

The Forgotten Output - Data for ClinicalTrials.gov and EudraCT

PHUSE EU CONNECT 2023 - SM10

Erik Brun

I can't forget, but I don't remember what
Leonard Cohen

Agenda

- ClinicalTrials.gov & EUDRACT. The why, what, and when.
- Why this output surprisingly often is forgotten.
- What is the statistical programmer's role in this?
- Why does the results rarely match the study report exactly? A deeper dive into the quirks of the disclosure sites
- Our solution
- Conclusions

An abstract painting on the left side of the slide. It features a dense cluster of small, rounded shapes in various colors (red, yellow, blue, green) against a background of soft, blended colors like pink, orange, and blue. The overall effect is that of a crowd or a collection of many small elements.

Disclosure sites - The why, what, when and which

- The purpose of the disclosure sites are to provide information about clinical research studies to all interested parties
- The information includes amongst other:
 - The sponsor of the trial
 - The trial design
 - Disease(s) studied
 - Subject disposition
 - Baseline characteristics
 - Endpoints, including results
 - Tables of SAEs and frequent non-serious AEs.
- Typically information must be submitted to these sites 12 months after last participant last visit, but special rules apply for certain studies
- The two main disclosure sites are <https://eudract.ema.europa.eu/> and <https://clinicaltrials.gov/>. There are other such sites, but as I have no experience providing data for them, so they are beyond the scope of this talk.

Why is this output forgotten by almost everyone?

- Honestly, I don't know. I can just see that it is.
 - Here are some possible reasons:
 - The delivery is typically some time after the study report, when everybody have been allocated to new projects
 - It is not a CDISC or similar high status standard
 - It flies under the radar because it's a relatively small task
 - Lack of awareness in the organisation that it is a requirement
 - When a study is outsourced it can become the forgotten output as well.
 - Sometimes it's not in the workorder
 - Sometimes it's not delivered

