ANONYMISATION OF CLINICAL REPORTS BY CRO: IMPACT FROM THE GDPR

30 Nov 2017

Uwe W Fiedler, Chief Privacy Officer

DISCLAIMER: **The views expressed in the workshop slides are purely those of the presenter and may not in any circumstances be regarded as providing legal advice or stating an official position of PAREXEL International Corporation or the Association of Clinical Research Organizations (ACRO)**

PAREXEL®

YOUR JOURNEY. OUR MISSION.®

© 2017 PAREXEL INTERNATIONAL CORP.

AGENDA

- **CROs acting as data processors under GDPR**
- **Anonymisation of pseudonymized data = processing of personal data?**
  - Pseudonymized data = Personal Data?
  - Patient ID = directly or indirectly identifiable personal data?
- **Possible solution – Anonymisation Code of Conduct**
- **Appendix 1 - How personal are pseudonymized study data?**
- **Appendix 2 - Loss of pseudonymized study data = data breach under GDPR?**
- **Appendix 3 - References**

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

CRO ACTING AS DATA PROCESSOR UNDER GDPR

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

GDPR REQUIREMENTS FOR DATA PROCESSORS

- **Typically sponsor companies contract CROs as data processors**
- **The EU General Data Protection Regulation (GDPR) clarify:**
  - **Article 28 - Processor**

    3. Processing by a processor shall be governed by a contract or other legal act under Union or Member State law, that is binding on the processor with regard to the controller and that sets out the subject-matter and duration of the processing, the nature and purpose of the processing, the type of personal data and categories of data subjects and the obligations and rights of the controller. That contract or other legal act shall stipulate, in particular, that the processor:

       a) processes the personal data only **on documented instructions** from the controller, including with regard to transfers of personal data to a third country or an international organisation, unless required to do so by Union or Member State law to which the processor is subject; …;

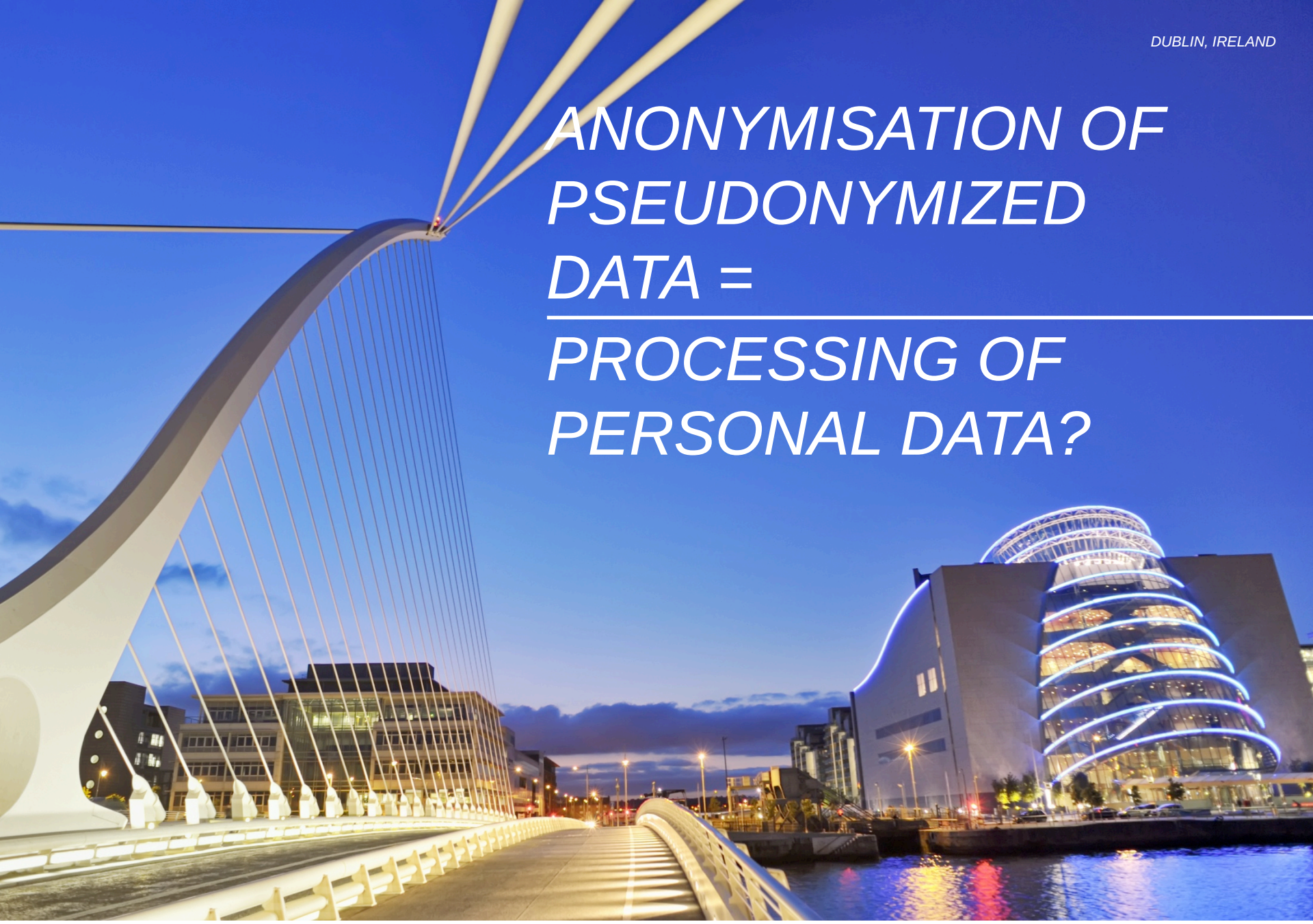
    10. Without prejudice to Articles 82, 83 and 84, if a processor infringes this Regulation by determining the purposes and means of processing, the processor shall be considered to be a controller in respect of that processing.

