

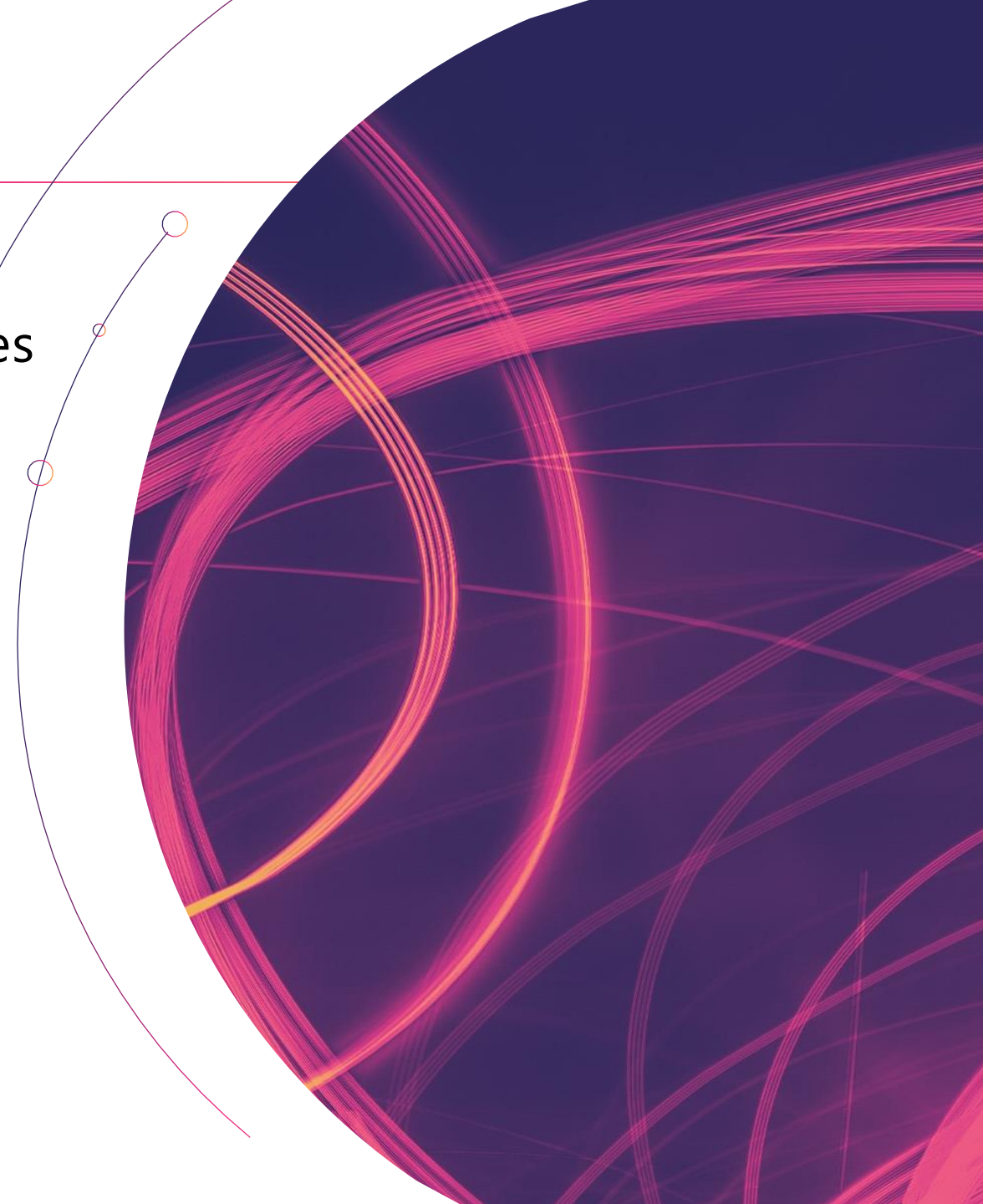
How Central Statistical Monitoring Drives a Risk-Based Approach to Clinical Trial Monitoring

Charles Warne, Cytel

*PHUSE APAC Connect, Hyderabad
20 February, 2026*

Agenda

- What is CSM and why is it needed
- When to apply CSM, how often, what analyses
- Role within central monitoring and RBQM
- CSM operating model and team structure
- Case studies
- Challenges to implementation
- Conclusion



Definition

➤ Central Statistical Monitoring (CSM):

*The **statistical detection of anomalies** in accumulating clinical trial data to identify groups (i.e. sites, patients, countries) that are **performing differently to the rest**, where the reasons for such differences may relate to issues in **data quality, site conduct or safety**.*

Origins of CSM

- Statistical methods first proposed by Buyse et al. (1999) to detect fraud:

STATISTICS IN MEDICINE
Statist. Med. **18**, 3435–3451 (1999)

THE ROLE OF BIOSTATISTICS IN THE PREVENTION, DETECTION AND TREATMENT OF FRAUD IN CLINICAL TRIALS[†]

