A Central Laboratory Interlaboratory Comparison Program to Assess the Comparability of Data of Forty-one Tests from Four Regional Laboratories Involved in Global Clinical Trials over a Twelve Month Period

Christopher R Daniels PhD[1], Venkat R Rajagopal MS[1], Traci Turner MD[1]

[1]Medpace Reference Laboratories, Cincinnati, Ohio, US

Objective: The objective of this study was to develop and support a process for assessing the comparability of data used in global clinical trials from four individual laboratories, wholly owned by the same central laboratory entity, and that the same tests on the same samples are under statistical control and acceptable limits of variation.

Relevance: Data from central laboratories have been key for assessing safety, tolerability, and efficacy of new drugs in clinical trials. With the increasing complexity and global scale of many clinical trials, it is important to maintain harmonization among the regional laboratories as part of a central laboratory participating in the same study. This is usually accomplished by maintaining standardized operating procedures (SOPs) and identical assay platforms, calibrators, and quality control material across all laboratories. However, continuous monitoring of the same samples run at all laboratory locations may not be common practice.

Methodology: The interlaboratory comparison program was established in 2006 and involves locations in the US, Europe, Singapore, and China. The program involves multiple platforms to assess over 40 analytes tested globally, encompassing several therapeutic areas including Oncology, Cardiometabolic, Infectious disease and others. Pooled serum/plasma/urine samples were aliquoted, frozen at -70ºC, and distributed quarterly to each laboratory for weekly analysis. Here we present the dataset for forty-one tests using multiple platforms. Samples were analyzed weekly on the same day at each laboratory for twelve months (N= 520). College of American Pathologist evaluation limits, the Westgard Biological Variation Database, and the Royal College of Pathologist of Australasia Quality Assurance Program were sources used to establish acceptable bias criteria. The percent and absolute bias was calculated for each result using the US laboratory as the reference laboratory. The percent bias for each week and month was also calculated. A Bland-Altman plot was created between each laboratory and the reference laboratory for the six month period, and a student T-test run using a relative bias limit for each test and a significance level of 5%.

Results: Overall, during the twelve month period, all forty-one tests had a twelve month mean bias within the acceptable bias limit for each individual test compared to the reference laboratory. When comparing the mean weekly bias for each tests from all laboratories, any week where the bias was outside acceptable limits, investigation and corrective action was undertaken to determine the source of the error. Examples include:

- One week during the six month period three chemistry tests from the same laboratory had a mean weekly bias outside the acceptable limit. After investigation it was determined that a pre-analytical issue with thawing/mixing of frozen samples was the cause of the bias.
- A negative bias was observed in US laboratory for triglyceride. However, subsequent review of CDC Lipid Standardization Program Part III data from all laboratories globally demonstrated acceptable performance.
- Inconsistencies in the reporting of results across laboratories, specifically the technical decision to not report data deemed biologically implausible, were noted after monthly review of calcium data, and all technologists were re-educated on handling repeat analysis and appropriate consultation of Laboratory Directors.

Conclusion: An interlaboratory program where frequent monitoring of identical samples run at all laboratories involved in clinical trials can provide valuable information into the harmonization of data reported by the central laboratory, and help mitigate pre-analytic, analytic, and post-analytic issues that may arise when assessing data used in the development of new therapeutics.

Introduction

- Data from central laboratories are key for assessing safety, tolerability, and efficacy of new drugs in clinical trials.
- The Interlaboratory Comparison Program was set up among all laboratories wholly owned and purpose built by Medpace Reference Laboratories (MRL), a global central laboratory, in 2006. The locations include Cincinnati, OH (US), Leuven, Belgium (BE), Singapore (SG), and Beijing, China (CN).
- Continuous monitoring of the same samples run at all laboratory locations are important to make sure no matter where the clinical trials samples are run similar results are obtained.