  - **Article 29 - Processing under the authority of the controller or processor**

    The processor and any person acting under the authority of the controller or of the processor, who has access to personal data, shall not process those data except on instructions from the controller, unless required to do so by Union or Member State law.

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL



ANONYMISATION OF PSEUDONYMIZED DATA = PROCESSING OF PERSONAL DATA?

PAREXEL

PSEUDONYMIZED DATA = PERSONAL DATA?

- **Art 4 GDPR clarify:**

(1) ‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

(5) ‘pseudonymisation’ means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person;

- **It therefore seems as the processing of pseudonymized data means the processing of indirectly identifiable personal data (please see Appendix 1 why this seems to be less clear in regard to pseudonymized study data)**
- **Consequently the data processor could only process pseudonymized data based on documented instructions from the data controller**

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

PATIENT ID = DIRECTLY OR INDIRECTLY IDENTIFIABLE PERSONAL DATA?

- **EMA guidance document define Patient IDs as direct identifiers:**

(Page 43) 5.3.2.1. Anonymisation of Direct Identifiers
Direct identifiers are elements that permit direct recognition or communication with the corresponding individuals. Direct identifiers generally do not have data utility, with the exception of the patient ID.

(Page 66) 1.2.2. Identification of data variables (direct and quasi identifiers)
- Describe direct and quasi identifiers in the clinical reports
  - Direct identifiers, e.g. patient ID
  - Indirect identifiers, e.g. age
- De-identification

Direct identifiers
  - Provide information on the redaction of direct identifiers, e.g. patient name, address if present in the reports
  - Regarding patient ID, provide information on whether it has been redacted or recoded and the resulting impact on the risk of re-identification

- **The Breyer decision of the European Court of Justice defined dynamic IP addressees as indirect identifiable personal data**

➢ **It seems to me as a Patient ID would therefore be an indirect identifier and not a direct identifier as the Patient Identification List is necessary to link the study subject name to the Patient ID**

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

ANONYMISATION OF PSEUDONYMIZED STUDY DATA = PROCESSING OF PERSONAL DATA?

- **The GDPR does not clarify if the anonymisation of personal data means the processing of personal data but defines, for example, the erasure or destruction of personal data as processing of personal data**
- **The EU Data Protection Authorities (WP 29) clarify in the Opinion 05/2014 on Anonymisation Techniques (WP 216 adopted 10 April 2014):**

… Anonymisation constitutes a further processing of personal data; as such, it must satisfy the requirement of compatibility by having regard to the legal grounds and circumstances of the further processing. …

2.2.1. Lawfulness of the Anonymisation Process

First, anonymisation is a technique applied to personal data in order to achieve irreversible de-identification. Therefore, the starting assumption is that the personal data must have been collected and processed in compliance with the applicable legislation on the retention of data in an identifiable format.

