Ethics in clinical trial regulation: ethically relevant issues from EMA inspection reports

RDLC Bernabe[1,3.4], GJMW van Thiel[1], NS Breekveldt[2], CC Gispen-de Wied[2], JJM van Delden[1]

[1] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, the Netherlands
[2] Medicines Evaluation Board, the Netherlands
[3] Centre for Medical Ethics, Institute of Health and Society, Faculty of Medicine, University of Oslo, Norway
[4] Corresponding author, r_bernabe@yahoo.com

WORD COUNT: 4893 (including tables)

Abstract

Background

Within the EU, regulators are obliged to take ethical issues into consideration during marketing authorization deliberation. The goal of this manuscript is to identify what kinds of ethical issues regulators encounter during marketing authorization application deliberations, and the incidence of these ethical issues.

Methods

We used an EMA-provided Excel file that contains all the GCP non-compliance findings from all inspection reports from 2008-2012. There were 112 medicinal products and a total of 288 clinical trial sites. There were a total of 4,014 GCP non-compliance findings. We extracted the findings that were ethically relevant using NVivo 10.0 and created categories for the ethically relevant findings (ERFs). We took note of the incidence of ERFs for each category and extracted the inspectors' gradings of these findings. We also looked at the mean and the maximum number of ERFs per grading per medicinal product application, as well as the number of medicinal products with at least 1 ERF and those with at least major ERFs.

Results

With multiple coding, there were 1685 ERFs. ERFs were present in almost all of the medicinal products (97.3%). The majority of ERFs were graded major. At least major ERFs were present in almost all medicinal products with ERFs. The categories with the highest number of ERFs were protocol issues, patient safety, and professionalism issues. In terms of the density of combined critical and major findings, monitoring and oversight, protocol issues, and respect for persons top the list. We also showed that on average, there were 7.54 major and 2.95 critical ERFs per medicinal product application, though ERFs can increase to 30 major and 12 critical.

Conclusion

Regulators regularly encounter ERFs that at least “might adversely affect the rights, safety or well-being of the subjects”. It remains to be explored how regulators respond to these ethical issues.

Keywords: GCP non-compliance findings, clinical trials, ethically relevant findings, GCP inspection, marketing authorization application, research ethics

BACKGROUND

Within the EU, the ethical compliance of clinical trials submitted to the European Medicines Agency for marketing authorization application is a must. This is made explicit in the document, *Points to consider on GCP inspection findings and the benefit-risk balance*[1]*,* which says,

> The EU legislation requires not only valid clinical data for the scientific evaluation of the benefit-risk balance, but also ethical conduct of the clinical development programme in order to ensure that the rights, safety and well-being of the trial subjects are protected. GCP inspection findings – even if not directly influencing the benefit-risk balance - will still be important if they raise serious questions about the rights, safety and well-being of trial subjects and hence the overall ethical conduct of the study. It is an obligation of clinical assessors, rapporteurs and the CHMP also to assess the ethics of a clinical development programme, and major ethical flaws should have an impact on the final conclusions about approvability of an application. Consequently, ethical misconduct could result in rejection of the application.

The document specifically places an obligation to "clinical assessors, rapporteurs, and the CHMP"[1] to assess the ethical aspect of a marketing authorization application (MAA) and to incorporate this assessment in their deliberation processes. To date, we know much about the processes and issues of research ethics committees because of the rich literature about them, but little information is available in the public domain on the kind of ethical issues encountered by the regulators mentioned above and how the issues are evaluated and taken into consideration in MAA processes. Knowledge of the latter is indispensable to allow for future discussions not only on the handling of these ethical issues by the said regulators but on the improvement of the ethical aspects of clinical trial sites and the content of educational and other related preventive activities for the sponsors, who are primarily responsible for the ethical conduct and practices in their clinical trials. This information may also provide input for the discussions on the practical steps to be undertaken during the provision of guidance and advice in the medicine development phase and during the evaluation of Marketing Authorisation Applications as laid down in the *Reflection paper on ethical and GCP aspects* for the oversight and conduct of all clinical trials submitted to the EMA for MAA, whether these trials are within the EU/EEA or outside of it.[2]