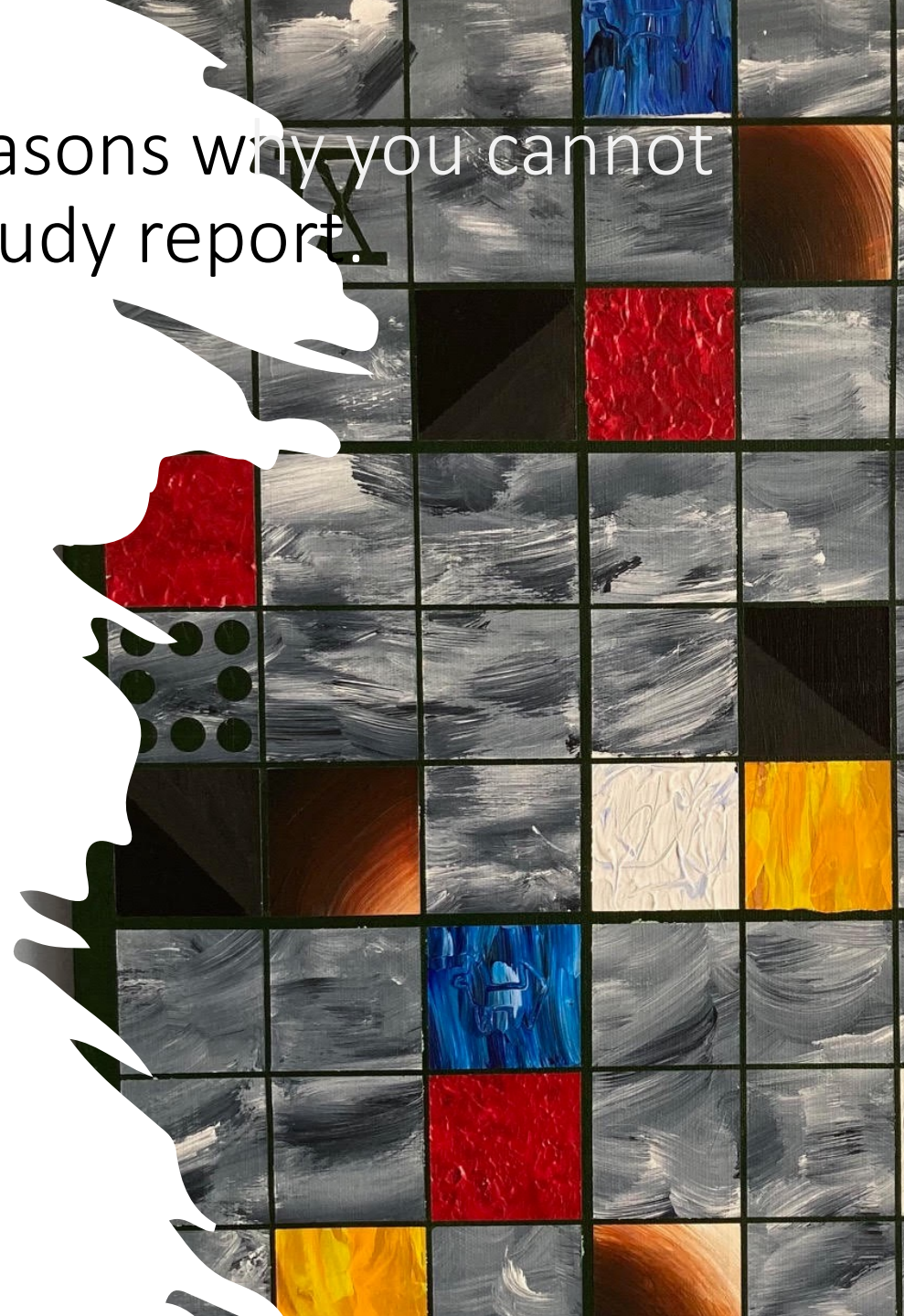


What is the statistical programmer's role in this?

- There are two ways to provide data the disclosure sites - Data entry or xml-upload.
- A combination is also possible. Use xml for tabular information, in particular AE tables. The study design and similar information will be entered by medical documentation
- The good news for the programmer is that the information needed for the disclosure sites is so similar that the data needed can be gathered in one go.
- The bad news is that tagsets etc. are not similar.

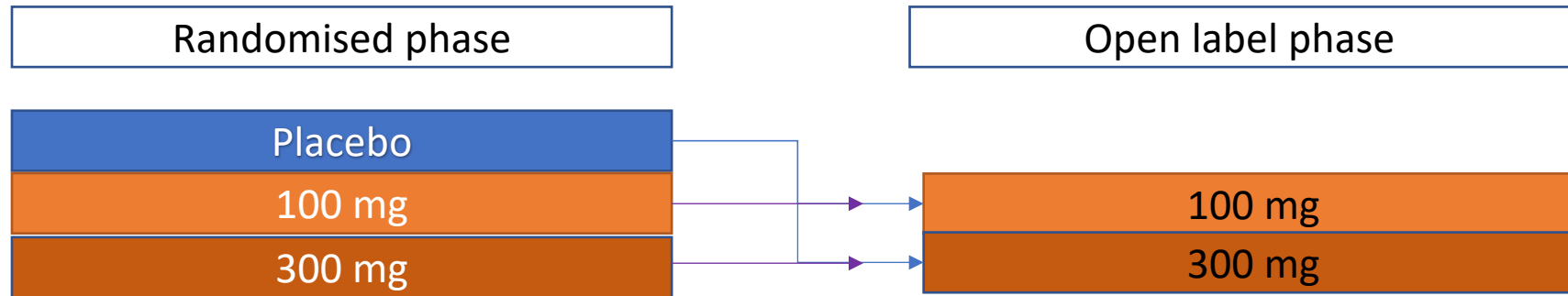
The quirks (by ascending quirkiness) – The reasons why you cannot just use the results from the clinical study report.

