CTTI Pregnancy Testing in Clinical Trials: Summary of Day 1

Evan R. Myers, MD, MPH
Goals for the Meeting

• Review
  • Why we do pregnancy testing in clinical research
  • Methods for doing pregnancy testing in clinical research
    – Types of tests
    – When to test
    – How are decisions about methods being made now?
  • “Comparative effectiveness” of different methods

• Feedback and input
  • What important general principles should be considered in designing pregnancy testing protocols?
  • What information/guidance would be most useful to the research community, and, ultimately, research subjects?
  • What resources would be most helpful for helping disseminate information/guidance?
  • Are there major evidence gaps that should be addressed through specific research?
Overview

• Key Points from Presentations
• Key Points from Breakout Sessions
Session I

Topics
• Rationale for Pregnancy Testing in Research
• Technical Aspects of Pregnancy Testing
• FDA and Pregnancy Testing

Questions to Consider
• Is an approach that tries to define the acceptable risk of a false negative test on a study-by-study basis reasonable?
• What criteria should a specific test meet in order to be considered for use in clinical research?
• How should those criteria be demonstrated, and who should document it?
• Is the use of home pregnancy testing ever acceptable, and, if so, under what circumstances?
I: Why We Test

• Females of reproductive potential participate in all phases of clinical research
• High background risk of fetal loss, congenital anomalies
• Teratogenicity can happen any time during pregnancy
• Even with preclinical testing, most late phase trials start with high degree of uncertainty about specific risk of teratogenicity
• Focus on clinical trials is on
  – Preventing exposure
    • Contraception
    • Pregnancy testing before enrollment and study intervention
  – Minimizing duration of exposure
    • Intermittent pregnancy testing after enrollment
II: Measuring hCG

- Multiple variants of hCG → potential impact on sensitivity and specificity of testing
- Lack of standardization in assays → variation in measured concentrations, detection of clinically relevant hCG variants
  - Variability in quantitative measurements
  - “Variant hook” → false negatives in later pregnancy
    - Consistency choice of test during study
- Analytical sensitivity varies with brand (home and POC)
- Cutoffs often not in agreement with manufacturer’s claims/package insert
  - Usually lower levels than reported
II: Measuring hCG

• Timing of testing relative to start of pregnancy affects how early hCG can detect pregnancy
  – Patient estimate of when menses will occur may be inaccurate to normal variation in cycle length

• False positive hCGs can occur
  – Interfering antibodies
  – Pituitary hCG (ovulation or perimenopause)
    • FSH can help discriminate
  – Exogenous hCG
III: FDA Regulation of Pregnancy Tests

- Class II device
  - Approved under 510(k)
  - “Substantial equivalence” to predicate device
- Cutoff for claims
  - Concentration that yields 50% positives and 50% negatives
- Newer tests include examination for variant hook effect from β-core fragment
- Typically 100 subjects used to establish clinical “substantial equivalence”
- Few reports of inaccurate results in postmarketing surveillance
Session II

• Topics
• Current practices
  – One sponsor’s experience
  – Survey results

• Questions to Consider
• Is the evidence that there is variability in current approaches to pregnancy testing strong enough to justify attempts to create greater consistency?
• Are there best practices that we can point to?
• What are the trade-offs between standardization and flexibility?
IV: Industry Perspective

• No plan for minimizing pregnancy exposure is perfect
  – Estimates from Phase III studies <0.1 percent/cycle

• Pregnancy tests intended for use in diagnosis (suspected pregnancy), not screening
  – Expect different positive and negative PVs
  – Need for confirmatory testing increases burden on subjects, investigators, and sponsors
IV: Industry Perspective

• At one sponsor, all interventional clinical studies enrolling females of childbearing potential require pregnancy testing
  – Prior to enrollment unless intent to enroll pregnant subjects
  – Post-enrollment periodic testing, with some exceptions
    • Intervention withheld for positive or indeterminate results
    • More rigorous protocol for known teratogenic risk
  – Most at end of treatment or early withdrawal
IV: Industry Perspective

• Outcomes of pregnancy testing depend on a complex process that involves more than just sensitivity/specificity and timing of test
  – Communication of results, actions taken based on results also key

• Guidance for pregnancy testing should be
  – Evidence-based (as much as possible)
  – Separate standards, recommendations, best practices and provide rationale
  – Need to consider operational feasibility at all levels
V: Survey Results

• 50% of respondents ➔ maximum acceptable risk of pregnancy < 1/10,000
  – Choices for testing options inconsistent with that standard
• NPV consistently rated most important consideration, followed by patient burden
• Variation in type of testing by risk to fetus, but home testing OK for 5-10% of respondents
• Most recommended
  – Continued post-enrollment testing unless very short duration study
  – After study, depending on PK
• Free text responses ➔ some consider age, contraceptive method
Session III

Topics
• Comparing estimated outcomes of different testing strategies

Questions to Consider
• Is this a useful approach?
• If so, are there ways to make the model more accurate and useful?
• If modeling results are useful, what is the best way to provide access to them (e.g., presentation of results for common scenarios vs. allowing users to run their own scenarios?)
VI: Model Parameters

- Subject Age
- Hysterectomy Status
- Menopausal Status
- Menstrual Cycle Characteristics
- Age-specific Contraceptive Method use
- Pregnancy Outcome Probabilities
- Contraceptive Effectiveness (Typical Use)
- hCG Levels in Non-pregnant women
- hCG Levels in Pregnancy
- Sensitivity of hCG assays
- Probability of detecting symptoms in the absence of testing
VI: Model results

• Fewer pregnancies with age
  – Fewer women of childbearing potential due to menopause, hysterectomy
  – Greater use of highly effective methods (particularly sterilization)
  – Lower probability of getting pregnant

• Fewer detected pregnancies when testing not performed relative to menstrual cycle
  – 9-10 day window when ANY pregnancy test will be negative
VI: Model Results

• False positive results
  – Increase with age
  – Only when threshold for positive test 5-19 IU/L

• Estimated absolute differences in false negative rates relatively small
  – Young women → difference between 5 and 20 IU/L about 5/1000, decreasing to 3/10,000 in perimenopausal women
Summary of Breakout Sessions
Are there other factors that should be considered in the modeling approach and for resource development?

