

Machine Readable Data Not Required for EMA – Really?

David C. Izard
PhUSE 2019



February 24th–27th
Baltimore Marriott Waterfront Hotel
Baltimore

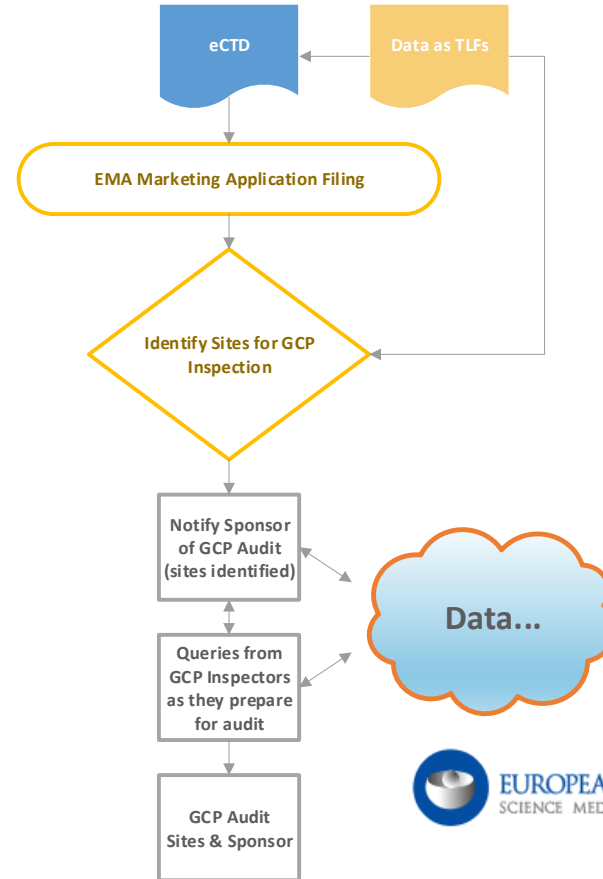
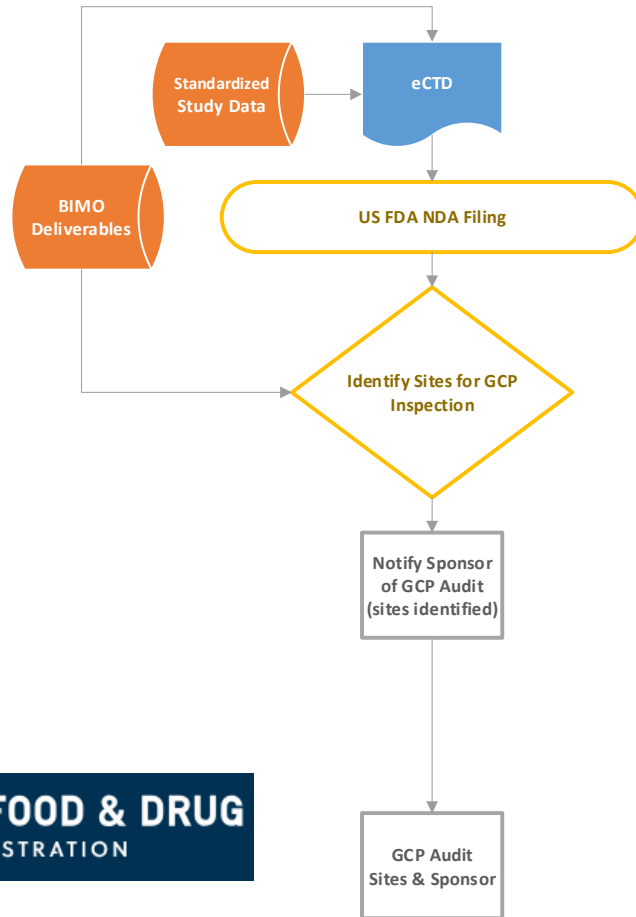
US 2019

The Clinical Data
Science Conference



-
- EMA does indeed have a mechanism for pulling in machine readable data to their review process!
 - Provided to support GCP Inspection during review cycle
 - Deployment of CDISC standards to these EMA deliverables is encouraged