In this context, the anonymisation process, meaning the processing of such personal data to achieve their anonymisation, **is an instance of "further processing"**. As such, this processing must comply with the test of compatibility in accordance with the guidelines provided by the Working Party in its Opinion 03/2013 on purpose limitation. …

PAREXEL

ANONYMISATION OF PSEUDONYMIZED STUDY DATA = PROCESSING OF PERSONAL DATA?

- **Other EU member state Data Protection Authorities such as the Irish one share this view:**

  Anonymisation and pseudonymization - When can personal data be anonymised?

  The process of making data anonymous is itself considered to be "processing" data, so if an organisation wants to anonymise personal data to bring it outside of the scope of the Data Protection Acts, it must be done fairly, in accordance with the Acts. The conditions for fair processing of personal data are considered in our guidance note on using personal data, which should be consulted prior to any such processing.

- **Consequently the anonymization of personal data means the processing of personal data**

- **Legal grounds for CROs would be the documented instructions from the data controller**

PAREXEL

POSSIBLE SOLUTION

ANONYMISATION CODE OF CONDUCT

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

CODE OF CONDUCT FOR THE PUBLICATION OF ANONYMISED STUDY DATA

- **Taken the examples of Appendixes 1 and 2 into account, it seems to be less clearer that pseudonymized study data covered by ICH GCP confidentiality obligations and Declaration of Helsinki are equally personal as "ordinary" pseudonymized data.**
- **As the GDPR contains heavy sanctions, GDPR compliance risk assessments may come to the conclusion that it would be preferable to play it safe by just defining pseudonymized study as personal data**
- **This could limit the opportunity of data processors to support data controllers in regard to the anonymisation of study data as data processors would have to follow documented anonymisation instructions**
- **This could also affect the scientific usefulness of the anonymised study data as it may not be possible for CROs to share their data utility experiences with the sponsor without becoming a joint data controller for the anonymisation of pseudonymized study data**
- **Could an anonymisation code of conduct help?**

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

CODE OF CONDUCT FOR THE PUBLICATION OF ANONYMISED STUDY DATA

- **The GDPR states:**

Article 40 Codes of conduct

1. The Member States, the supervisory authorities, the Board and the Commission shall encourage the drawing up of codes of conduct intended to contribute to the proper application of this Regulation, taking account of the specific features of the various processing sectors and the specific needs of micro, small and medium-sized enterprises.

2. Associations and other bodies representing categories of controllers or processors may prepare codes of conduct, or amend or extend such codes, for the purpose of specifying the application of this Regulation, such as with regard to:

   (a) fair and transparent processing;

   (b) the legitimate interests pursued by controllers in specific contexts;

   (c) the collection of personal data;

   (d) the pseudonymisation of personal data;

   (e) the information provided to the public and to data subjects;

   (f) the exercise of the rights of data subjects;

CODE OF CONDUCT FOR THE PUBLICATION OF ANONYMISED STUDY DATA

- **As the anonymisation of study data would be a form of further processing of personal data such Code could outline the legitimate interests pursued by data controllers in the specific context to publish anonymised study data**
- **The Code could also clarify the preferred anonymisation method (for small and midsize companies a list of data elements that should be deleted - analog to the US federal HIPAA approach – could be helpful)**
- **The Code could also clarify the minimum necessary data elements that should be kept to achieve the comparable scientific usefulness obligation**
- **Participating companies could be both data controllers as well as data processors**
- **The Code could clarify that data processors providing consultancy services in regard to the anonymisation of study data won´t become “unauthorized” data controllers if anonymising data in accordance with the Code**

➢ **In sum, such Code could improve the legal certainty and help to better inform the public and in particular the affected study subjects!**

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

THANK YOU

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

APPENDIX 1

PSEUDONYMIZED STUDY DATA = PERSONAL DATA?

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

HOW PERSONAL ARE PSEUDONYMIZED STUDY DATA?

- **On 19 October 2016 the European Court of Justice ruled in the decision about dynamic IP addresses (also known as Breyer case (C-582/14)) that:**

44 The fact that the additional data necessary to identify the user of a website are held not by the online media services provider, but by that user's internet service provider does not appear to be such as to exclude that dynamic IP addresses registered by the online media services provider constitute personal data within the meaning of Article 2(a) of Directive 95/46.

45 However, it must be determined whether the possibility to combine a dynamic IP address with the additional data held by the internet service provider constitutes a means likely reasonably to be used to identify the data subject.

46 Thus, as the Advocate General stated essentially in point 68 of his Opinion, that would not be the case if the identification of the data subject was prohibited by law or practically impossible on account of the fact that it requires a disproportionate effort in terms of time, cost and man-power, so that the risk of identification appears in reality to be insignificant.

