# Point-of-care nucleic acid diagnostics using isotachophoresis and isothermal amplification

Mark D. Borysiak

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:

Jonathan D. Posner, Chair

Daniel T. Schwartz

John C. Berg

Program Authorized to Offer Degree:

Chemical Engineering

University of Washington

## Abstract

Point-of-care nucleic acid diagnostics using isotachophoresis and isothermal amplification

Mark D. Borysiak

Chair of the Supervisory Committee:

Associate Professor Jonathan D. Posner

Department of Mechanical Engineering

Nucleic acid amplification tests (NAAT) have become the gold standard detection technique for many infectious disease diagnoses. NAAT offer high diagnostic sensitivity and specificity that lead to confident diagnoses, as well as relatively fast results and diverse functionality (*e.g.* quantification, multiplexing). Despite the advantages of nucleic acid amplification tests, they currently have limited point-of-care (POC) utility due to the need for multi-step procedures and complex instruments that require skilled personnel and well-equipped laboratories. In this thesis, we aim to create integrated nucleic acid amplification technologies that are better suited for POC applications by developing devices that integrate unique sample preparation and amplification techniques. Specifically, we use a powerful electrokinetic method called isotachophoresis (ITP) to rapidly extract and concentrate nucleic acids using two buffers and an electric field. We also replace traditional polymerase chain reaction (PCR) amplification with isothermal amplification methods that do not require thermocycling equipment.

We present two devices that use isotachophoresis and isothermal amplification to extract and amplify nucleic acids from complex samples. The first technology is named the NAIL device—Nucleic acid Amplification using Isotachophoresis and Loop-mediated isothermal amplification. NAIL uses ITP and loop-mediated isothermal amplification (LAMP) to extract and amplify nucleic acids from complex matrices in

less than one hour inside of an integrated chip. NAIL uses ITP to extract nucleic acids from whole milk samples in less than five minutes. Fluid actuation is performed by passive capillary barriers that act valves and air chambers that act as pumps when heated. LAMP amplifies nucleic acids at constant temperature and produces a bright fluorescent signal that can be detected using a mobile phone camera equipped with fluorescent filters. We show the limit of detection (LoD) of pathogenic *E. coli* O157:H7 cells from whole milk samples to be 1000 CFU/mL, which is a two orders of magnitude improvement to standard tube-LAMP reactions with diluted milk samples.

The second device is called ITP-RPA, which combines isotachophoresis with the isothermal amplification strategy recombinase polymerase amplification (RPA). This device offers improvements over NAIL because it simultaneously incorporates sample extraction with amplification, removing the need for any pumps, valves, etc. ITP-RPA offers detection of nucleic acids in less than 15–20 minutes, and uses low-cost glass fiber substrates that make it amenable to POC applications. We show the detection of nucleic acids spiked into serum and whole blood as we work towards creating a POC viral load monitoring diagnostic for HIV-1 antiretroviral therapy treatments. We discuss the design of ITP chemistry that can incorporate the reagents required for RPA, as well as how to adapt this chemistry for separating the reaction from sample inhibitors. We show the current LOD of the device to be 10,000 copies/mL in serum, which is one order of magnitude higher than the clinical threshold. Future work will aim to improve this value below the relevant threshold.

The introduction to these thesis discusses diagnostic devices with a focus on nucleic acid diagnostics. It also presents the descriptions and relevant principles for ITP and isothermal amplification. Chapter 2 reviews and summarizes relevant metrics for developing diagnostic devices in the laboratory, as well as transitioning these devices to the clinic. Appreciating and using these metrics will benefit test developers by providing consistent measures to evaluate analytical and clinical assay performance, as well as guide the design of tests that will most benefit clinicians and patients. Discussion topics include analytical laboratory statistics, diagnostic sensitivity/specificity and clinical thresholds, clinical metrics, and case studies from relevant commercially available tests.

Table of Contents

List of Figures

List of Tables

Citation to previously published work

Portions of this dissertation, Chapter 2 and 3, have appeared in the following peer-reviewed articles:

1. M. D. Borysiak, M. J. Thompson, J. D. Posner. “Translating Diagnostics from the Laboratory to the Clinic: Analytical and Clinical Metrics for Device Development and Evaluation,” *Lab on a Chip*, 2016, 16(8), 1293–1313.
2. M. D. Borysiak, K. W. Kimura, J. D. Posner. “NAIL: Nucleic Acid detection using Isotachophoresis and Loop-mediated isothermal amplification,” *Lab on a Chip*, 2015, 15(7), 1697–1707.

# Chapter 1: Introduction

## 1.1 Infectious disease diagnosis and point-of-care diagnostics

Diagnostic devices play an important role in giving physicians information about a patient's disease status for diagnosis and subsequent clinical management. The main goal of diagnostic devices is to identify ("rule-in") or eliminate ("rule-out") the likely cause of an illness in a patient presenting with symptoms.[1] Additionally, diagnostics may be used to give information about the prognosis of a condition, to monitor patients during the course of illness and/or treatment, or to screen otherwise healthy individuals for a disease that has not yet presented itself.[1]

The role of diagnostic devices in evidence-based decision making can be generalized by,

**What you thought before + New information = What you think now,**

where the diagnostic devices are providing the "New information" for clinical decision making.[1–3] "What you thought before" and "What you think now" can be thought of as probabilities of a patient having a particular disease before and after the "New information" is presented.[3] Figure 1-1 shows a schematic that demonstrates how the disease probability of the patient changes following diagnostic testing for a positive and negative result. Ideally, the post-test probability after testing is high or low enough that the clinician can take confident action, such as treating the patient or not treating the patient. The extent to which the new information changes the clinical decision depends on the accuracy of the diagnostic test.[3] While the goal of a diagnostic test is to be as accurate as possible, creating a perfect diagnostic test is extremely challenging, if not impossible, and there are numerous tradeoffs that inevitably occur in balancing accuracy with other test characteristics, including cost, ease-of-use, portability, patient comfort, and time-to-result.[4] How these different factors are balanced when developing a diagnostic test depend on characteristics of the disease that it is being developed for and the diagnostic's intended use during clinical management (*e.g.* screening test, confirmatory testing, triage testing, point-of-care vs. laboratory). Chapter 2 provides more detail about analytical and clinical statistical metrics for diagnostic development, with a specific focus on creating tests that meet clinical needs.

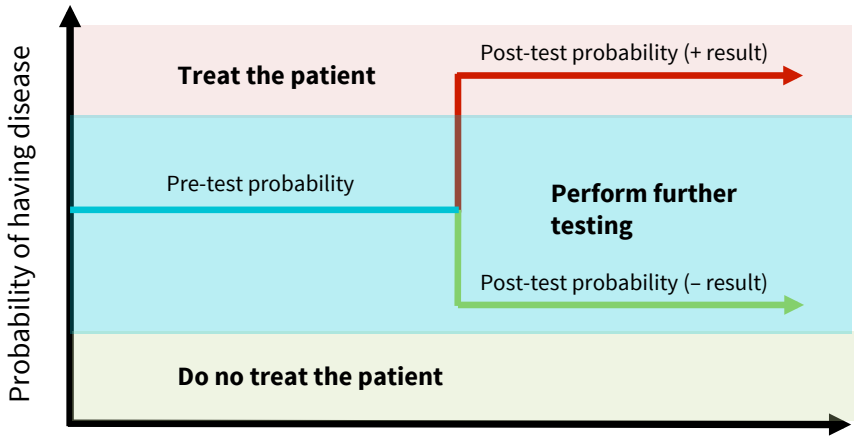



**Figure 1-1. Clinical decision-making after diagnostic testing**

The schematic shows a simplified version of evidence-based medicine decision making with a diagnostic test. The probability of the patient having the disease before diagnostic testing (pre-test probability) and after testing (post-test probability) for a positive (+) and negative (–) result is demonstrated by the horizontal lines. Ideally, the test result provides a post-test probability that can lead to confident clinical action (shaded regions) such as treating the patient (red shade) or not treating the patient (green shade). However, if the test does not provide conclusive enough results, further testing and decision-making is required (blue shade). The pre-test probability, post-test probability, and clinical decision thresholds are all test and disease dependent.

### 1.1.1 Existing diagnostic tests and the need for improved point-of-care tests

The first step in developing a diagnostic test involves choosing a proper biomarker that indicates the presence or absence of the disease of interest. Possible targets for diagnostic testing include indirect biomarkers such as antibodies that the body develops in response to a pathogen, or direct biomarkers such as a pathogen's antigen protein, nucleic acid, or direct visualization of the pathogen itself using microscopy and/or culture.[5] The detection of these targets can be achieved by techniques including cell culture, enzyme-linked immunosorbent assays (ELISA), nucleic acid amplification testing (NAAT), or lateral flow immunoassays. Cell culture and NAAT detect the pathogens directly through observation of the pathogen itself and the detection of the pathogen nucleic acids respectively. ELISA and lateral flow assays detect proteins that are either antigens from the pathogen (direct detection) or antibodies developed by the patient (indirect detection).[5] In general, direct detection of the pathogen offers higher confidence for patient diagnosis compared to indirect detection, but requires more resources for testing (*e.g.* equipment, time, cost, operator training). However, indirect detection may also have benefits in certain situations, such as monitoring protein levels related to conditions such as heart disease, detecting diseases long after

exposure, or observing immunity after vaccination. Figure 1-2 summarizes available testing strategies with general confidence in results, ease-of-use, and time to result for each method.

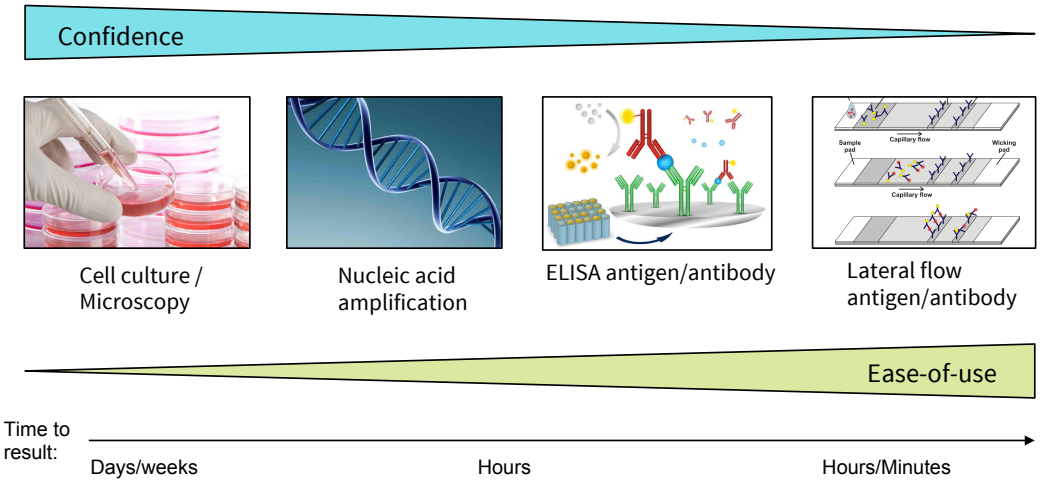



**Figure 1-2. Summary of diagnostic biomarkers (adapted from Peeling)[5]**

Confidence of results, ease of use, and time to result of various diagnostic methods. Direct detection methods (culture, nucleic acid amplification, ELISA antigen) offer higher confidence in results, but poor ease-of-use. Lateral flow tests are less complex, but do not give as confident of results. Our work focuses on developing a nucleic acid diagnostic device that has comparable confidence levels to traditional nucleic acid tests, but with greater ease of use and faster time to result.

Traditionally, cell culture, ELISA, and NAAT have been used as the gold standard techniques for many diagnostic applications. These techniques generally require well-equipped central laboratories with skilled personnel due to assay, instrument, and/or protocol complexity. For many infectious diseases, test results should ideally be provided during the initial visit so that clinicians can immediately make treatment decisions. The logistics around specimen collection, transport, and returning of results for laboratory restricted tests often leads to missed opportunities for appropriate therapeutic care and intervention.[6] The long turn-around time for laboratory testing is particularly problematic in low resource settings, where there is a significant amount of patient loss to follow-up. In these cases, patients may ultimately be disease positive according to the laboratory test, yet they may not receive the appropriate diagnosis and treatment because they leave the clinic after the initial visit and do not return.[6]

Patient loss to follow-up is an issue for many infectious diseases, but can be highlighted using HIV diagnosis as an example. A review of early infant diagnosis (EID) programs for HIV in low resource settings (LRS) noted the time to lab-based test results varies widely (from 9 days to 5 months) and in the majority of

studies, only half of families/caregivers returned for test results.[7] Another study in South Africa showed that only 69% of newly diagnosed HIV-infected adults received lab-based results within 90 days, and less than half of those began treatment within 12 months.[8,9] As a result, only 35% of HIV positive patients were receiving treatment within 12 months of their initial visit. Even in an urban environment, we have interviewed clinicians who have said they lose over 40% of their patients following testing at a clinic two miles outside of downtown Seattle, WA.[10]

POC diagnostic platforms that can rapidly and accurately detect infectious diseases are needed to improve patient care in peripheral settings.[11] POC tests would broaden access to routine disease screening and monitoring, leading to decentralized patient care.[12] Clinic-based POC testing would allow for earlier detection of disease, same day adherence counseling, and simplify patient management—potentially all at reduced cost.[12] The majority of commercially available POC tests currently use lateral flow technology for antibody or antigen detection. These tests have achieved success for certain applications (*e.g.* pregnancy tests, strep A, HIV-1 antibody/antigen) due to their simple operation, low-cost, and rapid results.[13–15] However, many of the tests, particularly for infectious diseases, are not useful in a clinical setting due to poor accuracy that typically stems from low diagnostic sensitivity.[16,17] Even the commercially successful Strep A test still requires a back-up culture test for negative test results, while the HIV antibody/antigen tests require a 2–6 week wait period after exposure.[18,15] For many POC lateral flow tests for infectious diseases, clinicians have limited confidence in the results provided by these tests, and they are not used to diagnose the majority of diseases.[17,19]

For the diagnosis of most diseases, the more rigorous and accurate cell culture, ELISA or NAAT are currently required. Cell culture has long been the gold standard for many disease diagnoses, but typically consists of multiple incubation steps (pre-enrichment, selective enrichment, selective and differential plating) before a biological, serological, or molecular confirmation test, meaning multiple days pass before results are given. These tests are still routinely used in the laboratory, but the time and technical complexity of culture make it unrealistic for POC adaption. Antibody ELISA tests require that enough days/weeks have passed for the patient to have developed antibodies in response to the pathogen. For diseases that have had time to develop antibodies, these tests can be useful diagnostic techniques, though direct detection of the pathogen is preferred in many cases due to the higher confidence in results.

Antigen ELISA tests and NAAT are both commonly used to a wide range of diseases, including, *e.g.* sexually transmitted, respiratory, and gastrointestinal diseases because they provide direct detection of the pathogen. The choice between antigen ELISA tests and NAAT will vary by disease target and the factors such as the intended biomarker, required diagnostic sensitivity and specificity, time-to-result requirements, and cost of testing.[20] NAAT typically provides improved sensitivity compared to antigen ELISA tests, but if relatively high levels of pathogen / antigen are present in the target sample, the ELISA test may provide a lower-cost and more straight-forward option.[20] Examples of diseases commonly detected with antigen ELISA include syphilis, HIV, and Lyme disease.[21] The development of automated NAAT platforms over the past 10–15 have led to NAAT increasingly being used as the gold standard, but cost, complexity, and availability remain drawbacks of NAAT. Currently, hundreds of diagnostic tests using NAAT are FDA cleared for patient diagnosis with the vast majority of these NAAT being performed in central laboratories, which prevents their use for POC applications, where significant clinical needs exist.

