Paper AR02

Advanced analytics for data quality assessment in countries and regions affected by crisis

Archana Nayak, Janssen R & D, High Wycombe, United Kingdom
Md Umarfaruque, ZS, Janssen R & D
Kris Lauwers, Janssen R & D, Beerse, Belgium
Weronika Stepczyk, Janssen R & D, Warsaw, Poland
Susan Meier, Janssen R & D, PA, USA
Marion Wolfs, Janssen R & D, NL
Bhargava Reddy, Janssen R & D, PA, USA

ABSTRACT

Due to unprecedented events in recent years, the conduct of clinical trials has been increasingly affected by crisis at country and regional level. The inability to conduct study visits, difficulties in recording high quality data and delivering effective quality oversight have the potential to significantly impact trial integrity and patient safety. In this paper, we highlight internal methodology implemented as part of Central Statistical Surveillance. Our tool identifies and compares viewpoints among different countries signals and assesses their risk magnitudes to indicate potential areas of high, medium, and low risk at various levels of the study data. These insights provide opportunities to apply appropriate corrective and preventive actions, thereby ensuring higher data quality and dynamic monitoring of ongoing study data.

INTRODUCTION

The Central Statistical Surveillance (CSS) application helps in the comprehensive examination of clinical trial data by utilizing statistical models to detect naturally occurring data patterns and relationships. It represents a collaborative effort between the Risk Management - Central Monitoring team and the Advanced Insights team within Janssen. CSS was initially launched in 2019 and is an internally developed capability as part of the Analytical Risk-Based Monitoring (ARBM) offering.

CSS provides a crucial layer of oversight for clinical trial data, helping to mitigate potential issues with data quality and thus safeguarding subject safety across multiple therapeutic areas in ongoing clinical trials. The CSS team supports diverse study teams, enabling near real-time analysis of clinical data on an ongoing and dynamic basis. In summary, the CSS team supports approximately 80+ Janssen studies spanning 8 therapeutic areas, with more than 190+ data processing cycles.

The overarching scope of CSS allows for a comprehensive examination of clinical trial data through the application of statistical models, enabling the detection of naturally occurring data patterns and relationships while identifying outlier sites that deviate from expected patterns.

Some key outcomes of CSS include uncovering unexpected overall values, identifying unexpected data patterns in subject site visits, assessing seasonal influences through Timepoint Analysis, and comparing adverse events at the site level. Examples of these outcomes are highlighted below.

PROBLEM STATEMENT

Various conditions, such as natural disasters, political unrest, public health emergencies, regulatory changes, economic crises, infrastructure failures, transportation disruptions, natural resource scarcity, security concerns, environmental pollution, localized healthcare capacity strain, and cultural or societal factors, have produced distinct local and regional effects specific to each country, differing from elsewhere in the world. Additionally, the COVID-19 crisis had varying impacts on different countries, largely influenced by country-level regulations. These factors have underscored the necessity for analyzing clinical trial data while taking into account country-level information. Such analysis is crucial to uncover potential local impacts on the progression of ongoing clinical trials within each respective country (George et al, 2023).

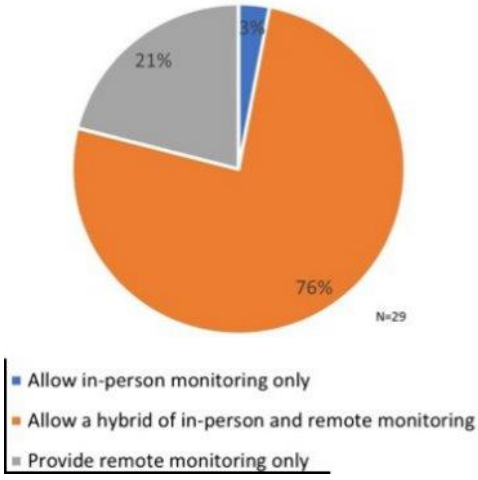


Figure: Covid crisis and other factors impacting country level data patterns (left) Graph credit from George et al, 2023 (Right) showing which illustrates an example about of impact of the COVID-19 pandemic and its quantitative impact on cancer center clinical trial operations highlighting the need for robust remote monitoring tools and capabilities

METHOD

Mainly in this paper, we describe our method of detecting country level data anomalies, analyzing data at country level allows for collective insights that drive innovative decisions for corrective and preventive corrections at larger scale that was previously examined only by site level comparisons. Here, we highlight the construction of programming and visuals of the dashboard along with its impact on ensuring high quality and dynamic monitoring of ongoing clinical trials in 4 main checks as follows:

- Data quality: To identify data quality issues at country and/or region level
- Data Pattern: To Understand country-specific behavior, country wide practices, etc.
- Local influence: To assess potential impact of local and regional crisis events on overall study data quality
- Systemic issues: To gain insight into country-level systemic issues including faulty batch of equipment, etc. by aggregating data at country level.

In order to determine the overall score for each country based on the individual site risk scores, we have developed a unique approach. Let's assume that we are studying η countries. In this case, the country level score can be expressed as follows:

$$\varphi_i = \frac{\eta}{\psi} \sum_{j \in \xi_i} \vartheta_j \chi_j$$

$$\varphi_i = \sum_{j \in \xi_i} \frac{\vartheta_j}{\left(\frac{\psi}{\eta}\right)} \chi_j$$

φ_i: $Risk\ score\ of\ country\ i$

η: $Total\ number\ of\ country$

ξ_i: $list\ of\ sites\ in\ country\ i$

ψ: $Total\ number\ of\ subjects\ enrolled\ in\ study$

$$\vartheta_j: Total\ number\ of\ subjects\ on\ site\ j$$

$$\chi_i: Risk\ score\ of\ site\ j$$

Here we have defined the risk score of sites as below

$$\chi_i = \begin{cases} 1 & \zeta_i = High \\ 0.5 & \zeta_i = Medium \\ 0 & \zeta_i = Low \end{cases}$$

$$\zeta_i: Risk\ category\ of\ site\ i$$

We have structured the scores in this manner to illustrate that penalizing a country due to the presence of one subject in 'High' risk sites is equivalent to penalizing two subjects in 'Medium' risk sites, with no penalization applied to subjects in 'Low' risk sites.

