Guideline for using VigiBase data in studies

The information set out in this document should be considered carefully when using data from VigiBase, the WHO global database of individual case safety reports, for research intended for scientific publications.

This guideline provides important information on the:

- background and content of VigiBase

- interpretation of VigiBase data

- standard terminology used when publishing on VigiBase data

This guideline complements the Uppsala Monitoring Centre (UMC) Caveat document, which must be read and adhered to.

Recipients of data are fully responsible to ensure the scientific integrity of their analyses involving VigiBase data. UMC welcomes scientific communication and debate around analyses involving VigiBase data.

If you have any further questions on what is covered in this document, please do not hesitate to contact the UMC (info@who-umc.org).

Background

The WHO Programme for International Drug Monitoring (WHO PIDM) was established in 1968 as a result of the thalidomide crisis of the early 1960s. WHO PIDM currently has around 140 member countries (November 2020). In each participating country, the ministry of health, or equivalent, has appointed a national centre for pharmacovigilance that collects and manages Individual Case Safety Reports (ICSR) and is the national point of contact. These reports are transferred electronically to VigiBase, the WHO global database of ICSRs ('report' denotes ICSR in this document).(1) VigiBase is maintained and developed by the UMC and members of the WHO PIDM can access and analyse this common resource using VigiLyze, a signal detection and management tool provided by UMC.

Purpose of VigiBase

VigiBase is used to identify novel medicine-related problems and to gain knowledge on specific features of adverse drug reactions (ADRs) such as time to onset, course of the event, and outcome. VigiBase is also used as a source of reference in evaluations of specific drugs and/or adverse reactions. The primary users of VigiBase are national pharmacovigilance centres and the UMC. Other stakeholders that wishes to use the data in VigiBase for public health benefit and for purposes that harmonise with the purpose of the WHO PIDM can access the public level of data, openly available at www.vigiaccess.org, or per request through a customised search or the full dataset in VigiBase extract case level. Recipients of data need to have appropriate experience and knowledge of working with spontaneous reporting databases; members of the medical, pharmacy or nursing professions who in the course of his or her professional activities may prescribe, administer or dispense to an end-user a medicinal product are eligible to request data.

Protection of personal data

According to WHO policy and UMC guidelines, reports sent from the WHO PIDM member countries to VigiBase are anonymized, but they are still to be considered sensitive due to the nature of the data (see the UMC Caveat document). Publication of identifiable data is not permitted, e.g. reporting individual age, sex and country information in connection is considered sensitive and may not be presented in a publication or communication to a third party. Age and country information is currently not available in the public level of VigiBase data.

VigiBase content

The reports in VigiBase contain qualitative information on patients reported with problems in connection to using medicines and include patient demographics, reported drugs and events, and additional information relevant to the case; the amount of information in each report varies between reports and submitting countries.

Terminologies: To enable efficient analysis, content of the reports, especially drug and reaction information, need to be coded using standard terminologies. The standard terminology for ADRs in VigiBase is the Medical Dictionary for Regulatory Activities (MedDRA) and for medicinal products it is WHODrug. These terminologies are both hierarchically structured, enabling analysis on different levels of aggregation. They also include a standard set of queries (Standardised MedDRA Queries, SMQs, for MedDRA and Standardised Drug Groupings, SDGs for WHODrug) that enables search and analysis by other properties than only the hierarchy of the terminology. VigiBase drug and reaction coding is updated with each new release of these terminologies, 2 times per year.

Variations in strength of causality: The likelihood that the reported event was caused by the medicine varies from report to report. Some countries collect only suspected ADRs with at least a possibility of a causal relationship between the drug and reported event, whilst for example the United States (contributing over 40% of total reports in VigiBase) collects "…any adverse event associated with the use of a drug in humans, whether or not considered drug related…".(2)

Heterogeneity: The reports in VigiBase originate from multiple sources: different countries and different types of reporters such as physicians, pharmacists, other health care professionals, and patients. A small proportion of reports in VigiBase have been generated from clinical studies or intensive monitoring programmes.

Completeness and quality of data

For a report to be valid according to the standard transmission format ICH[1] E2B, the minimum requirements are: an identifiable patient, an identifiable reporter, an adverse reaction, and a suspected drug. Reports from each national pharmacovigilance centre are anonymous but centres

---

[1] ICH=International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use

should validate the existence of both a patient and reporter before transfer to VigiBase. The minimum administrative information for processing and identifying a report in VigiBase is a sender's unique case identification number, worldwide case identification number, sender identifier, and date of receipt of most recent information. Each individual report contains a varying amount of additional data. Case data may be missing for a number of reasons; because it was not available to the reporter, was not considered relevant to the case, or was not transmitted in full to VigiBase by the national pharmacovigilance centres, e.g. because of national regulations. Standardised electronic formats for exchange of reports have changed over the years, generally to include more information. Currently VigiBase accepts 3 different standard formats and necessary data transformation will take place. VigiBase data structure resembles most closely the ICH E2B(R2) format, but includes additional data elements to enable data validation and relevant calculations.

