REFERENCE INTERVAL DETERMINATION IN THE COAGULATION LABORATORY

Donna D. Castellone, MS, MASCP, MT(ASCP)SH
New York Presbyterian-Columbia
Disclosures

- *I have nothing to disclose in relation to this presentation*
- *No products will be discussed in my presentation.*
New York Presbyterian-Columbia Medical Center

Dcastell@nyp.org
Objectives:

- Identify components in coagulation testing that will impact reference intervals
- Analyze data collection and determine the statistics to best represent your patient population that you are testing.
- Differentiate between verification and validation of a reference interval.
- Apply recommendations from guidelines for specific coagulation tests.
Reference Intervals

- One of the most important studies conducted in the laboratory.
- Up to 80% of medical decisions are made based on laboratory test results.
- Problematic because many laboratories lack the time, resources, finances and in many cases the expertise to conduct these studies.
- Many reference intervals are obtain from either package inserts or from publications.
Current Laboratory practices:

- College of American Pathology surveyed 500 laboratories and noted that 78% used manufacturers published ranges.
- Q probe studies of 163 laboratories revealed only 5.5% of labs could recall the year they revalidated their aPTT reference interval.
- In another cohort of 116 laboratories, 42% established their own RI for PT and aPTT, however only half did it for every change of reagent lot.

Issues that complicate this process in coagulation

- Unstable reagents, not well standardized, testing unstable proteins
- Reagents have different sensitivities to factors based on the concentration and type of phospholipids
- Different activators
- Many different Instrument/reagent combinations
- Preanalytical variables
- Most important is the selection of subjects and controlling pre-analytical variables
Selection of Subjects
What is “normal”?  

- Adults  
- Pediatrics  
- Geriatrics  
- Ethnicities  
- Gender
What is a reference interval?

Reference Intervals are defined in relation to a healthy population to include the values in which 95% of apparently healthy individuals would fall and in which 2.5% of results in the lower range are out of the RI and 2.5% of values in the upper range will be out of the RI.
Selection of subjects

- A questionnaire should be used to capture relevant information including health status, age, gender, ethnicity and any medications - including OTC, vitamins and supplements
- Equal amounts of male/female
- Representative of patient testing population
- A priori or a posteriori approach
- The more defined the population the better the outcomes
What about normal patients?

- Apparently normal patients should not be used
- May have acute phase reactants (I, VIII, vWF and fibrinogen)
- Inversely proportional—corresponding tests will be shorter, falsely shortening your RI
- Outcome: falsely classifying patients as abnormal
- Unnecessary work-ups
Biological Variation

■ Defined as the variation seen in an individual subject when they are measured for an analyte repeatedly over time.
■ Variations can be influenced by stress, APR, circadian rhythms and seasonal variations
■ Difference is the between subject or inter- individual variation
■ Studies are limited
■ Most RI testing is conducted in the morning, patient blood draws occur throughout the day- clinic samples tend to arrive later in the day
■ Does this impact RI reflection on patient values.

Studies on Biological Variation:

■ Chen et al: 16 coag tests (PT, aPTT, fib, TT, FDP, II, V, VI, VII, VIII, IX and X)
■ n=31- tested on days 1,3, and 5 at 8am, 12pm and 4pm.
■ All screening CVI < 5% exception of fibrinogen
■ CVI for coagulation factors >5%
■ Lupus, AT, PC, S, Vw (n=25) CV < 3%

What information from Biological variations

- Can aid in the evaluation of results obtained in RI
- If a large variation in results are seen for a particular test, it may be worthwhile to understand if the results may be caused by known BV for that analyte.
Age specific ranges

- Newborn and childrens’ hemostatic systems are immature in comparison to adults and older children
- Basing pediatric coagulation results on adult ranges can cause a result to be misclassified as abnormal
- Generate a diagnosis or treatment that may not be warranted.
Problems with these ranges:

- Obtaining blood from neonates and children problematic
- Parental consent
- Insufficient volumes for testing
- Laboratories used published data
- Manufacturers published ranges
- Ranges should be adapted from the same instrument/reagent combinations
Pediatrics Reference Intervals:

- Study of 218 healthy children stratified by age:
  - II, IX, XI and XII significantly decreased in the youngest children (< 12 months)
  - PC and PS decreased in young childhood
  - Highest levels vWF in youngest children, but not FVIII

- ARUP study - n=902 7-17 yrs old: each group n=164
  - PT testing 1 second longer than adults, aPTT not significant
  - Also confirmed age dependent ranges in FVIII, IX and XI and vW testing (activity and antigen)

Data mining; using EMR’s

- Hard to obtain populations - geriatric and pediatric
- Use of EMR - larger amount of data - stratification for age groups
- Global reference intervals over a large regional area
- Expanded sample size
- More challenging in coagulation due to different instrument/reagent combinations
- Difficult to control pre-analytical variables.

Gollomp K, Arulselvan A, Tanzer M, Shibutani S, Lambert MP, Blood. 2015;126:4450
EMR ranges: Pediatrics (n=265)

- Patients (n=265) were excluded based on diagnosis or medications that could impact coagulation testing results
- Established PT = 11.6-13.8 sec
  post EMR review: PT = 12.9-13.9 sec
- aPTT = 22-35 sec
  Post EMR review aPTT = 25-35 sec
- No difference in gender was found
- Age: 2-11 yrs. PT = 12.5-13.6 sec
  12-23 yrs. PT = 13.05-13.9 sec with no significant difference in aPTT.
- Important to understand the CV of the individual test to determine if these results are significant
- Results may be statistically significant, but not clinically significant.
Number of subjects
Validation versus Verification

- Validation requires a minimum of 120 subjects
- Verification can use as little as 20 subjects to demonstrate test performance from a previous claim.
- Statistical evaluation of a RI are based on the number of subjects used.
- Most common tests PT, aPTT, TT, Fibrinogen and D-dimer
- Performed:
  - Change in reagent
  - Lot number
  - Instrument
  - Collection system

CLSI document H47-A2; 2008.
Statistics:

- The RI is the range between an upper and lower limit which represents a percentage of the population tested.
- The mean and the standard deviation (SD) can be calculated. The SD is the spread of data around the mean.
- The more dispersed the data, the higher the deviation.
Confidence interval

- Measures the level of uncertainty
- 95% CI: ranges are measured at 2.5 and 97.5 percentiles of the distribution of results – contains the true mean of 95% of the population
- 99% CI – ensures 99% certainty
- The higher the confidence levels at the wider the CI
- Increasing the CI from 95-99% to ensure the interval contains the population mean will increase the sample size
- Using a smaller number of RI data should be statistically evaluated to see if they fit normal distribution or if it is skewed.
- Skewed data needs to be normalized by log transformation prior to calculating and converting it back
Statistical methods for evaluation:

1. Parametric method: used when population is normal or Gaussian, if not a statistical transformation to normalize the data is applied

2. Non-parametric method: used when careful subject selection and sufficient number (120) data is collected, doesn’t require laws of probability

3. Robust method: used in a limited sample size without requiring a Gaussian distribution- measures the position and dispersion instead of mean and SD. Sorts the data from lowest to highest in equal parts and looks at how far values are distributed from the center
Gaussian distribution

Skewed Data
Transformation of data

- Replacement of a variable by either the square root or the logarithm of the variable changes the shape of a distribution or relationship.
- Done when data is skewed to transform the data into symmetrical distribution.
Outliers

- Method by Dixon: Looks at the difference in the point which appears to be an outlier and the next observation - (D)
- This is divided by difference of the lowest observation and the highest observation (R)
- If the D/R ratio is greater than 1/3, this point may be an outlier.
- In a sample size of 20, allowed 2 outliers
- If > 2 outliers, must test an additional 20 normal
- If another 2 outliers other sources of errors (reagents, analyzers, biological variations) should be investigated
- A full RI may have to be conducted.

