Nira Pollock, M.D., Ph.D., D(ABMM) Associate Medical Director, Infectious Diseases Diagnostic Laboratory, Boston Children's Hospital
Division of Infectious Diseases, Beth Israel Deaconess Medical Center
Assistant Professor of Medicine, Harvard Medical School

Click to view Biosketch and Presentation Abstract
or page down to review presentation
A paper-based test for rapid visual measurement of ALT in fingerstick and venipuncture samples

Nira Pollock, M.D., Ph.D., D(ABMM)

Associate Medical Director, Infectious Diseases Diagnostic Laboratory, Boston Children's Hospital
Division of Infectious Diseases, Beth Israel Deaconess Medical Center
Assistant Professor of Medicine, Harvard Medical School
Motivation:

- Blood tests for monitoring liver status via measurement of transaminases (ALT, AST) are standard in developed nations, but often unavailable in resource-limited settings
  - Standard-of-care testing: venipuncture, transport to central laboratory, centrifugation, testing on large automated platform
  - Platforms are expensive and require highly trained technicians
  - Centralization of testing can introduce major delays
Motivation….

- Testing is particularly important for patients taking drugs which can cause liver injury (e.g. HIV and TB medications)
  - TB therapy: incidence of clinically relevant hepatotoxicity is 2-33%
    - Risk increased by many factors (HBV, HCV, ETOH, age, etc.)
  - HIV therapy: Nevirapine of particular concern; rates of DILI can exceed 13%
  - Simultaneous HIV/TB Tx can generate additive risk
  - Idiosyncratic DILI also a concern
Motivation....

- A cheap, accurate, point-of-care (POC) test for measurement of “liver function” would have a dramatic impact on patient care in the developing world
Development of a paper-based POC fingerstick LFT

- Collaboration with Dr. George Whitesides (Harvard Chemistry) and Diagnostics For All (DFA, Cambridge, MA)

- Based on layered patterned-paper technology originally developed in the Whitesides laboratory and furthered by DFA

- Partnership formed (DFA/Pollock-BIDMC/Whitesides) in 2009 to further develop and test device

- CIMIT grant 2010 (Pollock, PI)
Analytical Performance Evaluation

- Test linearity
  - linear across the target clinical range, i.e. 40 to 200 U/L

- Limit of detection
  - ALT, 53 U/L; AST, 84 U/L

- Repeatability
  - %CV values <10% for both AST and ALT tests in serum and blood samples

Pollock et al, Science Translational Medicine, 2012
Pre-clinical evaluation: BIDMC stat lab

- 223 clinical venipuncture specimens (WB, serum)
  - WB, serum drawn simultaneously within prior 5h for routine clinical testing
  - Serum already tested by automated method (Roche)
- 30 uL whole blood or serum applied to each device
- Read 15' later (U/L, to nearest 10 U/L) by three readers blinded to automated results
- Paper-based test results compared to results of automated serum transaminase testing as gold-standard
- “bin placement accuracy”: whether the result of the paper-based assay was in the same bin (<3× ULN, 3–5× ULN, or >5× ULN) as the gold-standard result

Pollock et al, Science Translational Medicine, 2012
Pollock et al, Science Translational Medicine, 2012
Additional evaluation

- Evaluation of analytes with potential for assay interference
  - No significant interference:
    - bilirubin (<10 mg/dl)
    - cholesterol (<500 mg/dl)
    - glucose (<1000 mg/dl)
    - lactate (<200 mg/dl)
    - urea (<100 mg/dl)
    - creatinine (<15 mg/dl)
    - hemoglobin (<120 mg/dl)
  - Some interference from pyruvate (>0.2 mM → falsely elevated ALT) and ascorbic acid (>3 mg/dl → falsely low ALT)

- Test performed well in specimens from clinically diverse population (including many patients who were critically ill with multiple abnormal lab values)

- Device performed well with fingerstick whole blood from 10 healthy volunteers

→ NEXT STEP: FIELD TESTING IN TARGET PATIENT POPULATION AND ENVIRONMENT
Preliminary Field Evaluation