This is the first of a series of manuscripts planned on this topic. In another manuscript, we shall look at how ethical issues brought up in inspection reports are handled by clinical assessors, rapporteurs, and the CHMP during marketing authorization deliberations. The modest goal of this manuscript is simply to identify the kinds and the incidence of ethical issues clinical assessors, rapporteurs, and the CHMP encounter in MAA. We chose to identify this data by reviewing GCP inspection reports. Inspection reports are the summaries of the GCP non-compliance findings of GCP inspectors after a "for cause" or routine inspection of a clinical trial site. Reports from different sites are usually integrated and submitted to the CHMP assessors for their consideration during the evaluation of a MAA or an application for extending the indication to an already authorized product[1]. Because these reports are dedicated to GCP non-compliance findings, and ethical issues form part of those findings, inspection reports are considered a valuable source of the ethical issues that regulators encounter during MAA

deliberations. For a complete account of the inspection process starting from CHMP request, please refer to Figure 1.

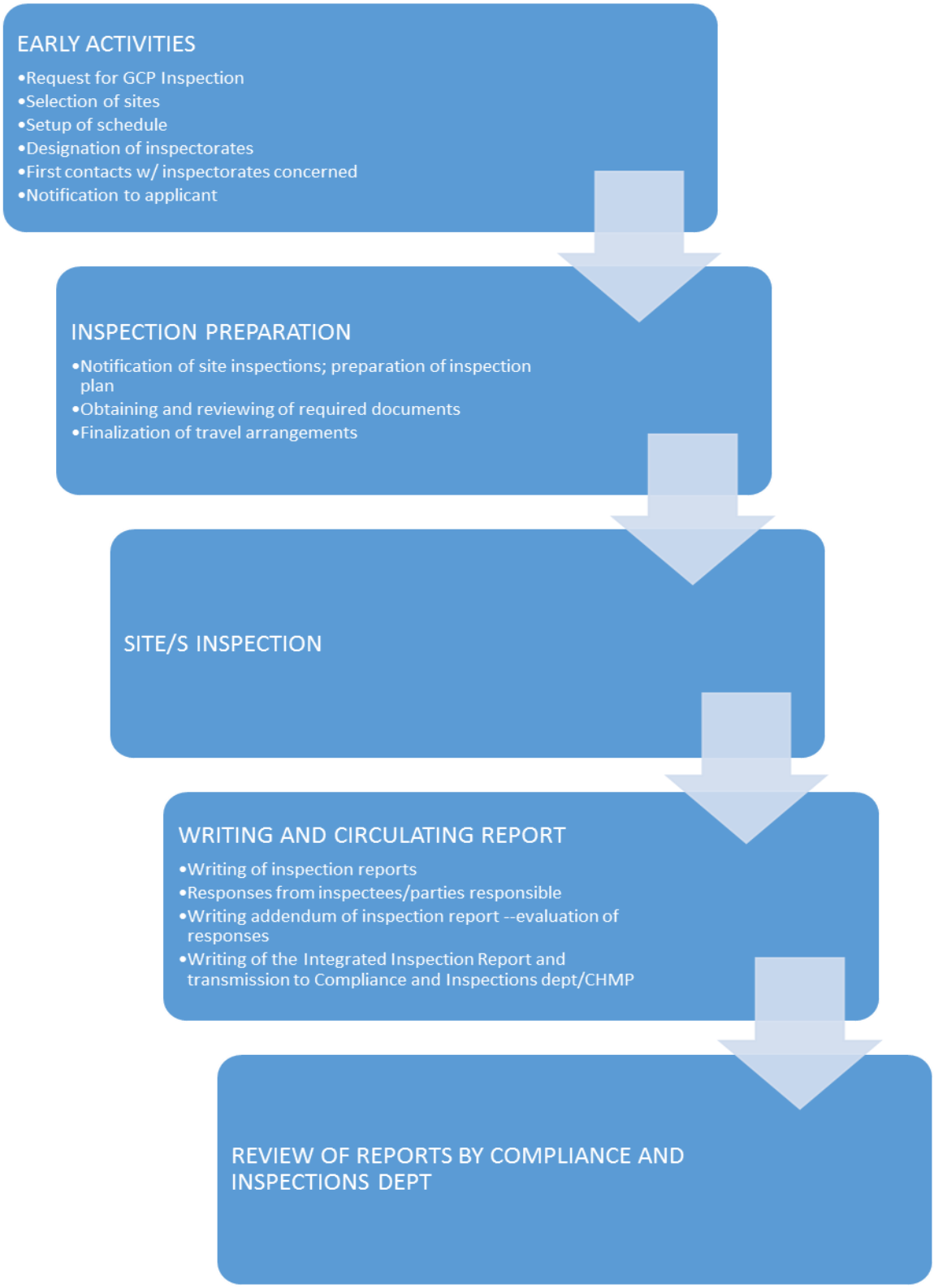

Figure 1: Process of inspection activities related to CHMP request[3,4i]

METHODOLOGY

GCP issues may also be ethical issues since GCP is an "international *ethical and scientific* quality standard for designing, recording, and reporting trials that involve the participation of human subjects[5]" (italics mine). Unfortunately, which of the GCP issues are ethical issues are not identified in a consistent manner in inspection reports. In this manuscript, we will identify the ethical issues regulators encounter by identifying which of the GCP non-compliance findings in inspection reports are also ethically relevant (henceforth, *ethically relevant findings, ERFs*).

We will also identify the incidence of these ERFs. We used an EMA-provided Excel file that contains all the GCP non-compliance findings from inspection reports from 2008-2012. This file contains the following information: request reference number, product invented name, site reference number, deficiency category name, deficiency subcategory name, deficiency grade, and deficiency description. There were 112 medicinal products and a total of 288 clinical trial sites. From all these sites, there were 4,014 GCP non-compliance findings.

