Genetics Salon: Proprietary DNA Analyses in the Courtroom

William Gordon (Public Health Genetics)
4/5/18

Overview:

- NYC's DNA Lab has been analyzing difficult (trace or mixed) forensic samples for NYPD as well as 50+ other jurisdictions. Two unique methods have been employed:
    - "High Sensitivity Testing" for trace amounts; 3,450 cases
    - Forensic Statistical Tool (FST) for DNA mixtures: 1,350 cases
- Are these methods valid? Potential issues raised:
    - Questions surrounding reliability and statistical rigor have been kept in-house.
    - For High Sensitivity Testing, additional amplification cycles (PCR, a process that makes many copies of, i.e. "amplifies" DNA) increases "noise." Original sample may be as small as n = 4-6 (3 cells).
        - This is lower than the amount "promised" during the approval process
        - Minimum sample size has since been approximately doubled from the initial "promised' amount.
    - For FST, a single likelihood ratio is reported, representing the likelihood that an individual's DNA is present in a mixture. Is this number accurate, and what exactly does it mean?
        - Shapiro resigns over black-box nature of the FST software
        - Software is not (initially) made available to defense
        - Budowle argues that FST is not statistically "defensible" because it assumes that mixtures of similar size are equally contaminated or degraded, and are unrealistically "pristine"
        - FST assumes unrelatedness in mixture (i.e. no population structure); this is a potential issue for some populations (e.g. Hasidim)
        - Quantification step has 30% "margin of error"
        - Plain statistical issues with code have been identified (by Nathaniel Adams)
        - Developers and scientists are not in communication with one another
            - Changes and updates have been made without notification or validation
        - FST eventually replaced by the more widely accepted STRmix.
    - Co-inventor (Caragine) fired from lab for changing colleague's FST results.
- The more complex the analytical method, the less readily a judge or jury can appreciate the nuances and limitations of derived evidence.

Questions for starting discussion:

1. How can the power of a technically complex research method best be leveraged in legal proceedings in which the strength of an argument is based on a layman's understanding of the method (and its limitations)? Is the "magic" of DNA evidence a problem?
    a. Consider the Daubert Standard (https://www.law.cornell.edu/wex/daubert_standard): *Standard used by a trial judge to make a preliminary assessment of whether an expert's scientific testimony is based on reasoning or methodology that is scientifically valid and*

can properly be applied to the facts at issue. Under this standard, the factors that may be considered in determining whether the methodology is valid are: (1) whether the theory or technique in question can be and has been tested; (2) whether it has been subjected to peer review and publication; (3) its known or potential error rate; (4) the existence and maintenance of standards controlling its operation; and (5) whether it has attracted widespread acceptance within a relevant scientific community. See Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993). The Daubert standard is the test currently used in the federal courts and some state courts.

2. What should be done about previous convictions reliant on these methods? What about guilty pleas entered in the face of this evidence?
3. Should a technique used for analysing evidence be eligible for copyright or other "blackboxing?" What are the pros and cons? (Other softwares, such as TrueAllele and STRmix, are protected in the same manner)

Pieces:

- 12/18/17: New York City Moves to Create Accountability for Algorithms
  - https://www.propublica.org/article/new-york-city-moves-to-create-accountability-for-algorithms
  - Bill (returned unsigned): "A Local Law in relation to automated decision systems used by agencies"
- 10/20/17: Federal Judge Unseals New York Crime Lab's Software for Analyzing DNA Evidence
  - https://www.propublica.org/article/federal-judge-unseals-new-york-crime-labs-software-for-analyzing-dna-evidence
  - GitHub directory of FST: https://github.com/propublica/nyc-dna-software
  - Defense expert affidavit #1: https://www.documentcloud.org/documents/4112650-10-17-17-Unredacted-NA-Exhibit-C.html
  - Defense expert affidavit #2: https://www.documentcloud.org/documents/4112649-10-17-17-Unredacted-NA-Exhibit-A.html
- 10/11/17: (Q&A) Putting Crime Scene DNA Analysis on Trial
  - https://www.propublica.org/article/putting-crime-scene-dna-analysis-on-trial
- 9/25/17: ProPublica Seeks Source Code for New York City's Disputed DNA Software
  - https://www.propublica.org/article/propublica-seeks-source-code-for-new-york-city-disputed-dna-software
- **9/4/17: Thousands of Criminal Cases in New York Relied on Disputed DNA Testing Techniques**
  - **https://www.propublica.org/article/thousands-of-criminal-cases-in-new-york-relied-on-disputed-dna-testing-techniques**
  - **Start here! (Discussion will be largely based on this piece; if you only read one piece, read this!)**
- 9/12/16: (Related) DNA Dragnet: In Some Cities, Police Go From Stop-and-Frisk to Stop-and-Spit
  - https://www.propublica.org/article/dna-dragnet-in-some-cities-police-go-from-stop-and-frisk-to-stop-and-spit