### 1.1.2 Strategies and requirements for developing point-of-care diagnostics

There is currently a disconnect between the confidence a diagnostic test provides (high accuracy) and the resources that it requires to give a result (*e.g.* time, cost, personnel training, infrastructure).[5] Lateral flow tests uses minimal resources that are suitable for the POC, but typically do not provide high confidence results.[17,19] Laboratory tests such as culture, NAAT, and ELISA provide high confidence results, but require significant resources for testing.

A strategy for creating POC diagnostic tests thus involves developing methods to improve the poor accuracy of lateral flow tests or reducing the equipment and user-training required for more complex tests such as NAAT, ELISA, cell culture, microscopy, etc. In order to create POC diagnostics that meet clinician and patient needs, there are two general strategies: (1) improve the accuracy of lateral flow devices or (2) reduce the resources required to perform lab-based testing, as shown in Figure 1-3. The first strategy has been pursued by many researchers, including our own group.[22,23] Much of this research has focused on improving the limit of detection of lateral flow tests because the poor diagnostic sensitivity of these tests is often due to insufficient detection limits for low levels of target analytes. Various publications and reviews detail this area of research.[13,24–26] The work presented in this thesis will focus on the second strategy of reducing the resources required to perform lab-based tests, specifically NAAT. We have chosen nucleic

acid amplification as our focus for a number of reasons, including high diagnostic sensitivity and specificity of the lab-based nucleic acid amplification tests, faster results compared to cell culture, the versatility to multiplex and detect multiple diseases at once, the ability to provide quantitative results, and the potential to adapt the required steps into a format that is amenable to POC use.

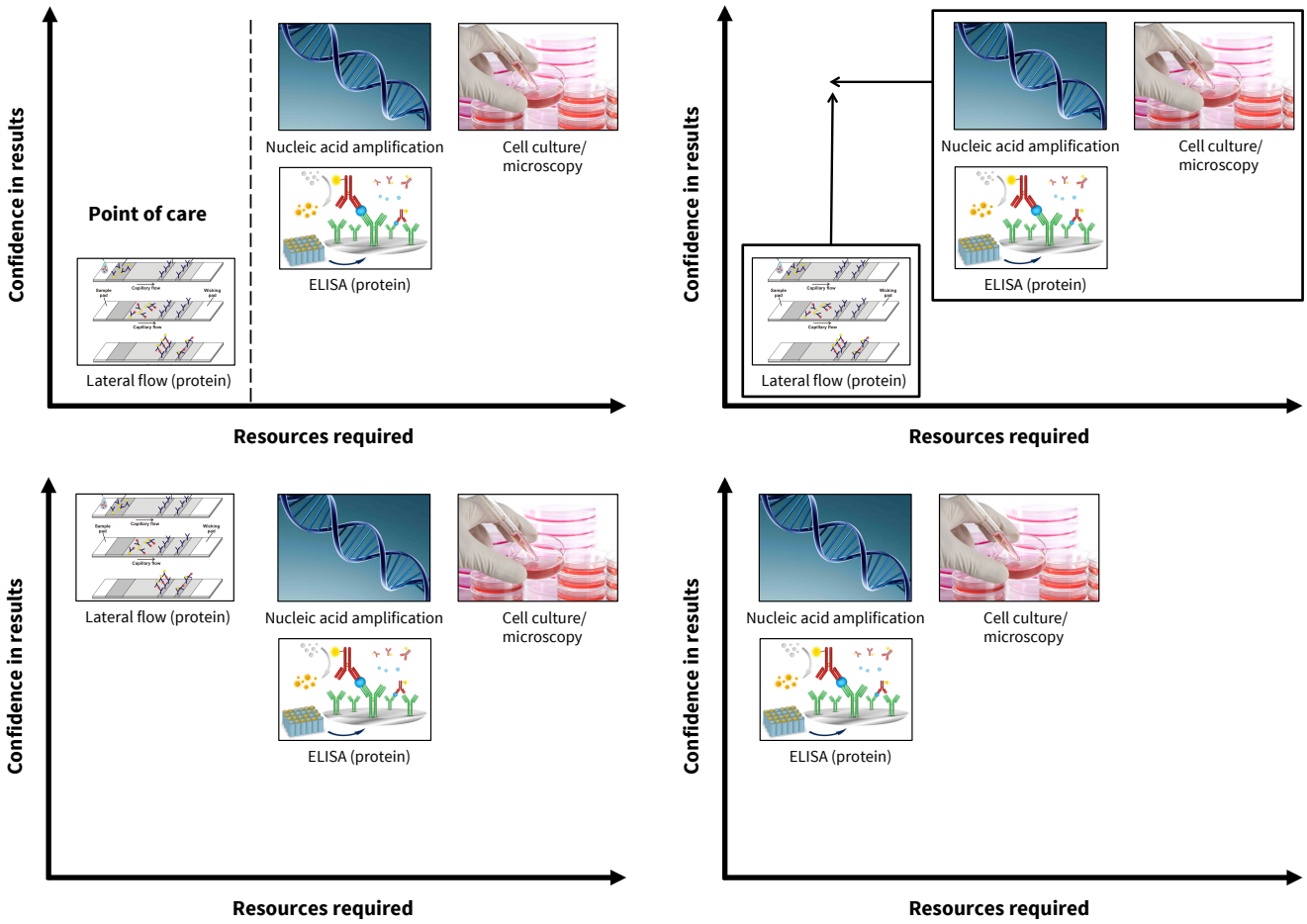



**Figure 1-3. Adapting diagnostic tests for the POC.**

(A) The confidence in test result and required resources are shown generally for common diagnostic techniques. Lateral flow tests require minimal resources and are well-suited for POC use, but often lack the accuracy for confident diagnosis. Laboratory-based tests such as cell-culture, nucleic acid amplification, and ELISA provide more confident diagnoses, but currently require resources that preclude their use at the POC. A simplified strategy is to either (B,C) improve the accuracy of lateral flow diagnostic tests or (B,D) reduce the resources required to perform laboratory-based tests. We will focus on reducing the resources required for nucleic acid amplification diagnostics, as discussed in the following sections.

For a diagnostic test to be used at the POC in the United States (and often in other areas, including low resource settings), it must receive a waiver under the 1988 Clinical and Laboratory Improvement Amendments (CLIA).[27,28] The CLIA regulations established quality control standards for laboratories

conducting diagnostic tests and specified oversight requirements including personnel qualification and training, proficiency testing, and routine quality assessments.[27,28] A CLIA waiver allows a test to be used by non-certified laboratories and un-trained personnel in settings such as physicians' offices, hospital rooms, or pharmacies. Because CLIA waived tests are not subjected to CLIA quality control regulations, they must "employ methodologies that are so simple and accurate as to render the likelihood of erroneous results by the user negligible".[27,28]

The FDA provides recommendations for CLIA-waived in vitro diagnostics (IVDs) which can be used to identify the characteristics needed to make a NAAT suitable for POC use. The first FDA recommendation for CLIA-waived devices is to "demonstrate [the device] is simple to use". While "simple to use" is a very general statement, the FDA goes on to define characteristics of "simple" tests, which include:[27,28]

- being a fully automated instrument or self-contained test,
- using unprocessed samples (*e.g.* whole blood, nasal swabs, urine),
- requiring only basic reagent manipulation (*e.g.* mix reagent A and B),
- needing no operator intervention during analysis steps, and
- producing clear results that do not require interpretation (e.g. unambiguously displaying "positive" or "negative").

The other key component of a CLIA waived device is "demonstrating insignificant risk of an erroneous result".[27,28] Because CLIA waived devices will be used outside controlled laboratory environments by users without specialized training, the tests must be developed to be insensitive to environmental and usage variation, and all known sources of error must be controlled. Sources of error may occur during all three phases of testing, including the pre-analytical phase (sample collection and identification), analytical phase (testing the specimen), and post-analytical phase (interpretation and recording of results).[27,28] The FDA and the Clinical and Laboratory Standards Institute (CLSI) EP-18 list a number of potential sources of error that must be considered including:[28,29]

- operational: poor timing of procedure or use of incorrect specimen/reagent type,
- reagent integrity: use of expired, contaminated, or improperly mixed reagents,

- hardware integrity: power loss, sample flow clog or bubble, electronics/mechanical failure, and
- environmental: heat, humidity, and other factors that may affect calibrations and/or results.

The requirements for a CLIA-waived test to be both “simple to use”, while also being insensitive to a significant number of potential perturbations—objectives that often oppose each other in practice—highlights the difficulty in creating POC tests. The World Health Organization (WHO) adds further requirements for POC tests in low-resource settings through their ASSURED criteria.[30] They state that POC tests should be Affordable, Sensitive, Specific, User-friendly, Rapid & Robust, Equipment-free, and Deliverable.[30] These criteria essentially encompass the requirements for CLIA-waived devices, but add equipment-free, rapid results, low-cost, and easily deliverable (*e.g.* portable, small, stable reagents) to the end user as additional objectives for test developers.

Taken together, the requirements for POC test developers include creating a device that gives highly accurate and rapid answers, is fully automated, maintains high precision in variable environments, uses minimal equipment, and has a low-cost. Creating an integrated diagnostic that performs sample-to-answer analysis with clinical specimens, while conforming to all of these requirements is a significant technical challenge for test developers. This thesis will discuss our progress towards developing a nucleic acid amplification diagnostic platform that aims to address these challenges and provide all required attributes for diagnostic testing at the POC.

## 1.2 Nucleic acid diagnostics

Over the past 10–20 years, nucleic acid amplification diagnostics have replaced immunoassays, cell culture, microscopy, and other diagnostic techniques to become the gold-standard of diagnosis for many diseases.[31–33] One of the main advantages nucleic acid amplification tests (NAAT) offer is the ability to detect disease targets with high diagnostic sensitivity and specificity, leading to results that give clinicians conclusive information about the disease state of the patient. Further, NAATs often offer much faster results than other techniques—they can typically give results in less than 2–4 hours, and certain tests may give results in less than 15–30 minutes. Additional capabilities include multiplexing for detecting multiple diseases simultaneously and quantitative results.[34–36]

As of July 2016, the Food and Drug Administration (FDA) has approved over 250 different nucleic acid diagnostic tests for detecting pathogens including influenza, *Mycobacterium tuberculosis*, *Staphylococcus*, HIV, and cancer targets with many more under commercial development.[37] The majority of these tests are based around polymerase chain reaction (PCR) nucleic acid amplification. PCR uses thermal cycling to denature double stranded DNA at high temperature (94–98°C), followed by hybridization of primers designed to anneal to target sequences at reduced temperature (50–65°C), and elongation at slightly raised temperatures (68–75°C), as shown in Figure 1.4.[38] The use of thermostable polymerase by Mullis and coworkers simplified the PCR temperature cycling process, leading PCR to become one of the most widespread molecular biology techniques used today.[39] Many PCR tests include quantitative detection, such as real-time polymerase chain reaction (qPCR) with Taqman probes or DNA intercalating dye, and/or the versatility to detect multiple targets using multiplexing strategies.[40] Antimicrobial resistance, virulence biomarkers and highly specific typing can also be identified using PCR.

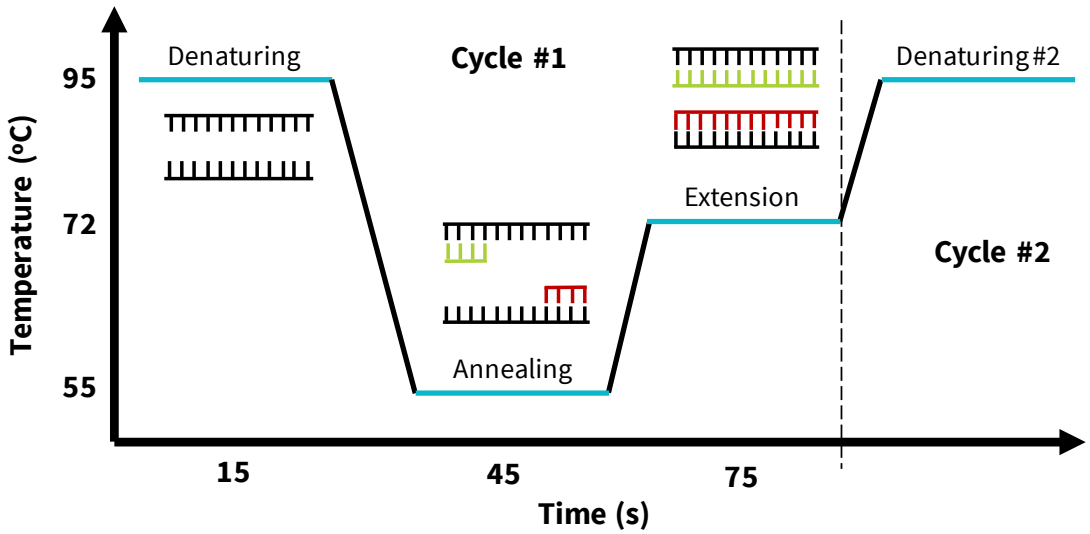



**Figure 1-4. Schematic of PCR thermocycling**

PCR requires heating to three different temperatures for denaturing of the target strand (95ºC), annealing primers to the target sequence (55–65ºC), and extension of the primers by a polymerase (68–75ºC). This process constitutes one cycle that doubles the amount of initial DNA. A typical PCR process consists of 35–40 cycles that can create over a million copies of the original target. Primers are typically 18–22 base pairs long and amplicons can range from 100–200 base pairs up to thousands of base pairs.

The required steps for nucleic acid detection include cell or virus lysis to release nucleic acids, extraction and purification of nucleic acids from contaminants in the sample, amplification of the target nucleic acid sequence, and detection of the amplified products. PCR may be inhibited by a number of substances found

in clinical samples, such as proteins, cations, and other biomolecules, requiring target nucleic acid to be highly purified before amplification. The majority of traditional purification methods involve multiple mixing, centrifuging, separation, and buffer exchange steps. These methods can largely be separated into two categories: liquid phase extraction (LPE) and solid phase extraction (SPE). One of the more common LPE extraction methods is the TRizol method developed by Chomczynski and Sacchi, which uses chaotropic salt lysis followed by phenol/chloroform extraction and ethanol precipitation of the nucleic acid.[41] This method produces high purity nucleic acids, but is very labor intensive and uses large quantities of toxic chemicals. Modifications have been made to this method to replace the hazardous chemistries, but still involve a series of bench-top procedures. SPE methods offer an alternative to LPE and use low pH, high ionic strength buffer to allow nucleic acids to bind to silica columns, glass beads, or other exchange carriers. The exchange carriers are then washed with ethanol to remove background impurities and the purified nucleic acid is eluted using a low ionic strength buffer. SPE kits are available from manufacturers such as Qiagen, which use spin columns that have become the standard bench-top method for nucleic acid extraction, but still require over an hour of time and >10 manual steps.[42]

After purification of the nucleic acid sample, the target sequence can be amplified and detected by a PCR platform. PCR's need for fast temperature cycling at high temperatures means that a thermocycler with regulated temperature control and rapid heat exchange is required for amplification. The detection of PCR results also typically uses laser fluorescence, which involves specialized light sources and optics. Altogether, PCR is an extremely powerful technique, but also requires significant time, money, and training investment that present a barrier towards POC diagnostic use.

The issues surrounding laboratory extraction and PCR operation have led researchers and industry to work towards developing systems that are more suited for the POC, which include near-patient testing in a hospital or primary care setting. The difficulty in adapting NAATs for POC settings focuses around creating a device that maintains laboratory characteristics such as high accuracy and reliability, while adding additional POC characteristics such as automated operation, portability, small footprint, internal quality controls, and low-cost. These challenges become considerably more difficult when trying to integrate all necessary steps for nucleic acid amplification from clinical samples. The presence of a handful of commercially available CLIA medium complexity and CLIA-waived NAATs highlight some of the advances

towards overcoming some of these challenges, but also current short-comings that still require continued innovation.