Based on the ethical framework we earlier developed[6], each GCP non-compliance finding was reviewed and those ethically relevant were extracted from the excel file using NVivo 10.0. Then, using the principles of Grounded Theory[7], nodes were initially created through incident coding and then through focused coding. Incident coding refers to the initial process of extracting and comparing incidents with incidents to allow the properties and patterns of the data to emerge.[7] These properties and patterns allow the researcher to create clusters called incident nodes. In this study, ERF incident nodes were extracted, compared, and clustered. These initial nodes were then sifted through focused coding, i.e., the process of synthesizing larger segments of the data and finalizing the categories of these data completely and inclusively.[7] From these focused codes, using the language of the ethical framework, the final categories and sub-categories were defined. The final ERF categories were as follows: informed consent, monitoring and oversight, patient safety, professionalism and qualification issues, protocol compliance and protocol issues, research ethics committees, and respect for persons. Note that these categories are not mutually exclusive, i.e., a GCP non-compliance may be placed under two or more different categories, which would explain why the examples we provide under one category may also have been placed under other categories. Findings that could fall into one of the categories but were not ethically relevant in nature were weeded out. For example, issues such as the lack of signature in certain pages of the informed consent form were not included, unless the inspector flagged it as a relevant ethical issue such as the possibility of the study not being (properly) explained to the participant.

We also extracted the inspectors' gradings of the relevant GCP non-compliance findings, i.e., critical, major, or minor. *Critical/major/minor* findings are "conditions, practices or processes that *adversely/might adversely/would not be expected to adversely* affect the rights, safety or well-being of the subjects and/or the quality and integrity of data".[4] Note that the gradings of the inspectors were taken at face value, i.e., we did not think it was necessary to scrutinize the accuracy of the inspectors' gradings as our intention is merely the identification, as well as the incidence, of the ethical issues encountered by clinical assessors, rapporteurs, and the CHMP during marketing authorization deliberations, as handed to them by the inspectors.

Lastly, we looked at the mean and the maximum number of ERFs per grading per medicinal product application (henceforth, medicinal product). We also accounted for the number of medicinal products with at least 1 ERF and applications with only minor ERFs.

RESULTS

Ethically relevant GCP findings

Of the total 4014 GCP findings, without multiple coding, there were 1452 (36.2%) ERFs. Since a finding may be coded more than once, there were 1685 ERFs in total (see Table 1).

