Technical description of vigiGrade™ Completeness score

Table of contents

List of abbreviations and definitions

Abbreviation or term	Explanation or definition
ADR	Adverse drug reaction
Dimension	Type of information accounted for in one or several text fields
Field score	Essential and valid information in one or several predefined text fields are rewarded the score"1". Reductions to the score are applied if information is missing or invalid.
ICD	International Statistical Classification of Diseases and Related Health Problems
ICH	The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use
ICH E2B	ICHs Efficacy guideline on data elements for transmission of ICSRs
ICSR	Individual case safety report
INDTIS	International Drug Information System, older reporting format versus E2B
ISO	International Standards Organization
lexicon	A group of terms related to electronic transmission of case reports and accepted by UMC to follow standard reporting.
MedDRA	Medical Dictionary for Regulatory Activities
numeric	Accepted numeric formats by UMC
One level penalty	The penalty is increased one level, i.e. from 0% to 10%, from 10% to 30% or from 30% to 50%.
Overlapping dates	If two dates are not specific enough to tell which of them occurred before the other the dates are said to be overlapping. E.g. “2014-10” and “2014-10-05”.
Text field	Data structure that holds a set of characters (marked out with "<" and ">" in this report)
Two level penalty	The penalty is increased two levels, i.e from 0% to 30%, from 10% to 50% or from 30% to 50%.
UMC	Uppsala Monitoring Center
VigiBase®	The WHO global ICSR database
vigiGrade™ Completeness score	A measure of the amount of clinically relevant information in an ICSR

Definition of vigiGrade™ Completeness score

The vigiGrade™ Completeness score ranges from 0.07 to 1 and is a measure of the amount of clinically relevant information in an ICSR as it appears in VigiBase®. A completeness score is calculated for each ICSR but is usually given as an average number for all ICSRs submitted from one country over time. The completeness score does not imply or reflect causality between a drug and an adverse event but focuses on information that is important to causality assessment (clinically relevant). A completeness score is calculated for an ICSR irrespectively of reporting format (INTDIS or ICH-E2B). ICSRs submitted in INDTIS format are made ICH-E2B compatible during the import process to VigiBase. In this technical description reporting format is referred to in line with ICH-E2B standard.

Field scores and calculation of completeness

Informative text on dimensions appearing in predefined text fields of the electronic report is rewarded a field score with penalties applied if information is missing or ambiguous. Dimensions accounted for in the vigiGrade Completeness score are presented in **Table 1**. Each text field where information is present receives the maximum field score 1. The more important the information is in the clinical assessment of a drug-reaction, the higher the penalty factor if information is missing. For example, completeness is reduced by 50% (multiplied by a factor 0.5) if "time-to-onset" is not available and by 30% (multiplied by a factor 0.7) if "age" is not specified (**Table 1**)

The completeness score is calculated from several field scores by a multiplicative model. Completeness is first computed for every reported drug-reaction pair appearing on the report (restricted to drugs listed as "suspected" or "interacting", hence excluding drugs listed as concomitant):

$$C = \prod_{i=1}^{10} (1 - P_i) = (1 - P_1) \cdot \ldots \cdot (1 - P_{10})$$

where "Pi" denotes the penalty (for penalties by dimension see **Table 1**) for the field score "i"(when no information is missing, the penalty is 0). Thus, the maximum completeness is 1 and the minimum completeness is: $0.5 \cdot 0.7^4 \cdot 0.9^5 = 0.07$.

The scores of all drug-reactions combinations on an ICSR are then aggregated to an average to yield a score for the corresponding ICSR:

$$C = \sum_{j=1}^{m} \frac{C^j}{m}$$

where "j" denotes the current drug-reaction combination and "m" denotes the total number of drug-reaction combinations on the report.

The multiplicative model how to calculate the vigiGrade Completeness score for an ICSR is illustrated in **Figure 1**.