47 Although the referring court states in its order for reference that German law does not allow the internet service provider transmit directly to the online media services provider the additional data necessary for the identification of the data subject, seems however, subject to verifications to be made in that regard by the referring court that, in particular, in the event of cyber attacks legal channels exist **so that the online media services provider is able to contact the competent authority**, so that the latter can take the steps necessary to obtain that information from the internet service provider and to bring criminal proceedings.

PAREXEL

HOW PERSONAL ARE PSEUDONYMIZED STUDY DATA?

- **Art 40 (2a) of the German Federal Medicines Act clarify:**

(2a) The person concerned shall be informed of the purpose and scope of the collection and use of personal data, especially medical data. The person concerned shall be informed especially of the fact that:

1. where necessary, the collected data:
   a) will be **kept available for inspection** by the supervisory authority or the sponsor's representative in order to verify the proper conduct of the clinical trial,
   b) will be passed on in a pseudonymised version to the sponsor or to an agency commissioned by the latter for the purpose of scientific evaluation,
   c) will be passed on, in a pseudonymised version, to the applicant and the competent authority for the marketing authorisation if an application for a marketing authorisation is filed,
   d) will be passed on, in a pseudonymised version, to the sponsor and the competent authority and subsequently by the latter to the European database in the event of undesirable events in connection with the investigational medicinal product,

- **As the German Federal Medicines Act does not define the term "pseudonymized version", this definition follows the definition of the German Federal Data Protection Act and this definition is materially equal to the definition of pseudonymization under Art 4 (5) GDPR**

PAREXEL

HOW PERSONAL ARE PSEUDONYMIZED STUDY DATA?

- **Investigators as physicians are bound to professional secrecy obligations that won´t allow them to disclose information protected by professional secrecy obligations without being authorized by the patient for doing this**
- **The German Federal Medicines Act clarifies that monitors and inspectors can only access such information onsite but could not take such information away without pseudonymizing them beforehand**
- **It should therefore be possible to argue that in regard to clinical research activities regulated by the German Federal Medicines Act, neither the sponsor nor the CRO would have a legal ground that would allow them to re-identify study subjects**
- **In addition, it won´t be possible for the competent authority to re-identify study subjects on request from the sponsor as solely the Investigator maintains the Patient Identification List**

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

HOW PERSONAL ARE PSEUDONYMIZED STUDY DATA?

- **It seems as the EU Data Protection Authorities (WP 29) shared this view in their Opinion 4/2007 on the concept of personal data (WP 136 adopted on 20 June 2007) as WP 29 argued in regard to Safe Harbor:**

… The pharmaceutical company has construed the means for the processing, included the organisational measures and its relations with the researcher who holds the key in such a way that the identification of individuals is not only something that may happen, but rather as something that must happen under certain circumstances. The identification of patients is thus embedded in the purposes and the means of the processing. In this case, one can conclude that such key-coded data constitutes information relating to identifiable natural persons for all parties that might be involved in the possible identification and should be subject to the rules of data protection legislation.

The issue of key-coded data in pharmaceutical research has been addressed within the Safe Harbor Scheme. FAQ 14 - Pharmaceutical and Medical Products reads as follows:

7. Q: Invariably, research data are uniquely key-coded at their origin by the principal investigator so as not to reveal the identity of individual data subjects. Pharmaceutical companies sponsoring such research do not receive the key. The unique key code is held only by the researcher, so that he/she can identify the research subject under special circumstances (e.g. if follow-up medical attention is required). Does a transfer from the EU to the United States of data coded in this way constitute a transfer of personal data that is subject to the Safe Harbor Principles?

7. A: No. This would not constitute a transfer of personal data that would be subject to the Principles."

PAREXEL

HOW PERSONAL ARE PSEUDONYMIZED STUDY DATA?

The Working Party considers that this statement in the Safe Harbor scheme is not inconsistent with the reasoning explained above in favor of considering such information as personal data subject to the Directive.

Actually, this FAQ is not sufficiently precise as it does not state to whom and under what conditions the data is transferred.

The Working Party understands that the FAQ refers to the case where the key coded data is sent to a recipient in the US (for instance, the pharmaceutical company), which receives only key-coded data and will never be aware of the identity of the patients which is known and will be known in case of need for treatment only to the medical professional/researcher in the EU, but never to the company in the US.