Diagnostic tests can be categorized as high complexity (central labs), medium complexity (hospitals), or waived (primary care) according to the 1988 CLIA amendments. The majority of clinical diagnostic nucleic acid amplification tests are considered high complexity by the United States. A few exceptions include Cepheid GeneXpert system, BD Diagnostics BD Max System, and Biofire Diagnostics FilmArray RP system, which are characterized as medium complexity.[27,43,44] Early in 2015, the Alere i was the first NAAT to receive a CLIA waiver. As of July 2016, there are three NAAT that have received CLIA waiver status: the Alere i, Roche Liat, and Cepheid Omni.[45–47]

The medium complexity and CLIA waived tests represent progress towards improving the usability of nucleic acid tests for POC diagnostic applications. The majority of these tests are miniaturized or lower throughput versions of larger lab-based systems with more compact pump-valve and thermal systems for SPE sample preparation and PCR thermocycling. The inclusion of miniaturized thermocyclers, laser fluorescence detection, and pump-valve fluidic systems result in relatively high cost, complexity, and maintenance that preclude the use of these machines in many settings. For example, the mentioned medium complexity systems cost at least $50,000 for the unit and over $70 per test, and the CLIA waived tests average approximately $10,000 for the platform and $50 per test. These tests are too expensive for the widespread deployment in primary care settings in developed countries, with further deliverability issues in low resource settings. Further, the Alere i and Roche Liat CLIA-waived platforms have only been approved for Influenza A/B and/or Strep A diagnosis as of July 2016. The Cepheid Omni has not yet been released after Cepheid announced a 2-year production delay due to issues with the amplification and detection modules that highlight the engineering complexity of these platforms.[48] Despite the improvements these platforms have made, opportunities still exist for continued development and innovation for lower-cost, more rapid, and less complex NAAT.

While there are numerous challenges associated with creating POC NAAT, we believe the most pressing challenges for continued improvement of these devices are (1) efficiently extracting nucleic acid targets from clinical samples, (2) reducing hardware required for amplification and detection, and (3) integrating all processes in a low-cost, automated manner.

Niemz et al. noted in a 2011 POC NAAT review that large-volume sample preparation in an integrated system is the greatest challenge for sustained implementation of these devices.[49] Their assertion still largely remains true five years later. Large volumes of clinical samples (0.1 to >1 mL) must typically be processed in order to achieve suitably low limits of detection. Handling and processing these physically (e.g. high viscosity) and chemically (e.g. presence of amplification inhibitors) complex samples in a device designed for POC settings is a significant technical challenge. Automation for SPE of nucleic acids from clinical samples and integration with downstream processes depends on the particular platform, but is typically achieved using various pump, vacuum, and valve-based mechanisms that often lead to higher complexity (and cost), maintenance, and risk of failure. Some alternative strategies being researched for sample preparation include centrifugal microfluidics, electrokinetic techniques including isotachophoresis, and paper-based extraction.[42,50,51] Centrifugal systems use a rotating disk at different velocities to move fluids through the multiple stages of sample preparation. Some examples include the LabDisk and LabTube systems.[52,53] Electrokinetic techniques use electric fields to extract and concentrate nucleic acids from samples based on their electrophoretic mobilities. Section 1.3 will discuss isotachophoresis, an electrokinetic separation technique, in more detail.[42] Finally, paper-based systems have been used to capture, purify, and elute nucleic acids using passive capillary flow that sometimes use additional mechanisms such as sliding or “origami” folding.[54–56] These systems are low-cost and can eliminate the valves and pumps of traditional fluid systems, but automation remains a challenge.

Following extraction of target nucleic acids, the system must perform amplification and detection. These two areas have experienced a number of advancements over the past 10–15 years that can be leveraged for POC development. For example, the growth in isothermal amplification chemistries has offered the ability to create NAATs that do not require thermocycling,[49,57–59] leading to reduced equipment requirements and (likely) cost for nucleic acid amplification. Further, many isothermal chemistries offer additional benefits such as greater tolerance to amplification inhibitors and faster reaction times compared to PCR. The growth of mobile phones and the hardware necessary for portable imaging have allowed for the development of simplified readers for fluorescence and colorimetric detection.[60–62] They also can be used for automated data analysis, storage, and transmission of results to cloud storage. Visual readouts

such as turbidity, colorimetric detection, and particularly lateral flow detection, have also gained popularity.[63–66]

In order to develop a NAAT diagnostic platform that can be well-suited for POC applications, we aimed to combine ITP sample preparation with isothermal amplification within an integrated device that performs all of the steps required for nucleic acid amplification. Our goal is to provide automated or semi-automated nucleic acid detection that can be translated for minimally trained users in low resource settings. Our work is differentiated from many of the current POC NAAT that simply seek to automate and miniaturize benchtop protocols such as solid phase extraction and PCR. We believe this work demonstrates progress towards creating NAAT platforms that have the potential to provide accurate, rapid, and low-cost patient diagnosis at the POC. Section 1.3 will discuss isotachophoresis (ITP) for sample preparation and section 1.4 will present isothermal amplification strategies. Our work towards integrating these technologies will be presented by our two developed devices in Chapters 3 and 4.

## 1.3 Isotachophoresis

Isotachophoresis (ITP) is a well-established nonlinear electrophoretic technique that is capable of separating and concentrating ions based on their electrophoretic mobility.[67] Electrophoretic mobility, $\mu_i$, is a proportionality constant that describes the relationship between an applied electric field, *E,* and the migration velocity, *V*, of an ion, *i*,

$$V_i = \mu_i E \quad . \tag{1}$$

ITP focuses sample ions between two groups of ions, the leading electrolyte (LE) ions and trailing electrolyte (TE) ions, which have mobilities that bracket the mobility of the sample ions of interest, $\mu_{LE} > \mu_{sample} > \mu_{TE}$. The LE has higher mobility than the sample ions and the TE has lower mobility than the sample ions. Nucleic acids have high electrophoretic mobility due to a highly negatively charged phosphate backbone (pKa~1.5), making ITP an advantageous method for purifying and concentrating nucleic acids from complex samples.[42] An ITP system usually consists of two reservoirs that contain TE and LE buffers, with a channel or substrate between the reservoirs to create a fluidic connection. At the start of ITP operation, one reservoir contains TE buffer, while the other reservoir and channel or substrate are filled with LE. The target sample can be mixed with the TE, LE, or placed in a region between the buffers in a

setup known as finite injection. When an electric field is applied via two electrodes, the TE ions begin to electromigrate into the channel, displacing the LE ions. The sample ions are able to overcome the slow moving TE ions, but cannot pass through the fast moving LE ions, which causes a small band of selectively focused sample to form. A schematic showing this process is shown in Figure 1-5.[68]

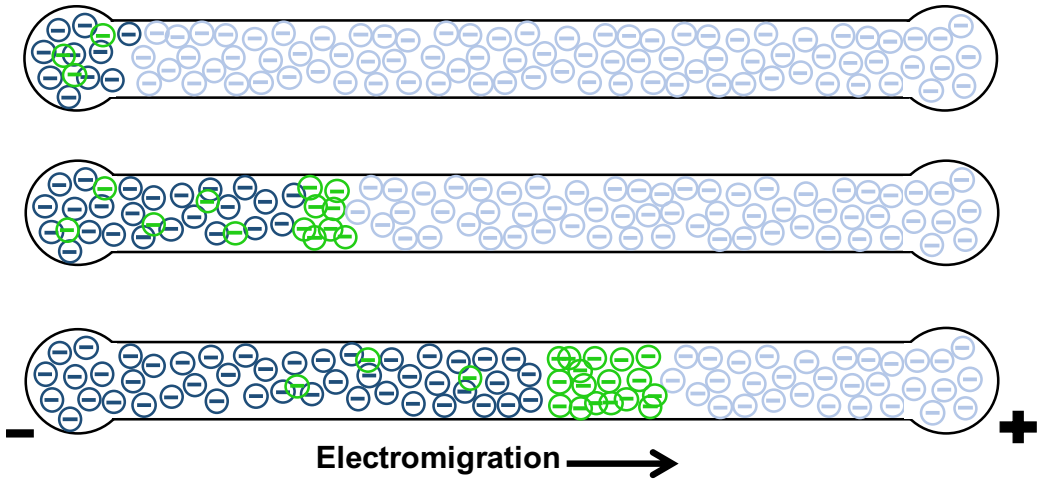



**Figure 1-5. Schematic of focusing analytes using ITP.[68]**
At the start of an experiment, the right reservoir and channel or substrate are filled with LE, while the sample and TE are mixed together in the left reservoir. Applying an electric field causes the TE (dark blue) to displace the LE (light blue) in the channel and the sample (green) to focus at the interface between the LE and TE.

The band of sample forms in the ITP zone, which is established at the interface between the LE and TE due to the difference in conductivity between the two ion groups. Conductivity can be described as a function of local ion concentration and mobilities,

$$\sigma = F\sum_{i}\sum_{z} z\mu_{i,z}c_{i,z} \ , \tag{2}$$

where $F$ is Faradays constant, $c_{i,z}$ is the concentration of ion $i$ with valence state $z$, and $\mu_{i,z}$ is the mobility of ion $i$ with valence state $z$.[67] The summation is performed over all species in a particular zone in order to determine the local conductivity of the zone. If a constant current is applied through the solution, the local electric fields of each zone can be determined for a given conductivity and channel geometry using conservation of current,

$$E_j = \frac{I}{A_j\sigma_j}, \tag{3}$$

where $I$ is the applied electrical current, and $A_j$ and $\sigma_j$ are the cross-sectional area and conductivity of zone $j$ respectively.[67] The name “isotachophoresis” comes from Greek roots meaning “equal speeds”. At steady-state, the conductivities and electric fields in the LE and TE zone will adjust so that each ion from the LE and TE is moving at the same speed. In practice, the LE zone has high conductivity and low local electric field, while the TE zone has low conductivity and high local electric field that create equal velocities. Figure 1-6 summarizes the mobility, conductivity, and electric field of each zone.

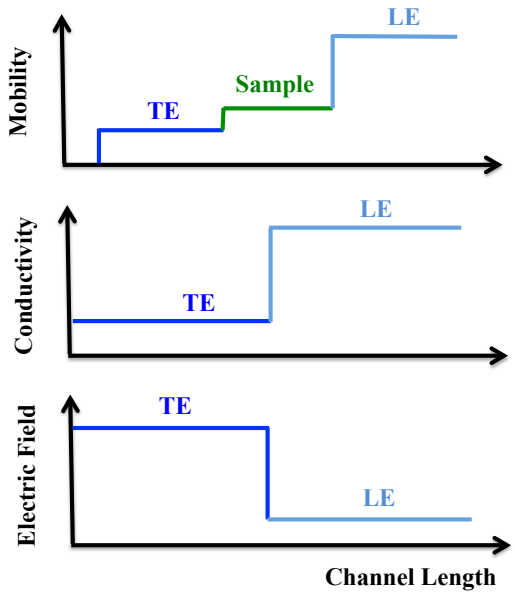



**Figure 1-6. Summary of mobility, conductivity, and electric field for ITP zones.**
ITP focuses samples based on mobility between a high mobility leading electrolyte (LE) and low mobility trailing electrolyte (TE). The conductivity differences between the LE and TE causes an electric field gradient to form when an electric field is applied, which is where the sample with intermediate mobility will focus. For low sample concentrations (less than micromolar concentrations), the sample does not significantly contribute to the solution conductivity or affect the electric field distribution. This ITP regime is known as “peak-mode” ITP and is the typical operating mode for nucleic acid extraction.

When the sample ions, such as nucleic acids, are in the TE zone, they experience high electric fields and have high velocity according to equation 1. In the LE zone, the sample ions experience low electric field and have slower velocity. Thus, the conductivity and electric field differences between the LE and TE cause the sample ions to migrate and focus in the region of the electric field gradient. The migration of sample ions into regions formerly occupied by other ions (*e.g.* LE and TE ions) is regulated by the Alberty and the Jovin conservation relations that describe adjusted ion concentrations for partially ionized species (the most interesting and useful case).[69] Unlike "linear" zone electrophoresis methods, such as capillary zone

electrophoresis, in which separating zones continually spread by diffusion or dispersion, ITP forms self-sharpening zones of analyte due to the non-linear electric field gradient.

The accumulation of analyte at the TE and LE interface can reach concentration effects up to 1,000,000 times the original analyte concentration under ideal operating conditions, with more common effects showing order 1000–10,000x concentration.[70] When target analytes are initially present at low concentrations (typically <μM), the focused analytes do not significantly contribute to the system conductivity. The analytes focus in an approximate Gaussian shape in an ITP regime known as “peak-mode”. For high initial analyte concentrations, the analytes will contribute to local conductivity and form plateau zones, known as “plateau-mode” ITP. The majority of nucleic acid sample preparation with ITP will occur in “peak-mode”, meaning the sample contribution to conductivity is negligible and the resulting ITP plug will be Gaussian shaped.

The ITP concentrating effect has been used to accelerate reaction rates using the high local concentrations of reactants within the ITP plug. Bercovici et al. showed 10,000x acceleration of nucleic acid hybridization reactions using the focusing effect of ITP.[71] Han et al. and Moghadam et al. showed increases in liquid-surface reaction rates by focusing one reactant within the liquid phase of the ITP plug and migrating the focused sample over a stationary reactant on a solid phase.[22,72] Han’s work showed increased sensitivity and and reaction speed for nucleic acid hybridization, while Moghadam demonstrated over 100x improved sensitivity for protein lateral flow binding assays. These works also provide analytical models that describe the physics behind ITP accelerated reactions.