Table 1 Dimensions accounted for in the vigiGrade Completeness score with penalties applied

Dimension	Description	Considerations	Penalty
Time-to-onset	Time from treatment start to the suspected ADR.	Imprecise information penalized if there is ambiguity as to whether the drug preceded the adverse event; with 30% if the uncertainty exceeds 1 month, 10% otherwise.	50% 30% 10%
Indication	Indication for treatment with the drug	Penalty imposed if information is missing or cannot be mapped to standard terminologies such as ICD or MedDRA.	30%
Outcome	Outcome of suspected ADR in the patient.	"Unknown" treated as missing.	30%
Sex	Patient sex.	"Unknown" treated as missing.	30%
Age	Patient's age at onset of the suspected ADR.	Age "unknown" treated as missing. 10% penalty imposed if only age group is specified.	30% 10%
Dose	Dose of the drug(s).	Penalty imposed if the total daily dose cannot be calculated from the included fields.	10%
Country	Country of origin.	Supportive in causality assessment since medical practice and adverse reaction reporting vary between countries.	10%
Primary reporter	Occupation of the person who reported the case (e.g. Physician, Pharmacist).	Supportive in causality assessment since the interpretation of reported information may differ depending on the reporter's qualifications "Unknown" penalized as missing information, whereas "Other" is not penalized.	10%
Report type	Type of report (e.g. spontaneous report, report from study, other).	"not available to sender (unknown)" treated as missing.	10%
Comments	Free text information.	Uninformative text snippets excluded.	10%

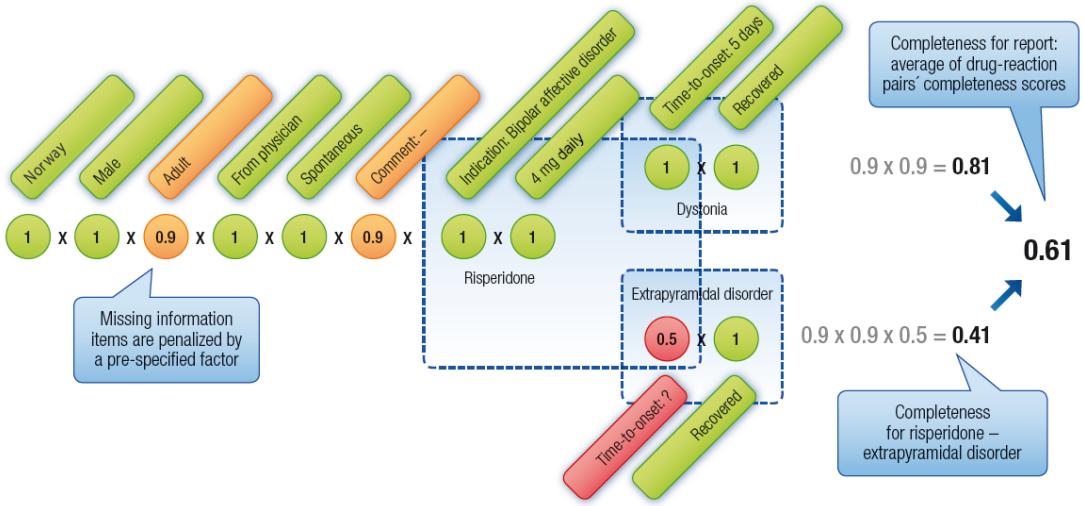

Figure 1 Example of how the vigiGrade Completeness score is calculated for an ICSR. Each text field where information is present is rewarded the field score "1". In this example there is no age specified, only age group "Adult", hence a penalty of 10% results instead of the full 30% penalty if "Age" is missing altogether. Which results in the field score 0.9 for the dimension "Age". In addition, all free text fields (dimension "Comments") are missing reducing the field score by 10% to 0.9. For the reaction-drug pair "Risperidone-Extrapyramidal Disorder" no information on the dimension "Time-to-onset" is given, hence reducing the field score for this combination by 50%, to 0.5. All field scores are multiplied on the reaction-drug combination-level, i.e. multiplying all field scores for the combination "Risperidone-Dystonia" gives the score 0.81 while the combination for "Risperidone-Extrapyramidal Disorder" results in 0.41. The mean of these two scores results is the vigiGrade Completeness score of the ICSR, which is 0.61.

