Notes:

- Problem Description
  - > 120 syndromes in NC – cannot predict every scenario that would want to be alerted to
  - TOA in NC DETECT – Methods
    - By hospital for a particular time period
    - Looking for abnormal clusters
    - Incorporate patients’ residence
    - Have clusters going back to May 2012

- Data provision
  - In NC, do release data for research – just requires DUA
    - However, would not be able to release all of data elements that would eventually want to use the solution on (e.g., triage notes)
    - Would be difficult to negotiate with state-wide hospitals
    - 1 year of raw data (no hospital identifiers – coded name for hospital) with DUA
  - Synthetic data?
    - Would it take a long time to developer?
      - Depends on what solution developers say they need
    - Would inject clusters to real data
    - Would need ~100,000 records for synthetic data set to be really useful

- Need:
  - Spatial precision
  - Temporal precision
  - Age group

- Develop a solution along with TOA or separately

- Discussion
  - Similar tool (WSARE)
    - Could only work with a couple fields at a time and was computationally intensive
    - Looking at multiple fields
  - Data could be hosted on iDASH
    - Advantage of using iDASH is that people can share tools
    - Could provide code in a public area
    - Data owner control?
      - Data owners have complete control
      - Invite members – would only invite people that have signed DUAs
- Would also have a DUA between iDASH and NC Department of Public Health
  - Compare results from algorithm that is developed with clusters identified via TOA (only looks at Time of Arrival by hospital)
  - This problem is a second level – this is a group that came in with an hour, are there any other commonalities
    - Easier option
  - Could also be a first level (separate approach) – looks at chief complaint for commonalities – not contingent on TOA
  - For TOA is there a time parameter – could it be longer (e.g., 1 day)
    - May not be worth worrying about longer intervals, because other methods will pick them up (FL experience)
    - May be worth it for other jurisdiction – e.g., 1 day, to detect things that are not covered by existing syndromes
    - Records of interest query in ESSENCE – selects records with similarities
  - How will I know if I am looking for the right things?
    - Could add to data
    - Can give people idea of what they are looking for
    - Options for enriching text
  - System wide or user-specific?
    - Start with system wide
    - Black list-style systems, but this is often user-centered process
    - May want to start with system wide and then add in option for users to de-select terms (may add to processing times)
  - NC DETECT has a solid group of users, so when it comes time to test, will be able to get feedback from all groups of users
  - Algorithm would possibly be run 2 times a day (ED data updated 2 times), assuming fast processing times, and would be added to reports provided to users
  - Is there some measure of computational power that user has access to?
    - Performance is a big issue
    - Once tool is developed, would be great if tool is made available, as well as the system requirements
  - How long before data set will be available?
    - Lana and Jenna will be doing data use agreements (takes ~2 weeks)
      - Everyone working on data must sign – really need to know all groups wanting to work on this problem (and everyone in groups that will work on)
      - If add someone, they will need to be added (amendment)
      - Description of how data will be kept
    - Amy will generate data (generating data easy – time consuming to inject clusters that want algorithm to find) - ~1-2 weeks
    - Data set can be used ONLY for this problem
  - We could add a data set to iDASH and start working out the process for joining group while working out system for making NC DETECT data available
  - Key question:
    - Expense with going simulation route and there is also skepticism, but also burden of getting real data?
      - Will address as group in NC – in terms of risks and mitigating risks
    - Real data with just chief complaints is consensus
Committee would help with publications (not necessarily as co-authors), but would assist in some way
  - Source code would need to be made available to NC DETECT (and beyond), so could embed in existing systems
    - Don’t want to eliminate COTS tools
  - Any solution developed in conjunction with the Technical Conventions Committee is open source

Next steps:

- E-mail Amy Ising by **March 8, 2013** if interested in working on this problem (to move forward with data use agreement process)
- Develop/clarify process for iDASH (Wendy)
- Edit problem template based on call (Amy)