- First field evaluation of this type of platform
- Collaboration between PATH, BIDMC/Harvard AIDS Initiative Vietnam, Hospital for Tropical Diseases (HTD; Ho Chi Minh City, Vietnam) and DFA

Pollock et al, PLOS ONE 2013
Study Objectives

- Evaluation of device performance for fingerstick testing in a target population at risk for DILI (HTD patients on HIV treatment, with high background rates of HBV/HCV co-infection)

- Designed to assess:
  - Operational feasibility
  - Inter-operator variability
  - Lot-to-lot variability
  - Device failure rate
  - Device accuracy

- Intention to utilize results to modify device for further field testing as needed

Pollock et al, PLOS ONE 2013
A

- Lamination
  - Hydrophilic paper
  - Plasma separation membrane
  - Wax in paper
  - Lamination

B

1. Clean finger
2. Prick finger
3. Get large blood drop
4. Collect blood using capillary tube
5. Apply sample to device
6. Blood cells filtered
7. Read results at 15 min

C

ALT Read Guide

- <3X ULN (Range: 0-119 U/L)
- 3-5X ULN (Range: 120-200 U/L)
- >5X ULN (Range: 201-400+ U/L)
Valid vs invalid test results

Pollock et al, PLOS ONE 2013
Study Design

- Device training → Pilot phase (50 subjects) → large-scale evaluation phase (600 subjects)
  - Targeted enrollment of patients with known HBV, HCV, or prior elevated ALT

- Paper-based ALT Test performed on fingerstick blood samples (safety lancet, capillary tube) and read independently by two HTD clinic nurses (N1 and N2)

- Paper-based ALT Test results compared with gold-standard ALT results (Roche Cobas platform) from clinical venipuncture samples obtained in parallel

- Paper-based test was scanned immediately after visual read for subsequent expert interpretation

Pollock et al, PLOS ONE 2013
Results

- 600 subjects enrolled in evaluation phase (105 women, 495 men)

- Two device lots used sequentially (Lot 1: first 218 subjects; Lot 2: next 382 subjects)
Inter-operator Variability

- Achieved strong inter-reader agreement overall
- >90% agreement for determining valid vs invalid [N1 vs N2 vs expert (gold standard)]
  - N1 vs GS: 93.3%; N2 vs GS: 90.6%; N1 vs N2: 95.6%

Sources of disagreement (A>>B):
A. Presence vs absence of hemolysis in the negative control zone
B. Interpretation of positive control zone

- 96.3% agreement for bin placement (N1 vs N2)
- Excellent correlation between N1 and N2 semi-continuous results (visual reads in U/L, rounded to nearest 10 U/L)

Pollock et al, PLOS ONE 2013
Inter-reader agreement
(N1 vs N2, U/L)

Pollock et al, PLOS ONE 2013
Device failure rates and lot-to-lot variability

- Rate of invalids due to hemolysis (seen in negative control zone): 8.7%
  - Lot 1: 21.1%
  - Lot 2: 1.6%
  **Variation thought to be due to faulty batch of plasma separation membrane (PSM) used in Lot 1 devices (degradation of PSM during storage)**
- Rate of invalids due to positive control failure: 0.7%
- Rate of invalids due to failure to fill: 0.2%

Pollock et al, PLOS ONE 2013
ALT results: Paper-based test vs. automated method

Overall bin placement accuracy:
84.3% (Nurse 1) and 83.6% (Nurse 2)

Pollock et al, PLOS ONE 2013
Operational Assessment: Feasibility

- Practical and operational issues assessed included:
  - **Ease of Use:**
    - Initially, intensive training required
    - Peer-to-peer training was promising (potential for cascade training)
    - With successful training nurses felt that they could easily use the device
  - **Sample collection/application and associated variables:**
    - Use of capillary tube allowed minimum required specimen volume to be obtained consistently
    - Observed no device failure due to insufficient sample volume

Pollock et al, PLOS ONE 2013
Conclusions

**Strengths:**
- Successful field study performed in target population/environment, with test performed by local health care personnel
- Excellent inter-reader agreement, demonstrating that reading process is feasible and reproducible and suggesting that inaccuracies in bin placement were due to the device, and not to the reading of the device