- Baseline characteristics
 - Study report – typically counts are based on treated patients
 - Disclosure – counts are based on randomised/enrolled patients
- The cut-off percentage for most frequent AEs must be an integer
 - Typically this is 5%, but for studies with drugs having a very good safety profile this number may be lowered, otherwise you might end up with a table only listing COVID-19
- Frequent AEs
 - Disclosure only counts **non-serious** treatment emergent AEs. Which means an event that occasionally is serious might dive under the cut-off percentage and thus being in the xml-file.
- EudraCT codes MeDDra coding
 - EudraCT for some reason applies 12-digit codes rather than the 8-digit MedDRA codes to the MedDRA terms.
- Sometimes treatment groups are pooled (in another way than the study report)
 - "Complex" studies. Both sites can only handle very simple studies (luckily they do it the same way).



A "complex" study design, and how it's presented at disclosure, an example

Actual study design:



At the disclosure sites the same study is presented as:

Random. Placebo
Random. 100 mg
Placebo-100 mg open label
100 mg-100 mg open label
Random. 300 mg
Placebo-300 mg open label
300mg-300 mg open label

Our solution – 6 quick steps

1) Clone the participants: ADAM.ADSL

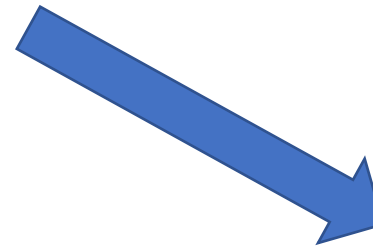
USUBJID	TRT01P	TRT02P
12345A-XX0001-1234	300 mg	300 mg
12345A-XX0001-5678	Placebo	100 mg



USUBJID	TRTP
12345A-XX0001-1234	Random. 300 mg
12345X-XX0001-1234	300mg-300 mg open label
12345A-XX0001-5678	Random. Placebo
12345X-XX0001-5678	Placebo-300 mg open label

2) Assign to "clone" ADAEE

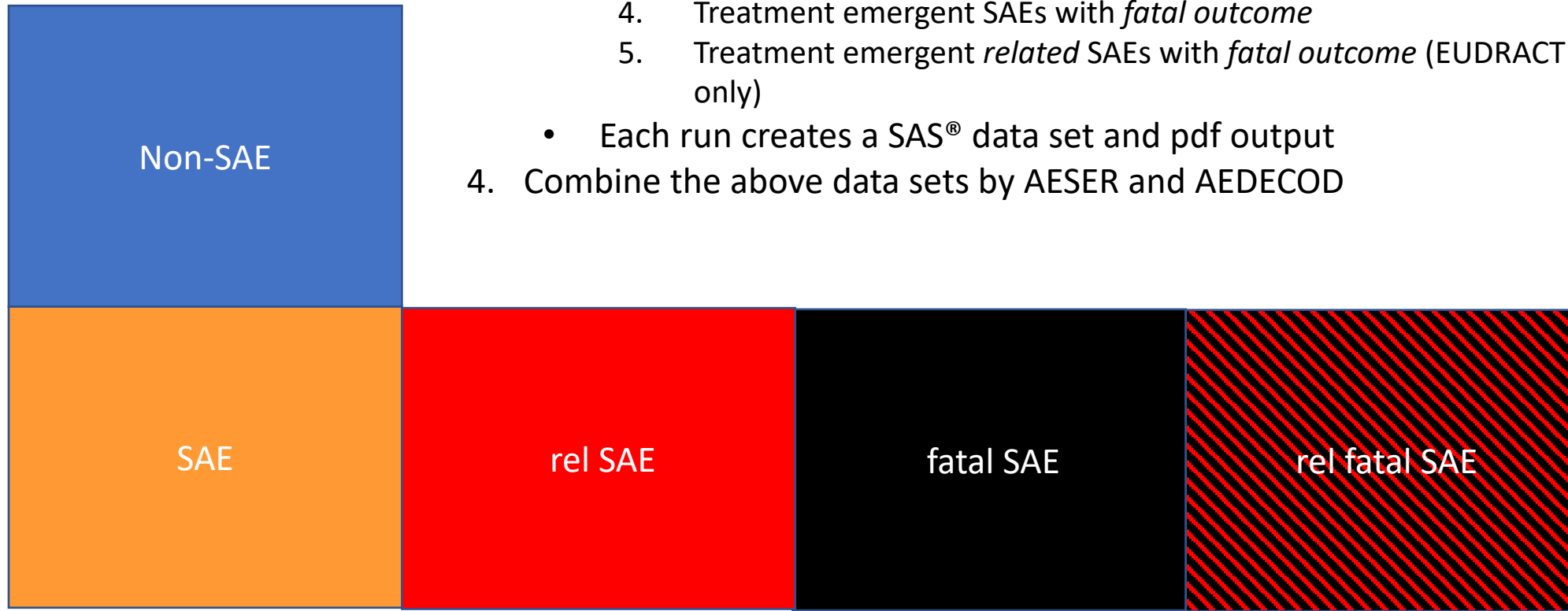
USUBJID	TRT01P	TRT02P	APHASE	AEDECOD, AREL, AOUT...
12345A-XX0001-1234	300 mg	300 mg	1	
12345A-XX0001-1234	300 mg	300 mg	2	
12345A-XX0001-1234	300 mg	300 mg	2	
12345A-XX0001-5678	Placebo	100 mg	1	
12345A-XX0001-5678	Placebo	100 mg	1	
12345A-XX0001-5678	Placebo	100 mg	2	
12345A-XX0001-5678	Placebo	100 mg	2	