By designing the TE ions to be faster than contaminants (assuming the contaminants are slower than the target), ITP can also separate the sample of interest from slower contaminating substances, effectively purifying the sample in a single step process that doesn’t require binding, buffer exchange, or moving parts. Classical ITP experiments focused mostly on separating and detecting small ions and molecules. Recently, greater research focus has been devoted to separating and concentrating trace biological molecules from complex samples.[42] Kondratova *et al.* used ITP to separate nucleic acids from biological samples as a downstream process in a series of extraction steps.[73] They used large glass and plastic capillaries to achieve high volume, high efficiency extraction of nucleic acids. However, they performed significant pretreatment prior to running ITP, resulting in negligible time and labor savings compared to traditional

benchtop methods. The Santiago group has since published a series of papers demonstrating the separation and purification of nucleic acids from complex samples including cell lysates, urine, and whole blood.[42,74–80] A few highlights include a chip with integrated heaters for cell lysis and a plastic microfluidic chip capable of processing 25 µL volumes of 1:20 diluted whole blood (1.25 µL of whole blood) with high extraction efficiency (>80%).[76,79] In all of these studies, they perform little to no pretreatment of samples prior to nucleic extraction by ITP. Purigen Biosciences is a startup currently working on creating isotachophoresis nucleic acid extraction platforms for commercial use.[81] To date, none of their work includes integrated amplification of target nucleic acids for detection and diagnosis—all post-processing was performed off-chip. Rogacs, Marshall, and Santiago offer a detailed review for nucleic acid purification using ITP and the following sections will detail a few key design considerations for designing an ITP system.[42]

### 1.3.1 Designing ITP electrolyte systems for ion separation

Designing an ITP system requires choosing the LE, TE, and counterion species and concentrations. For anionic ITP (focusing anions at the ITP interface), chlorine anions are the most common choice for LE selection. Chlorine anions have high electrophoretic mobility (~80 x $10^{-9}$ $m^2V^{-1}s^{-1}$) that allows them to establish a large electric field gradient, while also being biocompatible and readily available (*e.g.* as hydrochloric acid). The trailing electrolyte anion typically consists of a weak acid that has a mobility and pKa near the operating pH of interest so that the mobility can be tuned to the desired value for separation and extraction. Examples we have used in our work include Good's buffers and amino acids such as HEPES (pKa = 7.5, fully ionized mobility = 23.5 x $10^{-9}$ $m^2 V^{-1} s^{-1}$), tricine (8.26, 26.6 x $10^{-9}$ $m^2 V^{-1} s^{-1}$), TAPS (8.51, 22.1 x $10^{-9}$ $m^2 V^{-1} s^{-1}$), serine (9.33, 34.3 x $10^{-9}$ $m^2 V^{-1} s^{-1}$), and glycine (9.78, 37.4 x $10^{-9}$ $m^2 V^{-1} s^{-1}$).[82] Persat et al. list a number of electrolytes for the TE and counterion selection.[82]

Selection of pH is a critical step in designing an ITP system. A counterion is usually included in the LE and TE in order to maintain a relatively constant pH near the desired pH of the system. For anionic ITP, the ideal counterion is a weak base that has a pKa near the pH of the system for maximum buffering capacity.[82] We commonly use Tris (pKa = 8.06) in our system because it has strong buffering capacity in the relevant pH range for nucleic acid amplification (~pH 7.5–8.5), and is the counterion used in the majority of PCR and other amplification chemistries. The mobility of the leading chlorine anions is

independent of the system pH, while the trailing anion mobility is strongly dependent on pH.[83] As a result, the pH of the solution affects the separation and focusing of the ITP target due to changes in the trailing anion mobility. The mobility of a monovalent weak acid as a function of pH can be determined using,

$$\mu_{i,eff} = \mu_{i,-1} \frac{1}{1+10^{(pKa-pH)}}, \quad (4)$$

where $\mu_{i,eff}$ is the effective electrophoretic mobility of species *i*, $\mu_{i,-1}$ is the fully ionized mobility for the –1 valence state, $pK_{-1}$ is the pKa of the –1 valence state, and pH is the pH of the solution.[83] A table of common values can be found in Persat et al.[82] Figure 1-7 shows the absolute value of equation 4 plotted for five different anion species for pH 7.5 to 9. The selection of TE anion species and solution pH should be made so that the TE mobility is slower than the target to be focused by ITP and faster than the contaminants in solution to be separated. The mobility of nucleic acids is relatively high (approximately 40 x $10^{-9}$ $m^2 V^{-1} s^{-1}$) and pH independent above ~pH 3, which allows for a wide range of TE mobilities to focus nucleic acids.[84,85] Common contaminants such as proteins have mobilities that vary based on the isoelectric point of the specific protein or biomolecule and the pH of the solution. Based off our experiences, proteins in the pH range of interest for nucleic acid amplification (pH 8–9) can be approximated as having a mobility of 10 x $10^{-9}$ $m^2 V^{-1} s^{-1}$ for ITP separation design purposes.[22]

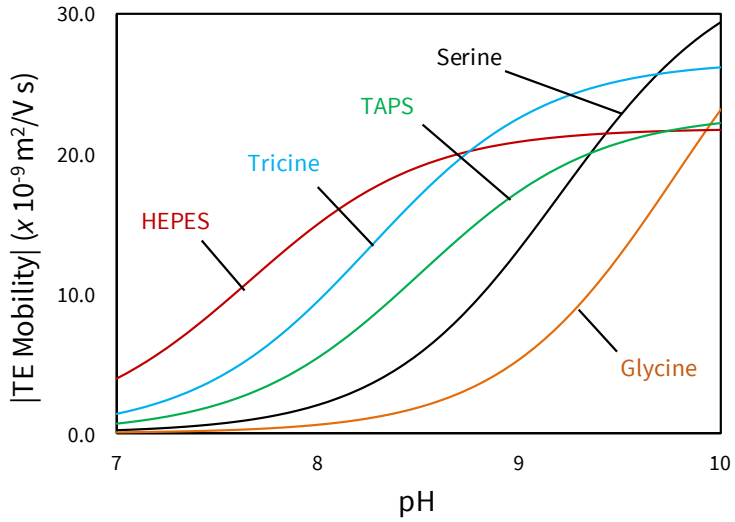



**Figure 1-7. Electrophoretic mobility of various anions for trailing electrolyte design.**

The absolute value of the electrophoretic mobility (x $10^{-9}$ $m^2 V^{-1} s^{-1}$) as a function of pH is shown for five TE anions.[83] The pH at which the anion mobility begins to increase is dependent upon the pKa value of the ion. Anions that increase at lower pH have lower pKa (e.g. HEPES) and anions that increase at higher pH (e.g. glycine) have higher pKa. The mobility is also dependent upon the fully ionized mobility of each anion,

though they do not drastically differ from each other. The pH for the majority of nucleic acid amplification applications ranges from 7.5–9, with the most common pH occurring near pH 8. The TE anion mobility (species and functional pH) should be to be slower than the mobility of the target to be focused and faster than the mobility of the contaminants to be separated from the sample.

In a typical anionic ITP system (the method used throughout this work), the weak-base counterion in the LE will migrate from the LE towards the TE during electric field application. During this migration, the weak-base counterion will maintain a concentration within 30% of its original concentration in the LE. Meanwhile, the trailing anions will enter the substrate to replace the leading anions that migrate from the TE towards the LE well. As a result, a zone is created where the trailing anion will be present within the substrate (previously occupied by the leading anion), but the weak-base counterion from the LE is still present due to its countermigration. In this zone, the trailing anion will interact with the weak-base counterion from the LE, meaning that the pH and resulting mobility of the trailing anion is dependent on the weak-base counterion in the LE. The region where the trailing anion interacts with the weak-base counterion is known as the adjusted trailing electrolyte (ATE) zone, where the ATE is defined as the TE zone that forms in regions previously occupied by the LE.[42]

The ATE zone is a critical for ITP operation because it determines the relevant mobility of the trailing anion for performing separation and concentrating of sample ions. Determining the pH and resulting mobility of the trailing anion in the ATE zone is most efficiently performed using Stanford Public Release Electrophoretic Separation Solver (SPRESSO) simulations, an open source software developed by the Santiago group at Stanford.[86] Figure 1-8 presents example SPRESSO simulations of the adjusted TE zone using Tris-HCl as a leading electrolyte system, which is the LE system used throughout this work. Figure 1-8A shows the initial concentration profile of the Tris and HCl in the LE, as well as the Tris and glycine in the TE before ITP has begun. In this case, the substrate can be thought of as the region occupied by the Tris-HCl (roughly 40–100 a.u. on the x-axis), while the glycine-Tris TE is present in a well (0–40 a.u.). Figure 1-8B shows the formation of the ATE zone as the glycine replaces the regions previously occupied by chlorine anions during ITP, which is simulating the TE moving across the substrate. In this region, the glycine interacts with counter-migrating Tris from the LE, leading to an increase in pH (9.1) and mobility magnitude ($7.0 \times 10^{-9}$ $m^2 V^{-1} s^{-1}$) in the ATE, as shown in Figure 1-8C and Figure 1-8D respectively. The pH and mobility of the original TE and LE zones are also shown (*i.e.* the Tris-HCl LE is initially pH 8.1

and the glycine-Tris TE is initially 8.6). This example shows the effects that the weak-base counterion in the LE can have on the mobility and separation due to the TE anion. The mobility magnitude in the ATE is approximately three times larger than the original TE solution (2.5 to 7.0 x $10^{-9}$ $m^2 V^{-1} s^{-1}$) as shown in Figure 1-8D. When choosing a trailing anion mobility, the mobility that should be considered is the ATE mobility, not the mobility of the trailing anion initially present in the TE well.

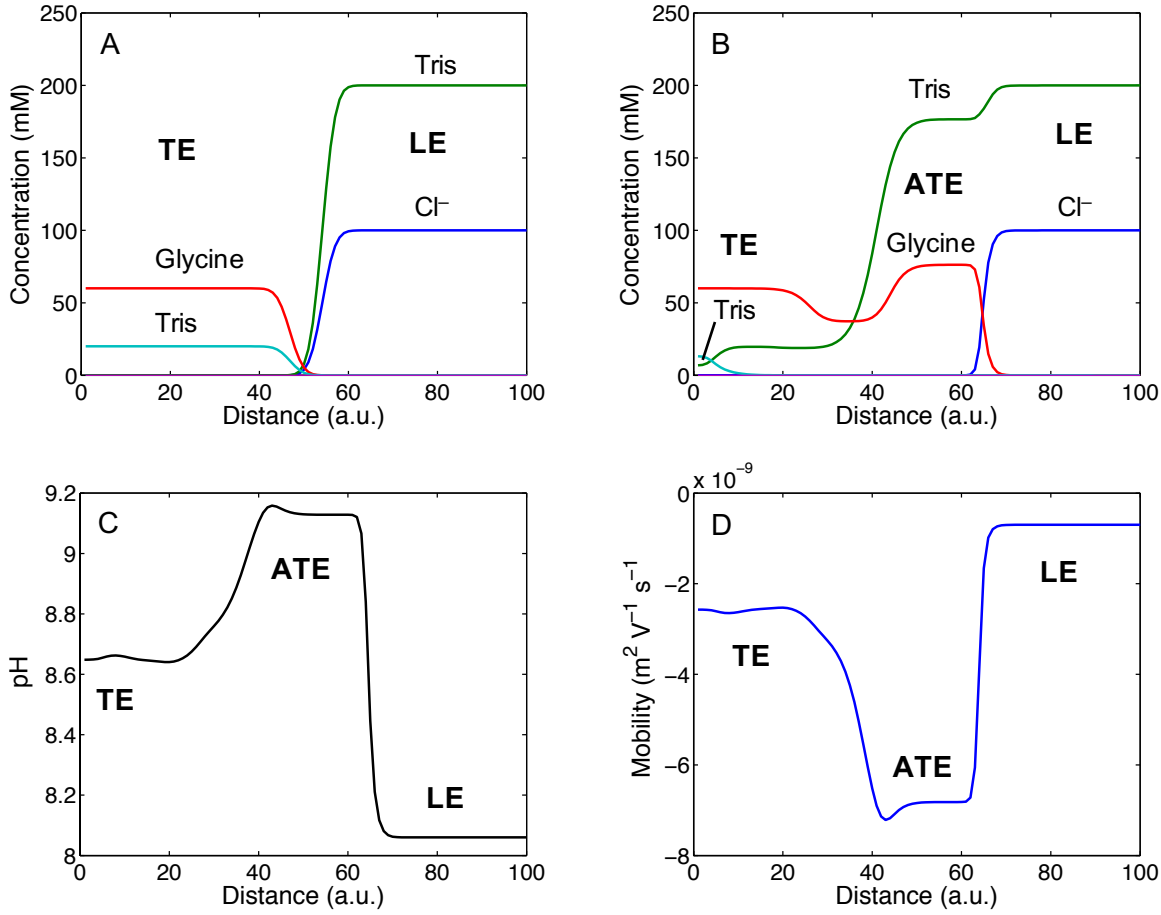



**Figure 1-8. Adjusted trailing electrolyte zone effects on trailing anion mobility**

SPRESSO simulations showing the development and affect of the adjusted trailing electrolyte (ATE) zone on the trailing anion mobility. (A) Initial concentrations of Tris, HCl, and glycine in the TE and LE prior to electric field application. (B) Formation of the ATE due to glycine displacing the chlorine anions in the substrate during ITP operation, while Tris counterions remain due to its countermigration. (C) The pH profile of the system including the pH in the TE, ATE, and LE regions. The ATE region experiences higher pH than the initial TE zone. (D) As a result of the higher pH in the ATE zone, the mobility magnitude of the glycine (TE) anions is larger than in the TE or LE zones. This ATE mobility is a critical parameter for determining which ions will focus or not focus in the ITP interface between the ATE and LE zones.

When designing an ITP system, the adjusted mobility in the ATE zone should ideally be simulated using SPRESSO software.[86] For a given weak-base counterion, a general heuristic is that the ATE pH

increases with increasing ratio of deprotonated weak-base counterion ($c_B$) versus protonated weak-base counterion ($c_{BH+}$), $c_B/c_{BH^+}$ in the leading electrolyte.[67] In other words, as the pH of the LE increases, the pH of the ATE zone and resulting mobility will increase. For example, a 100/150 mM HCl/Tris (pH 7.8) LE will give lower mobilities than a 100/200 mM HCl/Tris (pH 8.1) LE solution At a given ratio of $c_B/c_{BH^+}$ in the leading electrolyte, higher $pK_a$ values for the weak-base counterion will lead to higher pH in the ATE zone. For example, using Bis-Tris as the counterion (pKa 6.3) will give lower ATE mobilities than Tris (pKa 8.1).

### 1.3.2 ITP extraction, focusing, and separation capacity

Three important aspects of designing an ITP system include obtaining large concentrating effects, high separation between target ions and contaminants, and high extraction efficiency of target from the sample. The parameters that guide the design and optimization for these three ITP effects will briefly be discussed. All parameters are assumed to be operating in finite injection mode, where the sample mixture is placed between the trailing and leading electrolyte zones.

As mentioned previously, ITP systems are capable of concentrating analytes 1000–10000x the original concentration in the sample, a process referred to as “stacking” or “focusing”. The concentrating effect of ITP is strongly dependent on the TE conductivity and applied current density through the ITP system.[87] Khurana et al. give an approximate analytical solution for the concentration of the analyte in the LE-TE interface divided by the initial concentration of sample in the TE well, *i.e.*, how much ITP increases the concentration of the original target,[87]

$$\frac{C_s^{ITP}}{C_s^{TE,well}} = K\frac{j}{\sigma^{TE,well}}x \ . \qquad (5)$$

Equation 5 shows that the sample concentration in the ITP interface is proportional to the inverse conductivity of the TE solution in the well, $\sigma^{TE,well}$, and proportional to the current density, *j*. Additional terms in equation 5 include: *K* is a ratio of effective mobilities of LE, TE and sample species; $C_s^{TE,well}$ is the concentration of sample in the well; and *x* is the distance traveled by the ITP plug.[87] The key conclusions from this relation are that lower trailing electrolyte conductivities (lower concentration and/or lower mobility) and higher current densities lead to improved stacking due to higher electric fields in the trailing electrolyte that extract and focus sample at the ITP interface. Also, concentration of the analyte will increase with

distance traveled (or time as ITP is run). Practical operating limits for these design guidelines are discussed at the end of this section.

The ability of ITP to separate two ions from each other is dependent upon the effective mobilities of the ions in a relevant solution. If the goal is to separate out a contaminant by preventing it from entering the ITP interface, the relevant separation mobilities are the trailing electrolyte and contaminant mobilties, $\mu_{TE}$ and $\mu_C$. By designing the TE mobility to be faster than the contaminant through ion and pH selection, the ITP interface will eventually separate completely from the contaminants. The time required for separation for a given ion system is dependent upon the amount of electrical current passed through the system. Thus, the amount of charge (current x time) required to separate two groups of ions is an important parameter, which is known as the separation parameter,

$$Q_s = F \cdot \frac{N_A(\mu_A+\mu_R)+N_B(\mu_B+\mu_R)}{(\mu_A-\mu_B)} \quad (6)$$

where $N_i$ is the number of moles, and $\mu_i$ is the effective mobility for ions A and B to be separated, and counterion (R).[88] For contaminant separation, subscript A would be properties relevant to the trailing anion, and subscript B would be properties relevant for the contaminant. The properties should be evaluated in the ATE zone using SPRESSO or similar software.[86] Equation 6 shows that the separation parameter increases with the number of moles (larger volumes or higher concentrations) to be separated and decreases with larger mobility difference between the selected TE and contaminant. In other words, it requires more charge (and is often practically more difficult) to separate large volumes and concentrations of contaminants, as well as contaminants that have mobility very close to the ATE mobility of the trailing anion.

