QUALIANCE

Data Utility in Anonymized Clinical Study Reports (CSRs)

Technical Anonymisation Group (TAG) Meeting, November 2017, London

Jean-Marc Ferran (Qualiance) &
Sarah Nevitt (University of Liverpool)

Acknowledgement

- Thanks to the Researchers, Patient Organizations, Doctor Associations, Industry Representatives and Regulators, who engaged with us, for their time and invaluable input…

UNIVERSITY OF LIVERPOOL

QUALIANCE

Heard here and there…

- *”Only Pharmas access Policy 0070 documents to check what others are doing in terms of anonymization…”*
- *“Privacy is a secondary concern for Patients participating in Clinical Studies…”*
- *“Data must be available to ensure transparency of the information flow regardless of whether patients actually access the data themselves…”*
- *”Rare disease patients are concerned about Privacy as disease genetic roots may also affect their relatives…”*
- *”All data in CSR must be available in order to conduct efficient reviews”*
- *”Policy 0070 is not adapted for rare diseases”*
- *”It is more about transparency than utility…the submission documents are available to the public…”*
- *”What is Policy 0070, I cannot remember, we sent comments…”*

Data Utility & Policy 0070

- Policy 0070 "Phase 1" Guidance:
  - Data Controller must demonstrate that data utility has been considered and optimized
  - Data Utility is absent from section "3. Definitions"
  - Only reference to preserving results and conclusions

- A definition from the OECD*:
  - "A summary term describing the value of a given data release as an analytical resource. This comprises the **data's analytical completeness and its analytical validity**. Disclosure control methods usually have an adverse effect on data utility. Ideally, the goal of any disclosure control regime should be to maximise data utility whilst minimizing disclosure risk. **In practice disclosure control decisions are a trade-off between utility and disclosure risk**."

*: OECD: Organization for Economic Co-operation and Development

Insights from Policy 0043

Affiliation (per initial requests and appeals in 2016)

Affiliation	Number of requests received	In %	Number of pages released[6]	In %
Not-for-profit organisation	7	0.17	642	0.17
EU Institution (EC etc)	1	0.12	139	0.04
Regulator outside EU	2	0.24	0	0.00
EU NCA	4	0.49	103	0.03
Patients or Consumer	55	6.68	36388	9.55
Healthcare professional	24	2.92	16294	4.28
Academia/Research institute	66	8.02	120323	31.59
Legal	91	11.06	38463	10.10
Media	38	4.62	5960	1.56
Pharmaceutical industry	449	54.56	148013	38.86
Consultant	86	10.45	13044	3.42
Other	n/a	0.00	1542	0.40
Total	**823**	**100**	**380,911**	**100**

