	If any of the above validity criteria are not met, the EMA Validator returns the submitted data via the Catalogues' data management and notification system, accompanied by a justification. The Editor(s) then has/have the possibility to further refine and re-submit the study record, as deemed necessary. If the submission meets all the validation criteria, go to step 2b.	**EMA Validator**
2b	The EMA validator ensures that the study submission is *complete*, meaning that all relevant fields (e.g. sources of funding, study description, age groups, main study objective(s), scope of the study, data analysis plan, name of medicine (if applicable), medical condition to be studied (if applicable), estimated number of subjects) are filled in and the information is consistent. Additional to the mandatory information (marked by a red asterisk), the applicability of each data element depends on the type of study validated. The validation is performed against the submitted attachment files (e.g.: study protocol, study results where applicable) and/or publicly available information (e.g.: publications). An assessment of the *accuracy* of information and internal validity is also performed, using similar references as above. These checks aim to ensure high quality of the structured information provided with the study data.	**EMA Validator**

	The *timeliness* of the submissions of the imposed PASS are also assessed in accordance with the guidance set out by GVP Module VIII. Additionally, the uploaded documentation will also undergo validation. Documents containing internal comments will be returned for resubmission. Please note: No studies will be prioritised, including those that have been mandated. • If the submission is considered incomplete or inaccurate, the EMA validator returns the submission and requests the editor to make the necessary changes, or to provide further clarifications, go to step 3a. • If the submission is complete and accurate, go to step 2c.	
2c	If during the validation process it is found that minor corrections are possible and/or necessary, these may be performed by the EMA Validator. Minor corrections refer to adjustments made by the EMA Validator to maintain data consistency with the original submission, primarily from a data management perspective. These changes are meant to streamline the submission of information without altering its content. For example, the misplacing of information in the wrong field (e.g.: study scope data filled in the results section) would qualify as a minor change. Any other inaccuracies will be considered major and will be returned to the Editor(s) for revision. If minor corrections are necessary, the EMA Validator may rectify the study. Go to step 3c.	**EMA Validator**
3a	The Editor(s) receive(s) an e-mail notification regarding the validation outcome and the request to add further details or modify/correct the information previously provided. If the Editor(s) re-submit(s) the study with updated information via the system, go back to step 2b.	**Editor(s)**
3b	If the Editor(s) do(es) not provide updates within two calendar months, the publication of the study will be considered on a case-by-case basis.	**EMA Validator**
3c	The EMA Validator approves the study record, which triggers the automatic publication of the record in the Catalogue. Please note: if the study is finalised and the "date of final study report" has been filled in by the Editor(s), the system will automatically mark the study status as "finalised".	**EMA Validator**

9.4 Maintenance of Studies

Note: *For any study that is already published in the Catalogue and subsequently requests the ENCePP Seal, the ENCePP Secretariat performs the validation process of the study record following the EMA Validator steps as required. The ENCePP Secretariat grants the ENCePP Seal to the study if the necessary requirements are met.*

Step	Action	Responsibility
1	To support transparency and the exchange of information, the information on a study record should be kept up to date throughout the lifecycle of the study. Of particular importance, the maintenance of non-interventional PASS conducted pursuant to an obligation imposed by an EU competent Authority should follow the milestones and guidance described in GVP Module VIII. Updated study protocols in case of substantial amendments, progress reports and the final study report should also be entered in the register (as soon as possible and preferably within two weeks after their finalisation). A notification system is set up to provide support to the Editor(s) of studies by reminding them of the milestones described at a minimum, as follows: - on the planned date for the start of data collection, - on the planned date of the final study report. Please note: If the Editor(s) do(es) not provide updates for an already published study record, the record will not be unpublished.	**Editor(s)**
2	The Editor(s) access(es) the data management dashboard and updates the study record. Any updates to the record trigger a restart of the validation process, as described in the Validation and Publication of Studies Process (Section 9.3, step 2b).	**Editor(s)**

9.5 Validation and Publication of Institutions and Networks

Note: *For any new Institution/Network record requesting to become an ENCePP partner, the ENCePP Secretariat performs the validation process following the EMA Validator steps as required. The ENCePP Secretariat is verifying whether the Institution/Network can become an ENCePP partner based on the necessary requirements that need to be met.*

