Draft advice to the European Medicines Agency from the clinical trial advisory group on good analysis practice

29 January 2013

Introduction.

This is the fourth of five consultative groups related to the planned release of clinical trial data by EMA to third parties. The groups cover the following topics:

1. Protecting patient confidentiality
2. Clinical trials data formats
3. Rules of engagement
4. Good analysis practice
5. Legal aspects

The following report is made without individual attribution of opinions.

Protocols

Need for protocols

There was consensus that a formal protocol is desirable. Although reservations were expressed that intentional misrepresentation of the data cannot be wholly avoided the use of a protocol was generally seen as a tool that can facilitate the interpretation of the large majority of research and provide some defence against erroneous conclusions related to multiple analyses.

An important technical point was that writing a protocol requires a detailed knowledge of the data fields that are available but should be independent of the actual values observed. Hence, if the potential data recipients wished to follow best practice in an overtly verifiable way, the original protocols and data descriptions for the studies providing the datasets would have to be released to them prior to the datasets and a protocol finalised before the data were received.

Public access to protocols

A majority view was that protocols should be publically accessible. This was seen both as a way of confirming the pre-specification of hypotheses and of inviting constructive dialogue on the study design. However, some strong reservations were expressed based on the potential use of litigation by companies to prevent legitimate research.

Review of protocols

The people with most detailed knowledge of the clinical trial data will be those involved in their production. It was thus suggested that review of protocols by pharmaceutical company statisticians should be invited. Opinions for and against this suggestion were robustly expressed. A more moderate view was

also expressed by several people that review should be voluntary and that it would follow naturally if transparency was favoured.

Probably the strongest recommendation the EMA can make based on the views expressed is that researchers should seek opportunities to get informed review of their protocols. The point was made that transparency would be enhanced if all exchanges of views regarding the protocol were made public and suggestions for changes that were and were not adopted were recorded in an easily accessible for.

The point that central review of protocols (perhaps by EMA) might be desirable was made. It was acknowledged that this may not be possible with the current levels of resources available but it should be considered an aspiration.

A number of those present would have liked to see central review and approval as a prerequisite for data release. It was explained that this is not possible within the legal framework governing provision of the data.

Guidelines for analysis

The group discussed whether available guidelines on good practice in analysis and checklists for quality of analyses should be recommended for data recipients.

It was noted that such guidelines are already an accepted part of research and that appropriate use of them is expected by peer reviewed journals. Hence formal endorsement by EMA is unnecessary.

The comparative dearth of guidelines in secondary research was noted and EMA suggested that the ENCePP Code of Conduct and Guide on Methodological Standards in Pharmacoepidemiology might be worth considering. These will be circulated.

CONSORT fulfils a well-established role in research reporting and was mentioned in written comments from two members.

The general opinion was that EMA should note that researchers should be expected to be aware of relevant guidelines and apply them.

Open access to codes and interim datasets

It was explained that the intention of this point was to promote full transparency and verifiability of the analyses by ensuring that all datasets generated from the data supplied by the EMA and all computer code used to transform or analyse the data were made available. Given the appropriate computing environment, the whole analysis should be reproducible on the basis of these items.

There was strong support for this level of transparency and some additional points were made:

- The analysts might supplement the data from other sources. It would be important that these additional data were also revealed.

- Some analytical processes produce large datasets – for instance multiple imputation or bootstrapping. EMA opinion that all datasets required for replication of the analysis should be

made available. (Note: Repeated runs of appropriately large resampling or simulation exercises *should* give similar results but only the datasets as used will allow exact replication.)

A further point made with respect to transparency was that it would be desirable to have a list of all requests for data and names of the organisations/persons making the request. This falls more obviously within the remit of Rules of Engagement and it was noted that this point should be added to the agenda for Group 3.

ANNEX 1 List of attendees

European Medicines Agency
Hans-Georg Eichler – Medical director
Frank Petavy - Statistician
Jim Slattery – Chair
Urszula Piotrowska - Support

Other organisations

Roberto D'Amico	Academia
Kay Dickersin	Academia
Peter Doshi	Academia
Anthony Johnson	Academia
Kunal Merchant	Academia
Martin Posch	Academia
Karen Robinson	Academia
Jana Skoupa	Academia
Leslie Huson	Consultant
Paul Smith	Consultant
Alexis Clapin	Healthcare professional
Javier Garjón	Healthcare professional
Eugene Milne	Healthcare professional
Gisela Schott	Healthcare professional
Carla Souza	Healthcare professional
Aart Van der Molen	Healthcare professional
Roland Gordon-Beresford	Healthcare professionals' organisation
Christiane Abouzeid	Industry
Manfred Beleut	Industry
Helga Blasius	Industry
João Duarte	Industry
Eric Genevois-Marlin	Industry
Christoph Gerlinger	Industry
Merete Joergensen	Industry
Soeren Kristiansen	Industry
Hans-Juergen Lomp	Industry
Duncan McHale	Industry
Toby Lasserson	NGO
Adam Jacobs	Other/Unknown
Kieran Breen	Patients' organisation
Laila Abdel-Kader Martin	Payer / HTA
Ralf Bender	Payer / HTA
Mirjam Knol	Public health organisation
Regine Lehnert	Regulator
Mary Ann Slack	Regulator

Peter C Gøtzsche	Research institute
David Carroll	Student

Annex I - Comments from participants below may or may not have been made on behalf of the organisation they are affiliated with.