The score φ_i has been designed to reflect a specific principle: when a site's subject count surpasses the average subjects per country, the score is multiplied by values greater than one. In cases where a site falls into the 'High' or 'Medium' category, these enhanced scores serve as penalties substantial enough to categorize the entire country as 'High' or 'Medium' risk. This approach aligns with the hypothesis that a site's representation from a significant proportion of subjects should be a determining factor in assessing risk.

As part of our objective to categorize countries into distinct groups, we have established various threshold levels as outlined below:

$$\varrho_i = \begin{cases} High & \varphi_i \geq 0.66 \\ Medium & 0.33 \leq \varphi_i < 0.66 \\ Low & \varphi_i < 0.33 \end{cases}$$

$$\varrho_i: Risk\ category\ of\ country\ i$$

To express the equation in an interpretable format, we can state it as follows: If the number of equivalent high-risk subjects (defined as subjects associated with high-risk sites and half of the subjects in medium-risk sites) exceeds 66% of the average number of subjects per country, the country is categorized as "High risk." If the equivalent number of high-risk subjects falls between 33% and 66%, the country is classified as "Medium risk." Otherwise, it is labeled as a "Low risk" country.

Our subsequent goal was to determine the ranking of countries. This was achieved by initially sorting them in descending order based on their country risk scores and subsequently by the percentage of high-risk sites within each country.

$$Rank(C_i) = \min\{A \in Ordinal\ Number\ \ \forall\, Y \in C\ \ \varphi_Y \leq \varphi_i, \varpi_Y \leq \varpi_i\}$$

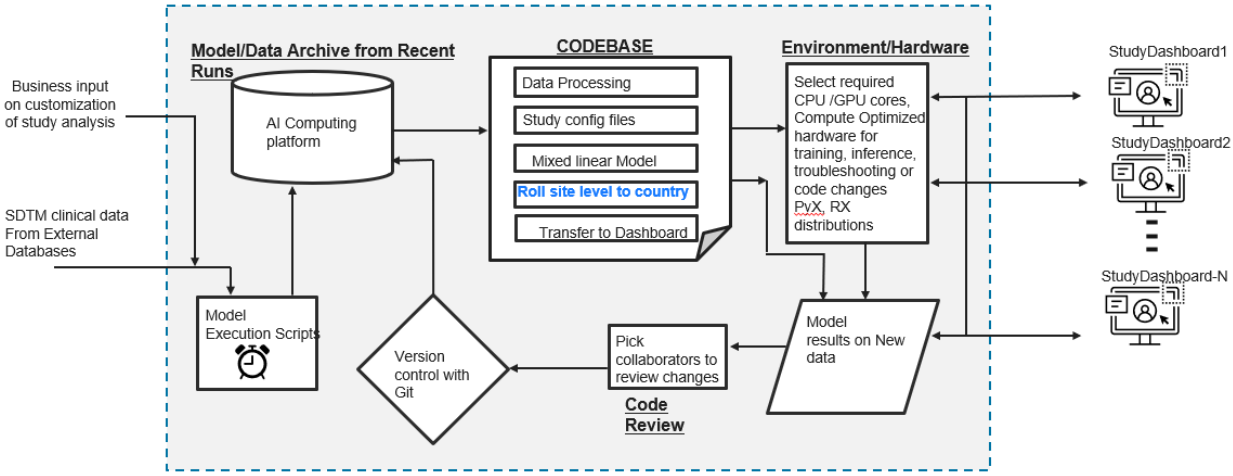
$$C_i: country\ i$$

$$\varpi_i: Percentage\ of\ high\ risk\ sites\ in\ country\ i$$

$$\varpi_i = \sum_{j \in \xi_i} \frac{\mathcal{H}(\zeta_j)}{len(\xi_i)}$$

$$\mathcal{H}(\zeta_i) = \begin{cases} 1 & \zeta_i = High \\ 0 & otherwise \end{cases}$$

Our data pipeline, as detailed in our past paper (Nayak et al, 2021), briefly, CSS analysis is primarily conducted in the SDTM CDISC data format. Data from multiple therapeutic areas can have heterogeneous data formats, so study-specific configuration files are constructed to process the data. The programming infrastructure, machine hardware, and computational memory requirements vary from one study to another. In this context, we have extended our pipeline to incorporate country-level details with extended model explained above.



CSS data pipeline [Nayak et al, 2021]

*Figure: Flow diagram for automation of data selection, pre-processing and transformation steps especially data pipeline (*Nayak et al, 2021*) extended with Roll up of site level to country pipeline highlighted in blue*

We have added a new tab to our CSS Spotfire dashboard, titled "Country Ranking Summary," which provides displays for various key metrics. These metrics include overall country rank, country signal panel, country signal details, and site average graphs to help visualize the metric.