Step	Action	Responsibility
1	In an initial step, the EMA Validator verifies the validity of the Institution/Network record provided by the Editor(s) via the Catalogues website. The EMA Validator assesses the below criteria: • **Uniqueness and Non-Duplication:** A check of the institution/network as a 'new' entity in the HMA-EMA Catalogues of RWD sources and studies is performed in the context of the other institutions/networks already registered in the Catalogue. An assessment of a potential record duplication is carried out and clarifications are sought from the Editor(s) when record duplication is suspected. The criteria describing what constitutes a duplicate record are adjusted and further aligned as further information is received via communication with the Editor(s). • **Comprehensiveness and inclusion of mandatory elements:** The record should contain at a minimum the complete set of mandatory metadata elements – the provision of this information is enforced technically; the EMA Validator performs an initial content validation of the mandatory fields, ensuring that the content is meaningful and correct, to the extent that correctness can be assessed.	**EMA Validator**
2	If the submission does not meet all the validation criteria described, the EMA validator returns the submission, using the Catalogue's data management and notifications system, adding a message in the Revision Log containing the justification for the return. The Editor(s) receive(s) an e-mail notification regarding the validation outcome of their submission and can further refine the information provided and resubmit the record, as needed. If the submission meets all the validation criteria, go to step 3.	**EMA Validator**

3	The EMA Validator verifies that the Institution/Network record submission is complete, meaning that metadata information expected on a given entity is submitted. The *accuracy* of the information provided is similarly verified in this step. The EMA Validator verifies the information submitted in its entirety, against publicly available information (e.g.: website of the institution/network). • If the submission is complete, go to step 5. • If the findings related to completeness or accuracy are minor and the EMA Validator can resolve them, go to step 5. • If the metadata submission is assessed as incomplete or inaccurate, the EMA validator returns the submission and requests the Editor(s) to make the necessary changes or to provide further explanations to the record. Go to step 4.	**EMA Validator**
4	The Editor(s) receive(s) the returned submission and request to add further details and/or to correct information. The Editor(s) now has/have the possibility to submit additional information. There are no prescribed timelines associated with this step, but it is recommended to submit the update within 2-3 weeks. If the Editor(s) submit(s) the updated record, go back to step 3. For institutions/networks where no updates are received, the information will remain in a 'returned' state and will not be published in the Catalogue. The process is halted until the Editor(s) update(s) the record and resubmit(s) for publication.	**Editor(s)**
5	If necessary, the EMA Validator performs minor corrections. Minor corrections refer to adjustments made by the EMA Validator ensuring that the structured data provided is following the rules set out, consistently, throughout the Catalogues. Minor corrections do not refer to changing the content provided by the Editor(s), but, matching existing information to the correct field where necessary, correcting typos, etc. This step is performed to streamline the data submission process wherever possible. Once corrections are performed, if any, go to step 6.	**EMA Validator**
6	The EMA Validator approves the Institution/Network record, which triggers the automatic publication of the record in the Catalogue.	**EMA Validator**

9.6 Maintenance of Institutions and Networks

Note: *For any Institution/Network record that is already published in the Catalogue and subsequently requests to become an ENCePP partner, the ENCePP Secretariat performs the validation process following the EMA Validator steps as required. The ENCePP Secretariat is verifying whether the Institution/Network can become an ENCePP partner based on the necessary requirements that need to be met.*

Step	Action	Responsibility
1	The Editor(s) is/are expected to keep the information of their record up to date and provide updates as necessary (e.g.: in case the institution's website link changes, if a new institution is added to the list of participating institutions for a network, etc). The Editor(s) may voluntarily update the Institution/Network within two (2) years of the last update, making use of the data management system available to registered users in the Catalogue (similar to the initial data submission process). Any update performed by the Editor(s) initiates a return to step 3 of the Validation and publication of Institution/Network Validation and Publication Process (Section 9.5). The Editor(s) will receive a notification when 2 years have passed since they last updated their record.	**Editor(s)**
2	If two years have passed since the last update, the Editor(s) receive(s) a notification via the messaging system, requesting them to update their record. • If the Editor(s) update(s) the institution/network record, go to step 3 of the Validation and publication of Institutions and Networks Process (Section 9.5). • If the Editor(s) do(es) not take any action after five years since the last update, go to step 3.	**Editor(s)**
3	The EMA Validator flags the Institution/Network as "outdated" and keeps the record public. End of process Note: In case the Institution/Network has been marked as "outdated" due to lack of an update for five years, the Editor(s) may update the record at any point after the five-year period and the "outdated" flag will then be removed.	**EMA Validator**

9.7 Unpublishing a record

1	The Editor(s) may request the unpublishing of their record, where this might be the best course of action (for example, for a data source that ceases to exist). The Editor(s) send(s) a request to the EMA Validator via the Service Desk to unpublish the record is made, providing sufficient justification for the request. To follow the un-publication of the data source record process, go to step 2.	**Editor(s)**
2	Upon receival of the Editor's request, the EMA Validator verifies whether adequate justification is provided for this request and whether the data source record is associated with any study record. The timelines to process this request may take up to two weeks or up to one month in exceptional circumstances (e.g.: high volume of data to be processed).	**EMA Validator**

References:

1. Proposal for a Regulation of the European Parliament and of The Council on the European Health Data Space, COM (2022) 197 final, 2022/0140 (COD). Available at: EUR-Lex - 52022PC0197 - EN - EUR-Lex