Comment form

Line Number	Comment and Changes proposed	Name	Affiliation
N/A	Comment received; awaiting permission to publish	Regine Lehnert	BfArM
N/A	Comment received; awaiting permission to publish	Manfred Beleut	Pareq Research AG
N/A	Comment Thank you very much for providing us with a summary of best practices for data analysis. I agree with you that analysis of any data should satisfy the requirements you mention. But I do not see why these requirements should be mandatory for studies performed on data obtained from EMA as they are not often applied for many published studies. I would like to split your points into three parts: • those which can limit or delay the possibility to obtain the data : point 1 and 2 • those dealing only with the quality of the study without any impact on the data availability: points 3 to 10 ; among these 8 points ○ Three points are dealing with transparency (points 3, 6 and 10)	Alexis Clapin	Neurologist, Paris

Line Number	Comment and Changes proposed	Name	Affiliation
	○ Other points are only dealing with quality of data analysis My proposal concerning the analysis of data obtained from EMA is: • Points 1, 2: Putting a protocol for review could be offered but should **not be mandatory** to get the data. If you decide to put a protocol, it should **not** be publicly available unless agreed by the requester. As a consequence, I consider that the name of the requester should not be made public. The requester could decide to make the protocol publicly available at any time, for example during the publication review process to demonstrate that analysis were planned (or not). These proposals are in line with what is done for other studies. • Points 3, 6, 10: All data or process allowing data checking should be made publicly available with the publication; this should be mandatory. For the point 3, it might be hard to implement such a project... I would consider logical to add in this category two other points: 1. Anonymous publication of results (on a website for example) should be forbidden. 2. If the results are published on a website out of the usual editorial process for scientific publication: discussion of the results should be made possible and points 6 and 10 should be satisfied. • Other points are "a plus" for publication; if you want to try publishing your study without using these rules: up to you! Requesting high standards for publication should be identical for all studies on data given by EMA or not. Of course, it would be beneficial for the requester to be informed of all good practices, rules...		

Line Number	Comment and Changes proposed	Name	Affiliation
	EMA reply: Thank you for taking the trouble to comment. I will incorporate your thoughts into the document but I attach some comments on your comments for your information. I see that you have signed up for Group 3 and I would ask you to address the question of whether names of requesters should be made public with that group as it falls in Prof Eichler's remit. Re: Points 1, 2: Putting a protocol for review could be offered but should **not be mandatory** to get the data. *This will definitely be the case* Re: the name of the requester should not be made public. *1. This falls within the remit of Group 3.* Re: Points 3, 6, 10: All data or process allowing data checking should be made publicly available with the publication; *Wholly agree* Re: this should be mandatory. *I do not think we have the legal authority to say this. It remains to be seen whether we can refuse further data to those who do not comply.* Re: Anonymous publication of results (on a website for example) should be forbidden. *'Strongly discouraged' is probably the limit of our powers*		

Line Number	Comment and Changes proposed	Name	Affiliation
	Re: If the results are published on a website out of the usual editorial process for scientific publication: discussion of the results should be made possible and points 6 and 10 should be satisfied. *We can maybe provide a forum but cannot force the owners of a website to accept outside comment.* Re: Other points are "a plus" for publication; if you want to try publishing your study without using these rules: up to you! Requesting high standards for publication should be identical for all studies on data given by EMA or not. Of course, it would be beneficial for the requester to be informed of all good practices, rules... *Yes!* Final comment: Ok for your comments to the two points I added to Sir Huson's points. I understand that it can't be mandatory. I will make the comment on requester 's name during CT3.		
N/A	Comment: As Gisella wrote, guidance for reporting systematic reviews and meta-analysis could be helpful. An extension of the PRISMA for systematic reviews of individual participant data (PRISMA-IPD) is going to be developed. http://www.surveymonkey.com/s/PRISMAIPDSURVEY	Javier Garjon	Servicio de Prestaciones Farmacéuticas

Line Number	Comment and Changes proposed	Name	Affiliation
	Meanwhile, the following resources are available: Chapter 19: Reviews of individual patient data. In: Higgins JPT, Green S (editors), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Stewart LA, Tierney JF, Clarke M. (Available from http://www.cochrane-handbook.org/) Meta-analysis of individual participant data: rationale, conduct, and reporting. Riley RD et al. BMJ 2010. (PMID: 20139215) EMA reply: This raises the question of whether we should single out particular guidelines as worthy of specific mention. The major point for doing so seems to me to be that nobody else has done so. At least, I am unaware of a complete 'Guide on Guidelines' although we included a fair few references in the ENCePP Methodological Standard – including a link to PRISMA. The cons are that (1) We would have to be selective as these documents are of variable quality (2) We would probably need to maintain and update it (3) It would have to be made clear that the agency does not necessarily endorse all the documents Final comment: Maybe, it would be practical to focus on the study types that most probably use the raw clinical trial data (systematic reviews of individual participant data, re-analysis of clinical trials, pharmacovigilance, modelling studies...) and try to identify widely adopted guidelines. Obviously, there can be innovative designs that are not covered by guidelines.		