GCP Inspection Process



Notice of EMA GCP Inspection Request



- The individual patient data listings for the patients. These may be restricted to the selected investigator site, but confirmation should be sought from the lead inspector prior to sending the documentation (including discontinued patients, protocol deviations, patients excluded from the efficacy analysis, demographic data, compliance and/or drug concentration data (if available), individual efficacy response data, adverse event listings, listing of individual laboratory measurements, etc.).
- Excel ® work book containing data listings for patients recruited at the selected site(s) (see EMA Q&A for further details –

Notice of EMA GCP Inspection Request

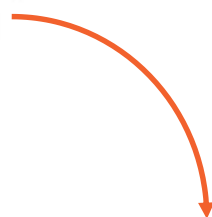


- The individual **patient data listings** for the patients. These may be restricted to the selected investigator site, but confirmation should be sought from the lead inspector prior to sending the documentation (including discontinued patients, protocol deviations, patients excluded from the efficacy analysis, demographic data, compliance and/or drug concentration data (if available), individual efficacy response data, adverse event listings, listing of individual laboratory measurements, etc.).
- Excel ® work book containing **data listings for patients** recruited at the selected site(s) (see EMA Q&A for further details –

Notice of EMA GCP Inspection Request



- Please note that no data should be provided until contact has been established with the reporting/lead inspector and the requirements for data listings have been discussed. Source data location list, see also Q&A: Good Clinical Practice on EMA homepage:
http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/q_and_a/q_and_a_detail_00016.jsp&mid=WC0b01ac05800296c5.



<https://www.ema.europa.eu/en/human-regulatory/research-development/compliance/good-clinical-practice/qa-good-clinical-practice-gcp>

EMA Website: Q&A for GCP



Navigation to Presenting Data to a GCP Inspection Team



Q&A: Good clinical practice (GCP)

GCP matters

7. How should data be presented when they are sent to the inspection team prior to a GCP inspection requested by the CHMP?

EMA Website: Q&A for GCP

Patient Data Listings in Excel Format



In connection with centralised applications, the Committee for Medicinal Products for Human Use (CHMP) often requests a good clinical practice (GCP) inspection of one or more sites to be performed. Prior to such GCP inspections, the European Medicines Agency (EMA) sends an announcement letter to the applicant in which – among others – a list of documents to be provided to the inspection team is presented. The data are used by inspectors for review in order to select patients and data to inspect. Among the requested documents are the individual patient data listings for the patients recruited at the sites to be inspected. Based on past experience, this request for data listings poses a significant number of problems and subsequently costs a lot of time for companies and inspectors, quite often resulting in listings of suboptimal quality.

The aim of this Q&A is to standardise and clarify the format of the data listings to be provided.

It is important to emphasise that the following guidance is the expected standard for most inspections; however, for some trials different, specific requests may be warranted. Consequently, no data should be provided until contact has been established with the reporting inspector and the requirements for data listings have been discussed.

In general, the following is expected:

- a. All data for the selected sites (and if requested for all sites in the trial) should be provided to the inspectors. That includes all case report form (CRF) data and in addition data which are not necessarily part of the CRF such as data provided to the sponsor by vendors (laboratory data, data from central evaluation of electrocardiograms (ECGs), imaging etc.), data from electronic patient-reported outcomes (ePROs) etc.
- b. Data should be provided as Excel spreadsheet line listings following the proposed formats and naming detailed below.
- c. All listings must have raw CRF data and any data derived or imputed from it that forms part of the data analysis.
- d. Both the CRF data set and the data set used for analysis, for instance as Excel exports from statistical analysis system (SAS), should be provided. They should be provided as exported to Excel document. Any process from raw data to analysed data should be explained if not explained in the statistical analysis plan.
- e. Paper copies should not be provided unless specifically requested by the inspection team.
- f. There should be a statement from the sponsor to confirm that the data provided is exactly the same as that submitted in the clinical study reports (CSRs) in the application (this will be checked at the inspection).

EMA Website: Q&A for GCP

Patient Data Listings in Excel Format

“Do you want data for just the selected sites or do you want it for all sites (the full set of study data)?”

“Yes, please provide the full set of study data, not just the sites being audited.”

In connection with centralised applications, the Committee for Medicinal Products for Human Use (CHMP) often requests a good clinical practice (GCP) inspection of one or more sites to be performed. Prior to such GCP inspections, the European Medicines Agency (EMA) sends an announcement letter to the applicant in which – among others – a list of documents to be provided to the inspection team is presented. The data are used by inspectors for review in order to select patients and data to inspect. Among the requested documents are the individual patient data listings for the patients recruited at the sites to be inspected. Based on past experience, this request for data listings poses a significant number of problems and subsequently costs a lot of time for companies and inspectors, quite often resulting in listings of suboptimal quality.