The amount of charge required to extract the target ions (*e.g.* nucleic acids) into the ITP interface from the sample can also be approximated using Equation 6. In this case, the goal is to design the system so that the target has higher mobility than the trailing anion so that it can enter into the ITP zone. For this situation, the values relevant to subscript A in equation 6 should be determined for the target, while the trailing anion properties are given by subscript B. Because the sample is typically at much lower concentrations than the trailing anion for peak mode ITP (relevant for nucleic acid extraction), equation 6 can be simplified to:

$$Q_S = \frac{N_{TE}(\mu_{TE}+\mu_R)}{\mu_{DNA}-\mu_{TE}}, \quad (7)$$

where all relevant properties are determined in the adjusted TE zone.

After determining the amount of charge required to separate out the contaminant from the system, as well as provide high efficiency extraction of target, the amount of charge provided by an ITP system must be determined. The amount of charge that an ITP system provides is known as the separation capacity, which can be evaluated experimentally by passing constant current and observing the time required for ITP operation ($Q_L = I \cdot t$), which is typically the point where the analyte reaches the sensor or end of substrate. It can also be determined analytically using,

$$Q_L = N_L F(1 - \mu_R/\mu_L) \quad , \quad (8)$$

where $N_L$ is the moles of leading anion and $\mu_L$ is the mobility of the leading anion.[88] The separation capacity can be thought of as the amount of time that a given current is passed before the ITP plug reaches its final destination, *e.g.* the LE reservoir.

In order for full theoretical separation and extraction to occur, the system must provide greater charge than is required by the separation parameter, $Q_L > Q_S$.[88] The separation capacity is dependent on the number of leading electrolyte ions in the system, *i.e.,* the concentration of the leading electrolyte ions times the volume of the channel or substrate. High concentrations of leading anions lead to lower electric fields in the LE, which allows more time for separation for a given applied current. Higher volumes necessitate larger channel geometries that also lead to lower electric fields and ITP velocities for a given applied current through equation 3. As a result, high efficiency separations of sample require high LE concentrations (though LE concentrations often have practical operating limits) and/or large channel or substrate volumes. As a first approximation, the substrate should hold at least equal volume of LE as the volume of sample to be separated. Paper-based systems hold significantly more volume than microchannel based systems, which make them an intriguing system for high volume and efficiency separations using ITP, as will be discussed in Chapter 4.

The general takeaways from this section when it comes to designing an ITP systems include:

- Lower conductivity TE solutions and higher current densities lead to better concentrating effects of the target in the ITP interface.[87]
- Large sample volumes and concentrations require more charge to be passed through the system in order to separate or focus ions of interest. The amount of charge required also depends on the relative difference in mobility between the adjusted TE anion mobility and the mobility of the contaminant and/or target.[88]
- Larger channel / substrate volumes and higher LE concentrations provide systems with higher separation capacity, meaning they pass more charge for better separation and extraction of sample ions.[88]

### 1.3.3 General procedure for ITP design and practical operating limits

In summary, the important parameters for high concentrating effects, separation of contaminants, and extraction efficiency are (1) the conductivity of the trailing electrolyte and current density, (2) the mobility of the trailing anion compared to the contaminant and target, (3) the volume of sample to be processed, and (4) the volume of the substrate and concentration of the leading electrolyte. The ITP system design procedure based on the described heuristics and our experience should be conducted as follows:

(1) Determine the mobility of the target to be focused and if needed, the contaminant to be separated. Nucleic acids can be estimated as $40 \times 10^{-9}$ $m^2$ $V^{-1}$ $s^{-1}$ and proteins as $10 \times 10^{-9}$ $m^2 V^{-1} s^{-1}$.

(2) Select the pH range of interest for operation. For nucleic acid amplification, the pH range should typically be near pH 8. Select an appropriate counterion with pKa near this pH for maximum buffering capacity, *e.g.* Tris for pH 8.

(3) Select an LE composition and concentration. HCl is typically used as the leading anion and a weak-base counterion can be added to adjust and buffer the pH, with a 2:1 base to acid ratio providing a pH close to the pKa of the base counterion. Trailing anions (*e.g.* chloride from HCl) typically range in concentration from 50–200 mM, with concentrations as high as 1 M. Any example of a common LE may be 75 mM HCl and 150 mM Tris at pH 8.1.

(4) Select a trailing anion that will provide a mobility faster than the contaminant of interest and slower than the target of interest at the operating pH range. An initial approximation can be made using equation 4 and Figure 1-7 by setting the pH equal to the LE pH. However, the mobility in the ATE zone should be determined using SPRESSO simulations with the selected LE system for a more accurate approximation. Slight adjustments in mobility can be made by changing the base:HCL ratio in the LE, with more drastic adjustments requiring changing trailing anion type.

(5) The concentrations of the trailing anion and weak-base buffering counterion should be chosen. Trailing anions commonly range from 10–50 mM, but can be set at over an order of magnitude higher or lower. The trailing anion and weak-base counterion concentration can be set to higher values for better buffering capacity and lower values for improved stacking (lower TE well conductivity). An example TE solution may be 20 mM HEPES, 40 mM Tris (better buffering capacity) or 10 mM HEPES, 5 mM Tris (better stacking).

(6) Determine the volume of sample to be extracted and approximate concentrations of contaminants and targets.

(7) Calculate the separation parameter for separation and extraction. As an approximation, the TE volume should be set at least equal to the sample volume to be processed, with larger volumes providing more conservative estimates of the separation parameter. Mobilities should be evaluated in the ATE zone.

(8) Determine the substrate volume required for a separation capacity that is greater than the separation parameter for separation and extraction for the designed system based off the chosen LE concentration. A separation capacity at least two times greater than the highest separation parameter is recommended.

At this point, the ITP system will need to be tested empirically and adjusted using the aforementioned guidelines. For example, if higher stacking is required, the trailing electrolyte concentration or mobility can be reduced for lower conductivity and/or the electrical current can be increased. If faster or slower trailing anion mobility is required for improved separation or extraction, the trailing anion species can be changed or the ATE zone mobility can be adjusted by changing the pH of the LE solution. Better separation capacity can be achieved using substrates with higher volume capacity or by increasing the LE concentration. While

these are general suggestions for ITP design, a number of practical operating limits will likely be encountered, including: (1) dispersion due to electroosmotic flow, (2) pH and ionic strength effects from complex samples, and (3) Joule heating due to poor heat dissipation of the applied electrical current.

Reducing the conductivity of the TE will lead to better concentrating effects, but will also likely induce dispersion or complete stoppage of the ITP plug migration due to electroosmotic flow (EOF).[89] EOF is the result of positive ion charges shielding the negatively charged surface within the electrical double layer. When an electric field is applied, the positively charged species counter-migrate against ITP and create a bulk fluid flow that can disperse the ITP plug. The EOF velocity, $U_{EOF}$, depends on the local EOF mobilities and axial location of the sample zone. Garcia-Schwarz et al. provide a detailed description of EOF, with their key result being that the adjusted velocity of the ITP zone is $U_{zone} = U_{ITP} + U_{EOF}$, which can be as expressed as:

$$U_{zone} = \mu_{LE} E_{LE} + [\left(\frac{L_{TE}}{L}\right) \mu_{EOF}^{TE} E_{TE} + \left(1 - \frac{L_{TE}}{L}\right) \mu_{EOF}^{LE} E_{LE}], \quad (9)$$

where $U_{zone}$ is the velocity of the ITP zone when taking into account EOF, $\mu_{EOF}^{TE,LE}$ are the mobilities of the substrate in the TE and LE regions, and $L_{TE,LE}/L$ is the fraction of substrate covered by TE or LE.[89] The first portion of equation 9 corresponds to the velocity of the ions due to electromigration, $U_{ITP}$, while the second portion corresponds to the opposing EO flow, $U_{EOF}$.

Because the electric field strength and surface electrical charge are greater in the TE compared to the LE, the key parameters that determine the magnitude of EOF are controlled by the $\left(\frac{L_{TE}}{L}\right) \mu_{EOF}^{TE} E_{TE}$ term in equation 9. The surface charge ($\mu_{EOF}^{TE}$) is characterized by the substrate zeta potential, which has data available for common substrates such as glass and plastic channels as a function of pH and ionic strength, but must be empirically measured.[90,91] For a given applied electric field, the important relationship controlling the electric field strength in the TE, $E_{TE}$, is the conductivity ratio of the LE compared to the TE, $\sigma_{LE}/\sigma_{TE}$ (as this ratio grows, the fraction of the applied electric field in the TE zone increases, resulting in higher $E_{TE}$). Finally, as TE replaces LE on the substrate during ITP operation, EOF increases due to the $\left(\frac{L_{TE}}{L}\right)$ term. As the $\left(\frac{L_{TE}}{L}\right) \mu_{EOF}^{TE} E_{TE}$ terms grow, the EOF increases will lead to dispersion of the ITP plug, a lower velocity of the ITP plug, or may even result in stoppage or flow reversal of ITP migration.

In order to reduce EO flow, the substrate zeta potential in the TE should be minimized, and the conductivity difference between the LE and TE should adjusted until EOF does not adversely affect ITP. Zeta potential suppression can be achieved by silanization treatments on glass or silicon, but most commonly is achieved by including high molecular weight polyvinylpyrrolidone (PVP), a bulky non-ionic polymer that can shield surface charge.[92] Further surface charge reduction techniques include increasing ionic strength in the TE (divalent ions have greater suppression effects) or reducing TE pH.[90] The zeta potential likely cannot be completely suppressed, so the final option tends to be limiting the conductivity difference between the LE and TE. As mentioned previously, decreasing TE conductivity increases stacking, but the practical limit is often met when the conductivity difference becomes too large and the EOF creates significant dispersion or stops plug migration. In this case, reducing the LE conductivity or increasing the TE conductivity is required until ITP migrates properly.

Another practical consideration for ITP design is the effect of complex sample ionic strength and pH. If the ionic strength of the TE has been carefully designed for stacking or the pH of the LE has carefully been designed for ATE zone mobility, perturbations due to additional electrolytes or pH changes from the sample may adversely affect ITP. The best method to account for these changes is to create LE and TE solutions with high enough ionic strength and buffering capacity that these perturbations do not significant affect the system.

The final limit to operation of ITP is Joule heating due to electrical current passing through the substrate. Joule heating can be calculated using $w = I^2/\sigma A^2$, where *w* is the heat dissipated per volume unit (W $m^{-3}$), *A* is the cross sectional area, and $\sigma$ is the conductivity of the solution.[67] This relation shows that the amount of heat generated increases four-fold with two-fold higher currents. It also shows that the heating is dependent on conductivity, which varies depending on whether the substrate is filled with LE or TE. Higher currents are desirable due for increased stacking and faster operation, but a practical limit is reached when Joule heating negatively affects performance. In microchannel or paper-based systems, Joule heating can lead to dispersion of the ITP plug or inhomogeneity of physical and chemical properties (*e.g.* mobility, pH, density) due to temperature gradients from the non-linear conductivity in the system.[88] In porous media based systems, Joule heating can also cause cause evaporation and/or the substrate to burn

in the case of nitrocellulose. However, Joule heating can also potentially be beneficial if leveraged to provide heating for amplification reactions, as will be discussed in Chapter 4.

## 1.4 Isothermal amplification

Nucleic acid amplification is a key step in nucleic acid detection assays. Nucleic acid amplification typically involves three main steps (1) denaturation, (2) annealing, and (3) extension. DNA is denatured thermally or biochemically so that primers designed to bracket the target sequence of interest can anneal. After primer annealing, a polymerase (protein that adds nucleotides to a nucleic acid strand) will extend the primers in order to create a copy of the original nucleic acid strand. Continuous cycling of this process results in as little as 5–10 copies being amplified to over millions of copies. Nucleic acid amplification improves sensitivity by enabling detection of a limited number of initial molecules. It is also highly specific because the primers can be designed to anneal only to nucleic acid sequences that are unique to the target of interest.

As discussed in section 1.3, PCR is the most common technique used for amplifying nucleic acids, but has limitations for POC diagnostic devices due to the need for energy and hardware intensive thermal cycling that can be difficult to incorporate into a compact, low-cost device. Isothermal amplification methods have been developed to replace thermocycling equipment with heating blocks, water baths, or other simple heating methods by amplifying at a single temperature.[93–96] Some isothermal amplification methods also do not require laser fluorescent detection, may produce results in as little at 20 minutes, and are not as prone to amplification inhibition as PCR.[66,97] These characteristics make isothermal amplification techniques advantageous for POC nucleic acid diagnostics.

There are a number of different isothermal amplification methods available, including nucleic acid sequence-based amplification (NASBA), loop-mediated isothermal amplification (LAMP), helicase-dependent amplification (HDA), rolling circle amplification (RCA), helicase dependent amplification (HDA), nicking enzyme amplification reaction (NEAR), and recombinase polymerase amplification (RPA).[93] While these methods all share the common trait of amplifying at a single temperature, each technique has unique characteristics including the needed enzymes and primers, temperature required (37–65°C), time for amplification, and detection of results. We will demonstrate the use LAMP (Chapter 3) and RPA (Chapter

4) for this thesis. Reviews detailing the various types of isothermal amplification are available for discussion on additional methods.[93]

LAMP is a sequence specific nucleic acid amplification protocol with three defining characteristics: (i) LAMP is performed at a single temperature of 55-65°C; (ii) LAMP produces up to 400 times more amplified DNA than PCR (~400 µg/mL for LAMP compared ~1 µg/mL for PCR); and (iii) LAMP is highly specific due to four different primers recognizing six distinct regions of DNA.[57] LAMP uses loop-creating primers and strand displacing polymerase to produce up to $10^9$ copies of target in less than one hour. Producing large DNA mass enables visual detection of LAMP products without expensive optical detection equipment.[65,66,97] LAMP has also been shown to have equal or better analytical sensitivity than qPCR (~5 copies/reaction) and is highly specific, with studies showing specific target amplification in the presence of 90 non-specific organisms.[98]

Simple visual detection of LAMP results is enabled by utilizing metal-ion chemistries that are similar to classical chemistry titrations of divalent cations using, *e.g.* EDTA and eriochrome black T. In these experiments, a metal ion (*e.g.* $Mg^{2+}$, $Ca^{2+}$) is initially bound to a dye that produces a distinct initial color in the solution. Upon titration of a chelating agent that more strongly binds the metal ion than the dye, the unbound dye is released into solution, causing a change in color of the dye and solution. This same principle can be used to detect LAMP reactions by the production of pyrophosphate divalent anions during amplification reactions.[65,66]

Calcein dye detection uses metal chelation chemistry to create high intensity fluorescent signals that can easily be imaged to detect LAMP reactions.[66] Initially, calcein is bound to low concentrations of manganese ions in the reaction solution to quench the fluorescence of the calcein dye. As pyrophosphate is produced during amplification, the pyrophosphate anions chelate the manganese ions so that the calcein dye may bind magnesium ions that are present in excess compared to manganese (100:1 Mg:Mn). Once bound to magnesium, the calcein fluoresces to indicate DNA amplification. The combination of simple visual detection, single temperature heating, and high sensitivity and specificity make LAMP well suited for POC diagnostics. We discuss the use of LAMP in our integrated diagnostic device in Chapter 3.