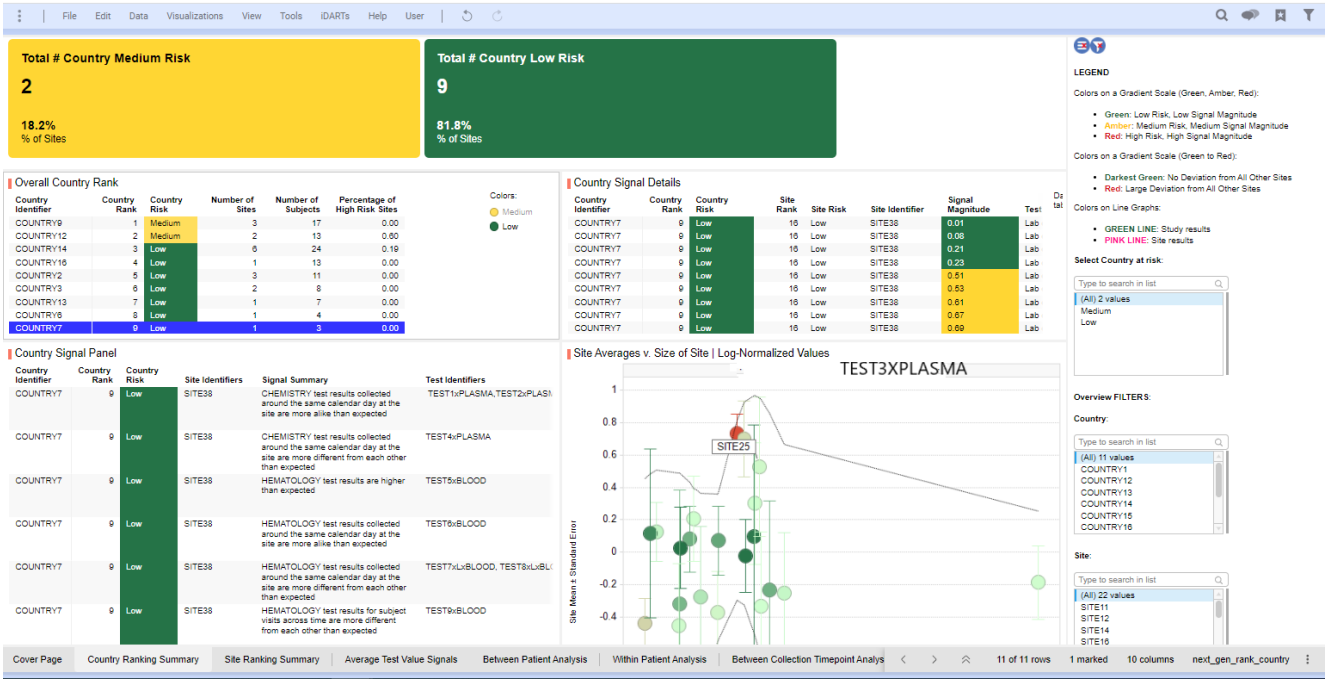


Figure: CSS dashboard with country tab showing the flow according to end-user functionality

CSS Country level capabilities

CSS pipeline is built in a modular manner employing the latest machine learning methods to investigate the study data at the various levels with brief description below:

A) Initial examination of data quality

The initial assessment of data quality involves evaluating the degree of variation in subject enrollment between individual countries and sites compared to all other countries and sites. This analysis seeks to understand how the study is progressing at specific sites or within countries in comparison to its overall progress. The objective is to assess the potential impact of crisis events at local or regional levels on the overall progression of study data in a clinical trial.

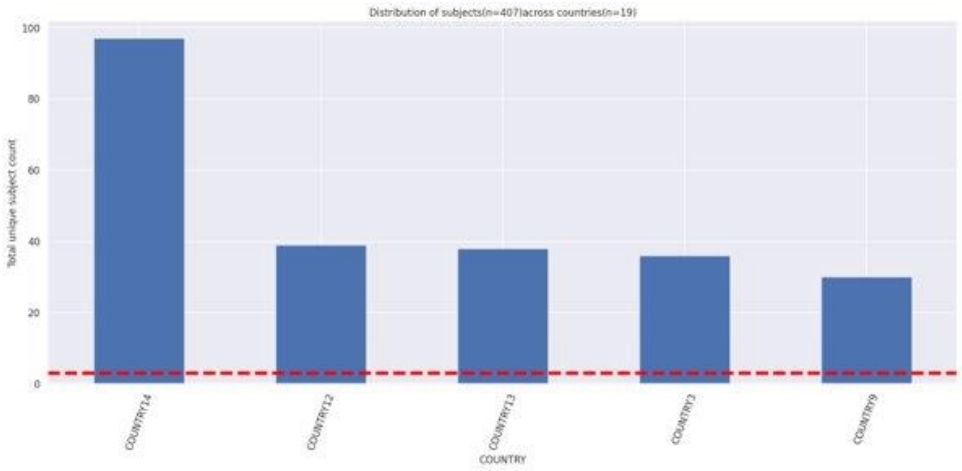


Figure: Country level distribution of subject enrollment and comparative counts assessment by red dotted line to show minimum data thresholds.

B) Examination of Visit progression between countries

It gives an estimate of the variability between subjects at site level within country and compares that with all other countries sites.

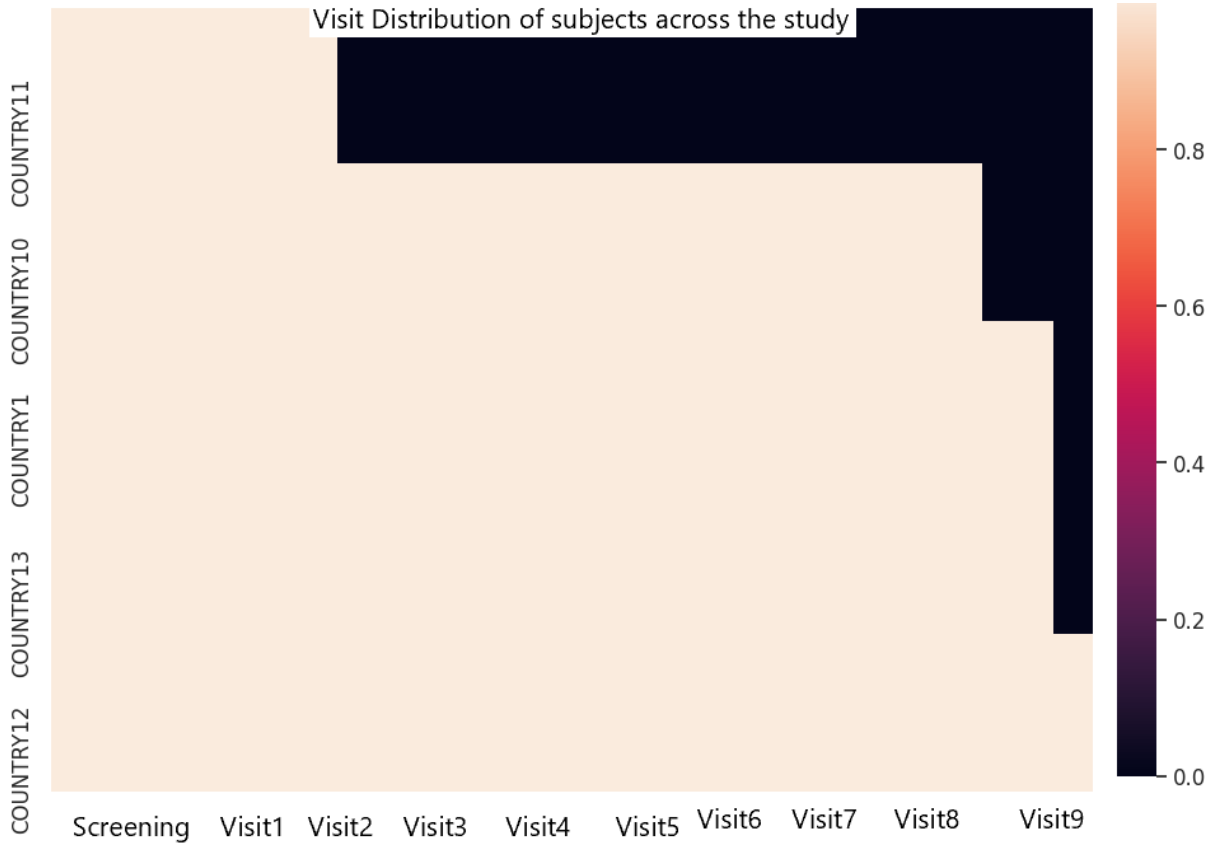


Figure: Results to compare between countries subject visit variance, individual country data records are compared to the average study norm. LEGEND: Red gradient within heat map plot indicates high subject visit counts compared to black gradient for low subject visit counts.