The aim of this Q&A is to standardise and clarify the format of the data listings to be provided.

It is important to emphasise that the following guidance is the expected standard for most inspections; however, for some trials different, specific requests may be warranted. Consequently, no data should be provided until contact has been established with the reporting inspector and the requirements for data listings have been discussed.

In general, the following is expected:

- a. All data for the selected sites (and if requested for all sites in the trial) should be provided to the inspectors. That includes all case report form (CRF) data and in addition data which are not necessarily part of the CRF such as data provided to the sponsor by vendors (laboratory data, data from central evaluation of electrocardiograms (ECGs), imaging etc.), data from electronic patient-reported outcomes (ePROs) etc.
- b. Data should be provided as Excel spreadsheet line listings following the proposed formats and naming detailed below.
- c. All listings must have raw CRF data and any data derived or imputed from it that forms part of the data analysis.
- d. Both the CRF data set and the data set used for analysis, for instance as Excel exports from statistical analysis system (SAS), should be provided. They should be provided as exported to Excel document. Any process from raw data to analysed data should be explained if not explained in the statistical analysis plan.
- e. Paper copies should not be provided unless specifically requested by the inspection team.
- f. There should be a statement from the sponsor to confirm that the data provided is exactly the same as that submitted in the clinical study reports (CSRs) in the application (this will be checked at the inspection).



EMA Website: Q&A for GCP

Patient Data Listings in Excel Format



“Do you want data for just the selected sites or do you want it for all sites (the full set of study data)?”

“Yes, please provide the full set of study data, not just the sites being audited.”

– Provide collected data as well as analysis (derived) data

In connection with centralised applications, the Committee for Medicinal Products for Human Use (CHMP) often requests a good clinical practice (GCP) inspection of one or more sites to be performed. Prior to such GCP inspections, the European Medicines Agency (EMA) sends an announcement letter to the applicant in which – among others – a list of documents to be provided to the inspection team is presented. The data are used by inspectors for review in order to select patients and data to inspect. Among the requested documents are the individual patient data listings for the patients recruited at the sites to be inspected. Based on past experience, this request for data listings poses a significant number of problems and subsequently costs a lot of time for companies and inspectors, quite often resulting in listings of suboptimal quality.

The aim of this Q&A is to standardise and clarify the format of the data listings to be provided.

It is important to emphasise that the following guidance is the expected standard for most inspections; however, for some trials different, specific requests may be warranted. Consequently, no data should be provided until contact has been established with the reporting inspector and the requirements for data listings have been discussed.

In general, the following is expected:

- All data for the selected sites (and if requested for all sites in the trial) should be provided to the inspectors. That includes all case report form (CRF) data and in addition data which are not necessarily part of the CRF such as data provided to the sponsor by vendors (laboratory data, data from central evaluation of electrocardiograms (ECGs), imaging etc.), data from electronic patient-reported outcomes (ePROs) etc.
- Data should be provided as Excel spreadsheet line listings following the proposed formats and naming detailed below.
- All listings must have raw CRF data and any data derived or imputed from it that forms part of the data analysis.
- Both the CRF data set and the data set used for analysis, for instance as Excel exports from statistical analysis system (SAS), should be provided. They should be provided as exported to Excel document. Any process from raw data to analysed data should be explained if not explained in the statistical analysis plan.
- Paper copies should not be provided unless specifically requested by the inspection team.
- There should be a statement from the sponsor to confirm that the data provided is exactly the same as that submitted in the clinical study reports (CSRs) in the application (this will be checked at the inspection).