Line Number	Comment and Changes proposed	Name	Affiliation
N/A	Comment received; awaiting permission to publish	Peter Gøtzsche	Nordic Cochrane Centre
N/A	Comment received; awaiting permission to publish	Mirjam Knol	RIVM, National Institute for Public Health and the Environment , Centre for Infectious Disease Control Netherlands
N/A	Comment received; awaiting permission to publish	Carla Sousa	European Association of Hospital Pharmacists
N/A	Comment: Roughly, the benefit of publication of clinical trial data at patient level can be classified in two types: (I) the opportunity for validation of the main results and (II) reuse of clinical trial data for secondary research. The recommendation of analysis standards needs to be tailored to the respective objectives of the analyses: I. **The opportunity for validating the main results and for investigating their robustness.** This would require access also to the clinical trial protocols (including amendments), statistical analysis plans together with software codes, data dictionaries and the study reports. For the validation purpose no prospective protocol seems to be necessary, as it will be guided by the original statistical analysis plan. However, the report on the validation should include all the necessary details to retrace the methodology applied. Standards should not be prohibitive and go beyond requirements for the	Martin Posch, Peter Bauer, Franz Konig	Medical University of Vienna

Line Number	Comment and Changes proposed	Name	Affiliation
	original applicant. Still, in the first place this validation remains the responsibility of regulatory authorities in time during the assessment procedure. II. **Reuse of clinical trial data for secondary research.** The scope of such research may range from quasi prospective research to full data mining, providing different levels of evidence. a) Higher levels of evidence will be provided by studies with a protocol written before access to the data. Meta-analysis based on individual patient data may be an example for such a higher level type of research. However, generally when planning secondary research projects, there will be study results already available, either from publications, reports or from other groups having access to the data, so that it will be difficult to exclude post hoc definitions of research objectives (e.g., resulting in hunting for significance). Early publication of protocols for secondary research, ideally before unblinded data of the phase III become available, would enhance credibility and persuasiveness of the planned secondary analyses. In any case, the protocol and resulting publications should clearly refer to time lines of data access and background knowledge available when formulating the research objectives - even though this will be difficult to verify independently. b) Full explorative discovery using data mining methodology provides a lower level of evidence but may reveal new and useful results. These results in general will have to be confirmed by further research and therefore such research projects have to be clearly identified as explorative when communicating results. The clinical trial protocol (including amendments) and the data dictionary should be accessible for all		

Line Number	Comment and Changes proposed	Name	Affiliation
	secondary research since it is hardly imaginable that anyone could otherwise perform and interpret analyses meaningfully. To avoid biased results due to incomplete access to data sets, administrative hurdles to get access to the data should be minimized. In general, it should be avoided that over-sophistication of protocol standards in secondary explorative research disguises its limited level of evidence. Therefore in all publications the nature of the research should be described clearly to assure an appropriate interpretation, for example by prominently indicating the source of the data and its secondary use. Overall, the quality of both, validation and secondary research undertakings, will depend on the availability of data sets containing original measurements (in contrast to heavily pre-processed analysis data sets). Proposed change (if any) For inclusion in document.		
N/A	Comment: The provision of a protocol template, including section headings corresponding to the recommended guidance documents, may help to ensure that the protocols follow the recommended guidance and provide sufficient detail. The goal of having prospectively defined analysis methods prior to seeing the data cannot be wholly achieved since the author of the protocol will have seen the results of the studies submitted for regulatory approval. As such, the protocol should include any results from the company analysis relevant to the specified analysis to identify what was known about the data prior to specifying the analysis.	Christiane Abouzeid	BioIndustry Association

Line Number	Comment and Changes proposed	Name	Affiliation
	EMA response: Changes made to document		
N/A	Comment Thanks for sending that. I think the report is a very fair summary of the meeting, and I have no new points to add to it. I do however, have just one comment on the report: in the section "public access to protocols", you say "some strong reservations were expressed based on the potential use of litigation by companies". If my memory serves correctly, that point was made by only one participant, and I don't think it was something that was seen by the group more generally as a legitimate concern. It's certainly my own strongly held opinion that transparency is important, and that making protocols publicly accessible should be a requirement for anyone wishing to receive data.	Adam Jacob	Dianthus Medical Limited