18.6%

45.4%

76.0%

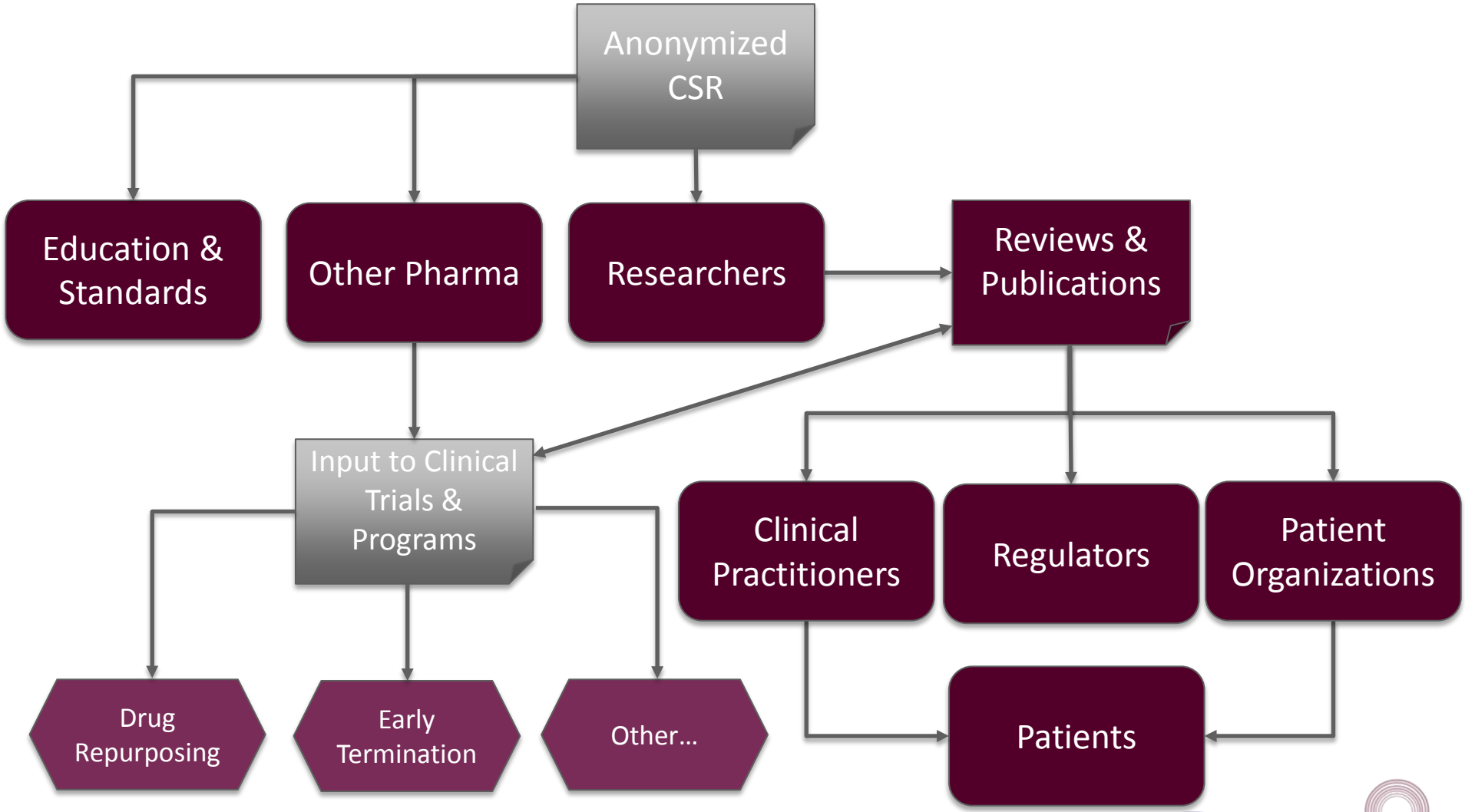
52,3%

Source: Annexes to the annual report of the European Medicines Agency 2016
http://www.ema.europa.eu/docs/en_GB/document_library/Annual_report/2017/05/WC500227333.pdf

UNIVERSITY OF LIVERPOOL

QUALIANCE

Data Consumers – Scenario 1
Simplified "Skill-based" Scenario



UNIVERSITY OF LIVERPOOL

QUALIANCE

Data Consumers - Scenario 2
Versatile use across Data Consumers

Anonymized CSR

- Researchers
- Other Pharma
- Clinical Practitioners
- Patient Organizations
- "Expert" Patients

- Better Understanding of Disease
- Reviews & Publications
- Better Understanding of Treatments
- Clinical Trial Participation
- Input to Clinical Trials & Program
- Input to HTA Work

- "Non-Expert" Patients
- "Clinical Trial" Patients
- "Expert" Patients

UNIVERSITY OF LIVERPOOL

QUALIANCE

Data Consumer Needs - Researchers

Awareness & Barriers

- **Survey conducted among Cochrane authors** between June and September 2016 on the theme "How is academia using regulatory data?" [2]
  - **156 Respondents**
    - **Only 10% have used or requested regulatory data**
      - **80% of these respondents believe regulatory data must be part of a review**
    - 5% considered using regulatory data
      - **32% of these respondents believe regulatory data must be part of a review**
    - 85% have not considered using regulatory data
      - **38% of these respondents believe regulatory data must be part of a review**
  - **32% of respondents had no understanding of the regulatory process**
  - **12% of authors knew where to access regulatory data**
  - **67% of respondents who accessed and included data in their reviews mentioned barriers** when using data:
    - Restricted and limited data
    - Time constraints
    - Lack of experience

UNIVERSITY OF LIVERPOOL

QUALIANCE

Data Consumer Needs - Researchers

Awareness & Barriers

- An *"Interim guidance on the inclusion of Clinical Study Reports and other regulatory documents in Cochrane Reviews"* is being developed.