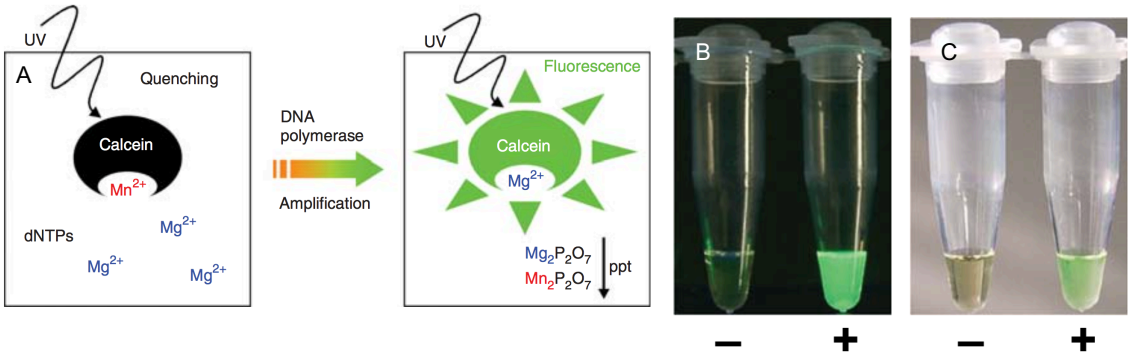



**Figure 1-9. Calcein detection of LAMP reactions (adapted from Tomita et al)**

Fluorescent calcein dye can be utilized to visualize LAMP reaction results using a metal-ligand chemistry. (A) Calcein dye is initially bound to manganese ions present at low concentration. Upon the production of pyrophosphate during amplification, the manganese is precipitated out. Magnesium ions present at high concentration compared to manganese bind the calcein to cause fluorescence. (B) An image of a negative and positive reaction under UV illumination with calcein chemistry. (C) An image of the same reactions without UV illumination. The reactions exhibit unambiguous differences for detection.

The work described in chapter 4 uses recombinase polymerase amplification (RPA) chemistry.[58] RPA is an isothermal amplification strategy that is well-suited for POC NAAT use due to its speed (<15 minutes in a tube), low incubation temperature (30–43 °C), under 10 copy/reaction sensitivity, and reagent stability.[58,99–103]. **Figure *1-10*** provides a brief overview of the RPA mechanism. RPA uses recombinase enzymes to coat single-stranded DNA, *i.e.* oligonucleotide primers, to form filaments, which can then scan double-stranded DNA sequences for homology. After finding the homologous sequence, the nucleoprotein filament can invade and create a short hybrid with a displaced strand. Single-stranded binding proteins bind to the displaced strand to stabilize it and prevent ejection of the inserted primer. Strand-displacement polymerase then performs primer extensions and continuation of this process results in exponential amplification of the target. The amplicon produced by the amplification process can be detected using a sequence specific probe that binds to a region of the amplicon. The probe contains a fluorescent molecule (*e.g.* fluorescein) and a quencher that prevents non-specific fluorescence. Upon binding to the amplicon, an exonuclease cuts a region of the probe between the fluorescent molecule and the quencher, creating a fluorescent signal that can be monitored. The main advantage of the sequence specific probe compared to an intercalating dye is that the sequence specific probe will not fluoresce in the presence of non-target nucleic acids, result background and non-specific signal. TwistDx has commercialized the RPA technology

and offers amplification kits that contain all of the necessary reagents to perform RPA.[104] Chapter 4 discusses the use of RPA in our paper-based nucleic acid extraction and amplification device.

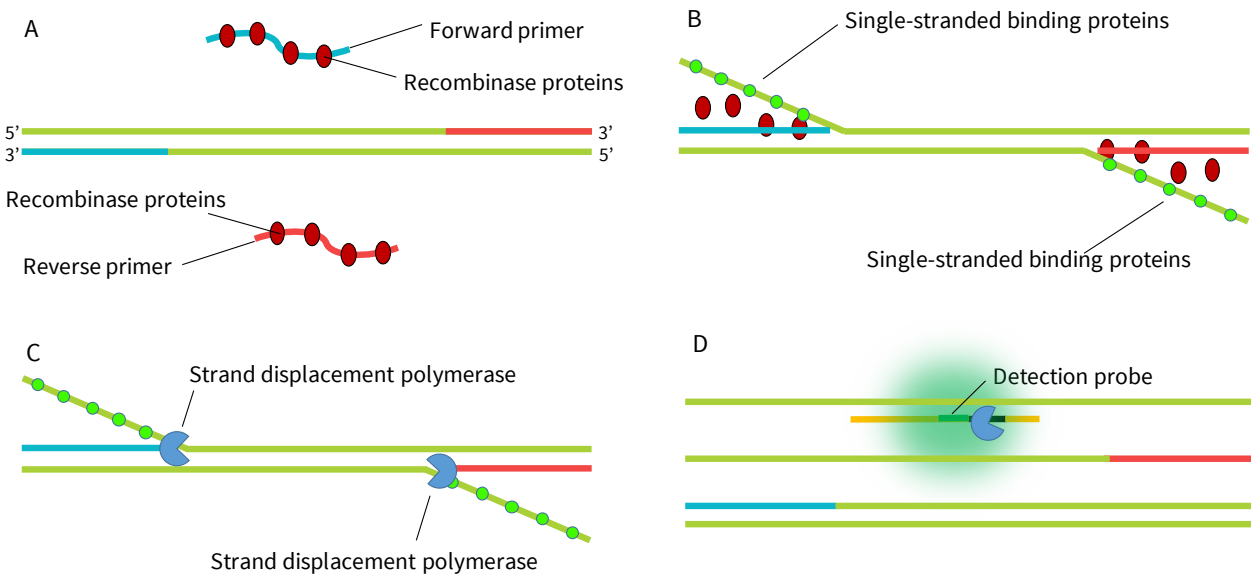



**Figure 1-10. RPA reaction mechanism**

(A) Recombinase proteins bind to the forward and reverse primer before scanning the target for the homologous sequence. (B) The recombinase proteins insert the primer into the target sequence and the displaced strand is stabilized by single-stranded binding proteins. (C) Strand displacement polymerase extends the primers to create a copy of the target nucleic acid strand. (D) A fluorescent probe binds to the produced amplicon, and upon binding, is cut by an exonuclease protein to remove the quencher and cause fluorescence for detection.

## 1.5 Objectives

In the research described in this dissertation, we aim to integrate ITP with isothermal nucleic acid amplification with the goal of creating diagnostic devices that have the potential to improve infectious disease diagnosis at the POC. Our work incorporates many different aspects related to diagnosis with nucleic acids, including sample preparation, target amplification, and detection.

The objectives for this dissertation include:

1. use scientific and engineering principles to rationally integrate ITP and isothermal amplification for POC nucleic acid diagnostics,
2. create diagnostic devices that reduce the number of user steps, instrumentation, and time to obtain relevant diagnostic information, and

3. provide proof-of-concept studies by finding the limit of detection from relevant samples including whole milk, blood serum, and whole blood.

We address these objectives by presenting the design and operation of two different nucleic acid diagnostics that combine ITP's powerful separation and concentration effects with isothermal amplification strategies. We use the ITP design principles presented in Section 1.4 to create ITP systems that can separate and focus nucleic acids from complex milk and blood samples, while integrating them with nucleic acid amplification and detection steps. The presented devices require limited user steps, equipment, and provide results in 20–60 minutes. We believe this approach has the potential to reduce the cost and complexity of POC NAATs.

In chapter 3, we present the Nucleic Acid Isotachophoresis Loop-mediated isothermal amplification (NAIL) device.[105] The NAIL device demonstrates the potential capabilities of combining ITP and isothermal amplification by extracting nucleic acids from complex samples in a single step, while using passive valves and pumps for fluid actuation to provide results with a cellphone in less than an hour. The limit of detection using the Clinical and Laboratory Standards Institute (CLSI) method is shown for *E. coli* O157:H7 cells spiked into whole milk samples.

In chapter 4, we show an improved version of the NAIL device by creating a device that performs single-step extraction and amplification simultaneously. In this work, we use ITP to extract nucleic acids from complex serum or whole blood samples and focus the nucleic acids with reagents required for amplification within the ITP interface. The result of this process is a single-step extraction and amplification method that can provide NAAT results from blood serum or whole blood samples within 10–20 minutes on a strip of glass fiber substrate. We describe the design steps and important parameters for creating the ITP-RPA reactions in an attempt to understand the defining features of the system and show some data that demonstrates relatively low limit of detections and the potential to provide quantitative results.

The following Chapter 2 shows our in-depth review of diagnostic development and evaluation metrics. Descriptions of these metrics are presented in other works, but we provide a concise and clear explanation of the meaning and use the metrics related to diagnostic development. For new diagnostics to achieve maximum impact on patient care, scientists and engineers developing the tests should understand the

analytical and clinical statistical metrics that determine the efficacy of the test. We believe that summarizing these metrics in a clear and concise way will benefit test developers by providing consistent measures to evaluate analytical and clinical test performance, as well as guide the design of tests that will most benefit clinicians and patients.

# Chapter 2: Literature and Tutorial Review: Analytical and Clinical Metrics for Diagnostic Device Development and Evaluation

## 2.1 Introduction

This chapter was published as a Tutorial Review in Lab on a Chip. It summarizes and discusses statistical metrics and procedures for developing and evaluating diagnostic devices. It is broken down into four sections, including analytical statistics, pre-clinical validation statistics, clinic metrics, and case studies that show the use of these metrics. Descriptions of these metrics are presented in other works, but we provide a concise and clear explanation of the meaning and use the metrics related to diagnostic development. We believe that summarizing these metrics in a clear and concise way will benefit test developers by providing consistent measures to evaluate analytical and clinical test performance, as well as guide the design of tests that will most benefit clinicians and patients.

As lab-on-a-chip (LoC) health diagnostic technologies mature, there is a need for scientists and engineers to understand standard metrics that define device performance, as well as how these metrics impact translation from the laboratory to the clinic. Terms such as “sensitivity" are often found in the microfluidic literature, but the lack of precise usage can cause confusion because terminology for analytical and clinical parameters may seemingly overlap, are used interchangeably, or are simply misused.[106] Understanding and correctly applying standardized metrics for both laboratory and clinical purposes will improve device characterization and reporting, as well as provide performance guidelines for scientists and engineers to create better tools for improving patient care. Specifically, engineers and scientists will have a common language to communicate performance of laboratory-based diagnostics between research groups, industry, and clinicians. Focusing on clinical metrics early in the development process will also ensure that devices are relevant to the end-users (e.g. clinicians) and reduce waste by focusing resources towards clinically relevant aspects of the device.

Diagnostic devices play an important role in giving physicians additional information about a patient’s disease status for diagnosis and subsequent clinical management. One goal of diagnostic devices is to identify (“rule-in”) or eliminate (“rule-out”) the likely cause of an illness in a patient presenting with symptoms, but they are also used to give information about the prognosis of a condition, to monitor patients during the course of illness and/or treatment, or to screen otherwise healthy individuals for a disease that has not yet presented itself.[1]

The role of diagnostic devices in evidence-based decision making can be generalized by,

**What you thought before + New information = What you think now,**

where the diagnostic devices are providing the "New information" for clinical decision making.[1,3] "What you thought before" and "What you think now" can be thought of as probabilities of a patient having a particular disease before and after the "New information" is presented.[3] The goal of diagnostics is to provide high enough probability that the disease is present or not present following testing that the clinician can make a decision whether or not to treat the patient. The extent to which the "New information" provided by the diagnostic result changes the potential diagnosis by the clinician depends on many features of the test. While the goal of a diagnostic test is to be as accurate as possible, creating a perfect diagnostic test is extremely challenging, if not impossible, and there are numerous tradeoffs that inevitably occur in balancing accuracy with other test characteristics. Scientists and engineers who understand some of the complexities of the metrics used to evaluate the accuracy of diagnostic tests may be able to provide clinicians with more useful diagnostic tools while considering all of the user needs. Further, using a standardized set of diagnostic performance metrics will allow researchers to connect the analytical and clinical performance of their devices, while unambiguously communicating and comparing the performance of developed devices through publications, and discussions with fellow researchers, clinicians, and industry.

In this tutorial review, we present and discuss analytical measures relevant to laboratory diagnostic development and clinical metrics important to patient treatment. The goal of this review is to (1) reinforce consistent protocols and nomenclature used to evaluate and describe desired analytical performance statistics during laboratory assay development,[107–110] (2) educate the LoC community on important clinical diagnostic statistical measures and how laboratory statistics impact these metrics,[111–113] and (3) encourage researchers to design and validate technologies with common clinical metrics in mind to ensure the devices meet the clinicians' needs.

This paper is broken into four sections that focus on relevant performance metrics for (1) analytical laboratory development, (2) pre-clinical device validation, (3) clinical trials and use, and (4) case studies of existing point-of-care (PoC) diagnostic technologies. Figure 1 summarizes the statistical metrics that will

be discussed with the approximate development stages where they apply. In general, section 1 and 2 metrics are determined in the laboratory with laboratory scale equipment and reference samples for preliminary assay development and validation. Section 3 metrics will most commonly be determined during a clinical trial that may include an FDA clearance application with possible iteration on section 2 metrics. Progressing through section 3 likely requires significant product development with quality control, packaging, etc. for regular clinical use, where the performance will be continuously validated over the long-term.[114]

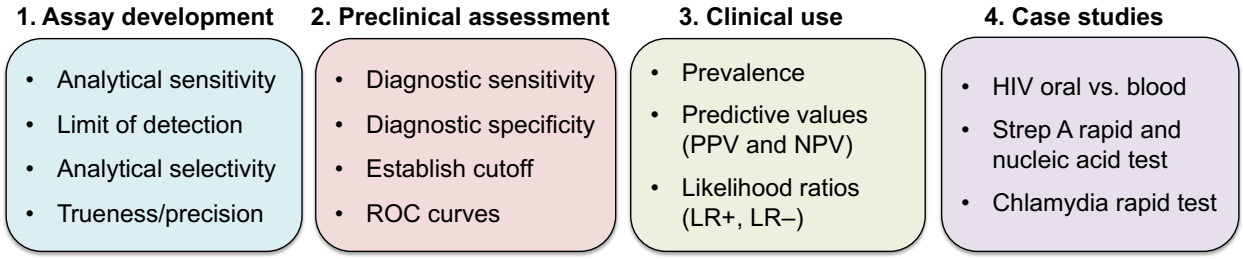



**Figure 2-1. Summary of the topics to be covered in the paper.**

The relevant statistics for general stages of diagnostic development will be explained and followed by example case studies utilizing the presented metrics. In general, stages 1 and 2 are likely relevant for laboratory development and validation, whereas stages 2 and 3 will begin the transition to clinical testing and implementation.

Statistical measures presented in section 1 include analytical sensitivity, analytical selectivity, limit of detection (LoD), and trueness/precision.[107–110,115,116] Section 2 discusses diagnostic sensitivity and diagnostic specificity, establishing clinical cutoffs, as well as receiver operator curves (ROCs).[111] Section 3 describes prevalence and pre-test probability, positive predictive values (PPV), negative predictive values (NPV), as well as positive and negative likelihood ratios (LR+ and LR–).[112,113] The final section presents brief case studies of existing diagnostic technologies and the importance of these metrics in understanding their role in diagnosis and clinical management.