Specification of formats and names:

- a. Naming of all files and Excel spreadsheets should be meaningful and self evident. Columns/name of variable in Excel worksheets could be called adverse event (AE), concomitant medications (CM), vital signs (VS) preferably in accordance with clinical data interchange standards consortium (CDISC) terms. Alternatively a 'translation' document should be provided to explain abbreviations.
- b. Data should be presented formatted, for example category values 1 and 2 as "yes" and "no" etc. If this isn't done then the format assignment to the data code must be provided.
- c. The data listings in Excel should ideally be consistent with the layout in the Clinical Study Report CSR, such that cross referencing the data is straight forward.
- d. Data types and field types should be appropriate for the specific data, for instance numerical data should always be numeric type and not character type, formats can be applied to the numeric data (e.g. dates). This is to allow the inspectors to do their own calculations.
- e. A copy in PDF format of the CSR listings per patient, for just the particular investigator site to be inspected should generally also be provided. The inspection team may also request to have paper copies brought to the investigator site for source data verification (SDV) purposes. The electronic data should not be an image, such that it is not possible to search, for example by date. There should be a statement from the sponsor to confirm that the data provided as copies in PDF format is exactly the same as that submitted in the CSR(s) in the application (this will be checked at the inspection).
- f. The data provided in PDF format and Excel worksheets should be set up for printing (e.g. print areas defined and suitable page arrangements set).

“Are you familiar with CDISC SDTM & ADaM data formats?”

“Yes”

“Should we provide our sets of SDTM & ADaM data to satisfy this requirement?”

“Yes”

Specification of formats and names:

- a. Naming of all files and Excel spreadsheets should be meaningful and self evident. Columns/name of variable in Excel worksheets could be called adverse event (AE), concomitant medications (CM), vital signs (VS) preferably in accordance with clinical data interchange standards consortium (CDISC) terms. Alternatively a 'translation' document should be provided to explain abbreviations.
- b. Data should be presented formatted, for example category values 1 and 2 as “yes” and “no” etc. If this isn't done then the format assignment to the data code must be provided.
- c. The data listings in Excel should ideally be consistent with the layout in the Clinical Study Report CSR, such that cross referencing the data is straight forward.
- d. Data types and field types should be appropriate for the specific data, for instance numerical data should always be numeric type and not character type, formats can be applied to the numeric data (e.g. dates). This is to allow the inspectors to do their own calculations.
- e. A copy in PDF format of the CSR listings per patient, for just the particular investigator site to be inspected should generally also be provided. The inspection team may also request to have paper copies brought to the investigator site for source data verification (SDV) purposes. The electronic data should not be an image, such that it is not possible to search, for example by date. There should be a statement from the sponsor to confirm that the data provided as copies in PDF format is exactly the same as that submitted in the CSR(s) in the application (this will be checked at the inspection).
- f. The data provided in PDF format and Excel worksheets should be set up for printing (e.g. print areas defined and suitable page arrangements set).

Patient Data Listings in Excel Format

“Are you familiar with the data definition file (define.xml) for SDTM & ADaM?”

“Yes”

“Can we provide this file in PDF format to serve as the ‘translation’ document?”

“Yes”

Specification of formats and names:

- a. Naming of all files and Excel spreadsheets should be meaningful and self evident. Columns/name of variable in Excel worksheets could be called adverse event (AE), concomitant medications (CM), vital signs (VS) preferably in accordance with clinical data interchange standards consortium (CDISC) terms. Alternatively a 'translation' document should be provided to explain abbreviations.
- b. Data should be presented formatted, for example category values 1 and 2 as “yes” and “no” etc. If this isn't done then the format assignment to the data code must be provided.
- c. The data listings in Excel should ideally be consistent with the layout in the Clinical Study Report CSR, such that cross referencing the data is straight forward.
- d. Data types and field types should be appropriate for the specific data, for instance numerical data should always be numeric type and not character type, formats can be applied to the numeric data (e.g. dates). This is to allow the inspectors to do their own calculations.
- e. A copy in PDF format of the CSR listings per patient, for just the particular investigator site to be inspected should generally also be provided. The inspection team may also request to have paper copies brought to the investigator site for source data verification (SDV) purposes. The electronic data should not be an image, such that it is not possible to search, for example by date. There should be a statement from the sponsor to confirm that the data provided as copies in PDF format is exactly the same as that submitted in the CSR(s) in the application (this will be checked at the inspection).
- f. The data provided in PDF format and Excel worksheets should be set up for printing (e.g. print areas defined and suitable page arrangements set).