MARC BUYSE^{1*}, STEPHEN L. GEORGE², STEPHEN EVANS³, NANCY L. GELLER⁴,
JONAS RANSTAM⁵, BRUNO SCHERRER⁶, EMMANUEL LESAFFRE⁷,
GORDON MURRAY⁸, LUTZ EDLER⁹, JANE HUTTON¹⁰, THEODORE COLTON¹¹,
PETER LACHENBRUCH¹² AND BABU L. VERMA¹³

for the
ISCB SUBCOMMITTEE ON FRAUD

SUMMARY

Recent cases of fraud in clinical trials have attracted considerable media attention, but relatively little reaction from the biostatistical community. In this paper we argue that biostatisticians should be involved in preventing fraud (as well as unintentional errors), detecting it, and quantifying its impact on the outcome of clinical trials. We use the term 'fraud' specifically to refer to *data fabrication* (making up data values) and

Motivation for CSM

➤ Types of data issues (adapted from Buyse, PHUSE):

Type	Typical examples	Intent
Fraud	Falsified data (<i>e.g. fabrication of ePRO data</i>)	Intention to cheat
Tampering	Fabricated data (<i>e.g. propagation of blood pressure</i>)	Deliberate
Sloppiness	Incorrect reporting (<i>e.g. under-reporting of AEs</i>)	Limited awareness
Errors	Technical problems (<i>e.g. miscalibrated thermometers</i>)	Unintentional

Less likely,
impacts credibility

More likely,
low/high impact
to data quality

CSM leads to higher data quality, lower cost, faster trials

- CSM can detect data issues that would be missed by a monitoring strategy that only relies on Source Document Verification [SDV], or a data cleaning strategy that only relies on standard DM checks
- SDV and standard DM checks are not designed to detect anomalies within/between patients, at site/study level, or in the multivariable space
- By flagging 'at risk' sites/patients, CSM supports a **risk-based approach** to monitoring and is a powerful tool for data quality oversight
- Benefits of CSM:

Smarter resource allocation

- Better return on investment (targeted on-site monitoring)
- Cost savings (reduced SDV, typically ~15% total budget)

Higher data quality

- Reduced timelines (faster enrollment & database lock)
- Reduced regulatory risk
- Cost savings (less re-work)

Evolution towards risk-based & centralised monitoring

➤ FDA (2013), EMA (2013), ICH (2016), MHRA (2022), ICH (2025)



CSM is a part of central monitoring

Central Monitoring

```
graph TD; CM[Central Monitoring] --> KRI[Key Risk Indicators (KRIs)]; CM --> QTL[Quality Tolerance Limits (QTLs)]; CM --> CSM[Central Statistical Monitoring (CSM)];
```

Key Risk Indicators (KRIs)

- Leading metrics for detecting risks to data quality, conduct, safety
- Simple summary measures: rates of serious/non-serious AEs, protocol deviations, missed assessments, early terminations, screen failures, auto-queries, eCRF visit to entry cycle time, etc
- Site-level
- 10-20 KRIs per study

Quality Tolerance Limits (QTLs)

- A special subset of KRIs focusing on critical risks
- Study-level
- 3-5 QTLs per study
- Thresholds defined from comparable studies

Central Statistical Monitoring (CSM)

- Data anomaly detection
- Advanced statistical methods applied to high-volume data
- Flag potential issues in data quality or site conduct that warrant attention
- Usually site-level, but can also be subject- or country-level

CSM – When to apply, how often, what analyses? (1)

- Eligible studies: Multi-center, ≥ 30 patients, ≥ 5 per site, ≥ 6 months duration
- Data sources: eCRF/EDC (raw), SDTM/ADaM, external data (labs, biomarkers, eDiary, PRO), IxRS/IRT, CTMS, queries, audit trail
- Timing of analyses:
 - Periodic (i.e. 3 monthly during enrolment, 6 monthly during follow-up)
 - Once before key milestone (Interim Analysis, Dry-run, Database Lock)
 - Risk-based: adapt frequency to observed findings

CSM – When to apply, how often, what analyses? (2)

- Comparisons: Site-level (one site Vs the rest), patient-level, country-level
- Supervised Vs Unsupervised:
 - Early approaches employed custom tests on a limited number of variables / endpoints / KRIs ("supervised") -> lower sensitivity
 - More recent methods use a brute force computer-intensive approach, exhaustive set of generic tests on all data collected ("unsupervised") -> higher sensitivity
- Analyses: Single variable (means, SDs within/between patients, proportions, counts, digit preference, calendar time), multiple variables (multivariate outliers/inliers, correlations, duplicated "professional" patients)
- Test p-values log-transformed to obtain overall risk scores for each site

Insights from CSM can drive RBM and support RBQM

Quality in clinical trials is “**the absence of errors that matter to decision making** – that is, errors which have a meaningful impact on the safety of trial participants or the credibility of the results”