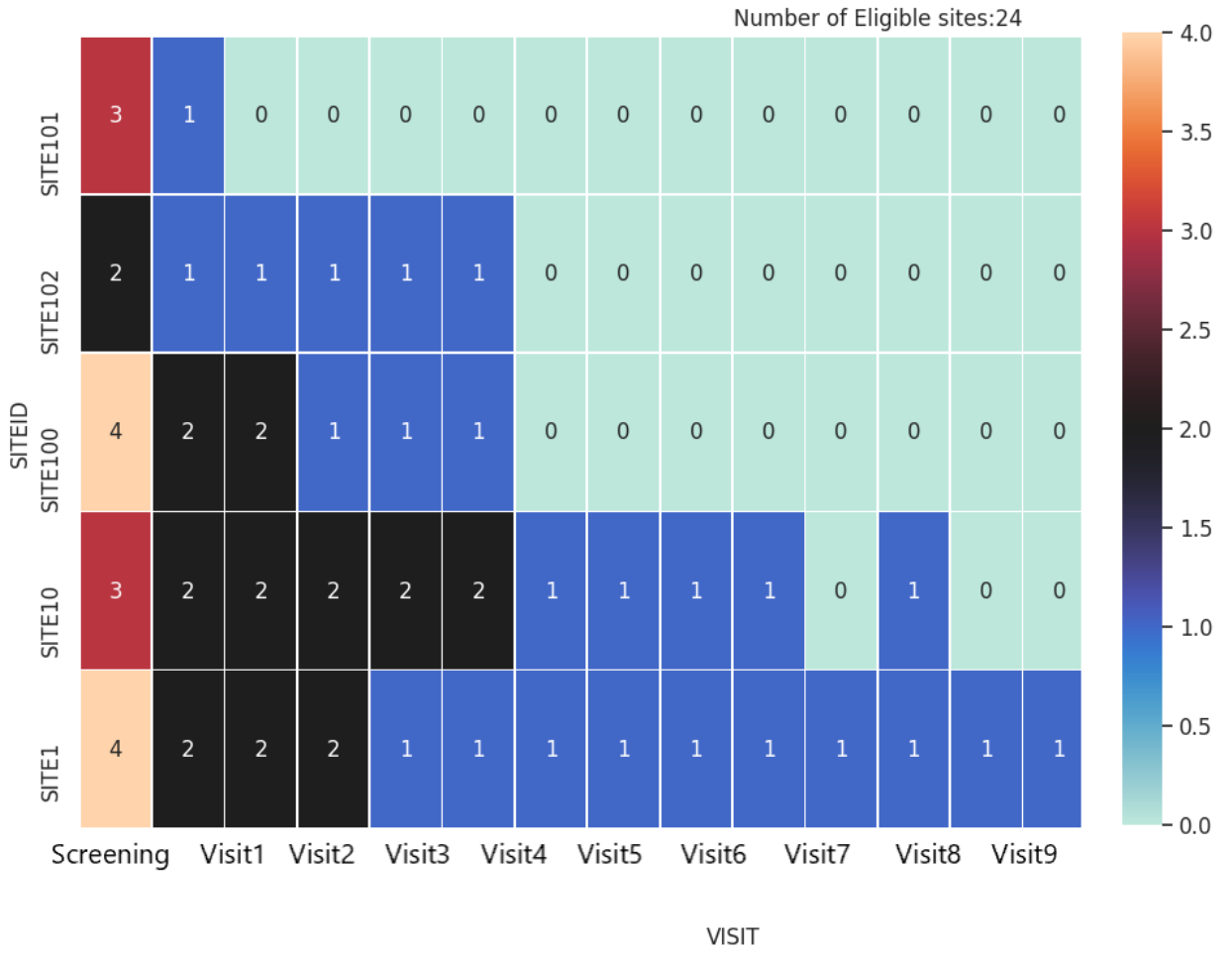


Figure: Subject enrollment at each country LEGEND: Dot Color: darker red color indicates high degree site subject enrollment at that country data differs from all country's sites in the study.

C) Examination of WITHIN SUBJECT ANALYISIS between countries

Within-subject analysis reveals differences in subject visits among countries in terms of variability in screen-failure rates, missing visits, and missing measurements for internal review. It provides an estimate of the variability between subject visits at the site level within those countries and compares it with all other sites in different countries.