- **It then seems as the WP 29 revised this understanding in their "Opinion 01/2016 on the EU – U.S. Privacy Shield draft adequacy decision" of April 13, 2016**

… 2.2.8 Pharmaceutical and medical products - Scope - The Privacy Shield considers that transfers of key-coded data from the European Union to the U.S. in the context of Pharmaceutical and Medical products do not constitute transfers that would be subject to the Privacy Shield (Annex II, III.14.g.). However, the transfer of key-coded data enjoys protection under European data protection law. This means that in practice the Privacy Shield cannot cover such transfers. The WP29 calls on the EU Commission to explicitly provide that the draft adequacy decision will not cover the transfer of key-coded data for pharmaceutical or medical reasons and as a consequence, such transfers must be covered by other safeguards, such as EU Standard Contractual Clauses or Binding Corporate Rules (BCRs). …

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

HOW PERSONAL ARE PSEUDONYMIZED STUDY DATA?

- **It then seems as the WP 29 revised this understanding in their "Opinion 01/2016 on the EU – U.S. Privacy Shield draft adequacy decision" of April 13, 2016**
- **In November 2012 the UK Data Protection Authority ICO stated in the "Anonymisation: managing data protection risk code of practice" that the disclosure of key-coded study data won´t mean the disclosure of personal data:**

Annex 2 – Anonymisation case-studies; Case study 1: limited access to pharmaceutical data.

In a clinical study, only key-coded data is reported by clinical investigators (healthcare professionals) to the pharmaceutical companies sponsoring the research. **No personal data is disclosed**. The decryption keys are held at study sites by the clinical investigators, who are prohibited under obligations of good clinical practice and professional confidentiality from revealing research subject identities. The sponsors of the research may share the key-coded data with affiliates overseas, scientific collaborators, and health regulatory authorities around the world. In all cases, however, recipients of the data are bound by obligations of confidentiality and restrictions on re-use and re-identification, whether imposed by contract or required by law. Given these safeguards, the risk of re-identification of the key-coded data disclosed by a pharmaceutical sponsor to a third party under such obligations is extremely low.

PAREXEL

APPENDIX 2 - LOSS OF PSEUDONYMIZED STUDY DATA

DATA BREACH UNDER GDPR?

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

DRAFT WP 29 GUIDELINE ON DATA BREACH REPORTING (WP 250)

- **On 03 October 2017 the EU Data Protection Authorities (WP 29) released proposed guidelines WP 250 “Guidelines on Personal data breach notification under Regulation 2016/679”.**

- **The proposed guideline is open to public comment until November 28, 2017.**

- **Excerpts from the draft guideline**

  3. The possible consequences of a personal data breach

  A breach can potentially have a range of significant adverse effects on individuals, which can result in physical, material, or non-material damage. The GDPR explains that this can include loss of control over their personal data, limitation of their rights, discrimination, identity theft or fraud, financial loss, unauthorised reversal of pseudonymisation, damage to reputation, and loss of confidentiality of personal data protected by professional secrecy. It can also include any other significant economic or social disadvantage to those individuals14.

  …

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

DRAFT WP 29 GUIDELINE ON DATA BREACH REPORTING (WP 250)

- **Excerpts from the draft guideline**

  Ease of identification of individuals

  An important factor to consider is how easy it will be for a party who has access to compromised personal data to identify specific individuals, or match the data with other information to identify individuals. Depending on the circumstances, identification could be possible directly from the personal data breached with no special research needed to discover the individual's identity, or it may be extremely difficult to match personal data to a particular individual, but it could still be possible under certain conditions. Identification may be directly or indirectly possible from the breached data, but it may also depend on the specific context of the breach, and public availability of related personal details. This may be more relevant for confidentiality and availability breaches. As stated above, personal data protected by an appropriate level of encryption will be unintelligible to unauthorised persons without the decryption key. Pseudonymisation, which is the process of de-identifying data so that a coded reference or pseudonym is attached to a record to allow the data to be associated with a particular individual without the individual being identified, can reduce the likelihood of individuals being identified in the event of a breach.

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

DRAFT WP 29 GUIDELINE ON DATA BREACH REPORTING (WP 250)

- **Excerpts from the draft guideline**

…

The European Union Agency for Network and Information Security (ENISA) has produced recommendations for a methodology of assessing the severity of a breach, which controllers and processors may find useful when designing their breach management response plan[31].