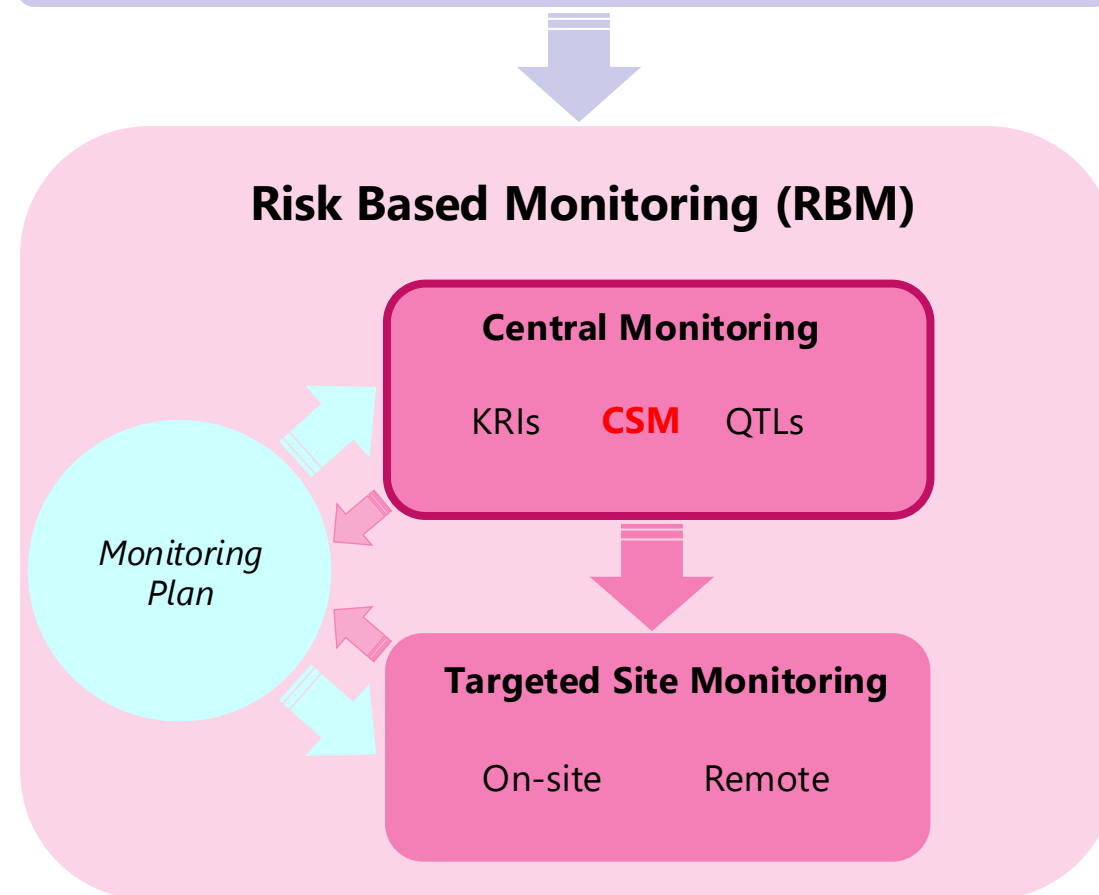
Quality Risk Management principles are embedded in RBM:

- Building **Quality by Design (QbD)** into trials
- **Critical Processes and Data** (“critical to quality factors”)
- Early and ongoing **Risk Assessment**
- Use of **Risk Indicators and Thresholds**
- Ongoing adjustment of monitoring activities based on the issues and risks that may arise throughout the study