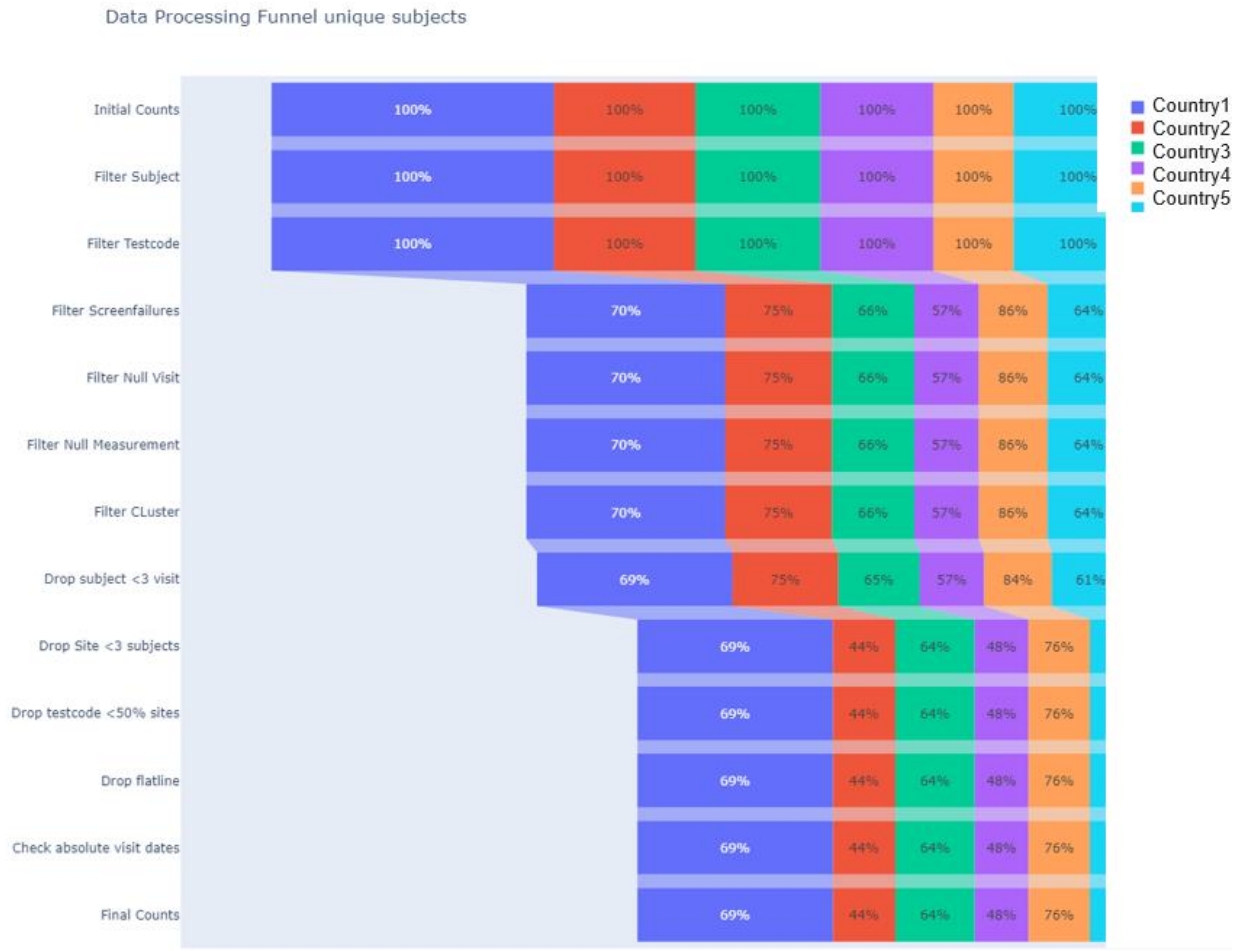


Figure : Results of within subject variance are different compared to the study norm.
LEGEND: Colored gradients: subject assessment results across all sites in between countries in a study; Highlights differences between screen-failure rates, missing visits, missing measurements, etc. included in the domain mapping file for internal review indicating differences in the pace at which progression of clinical study occurring in different countries.

Use case 1: Country Review functionality – study case 1, Country impacted by regional crisis in 2022

Country 1 was identified with a high-risk signal, suggesting the possibility of multiple sites exhibiting similar trends. Subsequent analysis of the highlighted parameters (Questionnaire data) was conducted and revealed that two sites displayed a lack of variance between subjects and across visits, resulting in significantly lower average results than expected throughout all study visits. These issues are followed-up accordingly and necessary resolution, feedback from the local study team is pursued for corrective and preventive actions.