ETHICALLY RELEVANT FINDINGS	# of ERFs (total GCP findings =4014)	CRITICAL (% of critical findings to the total ERF per category)	MAJOR (% of major findings to the total ERF per category)	MINOR (% of minor findings to the total ERF per category)
Protocol compliance or protocol issues	484	54 (11.1%)	285 (58.9%)	145 (30%)
Patient safety	351	30 (8.5%)	172 (49%)	149 (42.4%)
Professionalism and/or qualification issues	265	10 (3.8%)	133 (50.2%)	122 (46%)
Research ethics committees	210	16 (7.6%)	108 (51.4%)	86 (40.9%)
Informed consent	189	19 (10%)	83 (43.9%)	87 (46%)
Monitoring and oversight	176	26 (14.8%)	112 (63.7%)	38 (21.6%)
Respect for persons	10	0	7 (70%)	3 (30%)
TOTAL	1685	155 (9.2%)	900 (53.4%)	630 (37.4%)

Table 1: ERFs among the GCP non-compliance findings

The majority of ERFs were GCP non-compliance findings that were graded as major findings (53.4%). The categories with the highest number of ERFs and the highest number of critical and major findings are protocol compliance and protocol issues, patient safety, and professionalism and qualification issues. In terms of the density of critical and major findings, i.e., the percentage of combined major and critical ERFs given the total number of ERFs per category, monitoring and oversight (78.5%) tops the list, followed by protocol compliance and protocol issues (70%) and respect for persons (70%). In the subsection below, we provide the explanation for each category, with incidence broken down to sub-categories, and examples.

ERFs, incidence, and examples

Protocol compliance and protocol issues

Most ERFs were issues on procedures that were noncompliant with the protocol such as the deviation of a procedure from protocol-specified inclusion/exclusion or withdrawal criteria (81.6%), while more than a tenth were issues on the protocol itself such as internal contradictions in the protocol or the protocol contains procedures not in accordance with guidelines (15.7%). Table 2 provides the ERFs under this

category and Table 3 provides examples of protocol compliance and protocol issues that were graded as critical by inspectors.

ERF sub-categories	n =484	% of total n
Protocol noncompliance	218	45.0%
• Inclusion, exclusion or withdrawal criteria	177	36.6%
Issues with the protocol	76	15.7%
Lack of or insufficient protocol-related corrective measures	11	2.3%
Delayed implementation of protocol amendment	2	0.4%

Table 2: Protocol compliance and protocol issues

ERF subcategories	Examples of critical issues
Noncompliance	"Subject was included in spite of a severe fracture after granting of a waiver by European Medical Monitor." "The trial management allowed sites to be persistently noncompliant with the protocol as can be seen from the vast amount of minor or major protocol deviations"
Noncompliance/Lack of corrective measures	"The investigator was not aware of the stipulated procedures of treatment of breakthrough pain with study medication during titration phase and efficacy phase…patients were left untreated with severe pain for further 40 minutes after the 1st puff of study medication was taken. Patients had more unsuccessful pain relief than necessary because they had to stay in one titration step for 4 BTP episodes instead of proceeding after 2 unsuccessful BTP episodes… Neither was documentation available which shows that… corrective actions taken."
Issues with the protocol	"Female subjects older than 50 years can fulfil the ESR inclusion criterion defined in the protocol having an ESR within the normal range of the analytical method. Consequently subjects can become eligible with normal laboratory results as they pass the ESR selection criteria that is below the upper normal range for their age"

Table 3: Examples of Protocol Compliance and Protocol ERFs graded as "critical" by the inspectors

Patient safety

The majority (52.7%) of the ERFs were on reporting or recording of adverse events (AE)/serious adverse events (SAE), which may refer to not following reporting/recording requirements by the investigator, reporting SAEs as AEs or vice-versa, or non- or late reporting of AEs/SAEs. There were also other issues

such as treatment allocation codes or blinding issues when treatment allocation codes were not submitted or were submitted late to the site or when blinding can only be broken following sponsor approval (11.7%); and issues on the lack, delay, or insufficiency in the performance of safety procedures or follow-up measures such as when SAE/AE are not evaluated or evaluated late by the investigator (8%). Table 4 provides the ERFs under this category and Table 5 gives examples of patient safety issues that were graded by inspectors as critical.