Field score criteria presented by dimension

Time-to-onset

Time-to-onset is computed for every reported drug-reaction pair appearing on the report (restricted to drugs listed as "suspected" or "interacting", hence excluding drugs listed as concomitant)

Text fields:

Text fields applicable to dimension "Time-to-onset"	ICH-E2B data element
<ReactionFirstTime> *and* <ReactionFirstTimeUnit>	B.2.i.7.1a, B.2.i.7.1b
<DrugStartPeriod> *and* <DrugStartPeriodUnit>	B.4.k.13.1a, B.4.k.13.1b
<DrugStartDate> *and* <ReactionStartDate>, *or* <DrugStartDate> *and* <DrugEndDate>	B.4.k.12b, B.2.i.4b, B.4.k.14b

Field score:

A field score of 1 is generated if *one* of the following is true:

- <ReactionFirstTime> *and* <ReactionFirstTimeUnit> has valid values (according to lexicon and numeric) (see List of abbreviations and definitions, for a definition of lexicon)
- <DrugStartPeriod> *and* <DrugStartPeriodUnit> has valid values (according to lexicon and numeric)
- <DrugStartDate> *and* <ReactionStartDate> has valid values (correct dates) and no overlap of dates (for a definition of see List of abbreviations and definitions)

If none of the above criteria are fulfilled a score of 1 will be generated with the following reductions:

- Drug end penalty (one level penalty)(for a definition see List of abbreviations and definitions)
    - <ReactionStartDate> *and* <DrugEndDate> has valid values (correct dates)
- Incomplete date year penalty (two level penalty) (for a definition see List of abbreviations and definitions)
    - <DrugStartDate> *and*
      <ReactionStartDate> has valid values (correct dates) *and*
      both dates have the same year *and*
      one of the dates is incomplete but only includes year
    - <DrugStartDate>*and*
      <DrugEndDate> has valid values (correct dates) *and*

both dates have the same year *and*
one of the dates is incomplete but only includes year

- Incomplete date month penalty (one level penalty, cannot occur if combination already has an incomplete date year penalty)
    - <DrugStartDate> *and*
      <ReactionStartDate> has valid values (correct dates) *and*
      both dates have the same year and month *and*
      one of the dates is incomplete but includes year and month
    - <DrugStartDate> *and*
      <DrugEndDate> has valid values (correct dates) *and*
      both dates have the same year and month *and*
      one of the dates is incomplete but includes year and month

Note: it is enough to populate <ReactionFirstTime> and <ReactionFirstTimeUnit> for a reaction to generate a field score on all linked drug rows (not only the first suspected drug as stated in the ICH E2B guide).

Note: it is enough to populate <DrugStartPeriod> and <DrugStartPeriodUnit> for a drug to generate a field score on all linked reaction rows (not only the first reaction as stated in the ICH E2B guide).

Indication

Indication is computed for every reported drug appearing on the report (restricted to drugs listed as "suspected" or "interacting", hence excluding drugs listed as concomitant).

Text fields:

Text fields applicable to dimension "Indication"	ICH-E2B data element
<DrugIndication>	B.4.k.11b

Field score:

For each drug the score 1 is given if <DrugIndication> is valid. The score is reduced by 30% (multiplied by a factor 0.7) if information is missing or invalid. If more than one drug is reported, an average of all <DrugIndication> scores is calculated.

All drug indications reported according to ICD-8, ICD-9, ICD-10 or MedDRA are considered valid with exception for MedDRA code"10057097" ("Drug use for unknown indication") which is considered invalid.

Outcome

Text fields:

Text fields applicable to dimension "Outcome"	ICH-E2B data element
<ReactionOutcome>	B.2.i.8
<DrugRecurreAdministration>	B.4.k.17.1

Field score:

For each reaction a field score of 1 is generated if criteria a) *or* b) below is fulfilled. The score is reduced by 30% (multiplied by a factor 0.7) if information is missing or invalid.

a) <ReactionOutcome> has valid values (according to lexicon) except for "Unknown"

b) <DrugRecurreAdministration> has valid values (according to lexicon) except for "unknown" *and* the reaction(s) is listed in <DrugRecurAction> the score will be 1.

Sex

Text fields:

Text fields applicable to dimension "Sex"	ICH-E2B data element
<PatientSex>	B.1.5

Field score:

A valid value (according to lexicon, for a definition see List of abbreviations and definitions) in <PatientSex> will generate a field score of 1. The score is reduced by 30% (multiplied by a factor 0.7) if information is missing or invalid.

