

Table 2. Qualitative evaluation categories and criteria.

Evaluation category	Evaluation criteria
Characteristics	Principle Prediction (i.e. hazard versus potency [categories or continuous]) Publication Information sources
Input data	Test method (<i>in vitro</i> and <i>in chemico</i>) - read-out used - validation status - reproducibility - issues (e.g. IP, availability) <i>In silico</i> /expert system data/physicochemical properties - read-out used - availability - reliability - issues (e.g. IP, availability) Expert knowledge - input used - availability
Prediction algorithm	Principle Prediction (i.e. hazard vs. potency [categories or continuous]) Publication Information sources
Prediction algorithm	Type Availability Transparency Requirements for implementation (specific software) Self-learning Complexity Sequential information generation All inputs required? Predictivity: Sample size (total and for categories) Predictivity: Parameters (sensitivity, specificity, concordance)
Mechanistic relevance	OECD AOP key events covered Sequence of OECD AOP events considered Justification/discussion of the mechanistic relevance
Applicability domain	Chemical spectrum tested Limitations (solubility, surfactants, ...) Potential limitations for cosmetic ingredients (e.g. natural extracts cannot be processed by <i>in silico</i> approaches)
Practical aspects	Costs Can be conducted by CRO? Time required (per substance)