- September 2016: we conducted interviews with researchers who have published academic work using CSRs and regulatory data
  - Rationale
  - Methods used
  - Potential impact of Policy 0070 on data utility

http://www.phusewiki.org/docs/Conference%202017%20DH%20Papers/DH04%20Paper%20NEW.pdf

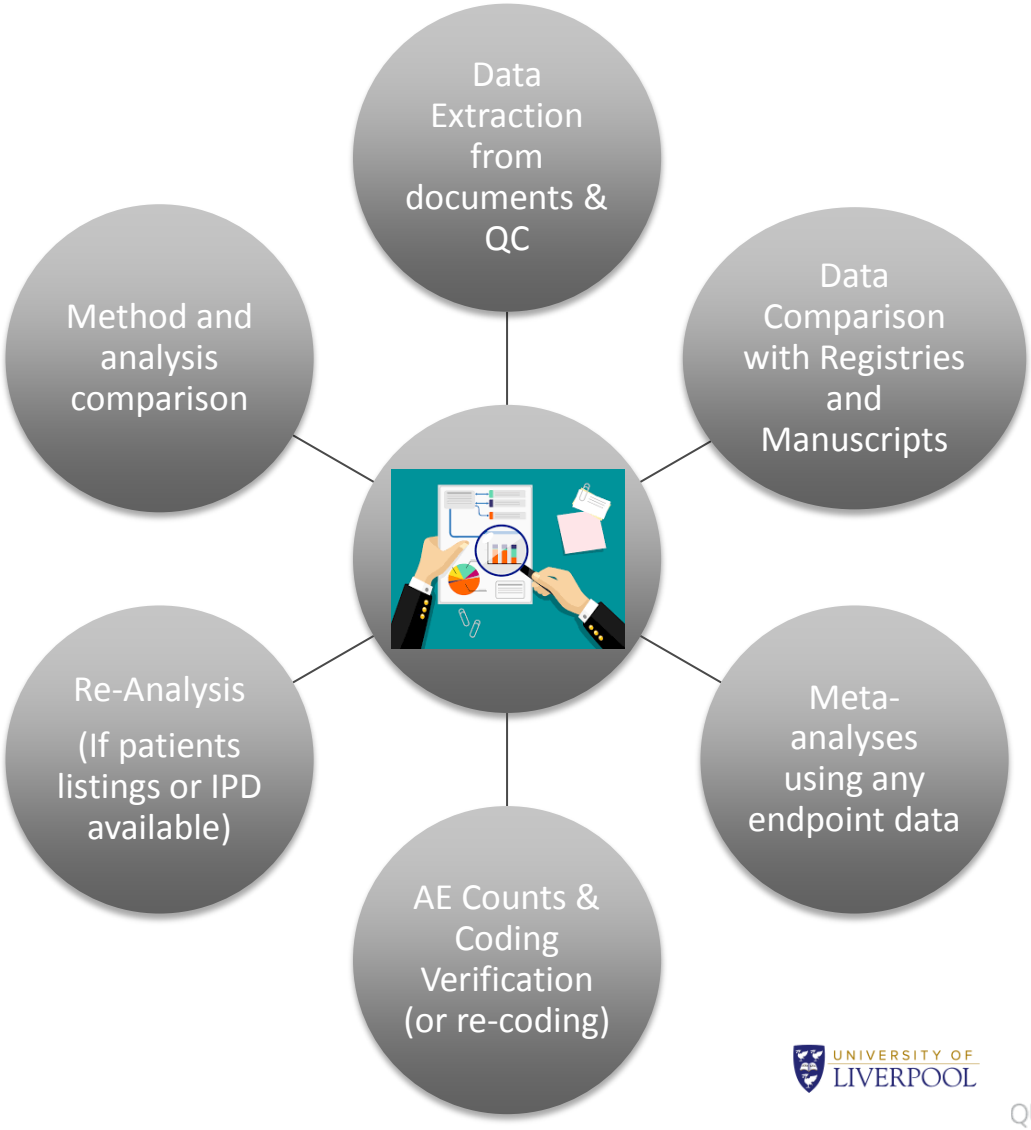
UNIVERSITY OF LIVERPOOL

QUALIANCE

Data Consumer Needs - Researchers

What Analyses & Tasks do researchers carry out using CSRs?