USUBJID	TRTP	AEDECOD, AREL, AOUT...
12345A-XX0001-1234	Random. 300 mg	
12345X-XX0001-1234	300mg-300 mg open label	
12345X-XX0001-1234	300mg-300 mg open label	
12345A-XX0001-5678	Random. Placebo	
12345A-XX0001-5678	Random. Placebo	
12345X-XX0001-5678	Placebo-300 mg open label	
12345X-XX0001-5678	Placebo-300 mg open label	

Our solution – 6 quick steps cont'

3. Using the counting mechanism from our standard tlf-macros.
 - 5 runs with changing selection criteria
 1. Non-serious treatment emergent AEs (5% cut-off)
 2. Treatment emergent SAEs
 3. Treatment emergent *related* SAEs
 4. Treatment emergent SAEs with *fatal outcome*
 5. Treatment emergent *related* SAEs with *fatal outcome* (EUDRACT only)
 - Each run creates a SAS® data set and pdf output
4. Combine the above data sets by AESER and AEDECOD



Our solution – 6 quick steps cont'

5) proc put

```
*** Open the tags for NonSAE/SAE reporting ***;
*** Writing to XML - Header information ***;
put '<importRecordCollection>';
put @4 "<importRecord uniqueStudyId=\""&studyid"\">";
put @4 "<exResultsInfo>";
put @8 "<result edifactNumber=\""&sponsor"&".&eddec"&"/s"\">";
put @8 "<trialInformation>";
put @12 "<sponsorProtocolCode>&studyid</sponsorProtocolCode>";
put @8 "</trialInformation>";

put @8 "<adverseEvents>";

*** THRESHOLD ***;
put @12 "<nonSeriousEventFrequencyThreshold>&NonSerPct_</nonSeriousEventFrequencyThreshold>";

*** ASSESSMENT METHOD (static) ***;
put @12 "<assessmentMethod>";
put @16 "<value>ADV_EVT_ASSESS_TYPE.non_systematic</value>";
put @12 "</assessmentMethod>";

*** DICTIONARY ***;
put @12 "<dictionary>";
put @16 '<otherName xsi:nil="true" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"/>';
put @16 "<version>24.0</version>";
put @16 "<name>";
put @20 "<value>ADV_EVT_DICTIONARY_NAME.meddra</value>";
put @16 "</name>";
put @12 "</dictionary>";

*** TREATMENT GROUPS ***;
put @12 "<reportingGroups>";

do y=1 to &num_treats;

  put @16 '<reportingGroup id="ReportingGroup-' y +(-1) ' ">';
  ttt = trim(scan("&all_treatments_fmt", y, "#"));
  put @20 "<title>" TTT +(-1) "</title>";
  put @20 '<description xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:nil="true" />';

  put @20 "<subjectsAffectedByNonSeriousAdverseEvents>" nonser{y} +(-1) "</subjectsAffectedByNonSeriousAdverseEvents>";
  put @20 "<subjectsAffectedBySeriousAdverseEvents>" sae{y} +(-1) "</subjectsAffectedBySeriousAdverseEvents>";

  put @20 "<subjectsExposed>" denom{y} +(-1) "</subjectsExposed>";
  put @20 "<deathsAllCauses>" ftl{y} +(-1) "</deathsAllCauses>";
  put @20 "<deathsResultingFromAdverseEvents>" ftlr{y} +(-1) "</deathsResultingFromAdverseEvents>";

  put @16 '</reportingGroup>';
end;

put @12 "</reportingGroups>";
```

