

We have not been able to find any appropriate reporting guidelines for descriptive analysis of retrospective harmonization procedures. While the number of project harmonizing and co-analysing data across studies increases, it would be important to develop such guidelines. We propose to use here the check list provided in the Maelstrom Research guidelines for rigorous retrospective data harmonization (Fortier I, Raina P, Van den Heuvel ER, et al. Maelstrom Research guidelines for rigorous retrospective data harmonization. Int J Epidemiol 2017;46:103-5.).

### Checklist helping to review the harmonization process

Step	Item	Description	
<b>Step 0: Define the questions and objectives</b>	1	The research question is well defined in term of population, exposure, comparator, outcome and timing.	√
	2	The protocol takes into account questions related to feasibility (e.g. data access, realistic timelines) and provides information required to guide the harmonization process.	√
<b>Step 1: Assemble information and select studies</b>			
Step 1a: Document individual study designs, methods and content	3	Study-specific information gathered allows understanding study designs, timeline, population characteristics, data contents, standard operating procedures and ethico-legal requirements to access data.	√
Step 1b: Select participant studies	4	Studies are selected based on explicit selection criteria.	√
<b>Step 2: Define variables and evaluate harmonization potential</b>			
Step 2a: Select and define the core variables to be harmonized (DataSchema)	5	The DataSchema variables are selected based on their relevance to answer the research question addressed, likelihood to be generated across a number of studies and, where relevant, input from experts.	√
	6	The DataSchema variables are clearly defined, including their specific nature, format, and acceptable level of heterogeneity.	√
Step 2b: Determine the potential to generate the DataSchema variables making use of study-specific data items	7	The potential (or not) for each study to create the DataSchema variables is assessed and documented.	√
<b>Step 3: Process data</b>			
Step 3a: Ensure access to adequate study-specific data items and establish the overall data processing infrastructure	8	If harmonization is possible, the study-specific data items required to generate the DataSchema variables are made available in a computing infrastructure allowing data processing.	√
	9	Quality of study-specific data items is assessed and considered adequate.	√
Step 3b: Process study-specific data items under a common format to generate the harmonized dataset(s)	10	Data processing is achieved using appropriate statistical models or processing algorithms.	√
	11	Harmonized data is generated and algorithms or models used to process data are documented.	√
<b>Step 4: Estimate quality of the harmonized dataset(s) generated</b>	12	Quality and consistency of the harmonized data is assessed. Where appropriate, statistical models are applied to evaluate heterogeneity and potential bias.	√
<b>Step 5: Disseminate and preserve final harmonization products</b>	13	Harmonized data is available to approved users	√
	14	All information required to understand harmonization procedures and to analyse the harmonized data are accessible.	√