## 2.2 Analytical laboratory statistics: analytical sensitivity/selectivity, limit of detection (LoD), and trueness/precision

The performance of a diagnostic test can be thought of in terms of analytical laboratory metrics and clinical metrics. Most scientists and engineers are more familiar with analytical statistics that include analytical sensitivity,[107] analytical selectivity,[108] LoD,[109] trueness,[116] and precision,[116] while clinicians are typically interested in diagnostic sensitivity and diagnostic specificity,[111] and the resulting predictive

values and likelihood ratios.[112,113] The first step in developing a diagnostic test is demonstrating feasibility, or the ability to detect the analyte under optimized conditions such as pure analyte in buffer. Next, the development and evaluation workflow usually consists of determining analytical metrics such as analytical sensitivity, LoD, analytical selectivity, trueness, and precision (section 1). These analytical metrics provide a wealth of information about the assay including the quantitative change in response for varying analyte amounts (section 1.1), the reportable range and linear response of the assay (section 1.1), the smallest analyte amount that can be detected (section 1.2), how interferents affect the response (section 1.3), and the trueness and precision of the assay (section 1.4).[109,117–119] Figure 2 shows a theoretical plot highlighting some of these metrics. In parallel with, or after analytical evaluation, a set of clinical samples may also be tested to estimate how laboratory performance translates to clinical performance with preliminary diagnostic sensitivity and diagnostic specificity determinations (section 2).[120]

It should be noted that the level of rigor required for these evaluations will depend on the stage of the assay. Laboratory and pre-clinical validation will likely require less rigorous evaluation than a test going through regulatory clearance, but the majority of these metrics can be applied to tests at many stages of development. For instance, sample sizes and replicates may be smaller for tests early in development, such as performing selectivity studies with 5–10 interferents as opposed to 50–100 interferents that may be required for medical-use clearance. Further, early tests may use pure analyte in clean buffer before progressing to contrived samples that consist of blank sample matrices (e.g. pooled serum, urine) with analyte spiked into them, certified reference materials (CRMs)[121], or patient samples. The important takeaway for microfluidic researchers is to apply these metrics where appropriate to their own work, while reporting detailed descriptions of their evaluations using consistent nomenclature and procedures.

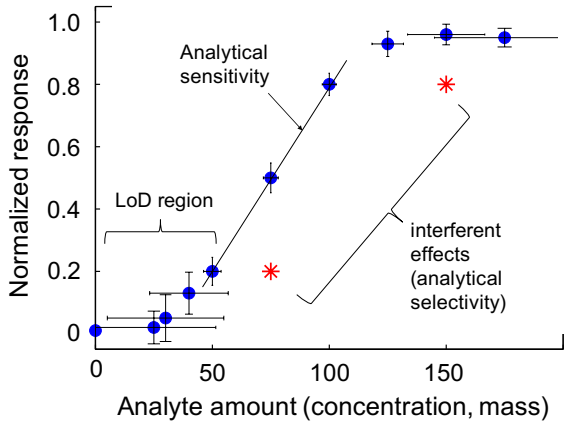



**Figure 2-2. Example of analytical metrics for lab development**

Representative data for analytical sensitivity, analytical selectivity, LoD, and precision for laboratory test development. The normalized response of the test is plotted versus the analyte amount. The measurement precision is characterized by y-axis error bars and the quantitative resolution is characterized by x-axis error bars. The analytical sensitivity of the procedure is the slope of the calibration curve, which is used to determine the quantitative resolution. We used a sigmoidal calibration function to determine the low and high level quantitative resolution x-axis error bars for demonstration purposes (Supplementary Figure S1), but only consider the reportable range to be the linear region of the response from 50–100 units/L due to the higher quantitative certainty in this region. The red stars show data collected in a matrix or sample that contains interferents that affect the response (suppression is shown here, but interferents can also enhance the signal). Finally, the limit of detection (LoD) region is bracketed, where blanks and multiple low level samples are measured so that the smallest analyte amount that is significantly different from the blank response can be determined.

### 2.2.1 Analytical sensitivity

Limit of detection and analytical sensitivity are often confused and used interchangeably to define the lowest concentration or amount of analyte that an analytical procedure can reliably detect.[109] This confusion mostly arises because the term “sensitivity” with or without a modifier has multiple alternative meanings with preferred uses in different fields.

Though there is some debate on the issue,[119,122] the preferred use of “analytical sensitivity” for journals such as Clinical Chemistry and industry standards organizations such as the Clinical and Laboratory Standards Institute (CLSI) is consistent with the International Union of Pure and Applied Chemistry (IUPAC) definition. They define analytical sensitivity as the ability of an analytical procedure to produce a change in signal for a defined change of the quantity being measured, $dR/dx$, where *R* is the measurement output response and *x* is the input variable of interest.[107,109,110] Analytical sensitivity is thus often visualized as the slope of the calibration curve.[107,109] In contrast, LoD is more commonly and precisely used only

to describe the lower levels of detection and as a result, is preferred for defining the detection of low level analytes.[109,123]

Analytical sensitivity plays an important role in describing the quantitative performance of analytical methods and is determined after constructing a calibration function that gives the relationship between the measured response and input quantity of interest.[107] In general, the calibration function is given by,

$$y = F(x) + e_y \ , \qquad (1)$$

where *y* is the response, *x* is the input variable of interest (*e.g.* analyte concentration), *F*(*x*) is the functional relationship between the analyte and response, and $e_y$ is residual error.[107] The simplest case is the linear calibration function, $F(x) = Ax + B,$ where *B* is the intercept and denotes the blank or baseline signal, and *A* is the calibration slope or analytical sensitivity which denotes the change in signal given a change in analyte concentration.

Determining the calibration function allows test developers to calculate the concentration of unknown analyte quantities within certain statistical confidence from the raw signal output of their device. A metric to define the analyte uncertainty given by an assay's calibration function is the quantitative resolution,

$$QR = \varepsilon_r / A, \qquad (2)$$

where $\varepsilon_r$ is the measurement response uncertainty and *A* is the analytical sensitivity of the established calibration function.[119] For a given measurement uncertainty (*e.g.* standard deviation), the quantitative resolution gives the corresponding uncertainty in the determined value of the analyte concentration using the calibration function. The uncertainty described by quantitative responses can be visualized as horizontal confidence intervals with the same units as analyte concentration (x-axis). Higher analytical sensitivity means that the procedure is more sensitive to changes in analyte amount and will give better quantitative resolution for a given measurement error.[119] In Figure 2 for example, the quantitative resolution for 50 units/L is approximately 3.75 units/L, while the 30 units/L data point has a quantitative resolution of 25 units/L due to slightly higher response error and lower analytical sensitivity (decreasing slope magnitude). These quantitative resolutions were calculated using measurement error bars and the analytical sensitivity from Supplementary Information (SI) Figure S1C with a sigmoidal calibration function.

The range of analyte concentrations over which the measurement response gives sufficient quantitative certainty is often referred to as the reportable range, measuring interval, or linear range.[117,120] The reportable range is most commonly the linear portion of the response due to the analyte amounts being directly proportional to response. Linear least-squares regression can be used to determine linearity with $R^2$ goodness of fit, but a number of assumptions must be met to ensure proper analysis, which are well described by other sources.[124,125] For curved calibration functions, the test developers can try to eliminate the cause of the non-linearity, linearize their data, or restrict their analysis to concentrations covered by the linear portion of the calibration, though the reportable range isn't required to be linear.[126] Possible non-linear calibration functions include quadratic, cubic, logarithmic, and sigmoidal relationships, but care should be taken to validate that these models best fit the data (*e.g.* using adjusted $R^2$ values)[125,127] and give quantitative results within acceptable limits of uncertainty over the entire reportable range.[119,126] For example, we use the sigmoidal calibration function in Supplementary Figure S1 to illustrate the effect of analytical sensitivity on quantitative resolution for all data points, but only consider the linear portion in Figure 2 due to large uncertainty outside of this range.

Higher analytical sensitivity for a given a calibration function results in better quantitative resolution. Test developers of quantitative tests should thus desire to have as high of analytical sensitivity as possible across the clinically relevant analyte concentration range. In practice however, there is likely a trade-off between analytical sensitivity and reportable range. For example, the reportable range for Figure 2 is approximately between 50 and 100 units/L, but if the experimenter determined a method to extend it between 50 and 150 units/L, the analytical sensitivity would likely decrease, assuming constant response dynamic range. Supplementary Figure S2 shows an example of this idea using a theoretical $C_t$ *vs.* log(*nucleic acid copy*) plot from qPCR measurements.

The calibration function and analytical sensitivity should ideally be determined using certified reference materials (CRMs) that can establish the relationship between the analyte and response.[107] CRMs are analyte and/or matrix standards that can be used to establish and validate calibrations or other analytical metrics. CRMs consist of the pure substance "primary" reference (the analyte) and matrix based "secondary" reference (pooled patient matrices).[121] These materials have well-established analyte concentrations, background interferent levels, and are often commutable, *i.e.* behave as closely as possible

to patient samples.[121,128] The standards may be obtained from standards institutes (*e.g.* National Institute for Biological Standards and Control); repositories such as the NIAID Cell, Tissue, and Organism Repositories; or private companies that produce certified reference materials. In cases where CRMs do no exist, are not available in a preferred matrix, or are not practical, in-house spiked samples with available analyte or primary CRM and available blank matrices or secondary CRMs can be created and used. Measurements establishing or validating the calibration curve should cover the entire range potentially used in practice (typically 0–150% or 50–150% of clinically relevant levels) and be performed at least in duplicate, but preferably triplicate or more in random order.[126]

### 2.2.2 Limit of detection

LoD provides the lower bound of the test dynamic range and will have an effect on the diagnostic sensitivity of the test. Determining the LoD means finding the lowest amount of analyte that can reliably be detected and differentiated from a blank sample with no analyte.[109] Figure 3A shows distributions of repeated measurements of a blank sample containing no analyte and a sample containing low-levels of analyte that overlap due to low measurement signal and random error. For example, in Figure 3A if a low-level sample is repeatedly measured, the low-level analyte distribution will overlap with the measurements from a blank sample containing no analyte. In this overlap region, there is the possibility of characterizing a blank sample as having analyte present, which is known as a false positive and characterized by Type I error ($\alpha$ error). There is also the possibility of characterizing a sample that contains low-level analyte as being a blank sample, which is known as a false negative and characterized by Type II error ($\beta$ error).

The LoD, or the amount of low-level analyte that can be reliably measured, is essentially determined by finding an analyte concentration that experiences set levels of Type I error ($\alpha$ error) and Type II error ($\beta$ error) that limit false positive and false negative measurements within certain statistical confidence. The International Organization for Standardization (ISO) recommends a default level of $\alpha = \beta = 5\%$ for the LoD.[109],[16,29] The CLSI EP17 document details an approach for determining the LoD of in vitro diagnostic tests with Type I and Type II errors set to 5% using a combined statistical and empirical method.[109] The strength of this method is that it uses statistics and measurements from both blank samples and low-level

samples to determine the LoD, as opposed to methods such as three standard deviations above the blank, which have been noted to "define only the ability to measure nothing".[131,132]

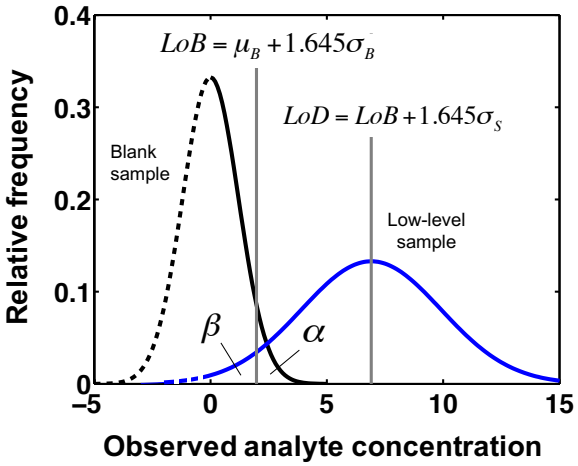



**Figure 2-3. Limit of blank and detection example**

Plot showing distributions of measurement replicates for blank samples and a sample containing low-levels of analyte. The overlap region due to measurement uncertainty causes some blank samples to be measured as containing analyte (false positive) and low-level samples to be measured as containing no sample (false negative). The LoB and LoD can be determined by setting the overlap error to a given certainty. The equations to determine the LoB and LoD for $\alpha = \beta = 5\%$ overlap error are given, with the LoB approximately equal to 2 units/L and the LoD approximately equal to 7 units/L. The dashed line indicates that some instruments do not report measurement values less than zero, resulting in a non-normal distribution that requires non-parametric statistics to calculate the LoD.[109]

The first step of the LoD procedure is to define the limit of blank (LoB) using measurements on blank samples. For a Gaussian distribution of blank values, the LoB is defined as,

$$LoB = \mu_B + 1.645\sigma_B \,, \quad (3)$$

where $\mu_B$ is the blank mean and $\sigma_B$ is the blank standard deviation. If the blank distribution cannot be determined to be Gaussian, such as when blank values are truncated at zero, a non-parametric procedure is required for determining the LoB. A discussion of the non-parametric procedure is shown in section 3 of the Supplementary Information, with more details in the EP17 document.[109]

After finding the LoB, the LoD can be determined by,

$$LoD = LoB + c_\beta \sigma_S, \quad (4)$$

where $\sigma_S$ is the sample standard deviation, $c_\beta = 1.645/(1-1/(4\cdot f))$, and $f$ is the degrees of freedom.[109] The sample standard deviation should be estimated using a pooled standard deviation (see section 3 of Supplementary Information for equation) from four to six low-level samples with concentrations ranging from approximately the LoB to 4 x LoB.[109] The degrees of freedom are the total number of samples tested minus the number of different low-level analyte concentrations tested ( $N_S - K$ ). For example, if 100 measurements are made on 5 different low-level analyte concentrations, e.g. 2, 4, 6, 8, 10 units/L, $f = N_S - K = 100 - 5 = 95$ and $c_B = 1.645/(1-1/380) = 1.649$. For large numbers of repeated measurements, the LoD simply becomes $LoD = LoB + 1.645\sigma_S$, because $c_B$ approaches a minimum of 1.645 for large *f*.

Figure 3B shows example frequency distributions with the LoB, LoD, and $\alpha = \beta = 5\%$ regions labeled. While $\alpha = \beta = 5\%$ default values are recommended by ISO, particular situations may require more stringent control over type I or type II error, where $\alpha$ and/or $\beta$ may be adjusted to provide better estimates for that situation.[123] The procedure for different type I and type II errors will be similar, except the $\alpha$ and/or $\beta$ levels will change the coefficients that the blank and low level sample standard deviations are multiplied by (the coefficient value is only 1.645 for $\alpha = \beta = 5\%$). The coefficients for desired confidence levels can be obtained from standard z-score tables. In any case, when reporting LoB and LoD values, the $\alpha$ and $\beta$ levels, as well as the statistical and experimental methods used to determine these values should be thoroughly documented and reasoned. For instance, if the test is to be used in complex samples with possible interferents, the measurement process should indicate whether the detection limits were determined in the presence of that complex sample because otherwise, the actual performance may fall short.[123]

LoD is a useful quantity because it defines the lower detection range of a given test, which will in turn impact the diagnostic sensitivity of the test. By using LoD to define lower detection limits, analytical sensitivity can be used in its most common form to describe the ability of assays to detect changes in analyte levels. The CLSI procedure to determine LoD uses variance from both blank samples and low level analyte measurements to determine the LoD within statistical confidence levels recommended by ISO. In general,

less noise from blank samples, minimizing the variance of low-level sample measurements, and increasing the number of replicate measurements will decrease the LoD. Further, using the CLSI method necessitates the determination of test variance at various low level concentrations, which is an important step towards determining the precision of the assay (see section 1.4). For more information concerning required number of samples for determining LoB and LoD, how to account for non-constant sample standard deviation between the LoD concentration and other low-level concentrations, an example LoD determination, and other details, we refer the reader to the CLSI EP17 document.[109]

### 2.2.3 Analytical selectivity

Analytical selectivity defines a test's ability to deliver signals that are free from interference in the presence of components other than the analyte in the sample[108,110] While use of analytical selectivity and analytical specificity may overlap or be used interchangeably, analytical selectivity is preferred because analytical specificity is considered an absolute term and cannot be measured.[108,110] The IUPAC quotes Gary D. Christian's view: "A specific reaction or test is one that occurs only with the substance of interest, while a selective reaction or test is one that can occur with other substances but exhibits a degree of preference for the substance of interest. Few reactions are specific, but many exhibit selectivity".[108,133] In other words, analytical specificity is absolute and exhibits essentially perfect selectivity. Selectivity is also preferred because it can be quantified to determine the extent of interference from compounds present in complex mixtures.[108,134,135]

Determining analytical selectivity is necessary during laboratory development and characterization because clinical samples will contain a wide range of interferents at varying concentrations due to patient-to-patient heterogeneity. Ensuring that a test preferably detects the target of interest in the presence of heterogeneous patient interferents may improve both diagnostic sensitivity and diagnostic specificity, which will give more accurate diagnostic results in patient samples. Interferents can both enhance or suppress typical test results in the absence of the matrix. In other words, interferents may produce some level of signal in blank samples, increase the signal for a given analyte amount in matrix-free solution, or reduce/eliminate the signal produced by a given analyte amount in matrix-free solution.