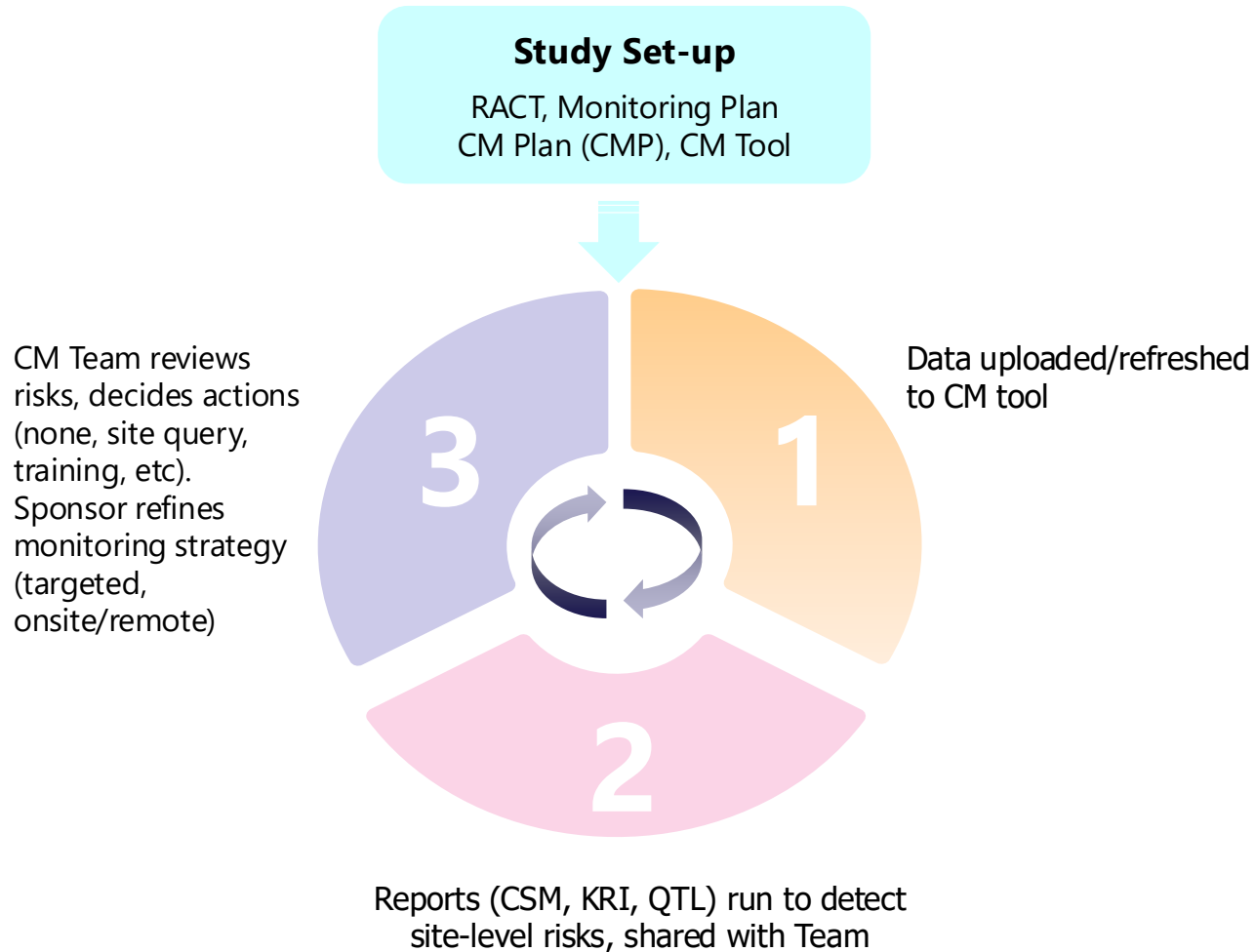


KRI = Key Risk Indicator; QTL = Quality Tolerance Limit

Risk Based Quality Management (RBQM)