Be careful, when a uploaded xml-file is containing errors the disclosure sites return few and very uninformative error messages

Our solution – 6 quick steps cont'

6) Baseline characteristics – no xml expected (yet)

Gender, race number of participants per treatment

Mean age

Subjects per country

Subjects per EUDRACT defined
age group

Note: Count enrolled/randomized
participants, not treated

Population of trial subjects		
Number of subjects	In utero	0
	Preterm newborn - gestational age < 37 wk	0
Country: Number of	Newborns (0-27 days)	0
Country: Number of	Infants and toddlers (28 days-23 months)	0
Country: Number of	Children (2-11 years)	0

Age continuous			
Units: years			
arithmetic mean (standard deviation)	42.6 (12.14)	42.3 (11.44)	43.2 (11.19)
Gender categorical			
Units: Subjects			
Female	91	94	89
Male	144	138	141
Race			
Units: Subjects			
American Indian or Alaska Native	0	1	0
Asian	1	0	1
Native Hawaiian or Other Pacific Islander	0	0	0
Black or African American	36	33	34
White	176	179	178
More than one race	0	0	0
Unknown or Not Reported	22	19	17

xml output

```
<importRecord uniqueStudyId="18898A">
  <euResultsInfo>
    <result eudractNumber="" sponsor="H. Lundbeck A/S">
      <trialInformation>
        <sponsorProtocolCode>18898A</sponsorProtocolCode>
        </trialInformation>
      <adverseEvents>
        <nonSeriousEventFrequencyThreshold>
          <assessmentMethod>
            <value>ADV_EVT_ASSESS_TYPE.non_systematic</value>
          </assessmentMethod>
          <dictionary>
            <otherName xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" version="24.0">
              <name>
                <value>ADV_EVT_DICTIONARY_NON_SERIOUS</value>
              </name>
            </dictionary>
          </dictionary>
        <reportingGroups>
          <reportingGroup id="ReportingGroup-1">
            <title>EPTI 100 mg</title>
            <description xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
              <subjectsAffectedByNonSeriousAdverseEvents>0</subjectsAffectedByNonSeriousAdverseEvents>
              <subjectsAffectedBySeriousAdverseEvents>0</subjectsAffectedBySeriousAdverseEvents>
              <subjectsExposed>299</subjectsExposed>
              <deathsAllCauses>0</deathsAllCauses>
              <deathsResultingFromAdverseEvents>0</deathsResultingFromAdverseEvents>
            </reportingGroup>
          <reportingGroup id="ReportingGroup-2">
            <title>EPTI 300 mg</title>
            <description xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
              <subjectsAffectedByNonSeriousAdverseEvents>0</subjectsAffectedByNonSeriousAdverseEvents>
              <subjectsAffectedBySeriousAdverseEvents>0</subjectsAffectedBySeriousAdverseEvents>
              <subjectsExposed>294</subjectsExposed>
              <deathsAllCauses>0</deathsAllCauses>
              <deathsResultingFromAdverseEvents>0</deathsResultingFromAdverseEvents>
            </reportingGroup>
          <reportingGroup id="ReportingGroup-3">
            <title>EPTI 450 mg</title>
            <description xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
              <subjectsAffectedByNonSeriousAdverseEvents>0</subjectsAffectedByNonSeriousAdverseEvents>
              <subjectsAffectedBySeriousAdverseEvents>0</subjectsAffectedBySeriousAdverseEvents>
              <subjectsExposed>288</subjectsExposed>
              <deathsAllCauses>0</deathsAllCauses>
              <deathsResultingFromAdverseEvents>0</deathsResultingFromAdverseEvents>
            </reportingGroup>
          <reportingGroup id="ReportingGroup-4">
            <title>EPTI 600 mg</title>
            <description xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
              <subjectsAffectedByNonSeriousAdverseEvents>0</subjectsAffectedByNonSeriousAdverseEvents>