Age

Text fields:

Text fields applicable to dimension "Age"	ICH-E2B data element
<PatientOnsetAge> *and* <PatientOnsetAgeUnit>, *or* <PatientBirthDate> *and* <ReactionStartDate>, *or* <PatientAgeGroup>	B.1.2.2a, B.1.2.2b, B.1.2.2b, B.2.i.4b, B.1.2.3

Field score:

A field score of 1 is generated if criteria a) *or* b) below is fulfilled. If only criterion c) is fulfilled the score will be reduced by 10% to 0.9. The score is reduced by 30% (multiplied by a factor 0.7) if information is missing or invalid.

a) <PatientOnsetAge> and <PatientOnsetAgeUnit> have valid values (according to lexicon and numeric) and is between 0 and 134 years.
b) The difference in years between the earliest <ReactionStartDate> and <PatientBirthDate> is between 0 and 134. Negative years (<ReactionStartDate> before <PatientBirthDate>) are considered invalid.
c) <PatientAgeGroup> has a valid value according to lexicon.

Dose

Dose is computed for every reported drug appearing on the report (restricted to drugs listed as "suspected" or "interacting", hence excluding drugs listed as concomitant)

Text fields:

Text fields applicable to dimension "Dose"	ICH-E2B data element
<drugstructuredosagenumb> *and* drugstructuredosageunit>	B.4.k.5.1, B.4.k.5.2
<drugcumulativedosagenumb> *and* <drugcumulativedosageunit>	B.4.k.5.6, B.4.k.5.7
<drugseparatedosagenumb> *and* <drugintervaldosagedefinition>	B.4.k.5.3, B.4.k.5.4
<drugtreatmentduration> *and* <drugtreatmentdurationunit>	B.4.k.15a, B.4.k.1b
<drugstartdate> *and* <drugenddate>	B.4.k.12b, B.4.k.14b

Field score:

The dose score is dependent on 5 sub-scores and correlating field scores:

- **amount score:** if <drugstructuredosagenumb> *and* <drugstructuredosageunit> have valid values (according to lexicon and numeric) this score is 1 otherwise 0.
- **total amount score:** if <drugcumulativedosagenumb> *and* <drugcumulativedosageunit> have valid values (according to lexicon and numeric) this score is 1 otherwise 0.
- **frequency score:** if <drugseparatedosagenumb> *and* <drugintervaldosagedefinition> have valid values (according to lexicon and numeric)
- **dose duration score:** if <drugtreatmentduration> *and* <drugtreatmentdurationunit> have valid values (according to lexicon and numeric)

- **drug duration score:** <drugstartdate> *and* <drugenddate> have valid values (correct dates) this score is 1 otherwise 0.

The dose score will be 1 if *one* of the following is true:

Amount score and frequency score has a value of 1
E.g. "15 mg", "once per day" = 15 mg/day

Total amount score and dose duration score has a value of 1
E.g. "200 mg", "10 days" = 20 mg/day

Total amount score and drug duration score has a value of 1
E.g. "200 mg", "10 days" = 20 mg/day

The score is reduced by 10% (multiplied by a factor 0.9) if information is missing or invalid.

Country

Text fields:

Text fields applicable to dimension "Country"	ICH-E2B data element
<PrimarySourceCountry> *or* <ReporterCountry>	A.1.1, A.2.1.3

Field score:

A valid country code will generate a field score of 1. The score is reduced by 10% (multiplied by a factor 0.9) if information is missing or invalid.

The country code is considered valid if it is reported as an ISO2 code (lexicon listed), two character country code.

Primary reporter

Text fields:

Text fields applicable to dimension "Primary reporter"	ICH-E2B data element
<Qualification>	A.2.1.4

Field score:

A field score of 1 is generated if at least *one* valid value (according to lexicon) in <Qualification> is reported. The score is reduced by 10% (multiplied by a factor 0.9) if information is missing or invalid.

Report type

Text fields:

Text fields applicable to dimension "Report type"	ICH-E2B data element
<ReportType>	A.1.4

Field score:

A valid value in <ReportType> will generate a field score of 1. The score is reduced by 10% (multiplied by a factor 0.9) if information is missing or invalid.

Values in <ReportType> are valid if lexicon listed with the exception "not available to sender (unknown)"

Comments

Text fields:

Text fields applicable to dimension "Comments"	ICH-E2B data element
<DocumentList>	A.1.8.2
<ResultsTestsProcedures>	B.3.2
<PatientMedicalHistoryText>	B.1.7.2
<LiteratureReference>	A.2.2
<NarrativeIncludeClinical>	B.5.1
<ReporterComment>	B.5.2
<SenderDiagnosis>	B.5.3b
<SenderComment>	B.5.4

Field score:

A field score of 1 is generated if at least *one* of the fields listed in the table above ("Text fields applicable to dimension "Comments") contains a valid text. The score is reduced by 10% (multiplied by a factor 0.9) if information is missing or invalid.