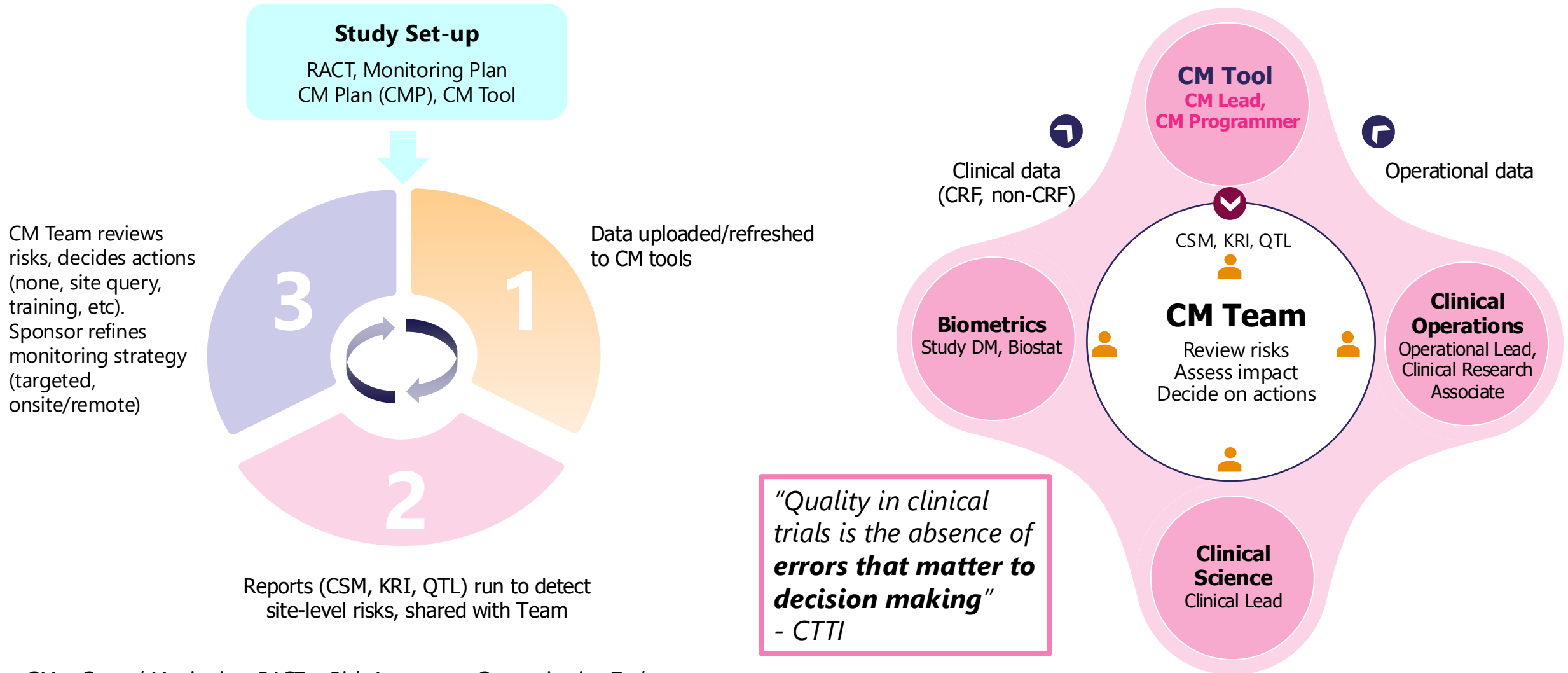


Operating model & multi-disciplinary team structure



CM = Central Monitoring, RACT = Risk Assessment Categorisation Tool

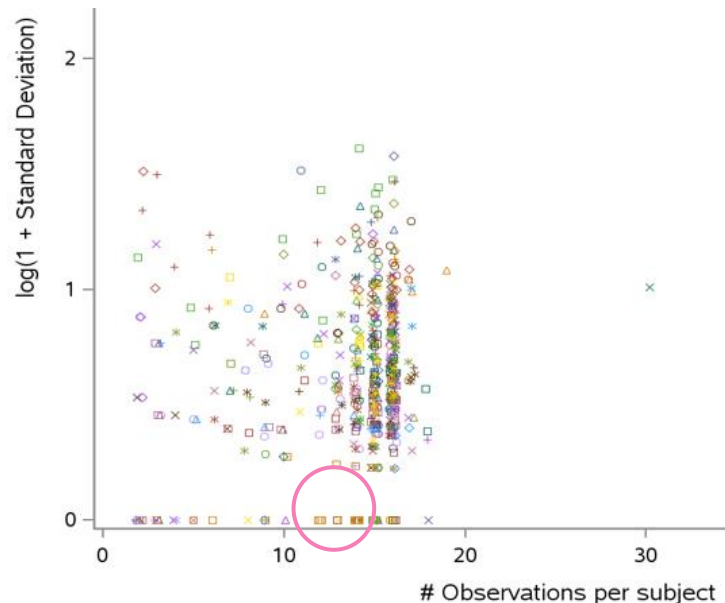
Operating model & multi-disciplinary team structure



CM = Central Monitoring, RACT = Risk Assessment Categorisation Tool

Case #1: Data propagation (tampering/sloppiness)

- Global Ph3 study subjected to CSM analyses
- Within-patient standard deviation (SD) revealed some outlying sites
- Many patients at Site 23 had Respiratory Rate replicated across visits (i.e. SD = 0)
- Root cause: Data copied forward ('propagated') from baseline



Site ID	Subject ID	# Obs. per subject	Standard Deviation	log(1 + Standard Deviation)
23	23001	14	0.0000	0
23	23002	15	0.2582	0.2296812
23	23003	15	0.2582	0.2296812
23	23004	3	0.0000	0
23	23005	15	0.0000	0
23	23006	14	0.0000	0
23	23007	10	0.3162	0.2747699
23	23008	15	0.0000	0
23	23009	14	0.0000	0
23	23010	14	0.0000	0
23	23011	13	0.2774	0.2447877
23	23012	15	0.0000	0

Site ID	Subject ID	Visit ID	Result	Units
23	23001	10	18 breaths/min	
23	23001	20	18 breaths/min	
23	23001	30	18 breaths/min	
23	23001	40	18 breaths/min	
23	23001	50	18 breaths/min	
23	23001	60	18 breaths/min	
23	23001	70	18 breaths/min	
23	23001	80	18 breaths/min	
23	23001	90	18 breaths/min	
23	23001	100	18 breaths/min	
23	23001	110	18 breaths/min	
23	23001	120	18 breaths/min	
23	23001	130	18 breaths/min	
23	23001	180	18 breaths/min	

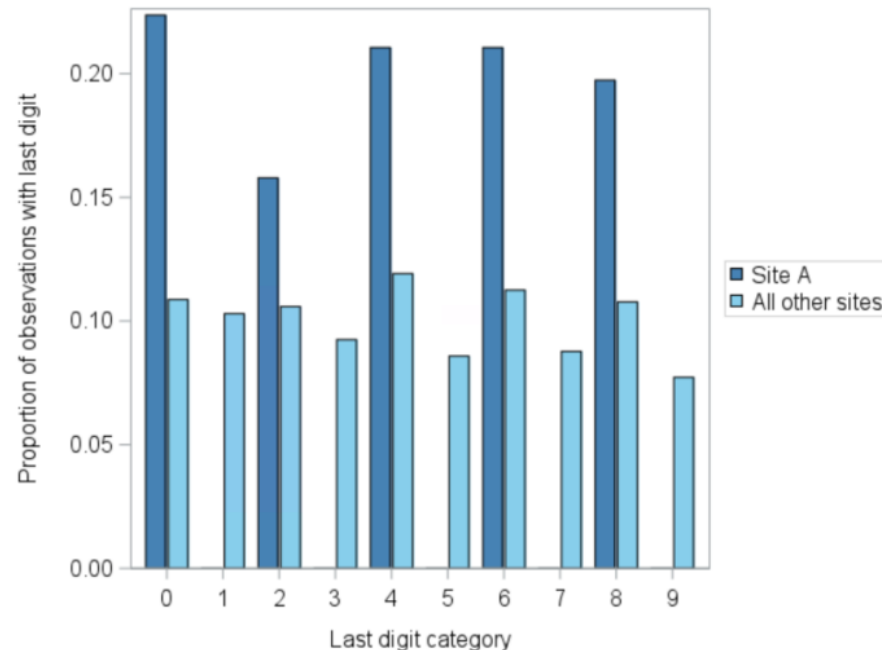
Note: Site and subject IDs are dummy coded

Case #2: Last digit preference (errors, tampering, fraud?)