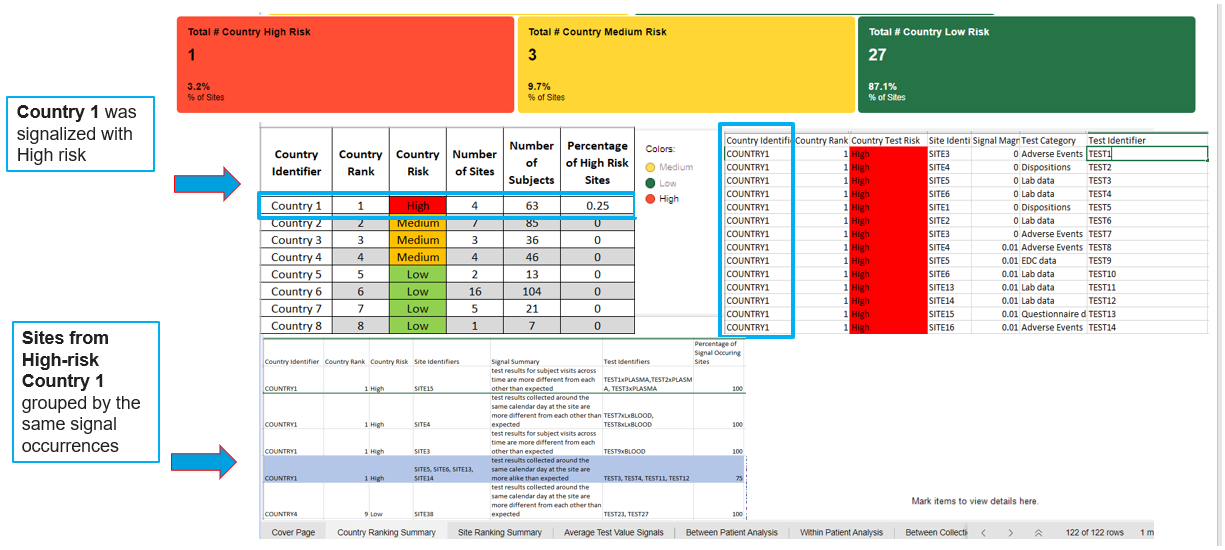


Figure: To show first use case example of high-risk country

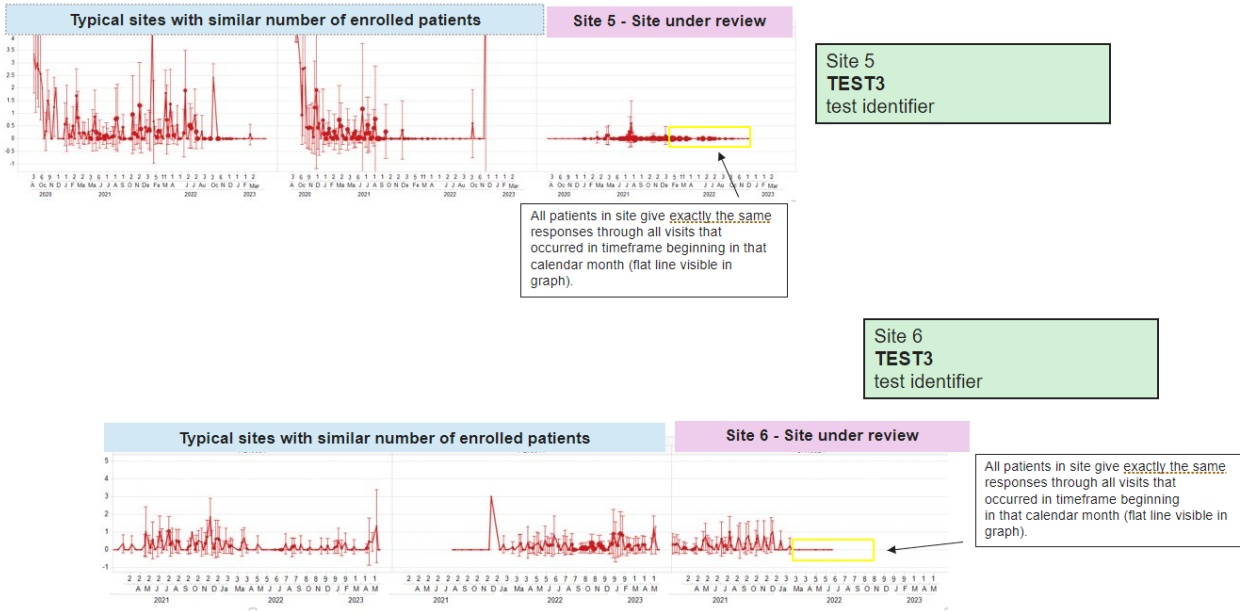
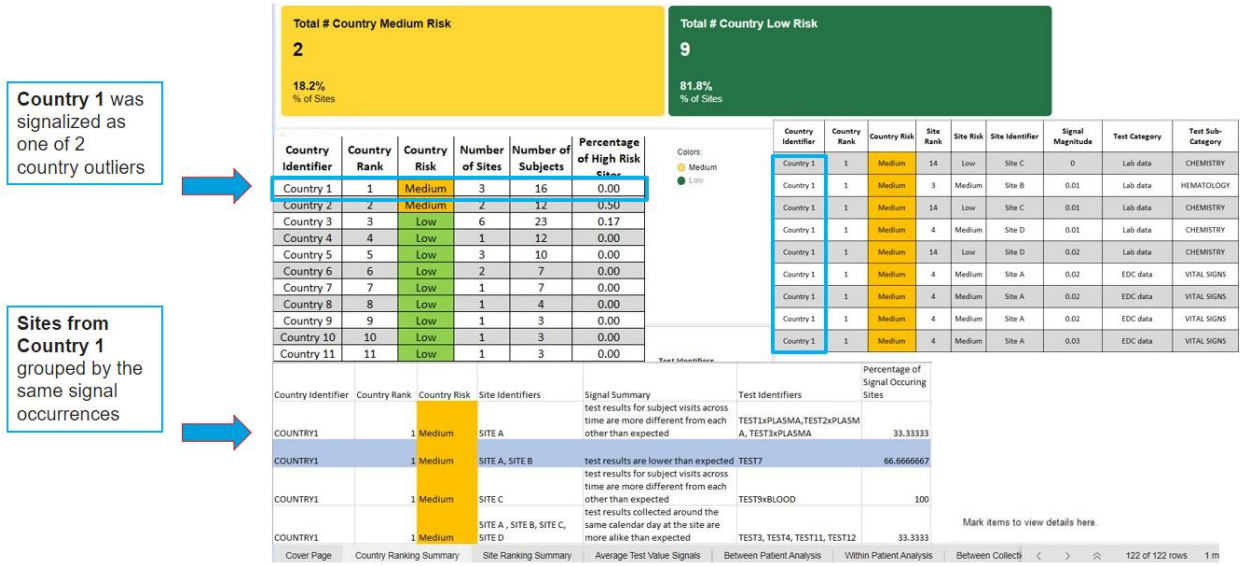


Figure: Both graphs show investigated sites (Site 1 in upper visualization & Site 2 in the bottom one) signalized with lack of variance for TEST3 (questionnaire related to efficacy assessment). In each graph the investigated site is marked in pink and named ‚site under review'. Two other sites in blue are other sites in this study that are similar in size and that are typical meaning that in these sites no issues with lack of variance were observed. We can see that in sites under review, during time period that is marked with yellow rectangle, there is no variance between subjects (no error bars) and across study duration visualized based on calendar dates. That means that the scoring has been exactly the same for all subjects during all visits starting with February/March in that respective year and this trend continued until the moment of preparing review.



Country Identifier	Country Rank	Country Risk	Number of Sites	Number of Subjects	Percentage of High Risk Sites
Country 1	1	Medium	3	16	0.00
Country 2	2	Medium	2	12	0.50
Country 3	3	Low	6	23	0.17
Country 4	4	Low	1	12	0.00
Country 5	5	Low	3	10	0.00
Country 6	6	Low	2	7	0.00
Country 7	7	Low	1	7	0.00
Country 8	8	Low	1	4	0.00
Country 9	9	Low	1	3	0.00
Country 10	10	Low	1	3	0.00
Country 11	11	Low	1	3	0.00

Country Identifier	Country Rank	Country Risk	Site Rank	Site Risk	Site Identifier	Signal Magnitude	Test Category	Test Sub-Category
Country 1	1	Medium	14	Low	Site C	0	Lab data	CHEMISTRY
Country 1	1	Medium	3	Medium	Site B	0.01	Lab data	HEMATOLOGY
Country 1	1	Medium	14	Low	Site C	0.01	Lab data	CHEMISTRY
Country 1	1	Medium	4	Medium	Site D	0.01	Lab data	CHEMISTRY
Country 1	1	Medium	14	Low	Site D	0.02	Lab data	CHEMISTRY
Country 1	1	Medium	4	Medium	Site A	0.02	EDC data	VITAL SIGNS
Country 1	1	Medium	4	Medium	Site A	0.02	EDC data	VITAL SIGNS
Country 1	1	Medium	4	Medium	Site A	0.02	EDC data	VITAL SIGNS
Country 1	1	Medium	4	Medium	Site A	0.03	EDC data	VITAL SIGNS