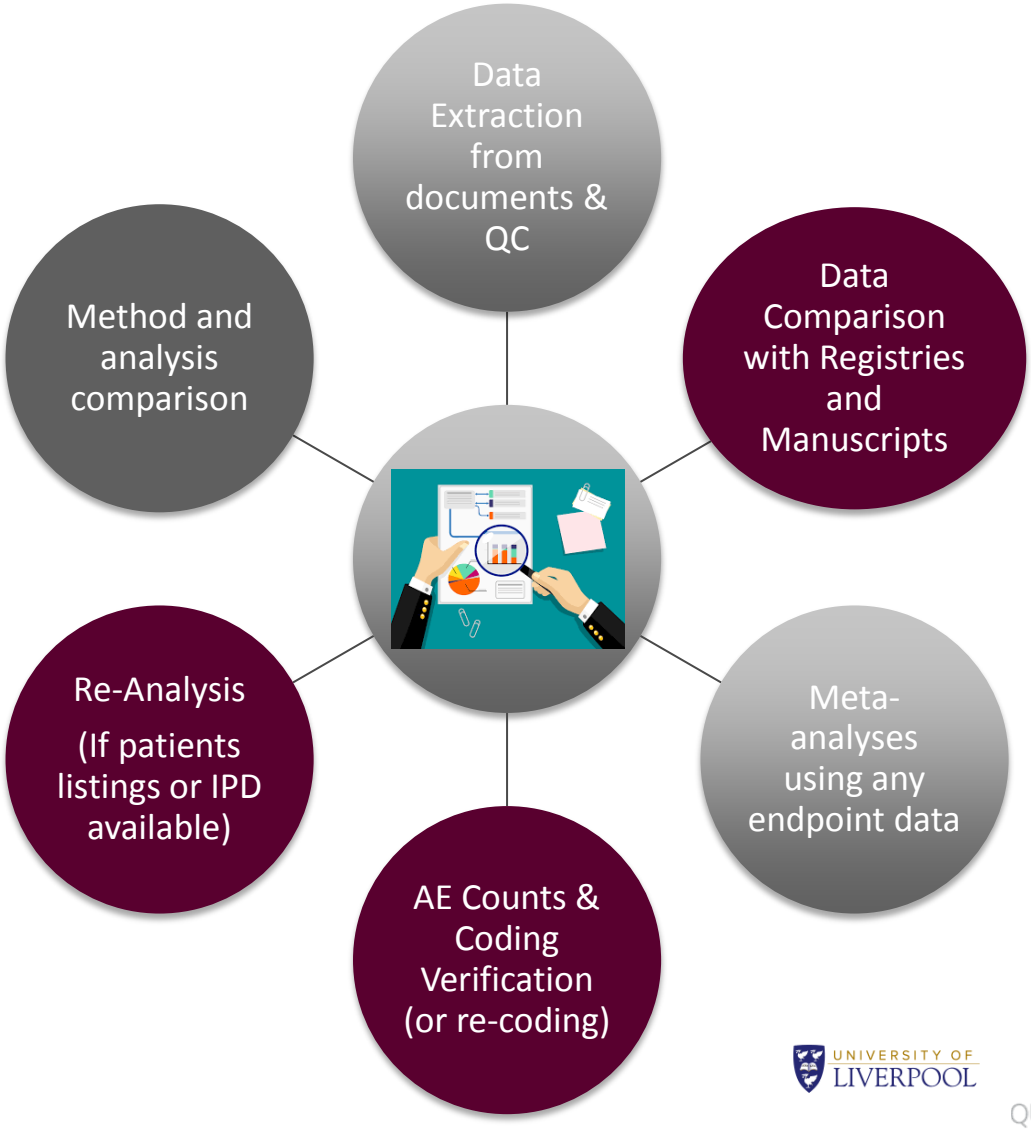
Purpose
Check for publication bias
Check for reporting bias
Systematic reviews
Novel analysis



Data Consumer Needs - Researchers

What Analyses & Tasks do researchers carry out using CSRs?