- Global multicenter trial subjected to last digit preference CSM analyses
- Outlying distribution for a lab parameter at one site with 76 observations across 5 patients
- Plot shows 'combed distribution' (dark blue) for Site A where **all** last digits were even numbers, statistically different to uniform distribution in remaining sites (light blue)
- Highly unusual, warrants further investigation at site to determine root cause (equipment calibration, conversion issue, tampering, fraud?)

All 76 lab parameter values at site A:

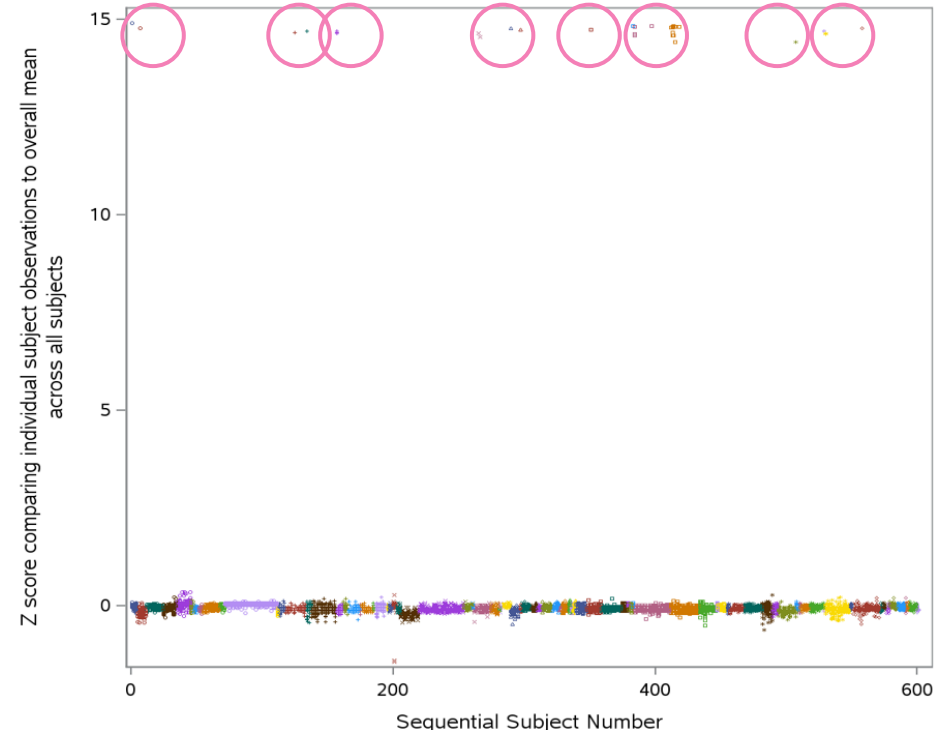
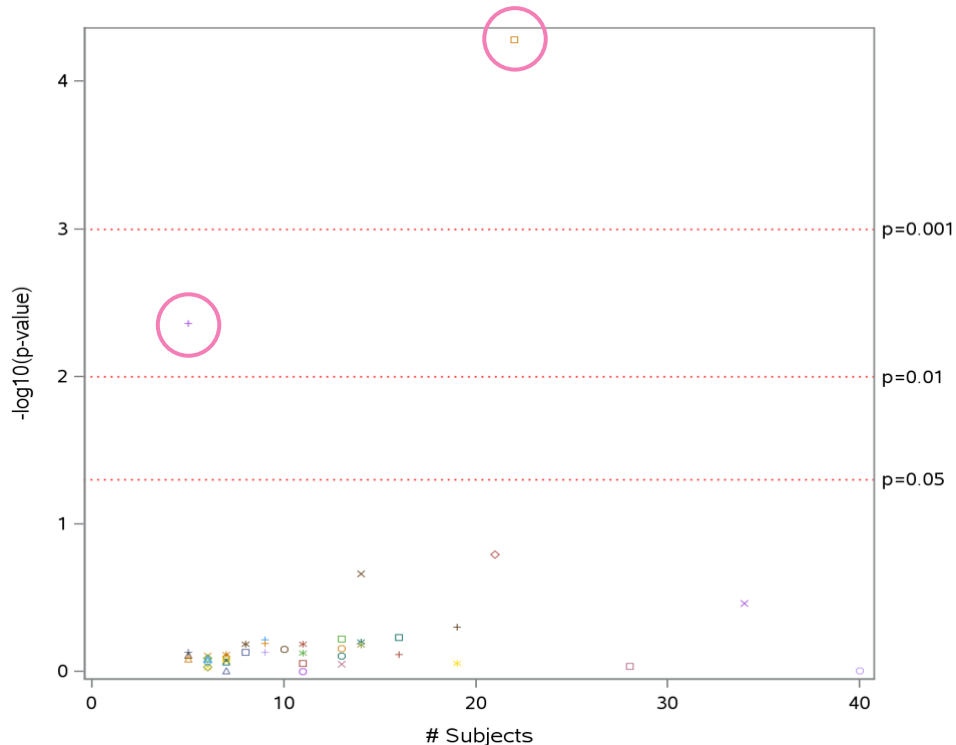
60	72	60	54	84	102	138	84
60	66	66	66	90	96	90	96
60	66	72	60	96	90	90	36
60	60	54	48	102	96	78	42
90	60	48	48	102	66	84	48
78	66	54	54	96	60	84	48
72	66	60	54	120	66	78	
66	54	54	54	108	84	78	
72	48	48	72	96	72	84	
90	72	48	78	102	84	84	