              <subjectsAffectedBySeriousAdverseEvents>0</subjectsAffectedBySeriousAdverseEvents>
              <subjectsExposed>294</subjectsExposed>
              <deathsAllCauses>0</deathsAllCauses>
              <deathsResultingFromAdverseEvents>0</deathsResultingFromAdverseEvents>
            </reportingGroup>
        </reportingGroups>
      </adverseEvents>
    </result>
  </euResultsInfo>
</importRecord>
```

```
<nonSeriousAdverseEvent>
  <term>Anaemia</term>
  <organSystem>
    <eutctId>100000004851</eutctId>
    <version>22</version>
  </organSystem>
  <assessmentMethod>
    <value>ADV_EVT_ASSESS_TYPE.non_systematic</value>
  </assessmentMethod>
  <dictionaryOverridden>false</dictionaryOverridden>
  <values>
    <value reportingGroupId="ReportingGroup-1">
      <occurrences>1</occurrences>
      <subjectsAffected>1</subjectsAffected>
      <subjectsExposed>299</subjectsExposed>
    </value>
    <value reportingGroupId="ReportingGroup-2">
      <occurrences>0</occurrences>
      <subjectsAffected>0</subjectsAffected>
      <subjectsExposed>294</subjectsExposed>
    </value>
    <value reportingGroupId="ReportingGroup-3">
      <occurrences>1</occurrences>
      <subjectsAffected>1</subjectsAffected>
      <subjectsExposed>298</subjectsExposed>
    </value>
    <value reportingGroupId="ReportingGroup-4">
      <occurrences>3</occurrences>
      <subjectsAffected>3</subjectsAffected>
      <subjectsExposed>288</subjectsExposed>
    </value>
    <value reportingGroupId="ReportingGroup-5">
      <occurrences>1</occurrences>
      <subjectsAffected>1</subjectsAffected>
      <subjectsExposed>145</subjectsExposed>
    </value>
  </values>
</nonSeriousAdverseEvent>
```

```
<seriousAdverseEvents>
  <seriousAdverseEvent>
    <term>Acute myocardial infarction</term>
    <organSystem>
      <eutctId>100000004849</eutctId>
      <version>22</version>
    </organSystem>
    <assessmentMethod>
      <value>ADV_EVT_ASSESS_TYPE.non_systematic</value>
    </assessmentMethod>
    <dictionaryOverridden>false</dictionaryOverridden>
    <values>
      <value reportingGroupId="ReportingGroup-1">
        <occurrences>0</occurrences>
        <subjectsAffected>0</subjectsAffected>
        <subjectsExposed>299</subjectsExposed>
        <occurrencesCausallyRelatedToTreatment>0</occurrencesCausallyRelatedToTreatment>
        <fatalities>
          <deaths>0</deaths>
          <deathsCausallyRelatedToTreatment>0</deathsCausallyRelatedToTreatment>
        </fatalities>
      </value>
      <value reportingGroupId="ReportingGroup-2">
        <occurrences>0</occurrences>
        <subjectsAffected>0</subjectsAffected>
        <subjectsExposed>294</subjectsExposed>
        <occurrencesCausallyRelatedToTreatment>0</occurrencesCausallyRelatedToTreatment>
        <fatalities>
          <deaths>0</deaths>
          <deathsCausallyRelatedToTreatment>0</deathsCausallyRelatedToTreatment>
        </fatalities>
      </value>
      <value reportingGroupId="ReportingGroup-3">
        <occurrences>0</occurrences>
        <subjectsAffected>0</subjectsAffected>
        <subjectsExposed>298</subjectsExposed>
        <occurrencesCausallyRelatedToTreatment>0</occurrencesCausallyRelatedToTreatment>
        <fatalities>
          <deaths>0</deaths>
          <deathsCausallyRelatedToTreatment>0</deathsCausallyRelatedToTreatment>
        </fatalities>
      </value>
      <value reportingGroupId="ReportingGroup-4">
        <occurrences>0</occurrences>
        <subjectsAffected>0</subjectsAffected>
        <subjectsExposed>294</subjectsExposed>
        <occurrencesCausallyRelatedToTreatment>0</occurrencesCausallyRelatedToTreatment>
        <fatalities>
          <deaths>0</deaths>
          <deathsCausallyRelatedToTreatment>0</deathsCausallyRelatedToTreatment>
        </fatalities>
      </value>
    </values>
  </seriousAdverseEvent>
</seriousAdverseEvents>
```