Purpose
Check for publication bias
Check for reporting bias
Systematic reviews
Novel analysis



Data Consumer Needs - Researchers

Researchers Interviews

- In summary, all of the authors stated that their analyses would not have been possible without access to CSRs.
- None of the authors raised any specific concerns about anonymized or redacted CSRs (in line with EMA policy 0070).
- All of the authors stated that some or all of their analyses or research would not have been possible if narratives and/or appendices (with participant listings) were removed from anonymised CSRs under EMA policy 0070.
- One author stated that: “Anonymized CSRs are ok, but the current EMA policy redacts important information about when the adverse events appeared as well as what they were…”

Data Consumer Needs - Researchers

Data utility in narratives

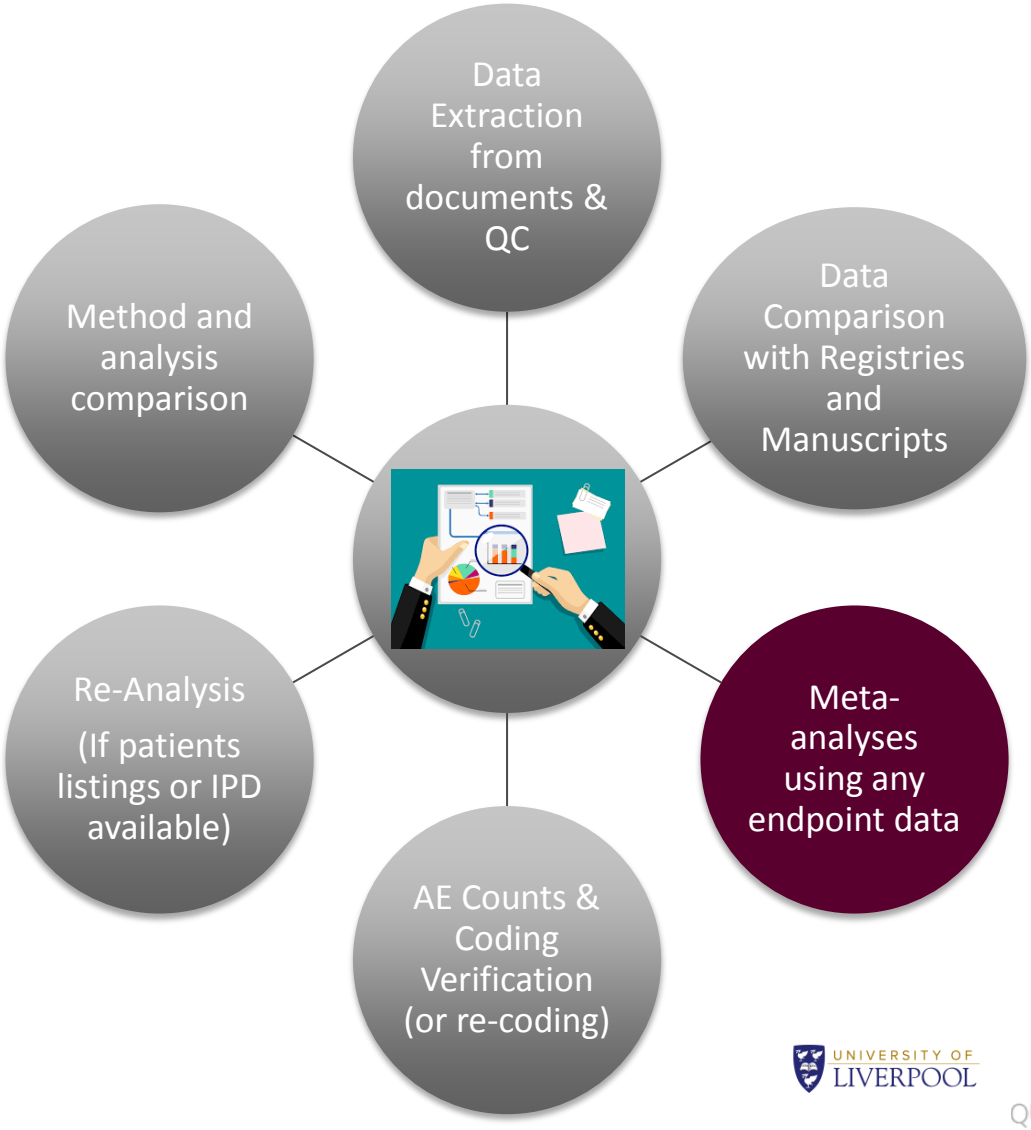
[4] Maund E et al. Coding of adverse events of suicidality in clinical study reports of duloxetine for the treatment of major depressive disorder: descriptive study (http://www.bmj.com/content/348/bmj.g3555)