last_ digit	site_ count	site_ proportion	overall_ proportion	p_value
0	17	0.224	0.109	0.0058
1	0	0.000	0.103	0.0044
2	12	0.158	0.106	0.2055
3	0	0.000	0.092	0.0659
4	16	0.211	0.119	0.0323
5	0	0.000	0.086	0.0423
6	16	0.211	0.112	0.0288
7	0	0.000	0.088	0.0019
8	15	0.197	0.108	0.0295
9	0	0.000	0.077	0.0064

Case #3: Incorrect units (sloppiness/errors)

- Global multicenter trial subjected to comparison of means CSM analyses
- Two sites with extreme mean temperatures (40.2 °C, 44.9 °C) compared to the rest (36.9 °C)
- Further investigation revealed a more systemic issue involving additional sites
- Root cause: some temperatures recorded in Fahrenheit but units fixed to Celsius in EDC

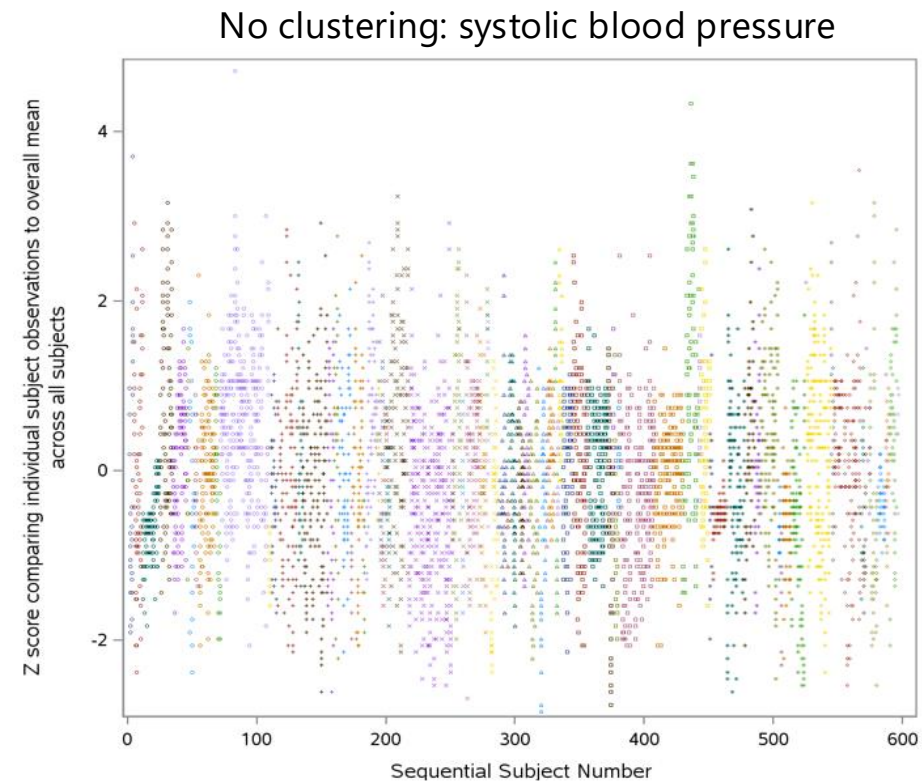
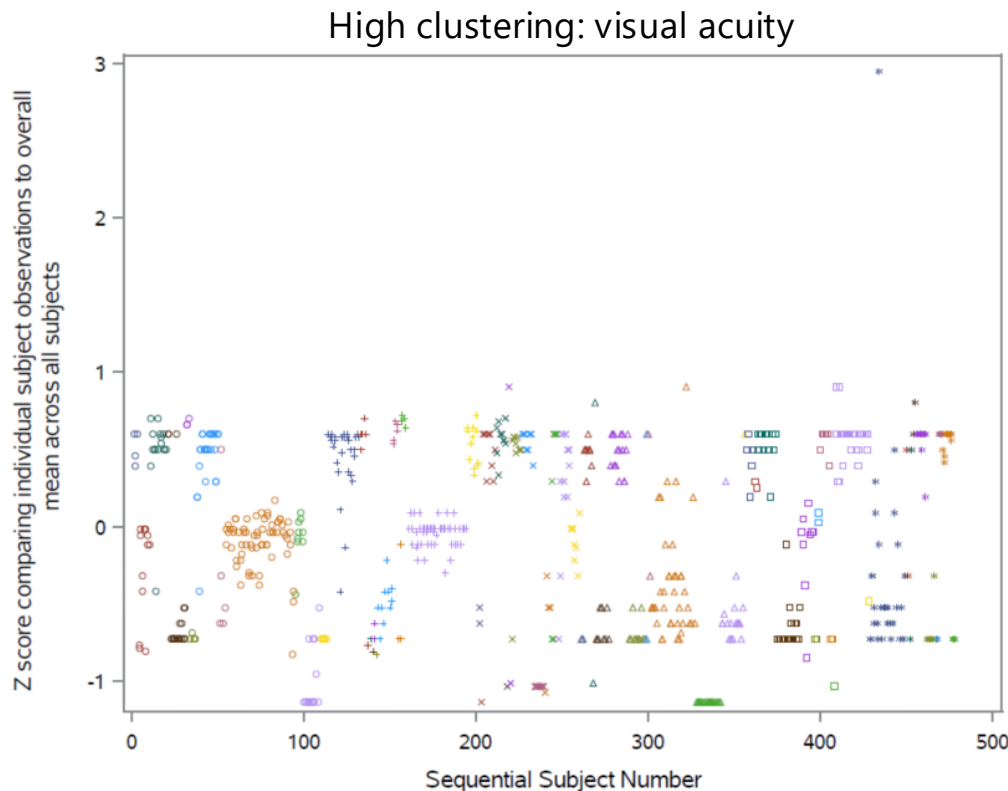