- Typical vs perfect contraception use
  - *Model currently uses “typical” use;*
- Duration of pre-enrollment use less than 12 months
  - *Can be readily modified, but remember that estimates based on Pearl index will underestimate failure rate early, overestimate later*
- Duration of counseling
  - *Could incorporate, but need evidence that to associated with effectiveness*
- Co-morbidities
  - Motivation to avoid pregnancy
    - *Could accommodate in a variety of ways*
  - Inherent decreased fertility
    - *Could include, but need estimates for how much a given condition affects fertility*
Are there other factors that should be considered in the modeling approach and for resource development?

- Potential litigation risk (even in a simplified way)
  - Can estimate overall likelihood of miscarriage and length of duration of exposure now
  - Would need way to estimate likelihood of
    - Miscarriage or congenital anomaly conditional to exposure
    - Likelihood of litigation given miscarriage/anomaly
- Consider the phase of the study (in determining level of risk)
  - Wouldn’t need to be incorporated in the model—model outputs risk of pregnancy based on population and testing protocol → sponsor/regulator decides whether that’s acceptable
- Factor in 9 day window/timing of testing
  - Already in model, can explore alternative strategies to incorporate impact of 9 day window (e.g., home LH kits)
Are there other factors that should be considered in the modeling approach and for resource development?

- **Cost**: Is opaque. Recognize that cost is a factor that needs to be considered. Not binary choice, but do what’s best for trial – leave to judgment of trial designer.
  - *Could include cost as a user-modifiable variable*
- More **analytic information** in model (more granularity).
  - *If on-line “TurboTax” format, can allow as much granularity as desired*
- **Feasibility** (cognitive dissonance). Risk is very low, but drives us toward infeasibility. Separate risks of test with risk of human pregnancy – find most accurate pregnancy test, use serum test when you need to.
  - *Might be benefit of including cost ➔ forces decision maker to see consequences of trying to achieve very low risk of pregnancy*
- Data on initial test, but not follow-up testing throughout the trial. Can **follow-up testing** be modeled? Data available? Risk to pregnancy if identified late -- what is acceptable risk? Can there be a risk threshold? Specify risk and then model? Benchmarking? Reflect what risks are.
  - *Follow-up testing in model, can be extremely flexible with both timing and choice of test*
What resources do you currently use to develop pregnancy testing protocols for clinical trials?

- Standard of care for specified patient population
- Investigator/opinion leader recommendations
- Literature review
- ICH guidelines
- Institutional (sponsor, hospital) standards
  - How to resolve if in conflict?
- Institutional experience
  - Cognitive bias
What additional resources are needed to help support the development of pregnancy testing protocols for clinical trials?

- Major evidence gaps
  - Risk of pregnancy in pre-approval studies
  - Risk factors for pregnancy in pre-approval studies
    - Leverage existing resources to get data
- General information on
  - Biology of reproduction and early pregnancy (background risk of miscarriage and congenital anomalies)
  - Biology of hCG (9 day window, levels during pregnancy, false positives, variants)
  - Performance characteristics of available hCG tests
  - Contraceptive effectiveness
    - Different levels for patients, study staff, investigators, sponsors
- Broad guidelines/recommendations/best practices based on above considerations
- Resources to ensure ongoing update of evidence, evaluation of model/recommendations

Expert Meeting—July 15/16, 2013, Bethesda MD
What are the patients’ perspectives on the proposed resources?

- Better information on (including degree of certainty)
  - Risks to fetus/embryo from
    - Exposure to study intervention
    - Maternal condition
  - Risks to mother from
    - Exposure to study intervention (i.e., risks changed because of physiologic changes in pregnancy)
    - Maternal condition (e.g., PAH, depression)
  - Risk of becoming pregnant during trial based on age and contraceptive method
  - Risks of pregnancy testing
    - False negatives and false positives
    - Additional burden above other required study activities
- Engage patients/potential subjects as partners early in the protocol development stage
  - PCORI?
How do you envision using these resources in designing pregnancy testing protocols for clinical trials?

- Define acceptable level of risk
  - *These resources will NOT do that—they will give estimates of WHAT the level of risk is.* Whether that risk is acceptable is a judgment, not a calculation
- Provide rationale for making decisions about testing
- Provide estimates of risk to allow testing to vary based on patient-specific risk
  - Across protocols
  - Within protocols

- Users
  - Background evidence (reproductive biology, testing performance, etc)
    - Patients, study teams, sponsors, regulators (tiered)
  - Interactive model
    - Study designers, regulators
What would be the most useful format for the potential resources (e.g., broad guidelines or recommendations, tables, webpage or smartphone application to enter trial specific data)?

- Broad guidelines in multiple formats (publication, guidebook, online)
- Background information in multiple formats (publication, guidebook, online)
- Interactive estimates
  - “TurboTax”—online
  - Smartphone app—could potentially do things like estimate risk of pregnancy given age, contraceptive method