A valid value is any text in the text fields except for:

- Equals "none"
- Equals "unknown"
- Equals "unk"
- Equals "no"
- Equals "none known"
- Equals "na"
- Equals "n/a"
- Equals "not reported"
- Equals "none provided"
- Equals "not provided."
- Equals "-"
- Only contains numbers, regardless of text length

Quality check of data: reaction start date consistency

Reaction start date consistency evaluates how well the reaction start date relates to other dates in the report. It is not part of the completeness score calculation, but still contributes with information on how complete the report is. The reaction start date consistency is a value of 1 or 0 and calculated per reaction.

The following rules must apply:

- <ReactionStartDate> has valid values (correct dates)
- <ReactionStartDate> must not be more than 135 years apart from <PatientBirthDate> (<ReactionStartDate> [Min] - <PatientBirthDate> [Max] <=135 years)
- <ReactionStartDate> (Max) >= <PatientBirthDate> (Min)
- <ReactionStartDate> (Min) <= <ReactionStopDate> (Max)
- <ReactionStartDate> (Min) <= <PatientDeathDate> (Max)
- <ReactionStartDate> (Min) <= UMC receive date
- Has a value in <ReactionStartDate> (Min) >= "1800-01-01"
- <ReactionStartDate> (Min) <= <ReceiveDate> (Max)
- <ReactionStartDate> (Min) <= <ReceiptDate> (Max)
- <ReactionStartDate> (Max) >= <DrugStartDate> (Min), at least one suspected or interacting drug)

If *any* of these checks fail the consistency gets a value of 0. If *all* checks fall out successfully the consistency gets a value of 1. The value is per reaction, i.e. a value is calculated for each reaction.

Missing dates are not included in the score. If a reaction start date is missing, the associated reaction start date consistency will receive no value (null and not 0!) and the reaction is marked as invalid (set to true). The same applies if no suspected or interacting drugs could be found. A rule that cannot be evaluated because of missing data is considered to be true.

Quality check of data: drug start date consistency

Drug start date consistency evaluates how well the drug start date relates to other dates in the report. It is not part of the completeness score calculation, but still contributes with information on how complete the report is. The drug start date consistency is defined to include suspected or interacting drugs and is a value of 1 or 0 calculated per drug.

The following rules must apply:

- <DrugStartDate> has valid values (correct dates)
- <DrugStartDate> must not be more than 135 years from <PatientBirthDate> (<DrugStartDate> [Min] - <PatientBirthDate> [Max] <=135 years)
- <DrugStartDate> (Max) >= <PatientBirthDate> (Min)
- <DrugStartDate> (Min) <= <DrugStopDate> (Max)
- <DrugStartDate> (Min) <= <PatientDeathDate> (Max)
- <DrugStartDate> (Min) <= UMC receive date
- Has a value in <DrugStartDate> (Min) >= "1800-01-01"
- <DrugStartDate> (Min) <= <ReceiveDate> (Max)
- <DrugStartDate> (Min) <= <ReceiptDate> (Max)
- <DrugStartDate> (Min) <= at least one <ReactionStartDate> (Max)
- <DrugStartDate> (Min) <= at least one <ReactionStopDate> (Max)

If *any* of these checks fail the consistency gets a value of 0. If *all* checks fall out successfully the consistency gets a value of 1. The value is per drug, i.e. a value is calculated for each drug.

Missing dates are not included in the score. If a drug start date is missing or no suspected or interacting drugs exist, the associated drug start date consistency will receive no value (null and not 0!) and the result will be marked as invalid (set to true). The same applies if no reaction start dates could be found. A rule that cannot be evaluated because of missing data is considered to be true.

Content and structure described in this report are based on:

ICH M2 EWG Electronic Transmission of Individual Case Safety Reports Message Specification, Document Version 2.3, November 9, 2000

ICH E2BM EWG Data Elements for Transmission of Individual Case Safety Reports, Revised E2B Step Four Document, Version 4.4.1

For further reading: 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 Jan; 37(1):65-77.