Site ID	Sequential Subject Number	Visit ID	Result	Units
140	157	10	97.6	C
140	158	10	36.2	C
140	159	10	36.2	C
140	160	10	36.4	C
140	161	10	36.2	C
381	412	10	98.1	C
381	413	10	97.6	C
381	414	10	36.8	C
381	415	10	96.6	C
381	416	10	36.7	C
381	417	10	37	C
381	418	10	98.2	C
381	419	10	36.7	C
381	420	10	36.1	C

Note: Site and subject IDs are dummy coded

Case #4: Site-level clustering (errors)

- Global multicenter trial subjected to CSM analyses
- High degree of site-level clustering found for visual acuity (ICC~0.9)
- Root cause: Different testing methods/equipment used across sites, lack of standardisation, training issue? At a minimum, clustering should be adjusted for in the analysis.
- Visualisations powerfully highlight the clustering when sorting by site/patient on X axis (seq. subj. number): i.e. compare the patterns seen for visual acuity (left plot) and systolic blood pressure (right plot)



Challenges to implementation

- Sufficient data needed before 1st CSM run:
 - Inherently low volume of data at the beginning of trials when corrective actions are most impactful (fewer patients and assessments)
 - Staggered site initiation (≥ 5 patients per site needed)
- Data may not be clean before CSM run -> adds noise
- Data that could potentially unblind a patient's treatment may need to be excluded (labs, biomarker, PK)
- Anomaly detection is **relative** and assumes most sites are performing well, if not, external QTL/KRI benchmarks may be more useful
- Valid reasons may exist for site differences: age, race, socio-economic, cultural

Conclusion

- CSM drives a risk-based approach to monitoring and quality management by **helping study teams identify where the risks are**, so they can target on-site monitoring resources effectively
- CSM achieves this:
 - Using statistical methods that detect anomalies more effectively than manual review methods such as SDV/SDR
 - Based on the totality of data collected (high volume -> better performance)
 - Using objective measures of comparison, applied in a standardised way across different types of clinical trials and disease areas
 - In an efficient and reproducible manner (via statistical programming code)

References

- Buyse M., et al. The role of biostatistics in the prevention, detection and treatment of fraud in clinical trials. *Statistics in Medicine* 1999; 18: 3435-3451.
- Buyse M. Role of Biostatistics in Fraud Detection, PHUSE presentation (slides).
- Clinical Trials Transformation Initiative: <https://ctti-clinicaltrials.org/about/ctti-projects/quality-by-design/>
- European Medicines Agency, Reflection paper on risk based quality management in clinical trials, 2013.
- Food and Drug Administration, Guidance for Industry: Oversight of Clinical Investigations – A Risk-based Approach to Monitoring, 2013.
- George S., Buyse M. Data fraud in clinical trials. *Clinical Trial Perspective* 2015; 5(2), 161-173.
- International Council for Harmonisation, Guideline for Good Clinical Practice E6(R2), 2016.
- International Council for Harmonisation, Guideline for Good Clinical Practice E6(R3), 2025.
- Medicines & Healthcare products Regulatory Agency, Guidance: Oversight and Monitoring Activities, 2022 (<https://www.gov.uk/government/publications/oversight-and-monitoring-of-investigational-medical-product-trials/oversight-and-monitoring-activities>)
- TransCelerate Biopharma Inc: <https://www.transceleratebiopharmainc.com/>
- Venet D., et al. A statistical approach to central monitoring of data quality in clinical trials. *Clinical Trials* 2012; 9(6), 705-713.

Thank you

charles.warne@cytel.com