Methodology

Samples

- Pooled serum, plasma, and urine samples were aliquoted, labeled with the appropriate sample information, frozen at -70°C, and distributed quarterly to each laboratory for weekly analyses on the same day.
- The pooled samples were from participants in clinical trials. All samples were received de-identified of demographic information.

Analytical methods

- Sample analysis was performed on 41 tests (Table1) at each Medpace laboratory location in weekly batches of 10 samples, and were analyzed on the same day in each laboratory. On the day of analysis, samples were thawed and thoroughly mixed. Appropriate Quality Controls were analyzed with each batch and results accepted based on global SOPs. Sample analysis took place over a 12 month period.
- Acceptable results were electronically transferred into ClinTrak Lab®, an in-house developed, clinical trial management system.

Statistical methods

- Summary statistics listing the bias criteria, number of samples and mean difference (absolute or relative) compared to the reference laboratory (MRL-US) were included. The P-values are from TOST t-test performed on absolute or relative difference.
- A Bland-Altman plot was created between each laboratory and the reference laboratory for the 12 month period (data not shown).
- Bias criteria were based on criteria from the College of American Pathologists, Westgard Biological Variation Database, and Royal College of Pathologists of Australasia.
- Some test and laboratory locations have an N of <520 for a given analyte due to insufficient quantity, instrument error, etc.

Results

- All tests run on the Beckman Coulter Chemistry Analyzers (Table 2), Roche Immunoanalyzers (Table 3), Siemens BNII Nephelometers (Table 4), by preparative Ultracentrifugation (Table 5), Stago Compact (Table 6), and Tosoh (Table 7) demonstrated acceptable equivalence when compared to the reference laboratory over a 12 month period. Representative examples of pre-analytic, analytic, and post-analytic issues detected over the twelve month period include:
- Monthly review of BUN data (Table 8) showed a mean bias for BE laboratory of -14.28%. Upon further review of other tests run on the same sample, a similar negative bias was seen (Figures 1, 2, 3). After investigation it was determined the cause of bias was due to a pre-analytical error, specifically improper mixing during the freeze/thaw process.
    - Technologists were re-educated on the proper freeze/thaw process when analyzing samples that have been stored at -70°C.
- A negative bias of ~4.5% for triglyceride was observed in the US laboratory compared to the other laboratory locations (Figures 4, 5, 6). The CDC Lipid Standardization program results for 2016 were reviewed for all four laboratories, and indicated good performance, with the bias averaging -2.29%, -0.51%, -0.57%, and -1.76% for the US, BE, SG, and CN labs, respectively when compared to CDC targets (Table 9).
- Reporting inconsistencies were noted between laboratories for calcium during the review period, with some labs reporting a numeric value, and other labs electing to not report a numeric value due to concerns the calcium values were biologically implausible, indicating a pre-analytic error affecting sample integrity. (Data not shown).
    - Corrective actions included re-education on the process to determine the validity of results, including, but not limited to, repeat analysis, delta checking, and consultation with the Laboratory Director as to clinical validity.

Conclusion

- An interlaboratory comparison program where frequent and continuous monitoring of identical samples run at all laboratories involved in clinical trials is conducted can provide valuable information into the harmonization of data reported by the central laboratory, and help mitigate pre-analytic, analytic, and post-analytic issues that may arise when assessing data used in the development of new therapeutics.

References

1. Clinical and Laboratory Standards Institute (CLSI) Estimation of Total Analytical Error for Clinical Laboratory Methods; Approved Guideline EP-21A Volume23 Number 20.
2. Westgard Biological Variation Database specifications, updated in 2014. https://www.westgard.com/biodatabase1.htm
3. National Cholesterol Education program (NCEP), Recommendations on Lipoprotein measurement by the Working Group on Lipoprotein Measurement. (Sept 1995) NIH pub:95-3044
4. National Glycohemoglobin Standardization Program (NGSP) (2009). CAP Survey limits lowered 2011.
5. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1: 307-310.
6. https://www.rcpaqap.com.au/wp-content/uploads/2013/06/chempath/Allowable%20Limits%20of%20Performance.pdf