AE xml output as displayed in EuDRACT

Adverse events	Frequency threshold for reporting	Serious adverse events	Prospective Confirmation (PC) Period - Risperidone	PC period - Olanzapine	DBT Lu AF35700, 10 mg	DBT Lu AF35700, 20 mg	DBT period, continued treatment from PC period
Adverse events information	Non-serious adverse events						
Timeframe for reporting adverse events		Total subjects affected by serious adverse events					
Assessment type	Total subjects affected by serious adverse events	subjects affected / exposed	14 / 708 (1.98%)	7 / 384 (1.82%)	6 / 234 (2.56%)	5 / 232 (2.16%)	5 / 230 (2.17%)
Dictionary used for adverse events reporting	subjects affected / exposed	number of deaths (all causes)	0	0	0	0	1
Dictionary name	Investigations	number of deaths resulting from adverse events	0	0	0	0	1
Dictionary version	Weight increased	Investigations					
Reporting groups	subjects affected / exposed	Electrocardiogram abnormal					
Reporting group title	occurrences all number	subjects affected / exposed	0 / 708 (0.00%)	0 / 384 (0.00%)	0 / 234 (0.00%)	1 / 232 (0.43%)	0 / 230 (0.00%)
Reporting group description	Nervous system disorders	occurrences causally related to treatment / all	0 / 0	0 / 0	0 / 0	0 / 1	0 / 0
	Akathisia	deaths causally related to treatment / all	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
	subjects affected / exposed	Injury, poisoning and procedural complications					
Reporting group title	occurrences all number	Intentional overdose					
Reporting group description	Headache	subjects affected / exposed	0 / 708 (0.00%)	0 / 384 (0.00%)	0 / 234 (0.00%)	0 / 232 (0.00%)	1 / 230 (0.43%)
	subjects affected / exposed	occurrences causally related to treatment / all	0 / 0	0 / 0	0 / 0	0 / 0	1 / 1
Reporting group title	occurrences all number	deaths causally related to treatment / all	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
Reporting group description	Somnolence	Vascular disorders					
	subjects affected / exposed	Hypertensive emergency					
	occurrences all number						

Conclusions

- Our solution is a relatively simple and fast solution, but it works
- No, typically the results from the study report do not work, slightly different definitions, like SAEs are not considered a subset of AEs.
- In fact, the results will only match the study report when the study is simple and there are no SAEs.
- "Complex" trials requires a little creativity, because there are rather limited possibilities in these systems
- Communication, especially with medical documentation, is important, particularly whether they plan to pool treatment groups or when the study design is "complex".
- No, this output is not a big task, but ensure that time and resources are allocated. Make this output a part of the standards package, it will reduce the chance of it becoming the forgotten output.
- Don't be afraid to plunge into the unknown, you will learn something new, and when no-one else around you know about that subject you are suddenly the expert.

Thank you for your
attention!

Questions?

Erik Brun, erik@lundbeck.com

LinkedIn: <https://www.linkedin.com/in/erik-brun-149a884/>

All art used are my own works (inquiries welcome 😊)