EMA Website: Q&A for GCP



Patient Data Listings in Excel Format

Make sure
define.xml / pdf
provides complete
decodes / controlled
terminology

Preserve data types
(character versus
numeric), particularly
when creating Excel
version of ADaM

Provide Excel
workbooks with
printing controls
in place

Specification of formats and names:

- a. Naming of all files and Excel spreadsheets should be meaningful and self evident. Columns/name of variable in Excel worksheets could be called adverse event (AE), concomitant medications (CM), vital signs (VS) preferably in accordance with clinical data interchange standards consortium (CDISC) terms. Alternatively a 'translation' document should be provided to explain abbreviations.
- b. Data should be presented formatted, for example category values 1 and 2 as "yes" and "no" etc. If this isn't done then the format assignment to the data code must be provided.
- c. The data listings in Excel should ideally be consistent with the layout in the Clinical Study Report CSR, such that cross referencing the data is straight forward.
- d. Data types and field types should be appropriate for the specific data, for instance numerical data should always be numeric type and not character type, formats can be applied to the numeric data (e.g. dates). This is to allow the inspectors to do their own calculations.
- e. A copy in PDF format of the CSR listings per patient, for just the particular investigator site to be inspected should generally also be provided. The inspection team may also request to have paper copies brought to the investigator site for source data verification (SDV) purposes. The electronic data should not be an image, such that it is not possible to search, for example by date. There should be a statement from the sponsor to confirm that the data provided as copies in PDF format is exactly the same as that submitted in the CSR(s) in the application (this will be checked at the inspection).
- f. The data provided in PDF format and Excel worksheets should be set up for printing (e.g. print areas defined and suitable page arrangements set).

Traditional Patient Data Listings in PDF Format

“Can we repurpose data listings from the CSR for this requirement?”

“Yes”

“Can we provide a single listing per domain (e.g., AE or VS) for each site?”

“Yes”

Specification of formats and names:

- a. Naming of all files and Excel spreadsheets should be meaningful and self evident. Columns/name of variable in Excel worksheets could be called adverse event (AE), concomitant medications (CM), vital signs (VS) preferably in accordance with clinical data interchange standards consortium (CDISC) terms. Alternatively a 'translation' document should be provided to explain abbreviations.
- b. Data should be presented formatted, for example category values 1 and 2 as “yes” and “no” etc. If this isn't done then the format assignment to the data code must be provided.
- c. The data listings in Excel should ideally be consistent with the layout in the Clinical Study Report CSR, such that cross referencing the data is straight forward.
- d. Data types and field types should be appropriate for the specific data, for instance numerical data should always be numeric type and not character type, formats can be applied to the numeric data (e.g. dates). This is to allow the inspectors to do their own calculations.
- e. A copy in PDF format of the CSR listings per patient, for just the particular investigator site to be inspected should generally also be provided. The inspection team may also request to have paper copies brought to the investigator site for source data verification (SDV) purposes. The electronic data should not be an image, such that it is not possible to search, for example by date. There should be a statement from the sponsor to confirm that the data provided as copies in PDF format is exactly the same as that submitted in the CSR(s) in the application (this will be checked at the inspection).
- f. The data provided in PDF format and Excel worksheets should be set up for printing (e.g. print areas defined and suitable page arrangements set).

EMA Website: Q&A for GCP

Patient Data Listings in Excel Format



Report format of the patient data listings

Unless otherwise agreed with the inspection team, the data should be collected in the groups defined below – each to be presented in a different Excel spreadsheet. The columns “study site ID”, “subject ID” and “treatment group” and where applicable “visit ID” and/or “visit Date” should be in all spreadsheets. Some data listings may – depending on the trial – belong to different groups, for instance “vital signs/physical exam” usually belongs in the safety data group; however, in a hypertension trial it is likely to be an efficacy parameter. If in doubt, please ask the lead inspector:

1. **Study populations & conduct data** – recruitment dates, analysis populations, protocol/GCP non-compliance (deviations/violations), withdrawals/completed and reasons for withdrawing/not completing/not being randomised and outcome.
2. **Subjects' data** – demographic/covariate data (birth date, age, gender, ethnicity, race), medical history (general and related to trial objectives, e.g. tumour details, disease history/measurements), eligibility (inclusion/exclusion), consent date(s).
3. **Treatment data** – stratification group, randomisation, treatment given (including kit number and batch number), dosing dates, dose, dose adjustments, data concerned with compliance with treatment, non-medicinal co-treatments as part of trial protocol (e.g. radiotherapy).
4. **Specific efficacy** – raw data, repeated assessments related to efficacy, imputed values presented such that their determination from raw data can be seen (changes in parameters compared to baseline, categorisation of data to form new endpoint, clarity on last observation carried forward (LOCF) when used).
5. **Safety data** – adverse event (AE)/ serious adverse event (SAE) – data from CRF log and also from SAE forms, repeated assessments related to safety, side effects captured in CRF (not AEs) with severity grading etc. Both data entered by site and coded data should be listed.
6. **Laboratory type data** – sample/scan/measurement date/time (nominal and actual from CRF)/settings and other necessary details, report data time, result (from laboratories), investigator review.
7. **Concomitant medication data** – both data entered by site and coded data should be listed.
8. **Subject questionnaire data** (if not part of efficacy) (e.g. quality of life (QoL)).