Efforts are being made to increase the quality of the reports within pharmacovigilance systems.(3) At UMC, work has been undertaken to specify the amount of information given on a report. Each report is assigned a score of the amount of information reported; this can help in pin-pointing trends in report quality over time in order to identify areas in need of improvement.(4) The score, called vigiGrade completeness, may also guide report reviewers in judging whether there is enough report information at hand for a problem to be properly investigated.

VigiBase can contain duplication of reports. Suspected duplicate reports can automatically be identified with vigiMatch, an algorithm that uses a statistical model to score pairs of reports, taking into account the amount of matching and mismatching information.(5,6) However, reports that are not in fact duplicates but otherwise related are difficult for the algorithm to handle. The algorithm can also miss duplicate reports when report information is minimal.

Updating information

VigiBase is continuously updated with new reports, reports can be deleted and reports are updated and replaced when additional information is received. When WHODrug and MedDRA are updated, new drugs and terms are introduced, may change and they may be rearranged within the terminologies, which will affect the coding on the reports. Therefore, a VigiBase search repeated at a later date may not generate data identical to the original search.

Study objectives and interpretation of data

As with all pharmacovigilance reporting databases, it is appropriate to use VigiBase data for generating but not for testing hypotheses. The data in VigiBase are subject to reporting biases, duplication, confounding issues and heterogeneity, over time and across regions. In addition, access to information on global use of a medicine can be limited. As a consequence, summary statistics must be interpreted with caution. Medical expertise should be sought to determine the clinical relevance of the findings and to highlight potential confounding issues.

VigiBase reports cannot be treated as a random sample from a population of patients, as one may do in a clinical trial or an observational study. In a clinical trial, the number of treated patients is known, as well as the number of patients with a certain reaction. This allows for an estimation of risk in the population selected for the trial. In VigiBase, the total number of reports for a specific drug is known, as is the number of reports with a specific ADR term coded to them. This allows for computation of a reporting rate, which is not the same as a risk estimate. Furthermore, the number of reports with a drug and ADR can be subject to lower or higher levels of under-reporting than with another drug with the same ADR. A typical mechanism for such a difference is positive reporting bias due to externally generated attention to a specific issue reducing the underreporting for it. Consequently, one must be very cautious when comparing such reporting rates between different drugs.

Data needing special consideration

It is important to gather information on whether there are specific factors to be considered when analysing details of VigiBase data. Please consult UMC staff.

Examples where special consideration is needed:

- **Information on causality assessments of the drug and event**. This information is not available in all reports, due to the diverse practices of the national pharmacovigilance centres, which also use a range of different methods for assigning causality, if at all. Therefore, an examination restricted to reports with causality listed as certain, for instance, should take this into account. It is important to note that causality can change over time as our knowledge expands, and that UMC does not validate the causality assessments specified on the reports received from the national pharmacovigilance centres.

- **Fatal outcome information** can be recorded as an ADR outcome, and/or in the 'seriousness' field, and/or coded with the ADR terminology used. Any information on the report suggesting fatal outcome, also including fields that may not be available to the data recipient, is given in the separate VigiBase field 'Died'.

- **Co-reported drugs**: Despite there being no other suspect, interacting or concomitant drugs specified in the report, other drugs might have been taken by the patient, but were not considered sufficiently relevant by the reporter to be recorded or were omitted for unknown reasons.

- **Concomitant drugs** may not necessarily have been used at the same time as the suspected drug and are often reported with limited details.

- **Changes in reporting terminology over time**: Reports originally coded in MedDRA started to be entered into VigiBase from November 2000 (with reports from USA being the first). WHO-ART, which has since been gradually phased out in favour of MedDRA, was a smaller terminology and restricted to ADR terms, whilst MedDRA is used also to specify indications. When reviewing VigiBase data retrospectively, it is important to bear in mind that the terminologies evolve; some MedDRA terms have not always been possible to report. This can introduce a bias of a higher reporting rate for certain terms during recent years compared to older reports, and vice versa. If the reports for the drug(s) under study have a different distribution in time compared to a reference group, this could lead to an over-representation of MedDRA terms that were not available in WHO-ART at the time of reporting.