ERF sub-categories	n=351	% of n
AE/SAE reporting		
Non or late reporting or recording of SAE	61	17.4%
Non or late reporting or recording of AE	50	14.2%
Other issues with AE SAE reporting	74	21.1%
Treatment allocation codes and blinding issues	41	11.7%
Lack, delay, or insufficient performance of safety procedures or follow-up measures	28	8.0%
Non identification, late or non reporting of relevant events besides AE and SAE	22	6.3%
Privacy and confidentiality	22	6.3%
Insufficient or lack of procedures for patient safety	21	6.0%
Unreasonable delay or non-provision of lab results	13	3.7%
GP late or not informed	7	2.0%
Expired or wrong medications	6	1.7%
Lack of safety-related preventive or corrective measures	5	1.4%
Insurance issues	1	0.3%

Table 4: Issues on patient safety

ERF sub-categories	Examples of critical issues
Privacy and confidentiality	"Patient national identification number was sent to the sponsor as well as the CRO and sponsor requested and stored information about subject names signatures."
Non-identification, late, or nonreporting of relevant events beside AE/SAE	"In total 46 out of 108 notifications concerning bone loss were provided by CRO to the investigator more than 58 days after receipt date. Some notifications were provided more than 1 year after."
Other issues with AE/SAE reporting	"Reporting requirements for AE/SAE were not followed by investigator."
lack of safety-related preventive or corrective measures; other issues with AE/SAE reporting	"The ongoing review of the safety data by an independent advisory committee as laid down by the sponsor in the approved clinical trial protocol was not performed. Additionally it seems as if such a review was not planned even in the beginning of the clinical trial as the clinical database was setup in a way that made it impossible to provide the sponsor with AE listings on an ongoing basis."

Treatment allocation codes and blinding issues	"numerous subjects were unblinded inappropriately"

Table 5. Examples of Patient Safety ERFs graded as "critical" by the inspectors

Professionalism and/or qualification issues

Almost a third of the ERFs were on the non-verifiable qualification of the investigator and/or trial staff (31.3%), while more than a quarter were issues on unofficial or not qualified investigators or trial staff members performing certain duties such as administering the informed consent or the assessment of SAEs/AEs (26.4%). Almost a quarter were issues pertaining to clinical trial investigators or staff members not adequately trained or informed for the trial (23%). Table 6 provides the ERFs under this category and Table 7 provides examples of professionalism and qualification issues that were graded by inspectors as critical.

ERF sub-categories	n=265	% of n
Unofficial or not qualified trial staff	47	17.7%
• Informed consent procedure not by qualified or delegated personnel	19	7.2%
• SAE AE assessment done by non-qualified staff	4	1.5%
Qualification not verifiable	83	31.3%
Staff not adequately trained or informed for the trial	61	23.0%
Negligence	24	9.1%
Dishonesty falsification or morally suspect behavior	15	5.7%
Insufficient or no trainings available or planned	8	3.0%
Delinquency	3	1.1%
Conflict of interest	1	0.4%

Table 6: Professionalism and/or qualification issues

ERF sub-categories	Examples of critical issues
Negligence	"The investigator did not ensure the accuracy and completeness of trail documents. Diaries were filled in by investigator or subjects relatives."
Unofficial or not qualified staff	"Obtaining informed consent from subjects and reporting of SAE were delegated to nonmedical staff."
Falsification or tampering	"The source data of one patient is not authentic and reliable. For another patient, there is strong suspicion that the source data are not authentic. Because of the seriousness of the observation the authenticity of the source data of all patients included in the inspected clinical trials at site 51 is questioned. The finding affects the rights of the patients and the quality and integrity of the data generated at this investigational site."

Table 7: Examples of Professionalism and Qualification ERFs graded as "critical" by the inspectors

Research ethics committees (RECs)

Almost half (46.2%) of the ERFs on this category were on the use of protocols that lacked approval or were approved late by the RECs while a tenth (10.5%) were on the delayed reporting of compulsory updates to the RECs. Closely nearby at 9.5% were issues on REC procedures or composition, such as the lack of a lay member or that the REC does not comply with GCP. Table 8 provides the ERFs under this category, while Table 9 provides examples of REC issues that were graded by inspectors as critical.