EMA Website: Q&A for GCP

Patient Data Listings in Excel Format

“Are you familiar with the Study & Analysis Data Reviewers Guides?”

“Yes”

“Can we provide those documents in order to communicate study data groupings?”

“Yes”

Report format of the patient data listings

Unless otherwise agreed with the inspection team, the data should be collected in the groups defined below – each to be presented in a different Excel spreadsheet. The columns “study site ID”, “subject ID” and “treatment group” and where applicable “visit ID” and/or “visit Date” should be in all spreadsheets. Some data listings may – depending on the trial – belong to different groups, for instance “vital signs/physical exam” usually belongs in the safety data group; however, in a hypertension trial it is likely to be an efficacy parameter. If in doubt, please ask the lead inspector:

1. **Study populations & conduct data** – recruitment dates, analysis populations, protocol/GCP non-compliance (deviations/violations), withdrawals/completed and reasons for withdrawing/not completing/not being randomised and outcome.
2. **Subjects' data** – demographic/covariate data (birth date, age, gender, ethnicity, race), medical history (general and related to trial objectives, e.g. tumour details, disease history/measurements), eligibility (inclusion/exclusion), consent date(s).
3. **Treatment data** – stratification group, randomisation, treatment given (including kit number and batch number), dosing dates, dose, dose adjustments, data concerned with compliance with treatment, non-medicinal co-treatments as part of trial protocol (e.g. radiotherapy).
4. **Specific efficacy** – raw data, repeated assessments related to efficacy, imputed values presented such that their determination from raw data can be seen (changes in parameters compared to baseline, categorisation of data to form new endpoint, clarity on last observation carried forward (LOCF) when used).
5. **Safety data** – adverse event (AE)/ serious adverse event (SAE) – data from CRF log and also from SAE forms, repeated assessments related to safety, side effects captured in CRF (not AEs) with severity grading etc. Both data entered by site and coded data should be listed.
6. **Laboratory type data** – sample/scan/measurement date/time (nominal and actual from CRF)/settings and other necessary details, report data time, result (from laboratories), investigator review.
7. **Concomitant medication data** – both data entered by site and coded data should be listed.
8. **Subject questionnaire data** (if not part of efficacy) (e.g. quality of life (QoL)).



Complete Set of Deliverables to GCP Auditor



Patient Data Listings in Excel format

- Full SDTM & ADaM data as Excel
- Define.xml/pdf for SDTM & ADaM
- Study Data Reviewer's Guide
- Analysis Data Reviewer's Guide

Traditional PDF Patient Data Listings

- PDF data listing by domain for each site
- Compiled into single PDF file w/ bookmarks
- Readme file that grouped listings by data type

A screenshot of an Excel spreadsheet titled 'SDTM [Read-Only] - Excel'. The spreadsheet is open to the 'Home' tab. The data is organized into columns A through I. The first row (row 1) contains the following headers: 'Description of Actual Arm', 'Actual Arm Code', 'Age', 'Age Units', 'Description of Planned Arm', 'Planned Arm Code', 'Date/Time of Birth', 'Country', and 'Domain Abbreviation'. The subsequent rows (rows 2 through 21) are empty, showing a grid of cells. The status bar at the bottom indicates 'Ready' and shows a zoom level of 100%.

-
- These deliverables are not all you will do...
...many queries from the GCP Inspector in advance of the site level inspection
 - Site inspections followed by Sponsor inspection
...and everything it stands for

-
- These deliverables are not all you will do...
...many queries from the GCP Inspector in advance of the site level inspection
 - Site inspections followed by Sponsor inspection
...and everything it stands for

Thank you!