- **Changes in reporting formats over time:** VigiBase includes three main report formats: INTDIS, being the original WHO reporting format, ICH E2B(R2) and ICH E2B(R3), the current standard. In certain database fields, the values which may be entered differ between these formats. Currently, data is transformed to a format most closely resembling ICH E2B(R2). An example of data transformation is the reporter type field ('notifier') where consumer reports in the INTDIS format are given as 'other' while in the E2B(R2) format they are 'Consumer/Non-health professional'. The reporter type 'other' in INTDIS is not limited to reports from consumers, so all reports in this category cannot be assumed to be consumers. Another example is that certain data (i.e. date of onset, dechallenge, rechallenge, outcome and causality assessment) can only be reported per report in INTDIS and not per event/drug as in E2B(R2). In addition, seriousness and

parent-child cases cannot be reported in structured format in INTDIS. An example for the transformation from E2B(R3) to E2B(R2) is that seriousness is given per event in R3, but per report in R2. Please consult UMC staff for complete information on the reporting format differences.

- **The 'datedatabase' given on VigiBase reports does not refer to when the suspected ADR occurred**. A time gap will exist between the onset of an event, when the ADR was recognized, when it was reported to the national pharmacovigilance centre, and when the report was transferred to and made accessible in VigiBase. The field 'first datedatabase' is the date when the report was first accessible in VigiBase. Reports that are replaced with updates will keep their 'first datedatabase' but get an updated 'datedatabase'.

- **Frequency of reporting to VigiBase may vary over time and between countries**. The variations may be due to delays in reporting from national pharmacovigilance centres (often due to technical problems) or database maintenance at UMC. These are examples of over- and under-reporting to VigiBase, over time:
  - **2000-2001**: a backlog of reports from the US was entered
  - **2007**: a new report import process was developed, resulting in the majority of reports received in 2007 being imported in 2008
  - **2011**: reports that in previous years had been rejected because of quality issues were made accessible in the system
  - **2014**: a new report import process was implemented, enabling reports from China to be accessible in the system. A substantial backlog of French reports was imported.

Standardised terminology and text to include in manuscripts

The following standard terminology and text should be adhered to when preparing the manuscript for publication. For a further glossary of pharmacovigilance terms, see the UMC webpage.

The first time VigiBase is mentioned, it should be followed by "the WHO global database of individual case safety reports". Thereafter, 'VigiBase' can be used (note that the "B" in VigiBase should always be capitalised). An appropriate reference in connection to describing VigiBase is: Lindquist M.

VigiBase, the WHO Global ICSR Database System: Basic Facts. Drug Information Journal. 2008;42(5):409-19.(1), and the UMC webpage for more updated information.

The general description of VigiBase should include information on the total number of reports and the number of contributing countries. In addition, the data extract should be time-stamped, specifying the time of data retrieval and any other restrictions in time.

The number of reports should be given as the number of unique case reports.

It should be stated that the drugs recorded on the reports have been coded according to WHODrug and that MedDRA (indicating which version) was used to specify the adverse effects in your study.

The three statements in the UMC Caveat document must be included in the publication, specifying the source of information, the heterogenicity of data and that the information does not represent the opinion of the WHO (the last statement can be placed in the acknowledgements section).

If there is no UMC-affiliated author in the study, it should be made explicit that UMC has provided the data but that the study results and conclusions are those of the authors and not necessarily those of the Uppsala Monitoring Centre, National Centres, or WHO.

References

1. Lindquist M. VigiBase, the WHO Global ICSR Database System: Basic Facts. Drug Inf J. 2008;42(5):409-19.

2. FDA U.S. Food and Drug Administration. CFR - Code of Federal Regulations Title 21. 2016 [cited 2016 Jun 10]; Available from: http://www.ecfr.gov/cgi-bin/text-idx?SID=77821c5f7ae204cd5251ed82639d876a&mc=true&node=se21.5.314_180&rgn=div8.

3. Lindquist M. Data quality management in pharmacovigilance. Drug Saf. 2004;27(12):857-70.

4. Bergvall T, Norén GN, Lindquist M. vigiGrade: a tool to identify well-documented individual case reports and highlight systematic data quality issues. Drug Saf. 2014;37:65-77.

5. Norén GN, Orre R, Bate A, et al. Duplicate detection in adverse drug reaction surveillance. Data Min Knowl Discov 2007;14:305-28.

6. Tregunno PM, Fink DB, Fernandez-Fernandez C, et al. Performance of probabilistic method to detect duplicate individual case safety reports. Drug Saf 2014;37:249-58.