- Comparing dictionary coded adverse events in summary tables to patient listings and narratives within CSRs
- Example: examination of CSRs of Duloxetine trials for events related to suicide
  - Coded events and narratives suggest different numbers of events
  - Coded events in summary tables may be misleading and not capture the true nature of the event
  - Different conclusions could be drawn
  - Authors suggest in this example, the narratives are more informative

Data Consumer Needs - Researchers

What Analyses & Tasks do researchers carry out using CSRs?

Purpose
Check for publication bias
Check for reporting bias
Systematic reviews
Novel analysis



UNIVERSITY OF LIVERPOOL

QUALIANCE

Data Consumer Needs - Researchers

Data utility in my personal research (Example 1)

[5] SJ Nolan et al: Lamotrigine versus carbamazepine monotherapy for epilepsy: an individual participant data review (https://www.ncbi.nlm.nih.gov/pubmed/27841445)

- Two first-line treatments for epilepsy
- Primary outcome: Time to Treatment Withdrawal
  - Reasons for withdrawal: treatment related or not?
  - Reasons for withdrawal: lack of efficacy? Adverse events? Other?
  - Coded data within IPD often not enough, free text 'verbatim' reasons for withdrawal often needed.
- Conclusions: Lamotrigine is superior in terms of withdrawals due to adverse events, carbamazepine is superior in terms of withdrawals due to lack of efficacy.
  - Allows a more personalised, patient-specific approach to treatment
  - Would not have been possible without IPD and free-text

Data Consumer Needs - Researchers

Data utility in my personal research (Example 2)

[6] SJ Nolan et al: Inhaled mannitol for cystic fibrosis (Cochrane Review) (https://www.ncbi.nlm.nih.gov/pubmed/26451533)

- Regulatory objectives versus Cochrane objectives
  - Demonstration of efficacy (e.g. lung function) versus meaningful outcomes for patients (e.g. quality of life)
- Published sources focus on efficacy (lung function), very limited information about Quality of Life
- Manufacturer of Mannitol provided additional unpublished summary data to allow detailed analyses of Quality of Life
  - Improvement in Quality of Life but increase in Burden of Treatment
  - Manufacturer helped with interpreting data, made extensive comments on our final report
  - Level of information required would be available in CSRs

Conclusions

- Data from Policy 0043 indicates that the **Researchers and the Pharmaceutical Industry** are likely to be the important data Consumers of Policy 0070 data.
- It cannot be excluded that **'expert'** patients may also access Policy 0070 data
- **Objectives and methods of researchers** accessing regulatory data have often related to assessment of reporting bias and re-analyses, but such data also has many benefits for novel analyses
- **More efforts on preserving Subject IDs and Dates in an anonymized format** should be considered to enable following patients through the narratives, any potential relationships as well as timing of events.
- **Reported terms (free-text)** and how such terms relate to coded terms should be considered for anonymization in the future.
- **Communication** between data consumers and data providers should be encouraged
- **PhUSE Data Transparency Working Group** is developing a **Data Utility Qualitative Scale** to evaluate Data Utility in Anonymized CSRs according to Data Consumers' needs.