ERF sub-categories	n=210	% of 210
Protocol (amendment) without or delayed approval	97	46.2%
Lack of or delayed reporting (of updates)	22	10.5%
Problems with REC procedure or composition	20	9.5%
Lack of or delayed approval on other documents	19	9.0%
Missing, late or no approval on IC (or amendment)	17	8.1%
Procedure prior to approval	15	7.1%
Conflict of interest of REC members	7	3.3%
Procedures or amendments requested or demanded by REC or RA not done	4	1.9%
Lack of or delayed approval of the trial itself	3	1.4%
No approval on trial site	2	1.0%
REC transparency	2	1.0%
Lack or delayed approval on organizational or personnel	1	0.5%
Patient enrolment w IC but w objection from REC	1	0.5%

Table 8: REC issues

ERF sub-categories	Examples of critical issues
Missing, late, or no approval on informed consent	"Some of the patient information and IC form were not approved in writing by the REC."
Late or no submission of protocol or protocol amendment	"Amendment 1 was partially implemented at the site although no approval was granted by the REC and by the Competent Authority" "Sponsor changed the recording and the reporting of "withdrawal syndrome" even though the protocol was very specific on this issue. Sponsor introduced the change without amending the protocol or notifying any IRB or competent authority."
Conflict of interest	"As a member of the local REC the PI was present during deliberation and voting procedure in the local REC."

Table 9: Examples of Research Ethics Committees ERFs that were graded as "critical" by the inspectors

Informed consent (IC)

A quarter of ERFs in this category pertain to relevant signature and date issues such as when trial subjects are routinely asked to sign as independent witnesses (25.4%). There were also a considerable number of issues with the IC procedure such as when securing of IC is conducted in a group setting 16.9%), as well as with the lack of or delayed IC or updated IC 13.8%Table 10 provides the ERFs under this category and Table 11 provides examples of IC-related issues that were graded by inspectors as critical.

ERF sub-categories	n=189	% of n
Relevant signature and date issues	48	25.4%
Issues w IC procedure	32	16.9%
No or delayed IC or updated IC	26	13.8%
Procedure prior to IC	15	7.9%
Non-approved version used	13	6.9%
Inaccuracies with the IC form	12	6.3%
Translation issues	12	6.3%
Incomplete information in the IC form or PI sheets	12	6.3%
No QC on IC	10	5.3%
Patients not informed or not informed on time	4	2.1%
Ample time to decide	3	1.6%
No system to update patient info sheet	1	0.5%
Signed IC but not (immediately) included	1	0.5%

Table 10: Informed Consent issues

ERF sub-categories	Examples of critical issues
Relevant signature or date issue; issues with IC procedure; procedure prior to IC	"The informed consent process was not performed properly since the investigator signed prior to the patients, screening started prior to IC, IC not correctly dated, independent witness was missing in most cases"
Inaccuracies with IC form	Various different versions of the ICF version 8.0 dated and ICF version 9.0 were provided to the site by the sponsor and used by the site. In some of these versions a physician not involved in this trial is given as contact person."
Relevant signature issues	"Trial subjects were routinely used as independent witnesses for signing the informed consent"

Table 11: Examples of Informed Consent ERFs graded as "critical" by the inspectors

Monitoring and oversight

Most of the ERFs concern inadequate monitoring or oversight such as when important signals were not identified by the monitors. At the same time, there were other issues such as the lack of or inadequate monitoring-related corrective measures such as when the sponsor neglects to address the deficiencies identified by the monitors (5.7%); or the lack of access to medical records of monitors or inspectors

(4.5%). Table 12 provides the ERFs under this category and Table 13 provides examples of monitoring and oversight issues that were graded as critical by the inspectors.

ERF sub-categories	n=176	% of n
Inadequate or insufficient monitoring or oversight	148	84.1%
Lack of or inadequate corrective measures	10	5.7%
Access to records	8	4.5%
Lack of detailed monitoring procedural agreement	3	1.7%
Lack of support for monitoring	3	1.7%
Unreviewed or delayed reviewing of monitor reports	2	1.1%
Conflict of interest among the monitors	1	0.6%
Problems with the monitors or monitoring team	1	0.6%

Table 12: Monitoring and oversight issues

ERF subcategories	Examples of critical issues
Inadequate or insufficient monitoring	"The monitoring was insufficient. The monitoring frequency during the double blind study phase of all but one patient was low and the quality of monitoring was insufficient (considering the high amount of discrepancies within the source documents and between SD and CRF). Therefore quality control and oversight of the trial by the sponsor was insufficient and no corrective actions could be identified."
Lack of support for monitoring	"The sponsor did not provide the CRA with adequate support and procedures for monitoring."
Lack of or inadequate corrective measures	"Monitor reported to the sponsor that all memory aids for visit 4 and 5 were present at the site seemed to be completed in the same way and had no signs of usage. There is no evidence of an adequate follow-up by the sponsor regarding this issue."