Transference

- Validation of a RI conducted by a manufacturer
- Smaller number of subjects n=20
- Compared to a larger study
- Keeping in mind the importance in coagulation using the same analytical test system and reagents
- If a laboratory changes methods and the method comparison is compatible, the RI can be transferred
Subjective validation

- All relevant processes of the original study including population demographics and geographics are described.
- Pre-analytical and analytical information must be provided.
- If the review is considered compatible with the testing laboratory, the RI can be used.
- The review needs to be purposeful and documented.
Recommendations & Guidelines:
Issues with Coagulation Testing:

- Testing is expensive
- Some procedures are labor intensive- platelet aggregation testing
- Is a reference interval required, or does there need to be a cutoff
- Can I use representative assays- if the method comparison from an aPTT matches to the previous lot, can the assumption be made that the aPTT based factor ranges are the same?
Overview by test:

**Routine Coagulation tests:**
- Initial validation (n=120)
- Careful selection of patient population
- Subsequent lot changes; verified n=20

**PT/INR**
- Verify n=20
- Calculation geometric mean which is more representative of lognormal distribution of PT.

**Platelet Testing:**
- All agonists, aggregation, secretion
- n=20
- No substance impacting platelet function for 14 days
- Tested over time to reflect between run variability

**Factor Assays:**
- Transfer n=20 (original study must be comparable)
- n=60 – larger number with more statistical power to discern population differences
Overview by test:

D-dimer:
- RI versus a cutoff
- RI is used to diagnosis conditions other than VTE; exclude subjects >60 yrs due to increased levels
- Cutoff is for exclusion of VTE - determine by manufacturer-FDA approved - imagining studies vs. d-dimer in patients with low to moderate pre-test probability

Lupus Testing:
- N=40- confirm Gaussain distribution
- Upper limit of RI determines cut off
- RI mean can be used to normalize assay
- DRVV screen, confirm, ratio and hexagonal phase
- Transform comparability from one lot to another
Plan, plan, plan the study
Pre-Analytical: 50-70% of errors occur here!

- Have a well written SOP that includes relevant stakeholders including phlebotomy, time frames and where and when the study will be conducted.
- Will this be a verification or validation? How many samples will be required? How much volume is required? Will additional volume be required for freezing?
- What is your institutional policy? Is there IRB approval? Do you have a consent form?
- Do you have a questionnaire, will the study be a priori or a posteri?
- Who will recruit/consent/draw the subjects?
- Do I have sufficient representation of males/females as well as ethnic groups that represent your testing population?
- What is exclusion criteria? Is this for only adults? Children? Pediatric?
- Are samples drawn for platelet testing—excluding patients on aspirin and aspirin containing products.
Analytical:

- Do you have sufficient reagents to conduct the study? How many sites? How many analyzers?
- Will reagent sensitivity be determined to determine if the reference interval for the PT/aPTT will be prolonged to reflect factor levels at 30% sensitivity?
- Are the analyzers working correctly? Should there be a preventative maintenance?
- Do the centrifuges produce platelet poor plasma?
- Will the study be conducted over a period of time to introduce analytical variables that are seen in patient testing? How many samples/day?
- Time frame: 4 hours aPTT, 24 hours unspun, capped RT for PT?
- Will all samples be run fresh? Will some be run frozen? Who will aliquot samples to be frozen?
Post Analytical

- What is considered an outlier?
- Who will do the statistical analysis? What type will be based on the amount of samples? Parametric, non-parametric?
- Geometric mean for the PT calculation of the INR?
- Report sign off
- Update LIS and downstream systems
- Alert clinicians to range changes.
RI requirements

- RI should reflect patient population testing
- If possible a full validation should be conducted (120)
- If not possible, the level of uncertainty the laboratory is willing to accept needs to be considered.
- Method for determining outliers
Conclusion

- Carefully planned and documented event
- The more defined your population and subject selection and the more controlled your pre-analytical variables are the better the outcomes
- If a RI is to be transference, should be compatible for reagent/instrument combination
- Ensure best possible results to provide tools for clinicians to diagnose coagulopathies and provide optimal treatment to patients