Country Identifier	Country Rank	Country Risk	Site Identifiers	Signal Summary	Test Identifiers	Percentage of Signal Occuring Sites
COUNTRY1	1	Medium	SITE A	test results for subject visits across time are more different from each other than expected	TEST1xPLASMA,TEST2xPLASMA, TEST3xPLASMA	33.33333
COUNTRY1	1	Medium	SITE A, SITE B	test results are lower than expected	TEST7	66.6666667
COUNTRY1	1	Medium	SITE C	test results for subject visits across time are more different from each other than expected	TEST9xBLOOD	100
COUNTRY1	1	Medium	SITE A , SITE B, SITE C, SITE D	test results collected around the same calendar day at the site are more alike than expected	TEST3, TEST4, TEST11, TEST12	33.3333

Mark items to view details here.

Cover Page | Country Ranking Summary | Site Ranking Summary | Average Test Value Signals | Between Patient Analysis | Within Patient Analysis | Between Collecti | 122 of 122 rows | 1 m

Figure: Country 1 was flagged with a medium-risk signal, indicating the potential presence of multiple sites exhibiting similar trends. Subsequent analysis of the highlighted parameters was conducted, revealing that two sites showed a lack of variance between subjects and across visits, resulting in significantly lower average results than expected throughout all study visits.

The issue was subsequently followed-up with local study team to apply corrective and preventive actions.

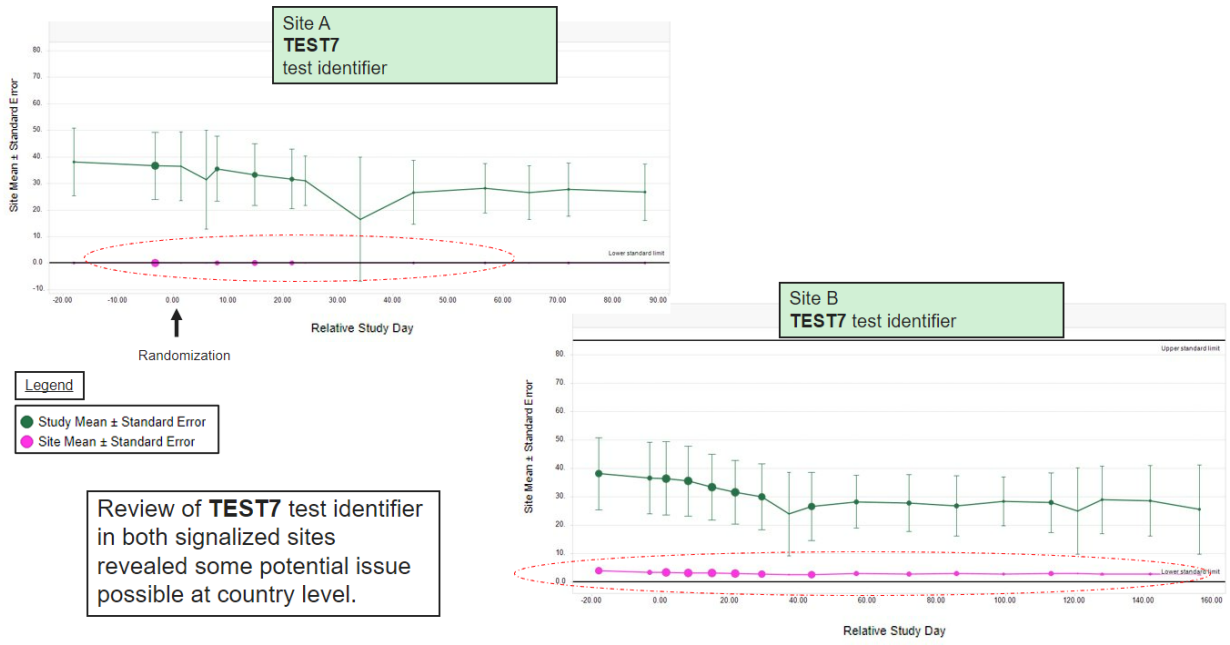
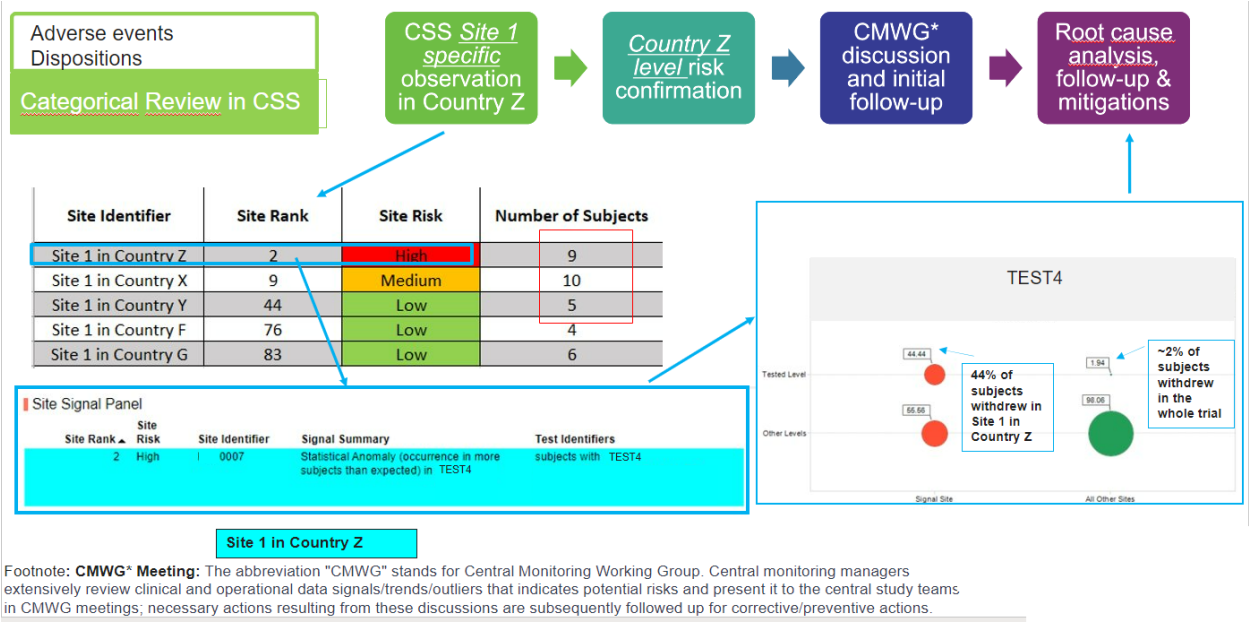


Figure: Both graphs display data from investigated sites within the same country (Site A in the upper visualization and Site B in the lower one). In each graph, the investigated site's average across the study duration is represented by the pink/purple line (also circled in red), while the study-level average is depicted by the green line. It is apparent that in both sites, the TEST7 test results are significantly lower than the study average (almost approaching 0).