Table 13: Examples of Monitoring and Oversight ERFs graded as "critical" by the inspectors

Respect for persons

Half of ERFs within this category pertain to the lack of adequate facilities such as when the number of beds in an intensive care unit is inadequate for the participants (50%). A fifth of the issues were on the lack of cultural sensitivity for multinational studies, for example, when sponsor country regulations are by default exported to the host country, i.e., without consultation (20%). Table 14 provides the ERFs under this category, none of which were graded critical.

ERF sub-categories	n=10	% of n
Adequate facilities provided	5	50%
Cultural differences	2	20%

End of trial issues	1	10%
Exploitation	1	10%
Undue inducement	1	10%

Table 14: Issues on Respect for Persons

On the number of ERFs

Though ERFs are spread in the 288 sites of the 112 medicinal products, some sites had more ERFs than others. Table 2 provides the average number of ERFs based on inspector grading and the maximum number of ERFs found based on the aggregate number of trial sites per medicinal product.

	MINOR	MAJOR	CRITICAL
Mean	6.36	7.54	2.95
Most number of ERFs based on aggregate number of sites per medicinal product	26	30	12

Table 15: Mean and maximum number of ERFs per product

On average, the aggregate sites of a medicinal product would have 7.54 major and 2.95 critical ERFs; however, as many as 30 major and 12 critical ERFs were observed in the aggregate sites of one medicinal product.

Lastly, we looked at how many of the 112 medicinal products had at least 1 ERF and how many of these products had at least major ERFs. The results are presented in Table 16.

	N of products (total=112)
Products with at least 1 ERF	109 (97.3%)
Products with at least major ERF	104 (92.9%)

Table 16: Products with at least 1 ERF and those with at least major ERFs

ERFs were present in almost all of the medicinal products (97.3%), and most of the time, these ERFs were at least major (92.9%).

DISCUSSION

In this manuscript we were able to provide an initial identification and incidence of ERFs in pharmaceutical clinical trials. In doing so, we hope to have made the ethical non-compliance more apparent, hence allowing for the possibility of future discussions over areas that sponsors and investigators need to improve on; how assessors should respond to ERFs and the content of educational activities for the sponsors, investigators, and RECs.

Our query shows that site inspectors frequently encounter ERFs during inspection, i.e., at least a third (36.4%) of GCP non-compliance findings is ethically relevant, and ERFs were present in almost all of the medicinal products (97.3%). Of the 1685 ERFs, there is a majority of major findings (53.4%), while 9.2% are critical. At least major ERFs were present in almost all medicinal products with ERFs (95.4% of all medicinal products with ERFs). Considering that the grading is inspector-dependent, we cannot say with

definiteness that inspectors and the regulators who are recipients of the inspection reports as a matter of fact encounter an objectively unquestionable unethical conduct. To provide that judgment means to look deeper at the rationale of the inspection finding and the grading. Nevertheless, considering the number of ERFs which are either major or critical, it seems that regulators quite often face issues that from a GCP perspective might at least affect "the rights, safety, and/or well-being of participants," i.e., issues that are ethically relevant and hence needing a judgment that necessitates ethical reflection. As part of the inspection process, inspectors will request a corrective and preventive actions plan (CAPA) from the sponsor or inspectee. Although CAPAs cannot affect the performance of the trial(s) concerned as these are usually already closed during inspection, the main purpose is to prevent recurrence of the problems in future clinical trials. Further, national competent authorities authorizing trials may follow-up on the findings as well.

Apart from amount of ERFs that regulators may be confronted with, we also showed the mean and the maximum number of ERFs per grading. This means that though it may be reasonable for a regulator to expect, on average, 7.54 major and 2.95 critical ERFs in a marketing authorization application, the range of ERFs can increase to 30 major ERFs and 12 critical ERFs.