References

[1]: **Overview of comments from stakeholders received as part of public consultation on EMA Policy 0070 of June 2013.** http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/general/general_content_000556.jsp&mid=WC0b01ac05809f363f

[2]: Hodkinson A, Dietz KC, Lefebrve C, Golder S, Jones M, Doshi P, Henegan C, Jefferson T, Boutron I, Stewart L. **The use of clinical study reports and other regulatory documents to enhance the quality of systematic reviews: a survey of systematic review authors** [submitted]. Summary available at: http://methods.cochrane.org/methodsinnovation-fund-2

[3]: Ferran JM, Nevitt S. **EMA Policy 0070: Data Utility in Anonymised Clinical Study Reports (CSRs).** Available at: http://www.phusewiki.org/docs/Conference%202017%20DH%20Papers/DH04%20Paper%20NEW.pdf

[4] Maund E, Tendal B, Hróbjartsson A, Lundh A, Gøtzsche PC. **Coding of adverse events of suicidality in clinical study reports of duloxetine for the treatment of major depressive disorder: descriptive study.** BMJ. 2014 Jun 4;348:g3555. http://www.bmj.com/content/348/bmj.g3555

[5] Nolan SJ, Tudur Smith C, Weston J, Marson AG. **Lamotrigine versus carbamazepine monotherapy for epilepsy: an individual participant data review.** Cochrane Database of Systematic Reviews 2016, Issue 11, Art No. CD001031. DOI: 10.1002/14651858.CD001031.pub3: https://www.ncbi.nlm.nih.gov/pubmed/27841445

[6] Nolan SJ, Thornton J, Murray CS, Dwyer T. **Inhaled mannitol for cystic fibrosis**. Cochrane Database of Systematic Reviews 2015, Issue 10. Art. No.: CD008649. DOI: 10.1002/14651858.CD008649.pub2. https://www.ncbi.nlm.nih.gov/pubmed/26451533

Thanks!

Jean-Marc Ferran

Consultant & Owner, Qualiance

Sarah Nevitt

Research Assistant, University of Liverpool

UNIVERSITY OF LIVERPOOL

QUALIANCE

Data Consumer Needs - Researchers

How to get data prior to Policy 0070?

Approach & Methodology

- **Questions:**
  - "Who are/will be the Data Consumers of Policy 0070 CSRs?"
  - "How can anonymized CSRs be used by legitimate Data Consumers?"
  - "What data entities are at stake and must be preserved in priority?"
  - "Is there a group of Data Consumers whose needs should be prioritized?"

- **Methodology:**
  - Review of academic literature based on industry CSRs
  - Review of Policy 0070 comments received during 2013 public consultation [1]
    - NOTE: "Publication and access to clinical-trial data" of 24. June 2013 has both CSRs and IPD in scope
  - Conduct of interviews across selected potential data consumers
  - Input from PhUSE Data Transparency WG

- **Out-of-Scope:**
  - Analysis of "non-legitimate" Data Consumers or "Plausible Attackers"