[31] ENISA, Recommendations for a methodology of the assessment of severity of personal data breaches, https://www.enisa.europa.eu/publications/dbn-severity

HOW IDENTIFIABLE ARE PSEUDONYMIZED STUDY DATA - ENISA GUIDELINE

- **ENISA guideline "Recommendations for a methodology of the assessment of severity of personal data breaches" of December 2013 - Annex 2 – Ease of identification (EI) scoring - Coding/Aliases**

  … Coding refers to the assignment of a unique ID number to each individual, e.g. in the context of a specific database. The use of aliases is a form of pseudonymisation, in the sense that a specific identifier (usually the individual's full name) is substituted by an alias (pseudonym). … Like in the case of unique identifiers, codes and aliases can be used to identify the individual as long as it is possible to link them to a reference database (e.g. linking the code/alias to the full name of a particular person).

  When identification is based on coding or use of aliases:

  - ***EI=0,25 (Negligible)** when the code/alias does not reveal and cannot be linked to any other personal data about the individual unless access to the reference database is obtained.* (e.g. pseudonamised study data processed by sponsor company)
  - ***EI=0,75 (Significant)** when the alias reveals some data about the individual (e.g. first name) and is linked to other personal data (e.g. the individual's email address).*
  - ***EI=1 (Maximum)** when the alias reveals the individual's full name or data from the reference database are also available.* (e.g. Patient Identification List processed by Investigator)

PAREXEL

HOW IDENTIFIABLE ARE PSEUDONYMISED STUDY DATA - ENISA GUIDELINE

- **Let`s use, as an example, a lost <u>unencrypted</u> laptop that contains copies of electronic Case Report Forms with pseudonymised study data.**
  - **If you then calculate the overall severity (SE) of the loss of key-coded study data as follows:**
    - Data Processing Context (DPC) = 4 because the data controller decided to define pseudonymised study data derived from medical records and containing Patient ID as equally sensitive to medical records containing full names of patients
    - Ease of identification (EI) = 0,25 because only the Investigator can lawfully link the Patient ID to the Patient Name
    - Circumstances of the breach (CB):
      - Loss of confidentiality = +0,5 as the lost data may have been disposed to an unknown number of recipients
      - Loss of integrity = 0 because there is no indication of incorrect or illegal use
      - Loss of availability = 0 because data being recoverable without any difficulty as the source data are still at the research site and the pseudonymised study data are already processed in the EDC system
      - Malicious intent =0 as loss happened by mistake (lost laptop)

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL

HOW IDENTIFIABLE ARE PSEUDONYMISED STUDY DATA - ENISA GUIDELINE

- **Calculation as follows: SE = DPC x EI + CB**
- **Then you would get a score of 4 x 0,25 + 0,5 = 1.5**
- **Taken the severity of a data breach (as outlined on page 6 of the ENISA guideline) into account you would end up with a low severity:**

Severity of a data breach		
SE < 2	Low	Individuals either will not be affected or may encounter a few inconveniences, which they will overcome without any problem (time spent re-entering information, annoyances, irritations, etc.).

- **It seems that based on the ENISA guideline, the loss of unencrypted pseudonymised study data won´t have significant adverse effects on individuals as long as the confidentiality of the Patient Identification List won´t be compromised by the incident.**

➢ **It therefore seems that pseudonymised study data could be closer to anonymized data as to personal data**

© 2017 PAREXEL INTERNATIONAL CORP.

TOKYO, JAPAN

APPENDIX 3 - REFERENCES

PAREXEL

REFERENCES

- **WP 29 Opinion 4/2007 on the concept of personal data (WP 136 adopted on 20th June 2017) - http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf**
- **WP 29 Opinion 05/2014 on Anonymisation Techniques (WP 216 adopted on 10 April 2014) - http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf**
- **WP 29 draft Guidelines on Personal data breach notification under Regulation 2016/679 (WP 250 Adopted on 3 October 2017) - http://ec.europa.eu/newsroom/document.cfm?doc_id=47741**
- **ENISA Recommendations for a methodology of the assessment of severity of personal data breaches (Working Document, v1.0, 20 December 2013) - https://www.enisa.europa.eu/publications/dbn-severity**
- **Irish Data Protection Commissioner - Anonymisation and pseudonymisation - https://www.dataprotection.ie/docs/Anonymisation-and-pseudonymisation/1594.htm**
- **UK Data Protection Commissioner - Anonymisation code of practice - https://ico.org.uk/for-organisations/guide-to-data-protection/anonymisation/**

© 2017 PAREXEL INTERNATIONAL CORP.

PAREXEL