Among the ERFs encountered by inspectors, the most common are within the categories protocol compliance and protocol issues, patient safety, and professionalism and qualification. These are also the most common sources of critical and major findings. However, in terms of the density of critical and major findings within a category, (i.e., the probability that a finding within a category is either critical or major, monitoring and oversight), protocol compliance and protocol issues, and respect for persons top the list.

It is noteworthy that protocol compliance and protocol issues are most common in terms of incidence and the second most common in terms of density of critical and major findings. This category may not only adversely affect the rights, safety, and well-being of the participants but also the quality and integrity of the data, i.e., it clearly covers both "scientific benefit-risk evaluation" (i.e., science) and "basic ethical principles"[1] (i.e., ethics). We can speculate that issues such as these are most likely to be identified and addressed by regulators because of the strong scientific and ethical element in them; however, we have yet to explore if this in fact is the case. On the other hand, it is also worth studying how ERFs that violate only the basic ethical principles without affecting the scientific benefit-risk evaluation affect the MAA deliberation processes. This would presumably be the case for issues related to informed consent, research ethics committees, and respect for persons..

With regard to limitations, the findings of this study are naturally limited by the small number of inspections done compared to the number of trial sites. As such, there is always the probability that there are more ethically relevant GCP non-compliances that we failed to identify. Also, there could be researcher-induced biases in extracting data, creating nodes, and interpreting the results. To minimize bias, two researchers (RB and GvT) were involved in the extraction of data and the creation of nodes, and three researchers (RB, GvT and JvD) in the interpretation of results. Since two researchers were involved in both data extraction and the interpretation of results, discussions were frequent, also to resolve conflicts and differences. Next, as earlier stated, we took the inspector reports at face value, i.e.,

we assumed the validity of the inspectors' gradings, and we did not take into consideration the responses of the study sponsors, the corrective and preventive action plan, and the integrated inspection report. We also did not look into the influence of ERFs in the assessment and evaluation processes of the CHMP, i.e., we did not touch on the relation between ERFs and the acceptability of the data of the marketing authorization application. This is something we shall do in the succeeding manuscript. Lastly, we were not provided with the location of the trial sites and hence we were not able to account for geographical nuances. These may be topics for future study. For our current purposes, we hope to have provided a concrete picture of the ethical issues regulators encounter, which hopefully would inform future deliberations on regulatory actions.

CONCLUSION

Based on inspection reports, inspectors, clinical assessors, rapporteurs, and the CHMP regularly encounter ethically relevant findings that at least "might adversely affect the rights, safety or well-being of the subjects", i.e., ethically relevant GCP non-compliance findings that are at least graded as major by inspectors. The most common ethically relevant findings are also likely to affect the benefit/risk balance of a marketing authorization application by affecting the integrity of the research data, such as issues related to monitoring and oversight, or protocol compliance. It remains to be explored how these are considered in the assessment reports.

CONFLICTS OF INTEREST

RDLCB received funding from Dutch Medicines Evaluation Board for this study. NSB and CCG are/were employed by the Dutch Medicines Evaluation Board.

REFERENCES

1. European Medicines Agency. *Points to consider on GCP inspection findings and the benefit-risk balance*. (2012).
2. EMA. *Reflection paper on ethical and GCP aspects of clinical trials of medicinal products for human use conducted outside of the EU/EEA and submitted in marketing authorization applications to the EU Regulatory Authorities*. (2012).
3. European Medicines Agency. *Procedure for coordinating GCP inspections requested by the CHMP*. (2014).
4. EMA. *Procedure for reporting of GCP inspections requested by the Committee for Medicinal Products for Human Use (CHMP)*. (2013).
5. EMA. *Classification and analysis of the GCP inspection findings of GCP inspections conducted at the request of the CHMP*. (2014).
6. Bernabe, R. D. L. C., Van Thiel, G. J. M. W. & Van Delden, J. J. M. What do international ethics guidelines say in terms of the scope of medical research ethics? *BMC Med. Ethics* **17,** (2016).
7. Charmaz, K. *Constructing grounded theory: a practical guide through qualitative analysis*. (Sage, 2006).