UNIVERSITY OF LIVERPOOL

QUALIANCE

Data Consumer Needs - Researchers

Sample of Articles based on Regulatory Data

Article	Data Source	Methods	Article Conclusion
Duloxetine - Maund 2014	9 CSRs inc. protocols as appendices obtained from the EMA. Journal articles Clinicaltrials.gov and Eli Lilly's online clinical trial registry were searched for trial results.	Data extraction, Data comparison, AE counts, Method comparison	"Clinical study reports contained extensive data on major harms that were unavailable in journal articles and in trial registry reports. There were inconsistencies between protocols and clinical study reports and within clinical study reports."
Gabapentin - Vedula 2009	20 CSRs available from Pfizer of which 12 were published in journals	Data extraction, Data comparison, Method comparison	"Selective outcome reporting for trials of off-label use of gabapentin. This practice threatens the validity of evidence for the effectiveness of off-label interventions. "
Reboxetine - Eyding 2010	Bibliographic databases, clinical trial registries, trial results databases, and regulatory authority websites up until February 2009, as well as unpublished data from the manufacturer of reboxetine (Pfizer)	Data extraction, Data comparison, Meta-analysis, Method comparison	"Reboxetine is, overall, an ineffective and potentially harmful antidepressant. Published evidence is affected by publication bias."
Lamotrigine versus carbamazepine – Nolan, 2016	We included 13 studies in this review. Individual participant data were available for 2572 participants out of 3394 eligible individuals from nine out of 13 trials: 78% of the potential data.	Meta-analysis	"Lamotrigine was significantly less likely to be withdrawn than carbamazepine but the results for time to first seizure suggested that carbamazepine may be superior in terms of seizure control. A choice between these first-line treatments must be made with careful consideration. We recommend that future trials should be designed to the highest quality possible with consideration of masking, choice of population, classification of seizure type, duration of follow-up, choice of outcomes and analysis, and presentation of results."
Tamiflu - Jefferson 2014	89 Clinical study reports from EMA and Roche (23 were used), trial registries, electronic databases, regulatory archives, and correspondence with manufacturers.	Data extraction, AE recoding, Meta-analysis, Method comparison	"The trade-off between benefits and harms should be borne in mind when making decisions to use oseltamivir for treatment, prophylaxis, or stockpiling. "
Paroxetine and Imipramine - Le Noury 2015	Reanalyse SmithKline Beecham's Study 329 using CSR publically available on the GSK's website and IPD on SAS CTDT including de-identified CRFs.	AE recoding, Re-analysis using IPD	"Neither paroxetine nor high dose imipramine showed efficacy for major depression in adolescents, and there was an increase in harms with both drugs. "
Orlistat - Schroll 2016	7 Publications and corresponding study CSRs provided by EMA	Data extraction, Data Comparison, AE counts	"We identified important disparities in the reporting of adverse events between protocols, clinical study reports, and published papers. Reports of these trials seemed to have systematically understated adverse events."

Data Consumer Needs
Patients & Patients Organizations

- **5 contacted:**
  - 3 Interviewed
    - EURORDIS - Rare Disease Europe
    - Genetic Alliance (UK)
    - BEUC – The European Consumer Organisation
  - EPC (European Patients Forum) answered that it was too early to provide input and would like to participate in the future
  - No answer from ECPC (European Cancer Patient Coalition)

- **UnderstandingPatientData.org.uk**
  - No current plan to develop guidance for patients to utilize Policy 0070 documents

UNIVERSITY OF LIVERPOOL

QUALIANCE

What can Patients Organizations do with Regulatory Data?

- **Use in HTA work**
- CSR data can be used to build stronger cases in HTA work and provide stronger arguments towards investment and research.

- **Understand better disease, treatments and provide input to clinical trials**
- **Placebo data** can be used to understand better the disease from this controlled and carefully monitored population.
- CSRs may help to inform better **Academic Strategic Clinical Trials** aiming at learning e.g. when to start and how to use better the available treatments.
- **Indirect comparison** of drugs could be supported
- CSRs could be used to support **Community Advisory Board** discussions.
  - Patient representatives interacting with sponsor on methods, ethics and logistics etc.

- **Drug Repurposing**
- In the case of drug repurposing (from e.g. a frequent to a rare disease area), having access to all previous data is of the utmost importance and would speed up the process.

Limitations of Policy 0070 expressed in the case of rare diseases

- In general, rare diseases require other policies than Policy 0070 to address earlier and faster access to data including access to IPD.
  - Limited available treatments for rare diseases – which clinical studies should patients join?
  - This would require phase I/II clinical studies CSRs to be available early while Policy 0070 enforces the publications of anonymized CSRs at the time of the drug submission and would not address this need.

Data Consumer Needs
Clinical Practitioners

- **The comments at the time were all supportive** of the policy and referred to recent examples in the media where initial data conclusion, and drug approvals, were overturned when additional data was further analysed.

- In recent follow up with a sample (n=5) of these practitioners they reflected on their reasons for submitting comments and that they were statements of support. **Furthermore, they confirmed that they have no direct interest in producing research and meta-analysis** to investigate data reports but rather **want reassurances that researchers are able to access reports** and data to continue to produce additional analyses.

Limitations

- The work presented is not based on surveys across a significant population of data consumers and only represent tendencies to explore further.
- Certain data consumers such as “Other Pharmas”, “Generics” or “Regulators” should be investigated further.
- This field is “merely a new born” and more experience from use and findings based on Policy 0070 data should be reviewed in the future.