Moreover, there is a lack of variance between subjects, indicated by the absence of error bars in the pink/purple line, and across all visits for this test. The pink/purple line remains flat with no fluctuations across the entire study duration (the x-axis denotes Relative Study Day, indicating the day within the study duration).

Country Review functionality – study case 3



Footnote: **CMWG* Meeting:** The abbreviation "CMWG" stands for Central Monitoring Working Group. Central monitoring managers extensively review clinical and operational data signals/trends/outliers that indicates potential risks and present it to the central study teams in CMWG meetings; necessary actions resulting from these discussions are subsequently followed up for corrective/preventive actions.

Figure: Site 1 in Country Z was flagged in the CSS dashboard due to a very high percentage of patients choosing to withdraw from the trial when compared to the distribution of this discontinuation reason across the entire study. Specifically, 4 out of the 9 patients enrolled (which is 44%) at Site 1 in Country Z opted to withdraw from the trial. In contrast, across the entire study, only 9 out of 465 subjects (2%) made the same decision.

A more in-depth analysis by CMM revealed a country-specific signal, as more withdrawals were observed in other sites within this country.

Country Review functionality – study case 3 follow up process

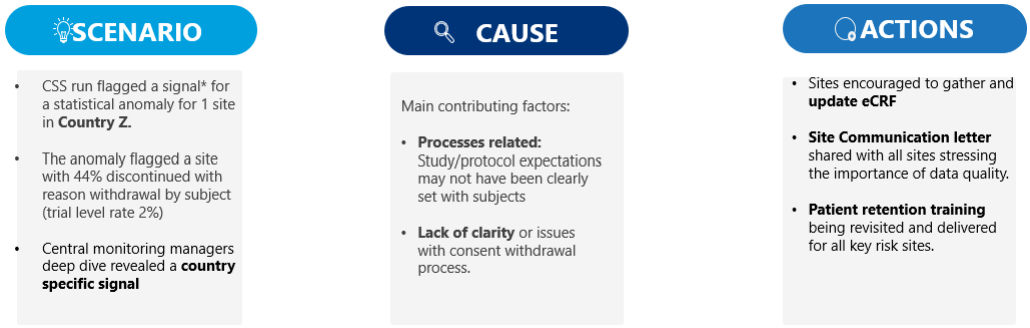


Figure: The signals identified from the CSS review were utilized for follow-up investigation, which emphasized that this country had a higher withdrawal rate, possibly due to issues such as unclear protocol and lack of clarity. Since then, corrective and preventive actions were implemented. These actions included updating the eCRF and site communication letter, ultimately improving patient retention through training.

CONCLUSIONS

In summary, our tool excels in identifying and comparing diverse data points across different countries, meticulously assessing their associated risk levels (categorized as high, medium, or low) in different levels of data dimensions. This robust functionality empowers users to proactively take corrective actions, bolster data quality, and facilitate real-time monitoring, making it especially valuable for detecting anomalies related to conflicts or pandemics.

Our tool offers in-depth, country-specific analysis that not only enhances operational efficiency but also streamlines the data review process, providing deeper insights into the information at hand. It enables a comprehensive evaluation of data quality across multiple levels, guiding users through tabular listings, facilitating dashboard reviews, automating data ingestion, simplifying pattern selection through intuitive dashboard interactions, and enabling near real-time data correction communication.

This aligns perfectly with our overarching objective of expanding the scope, accuracy, and efficiency of our CSS services.

REFERENCES

1) Lauwers et al, 2018: Central Statistical Surveillance in Clinical Trials. Kris Lauwers, Stefan Van Yper, Giordano Lo Regio, Janssen R & D
https://phuse.s3.eu-central-1.amazonaws.com/Archive/2018/Connect/EU/Frankfurt/PAP_AR04.pdf
2) Nayak et al, 2021: Programming infrastructure to automate process flow for Central Statistical Surveillance. Archana Nayak, Salvatore Saporito, Kris Lauwers, Stefan Van Yper, Mateusz Hinzmann, Bhargava Reddy, Janssen R & D
https://phuse.s3.eu-central-1.amazonaws.com/Archive/2021/Connect/EU/Virtual/PRE_AR03.pdfc
3) George et al, 2023: Quantifying the impact of the COVID-19 pandemic on cancer center clinical trial operations
https://pubmed.ncbi.nlm.nih.gov/37467065/

ACKNOWLEDGMENTS

Extending our sincere thanks to the excellent Janssen leadership team and all members of Advanced insights, Data science, Central Monitoring team, Data Management teams, Central Data science platform team, R&D Advance Computing- Data Analytics & Engineering teams within Janssen and FLEX staff supporting with the dynamic CSS application activities and ongoing collaboration.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Archana Nayak
Janssen R & D
High Wycombe, UK
Email: anayak32@its.jnj.com

Md Umarfaruque
Zeus,
Janssen R & D
umd@its.jnj.com

Kris Lauwers
Janssen R & D
Beerse, Belgium

KLAUWER2@its.jnj.com

Weronika Stepczyk
Janssen R & D
Warsaw, Poland

Susan Meier
Janssen R & D
PA, USA
smeier11@its.jnj.com

Marion Wolfs,
Janssen R & D
NL
mwolfs1@its.jnj.com

Bhargava Reddy
Janssen R & D
PA, USA
BReddy1@ITS.JNJ.com

Disclaimer

This presentation is for informational purposes only and does not represent professional guidance or advice. Any views and opinions expressed during this presentation are those of the presenters and do not necessarily reflect the views or policies of Janssen Research & Development, LLC, or any other company in the Johnson & Johnson Family of Companies.