**Areas identified for improvement:**
- Variability in lot-to-lot manufacturing
- Suboptimal bin placement accuracy (goal 90-95% in each bin, to allow use in management of HIV/TB therapy); improvement in dynamic range needed
- High rate of invalids seen with Lot 1 devices only, almost exclusively due to hemolysis

Pollock et al, PLOS ONE 2013
Questions at this stage

- This test is the first device of its class to come this far down pathway towards clinical use

- No clear precedents for performance standards

- How accurate must this device be to be clinically useful?
  - Should performance of a paper-based test be expected to match the performance of an automated platform?
  - Low cost/simplicity/POC vs expensive/automated/central lab
Device reoptimization, 2013

- New plasma separation membrane/treatment to reduce hemolysis rates
- Reformulation of assay chemistry to improve readout in the 3x ULN range
- Recalibration against an automated reference standard, Abaxis Piccolo Xpress
Clinical validation of optimized ALT test, 2014

- BIDMC Liver Center and Infectious Diseases outpatient clinics
- 96 ambulatory patients with varied baseline [ALT]
  - Median age 56 (range 22 to 79); 68% male
  - Median [ALT] on day of enrollment: 94 (range 18 to 752)
  - Sixty HCV+, 8 HIV+, 2 HBV+
- Fingerstick testing, DFA paper device
  - On-site reader blinded to baseline [ALT]
  - Images of completed devices captured with cell phone cameras and texted to blinded off-site reader
- Venipuncture serum obtained for routine clinical testing (Roche/Hitachi modular); subsequently captured and applied to DFA and Abaxis platforms (operators blinded to all other results)
- Paper test and automated reference standard results compared by direct correlation and Bland-Altman analysis
Direct correlations
Bland-Altman plots

Bias -23.6 U/L

Bias -18.4 U/L

Bias +4.5 U/L

Abaxis 8.6% lower than Roche
Remote reading of images captured by 2MP or 8MP cellphone cameras

Bias +5.5 U/L

Bias -6.2 U/L
Conclusions

- DFA paper ALT test is highly accurate for serum testing, matching reference method used for optimization (Abaxis) better than reference methods (Abaxis, Roche) matched each other
  - Abaxis systematically 9% below Roche (also seen in standardized proficiency testing)

- Systematic difference between ALT values measured in FS vs paired serum (FS < serum)
  - Existing FDA-approved platforms for FS testing (Roche Reflotron, Alere Cholestech) were only validated with FS samples from patients with normal to mildly elevated ALT (e.g. max 65 U/L)
  - Both of these tests currently off the market (thus preventing direct comparison)
  - Correction factor calculated using FS vs serum results on paper: \( \text{serum } [\text{ALT}] = 14.81 + 1.12 \text{ FS } [\text{ALT}] \).

- Remote reading of this device is feasible
Next steps

- FDA discussions in progress
  - Define protocols required to demonstrate safety and effectiveness in POC and potentially home-use settings

- ISO certification of DFA laboratory for production under GMP

- Completion of cost effectiveness analysis (Rajasingham/Linas)
Study Collaborators

- DFA
  - Jason Rolland
  - Shailendra Kumar
  - Sid Jain
  - Erin Coonahan
  - Ryan Schoeplein
  - Marcus Lovell-Smith
  - Patrick Beattie
  - Una Ryan

- BIDMC/HAIVN (Vietnam study)
  - Donn Colby (site PI)
  - The Anh Nguyen
  - Sariah Khormae
  - Huyen Nguyen

- BIDMC (2013-14 clinical study)
  - Radha Rajasingham (also CE study)
  - Nid Afdhal
  - Michael Curry
  - Rachel Baden

- HTD
  - Van Vinh Chau Nguyen
  - Vo Minh Quang
  - HTD HIV Clinic nursing team

- PATH
  - Bernhard Weigl
  - Sarah McGray (study coordinator)
  - Ken Hawkins
  - Cori Barfield
  - Kathy Tietje
  - Matt Steele

- Harvard Chemistry
  - George Whitesides

- MGH/Tufts
  - Farzad Noubary (stats)

- Boston Medical Center
  - Ben Linas (CE study)