Selection of interferents to test should be guided by an understanding of the chemistry and physics of the measuring system, as well as the clinical setting and use.[110,118,136] Possible sources of interferents include patient produced metabolites, treatment compounds (*e.g.* drugs, anticoagulants), substances ingested by the patient (*e.g.* food*,* alcohol, vitamins), substances added during sample preparation (*e.g.* preservatives), and compounds present in the sample matrix itself (*e.g.* proteins, hemoglobin, bilirubin, lipids).[118] The specific interferents may include competing non-target organisms (cross-reactants), impurities, degradents, and matrix components that can affect the chemical reaction, physical properties of solution, and detection of the signal.[110,118]

For many disease targets, particularly infectious diseases (e.g. tuberculosis, MRSA, chlamydia), the FDA has created "Guidance for Industry and FDA Staff—Establishing the Performance Characteristics of In Vitro Diagnostic Devices" documents that provide recommended organisms, matrix interferents, and other compounds to use for analytical selectivity determination and validation.[137] The CLSI EP-7 "Interference Testing in Clinical Chemistry; Approved Guideline—Second Edition" and a review from Dimeski also provide common interferents to consider for clinical samples.[118,136]

While test developers cannot reasonably expect to test every possible interferent at every possible concentration, critically analyzing the system in development and validating the test with the most likely and potentially detrimental interferents will increase confidence in the test performance with variable samples. Further, performing selectivity experiments in controlled lab-based settings prior to running clinical trials will increase the probability of the test performing well in heterogeneous clinical samples.

Different methods have been proposed for quantifying the analytical selectivity of an analytical procedure,[108,135] but here we will briefly summarize the procedure for testing interferents proposed by the CLSI EP7–A2 document, with a slight modification on the reporting of results.[138] The most common methods to test interference effects include: (1) "paired-difference testing" where interferents are added to controls one-by-one and tested for bias, (2) "interference screen" where many potential interferents at high "worst case" concentrations are all tested together, and (3) testing actual patient samples in comparison to a well-established reference procedure.[118] Options 1 and 2 can be performed using contrived samples in the laboratory. Option 3 requires obtaining patient samples for both the test case (*e.g.* a patient with the

disease taking a potentially interfering drug), and patient control samples (*e.g.* patients with the disease not taking the drug) that span the different analyte concentrations.[139]

For “paired-difference testing”, potential interferents should be tested at high and low concentrations for two different analyte concentrations and compared to a positive control sample.[118] The chosen sample used to create the test and control pools depends on the stage of development, as well as the intended clinical use. For instance, if interference testing is being performed for FDA clearance and the intended use is in blood serum, then the control sample should be standardized blood serum (possibly secondary CRM) that will have some average level of proteins, electrolytes, hormones, etc., and then possible interferents should be added to the test pool to observe variations from this control pool. On the other hand, if a test is in very early stage development and/or does not have a specific clinical sample in mind, then the control may just be an interferent-free buffer. The high and low levels used for interferents should take into account the maximum and minimum levels observed in patient samples and the CLSI EP14-A2 provides a number of recommendations for different interferents.[118]

To calculate the interference for each test, the observed difference between the mean sample and mean control can be calculated, $d_{obs} = \bar{x}_{test} - \bar{x}_{control}$, along with 95% confidence intervals for the observed differences. Figure 4 shows a representative plot of percent interference ($100 * d_{obs}$) with 95% confidence intervals for a number of interferents tested using “paired-difference testing”. This plot can be generated for any pair of the high and low levels of interferent and analyte, and then compared to a suitable positive control. If a clinically acceptable percent bias—*e.g.* 10% difference between the control and interferent—has been established *a priori*, the interference of each sample may be tested for clinical significance by simply determining if $d_{obs}$ is above the cutoff.[118]

For the data given in Figure 4, both interferents B and D would be determined to be clinically relevant for a 10% difference cutoff, with B suppressing the response and D enhancing the response. The CLSI EP7–A2 document only recommends to report clinically relevant interferents, but provides a power calculation to determine the number of replicates needed to determine significance for a given set of experimental conditions. However, interferents A and E still appear to have some level of interference and statistical testing with a null hypothesis that the interference is equal to zero would conclude that these interferents do indeed interfere (interferent C would be concluded to show no interference). Krouwer thus recommends

to report the percent interference for each compound if some level of interference is detected (with 95% confidence intervals) as well as the clinically relevant conclusion.[138] For more details, we refer the reader to the CLSI EP7-A2 document, as well cautionary words from Krouwer about the reporting of interference testing data.[118,138]

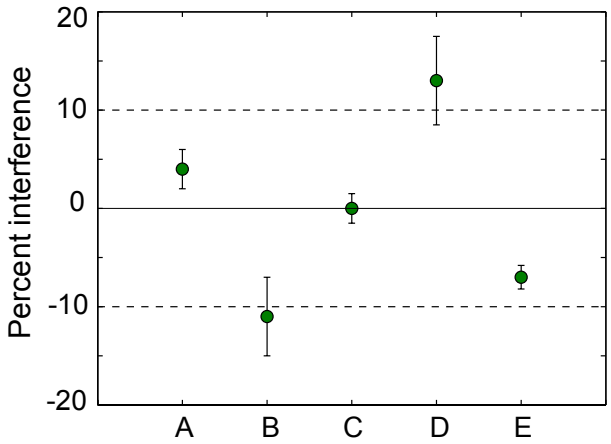



**Figure 2-4. Analytical selectivity example**

Example of the effect of interferents on the response with 95% confidence intervals. All interferents exhibit some level of interference except for interferent C. Interferents B and D exhibit clinically relevant interference according to the established 10% cutoff, denoted by the dashed line. The 10% cutoff was chosen arbitrarily and would be determined as the acceptable bias for a particular clinical application. Interferents A and E exhibit interference significantly greater than zero, but not above the clinically relevant threshold. This analysis can be performed for all levels of interferents (e.g. high and low level interferents) and analytes (e.g. high and low level analytes).

### 2.2.4 Trueness and precision

In order for a diagnostic device to be useful for clinicians, it must give both accurate and precise results. The accuracy of the device refers to the trueness of the result, *i.e.* how close the determined analyte value is to the actual analyte value, which can be used to predict the disease status of the patient.[140] Precision, on the other hand, refers to the closeness of independent measurements for a given set of experimental conditions, regardless of whether they are accurate or not.[141] Improving accuracy is achieved through reducing systematic errors that are described using bias, and improving precision is achieved through reducing random errors that are described using common statistics such as standard deviations.

Assays can be accurate but not precise, precise but not accurate, neither, or both. The common example describing accuracy versus precision is the dart board analogy where, for example, darts being clustered on a dart board at a location other than the bulls-eye indicate precise but not accurate, while darts clustered

at the bulls-eye indicate precise and accurate. Ideally, diagnostics are both accurate and precise, resulting in independent measurements that consistently give results with low variability, while providing analyte values close to the true value that correctly predict the disease status of the patient.

The accuracy or trueness of a test is often determined by comparing the results of the test in development to CRMs or samples analyzed by a high-quality reference test whose results are known to be correct.[117,126] For quantitative tests, the difference or bias between the new test and reference test or CRM can be evaluated and quantified using well-known methods including Bland-Altman difference plots and/or correlation comparison plots.[140,142] It is recommended that these comparisons be completed over time at varying concentration levels, and be conducted using sample matrices that are representative of intended-use patient samples. Statistical methods to calculate the disagreement or bias are given by a number of resources including Westgard, CLSI, and others.[3,117,140,143]

In cases where quantitative data are converted to a binary yes or no answer for diagnostic purposes, or for qualitative tests such as visually-read lateral flow assays, the trueness or accuracy is typically determined by simply comparing the positive versus negative test outcomes to the reference materials or CRM.[126] This comparison method is analogous to the clinical accuracy of the test using patient samples, which is described by diagnostic sensitivity and diagnostic specificity. Section 2.1 provides an in-depth discussion of diagnostic sensitivity and diagnostic specificity.

It should be noted if a reference test or quality CRMs are not available for the target indication, the evaluation of trueness becomes more challenging. In the laboratory, bias may be investigated by using primary CRMs spiked into relevant blank matrices or pure analyte spiked into secondary CRMs. However, caution must be exercised when using these samples because spiked analytes may differ from native analytes.[126] In these cases, low bias does not necessarily reflect high accuracy or trueness, but large bias with these samples should raise concern about the trueness of the test in more representative samples. For evaluating unknown patient samples without a reference method, we refer the reader to an FDA document discussing these issues.[144]

Determining the precision of a diagnostic test allows the developer to estimate the impact of possible sources of variability or imprecision for the assay including patient samples, calibrations, reagents,

operators, time, etc.[141,145,146] Device precision may be particularly important for microfluidic developers due to focus on point-of-care (POC) diagnostics. Lab-based diagnostic tests that use central measuring systems have well-established quality control procedures for diagnostic test systems, with most using Westgard type rules by trained laboratory personnel.[140,146] Further, many lab-based systems can test for precision in batches due to their high-throughput nature. POC devices on the other hand, are commonly single-use cartridges (with or without a reader) and may experience higher environmental variability, including the training of personnel using them. As a result, these devices often require internal controls or calibrations with each measurement and highlight the need to rigorously study the performance of these devices under variable conditions.

The imprecision of quantitative assays is characterized by familiar metrics such as standard deviations (SD) and coefficients of variance (CV).[141,146] For qualitative assays, precision testing using SDs and CVs can still be completed if the qualitative result is derived from a quantitative measurement such as optical density. If an assay is truly qualitative, reliability of positive versus negative results with consideration for the factors discussed below is sufficient. CLSI EP12 documents statistical methods for analyzing and reporting qualitative results.[147]

Precision may be referred to as repeatability or reproducibility, where repeatability generally refers to the variability of a test within a single laboratory, while reproducibility refers to the variability of the test between multiple laboratories.[141] Methods for determining the precision of the test are given by various sources including CLSI EP5, EP10, and EP15 documents (EP5 is summarized by Chesher), Westgard protocols, as well as FDA guidance documents.[116,137,145,146,148] The details of the protocols differ in exact details and are geared towards lab-based measurement systems and quality control validation, but a number of key considerations for precision testing can be generalized from these sources.

- **Within-run and within-day replicates.** Generally, 2–3 replicates of each experiment (within-run) should be repeated twice per day (within-day), *i.e.* aliquoting the same test sample into 2–3 different tests, waiting at least 2 hours and then repeating the procedure. If the instrument is capable of simultaneously running multiple samples, the 2–3 replicates may be run in a single trial. Single-use disposable tests require consecutive replicates or running multiple devices simultaneously.

- **The time period of the experiment (between-day).** The mentioned guidance documents all recommend studying between-day variability by testing samples over the course of multiple days (up to 20 days). Increasing the testing period helps to account for random error associated with factors such as laboratory temperature and humidity, stability of reagents and samples, and other possible sources of drift associated with time.

- **Various analyte levels**. The analyte test panel should consist of at least 2 and preferably 3–6 different analyte concentrations. These samples should consist of high and low-level analyte concentrations. The FDA guidance documents recommend testing at negative (no analyte), high negative ($C_5$), low positive ($C_{95}$—same as LoD), moderate positive (expected 100% positive results), and high positive (near the upper limit of the range of the assay) levels of analyte.

- **Different operators.** Ideally, the operators would receive the level of training administered to actual device operators. Use of multiple operators of appropriate training levels is particularly important for POC devices, where the end-user will not necessarily be laboratory staff. The FDA recommends using additional operators when testing POC versus laboratory tests. Further, for tests that require visual determination, such as lateral flow assays, using different operators can help account for variability when interpreting the results.

- **Patient or sample matrix variability.** Patient samples will contain non-uniform levels of interferents, which may affect the assay precision. In the case where multiple patient samples are not available, this variability can be attempted to be simulated using different CRMs or contrived samples that contain variable levels of interferents.

- **Lot-to-lot precision.** This variability may include different sets of test reagents, batches of physical devices, and/or different calibration lots.

- **Sites or laboratories.** Study of these different factors should occur at multiple laboratory or testing sites to account for inter-laboratory differences. For POC devices, these sites should include non-laboratory settings of intended use such as physician's offices.

The stringency of testing for each of these factors obviously depends on the device being developed and its stage of development. For example, testing 5 analyte levels with 3 replicates of different patient samples over 12 days using 3 operators at 3 different sites is the level of study required for FDA clearance and likely not feasible, nor necessary for the majority of researchers during development. However, awareness of these different sources of variability and conducting smaller-scale versions of these precision studies are important steps in determining the repeatability of device operation. For instance, time, replicate, and analyte level factor studies can be accomplished by the majority of laboratories. In any case, reporting the standard deviations or coefficient of variance results should specify the factors taken into account during study. CLSI EP5 and Chesher provide calculations for simple repeatability studies, while the FDA guidance documents provide sample tables for reporting precision studies, which can likely be adapted for smaller-scale laboratory purposes.[137,141,145]

### 2.3 Diagnostic sensitivity, diagnostic specificity, and establishing clinical cutoffs

The clinical accuracy of a diagnostic test is generally characterized by diagnostic sensitivity and diagnostic specificity, which characterize the ability of a new test (often called the 'index test') to accurately detect the presence or absence of disease from clinical samples compared to a "gold standard" reference test.[110,111] Diagnostic sensitivity measures the proportion of true positives compared to the reference test, while diagnostic specificity measures the proportion of true negatives compared to the reference test.[111]

#### 2.3.1 2x2 tables for determining sensitivity and specificity

The typical method for presenting comparison results between an index test and reference test is the 2x2 table, where the reference test is commonly assumed to give the actual disease state of the patient (disease positive, D+ or disease negative, D–).[1,3] Table 1A shows how index versus reference data is interpreted and Table 1B shows how the data is organized in a 2x2 table. The comparison results are given as True Positives (TP): both tests give positive answers, True Negatives (TN): both tests give negative answers, False Positives (FP): the new test gives a positive result, but the gold standard gives a negative result, and False Negative (FN): the new test gives a negative result, but the gold standard gives a positive result.

**Table 2-1. 2x2 tables for diagnostic sensitivity and specificity**

Classifying comparison results of new tests versus reference gold standard reference tests in a 2x2 table. (A) Table demonstrating how results are classified based on the answer provided by the new test and gold standard reference for 6 different example test outcomes. When the new test and reference test are both positive (+), the result is a true positive (TP). When the new test and reference test are both negative (–), the result is a true negative (TN). A new test positive (+) and reference test negative (–) is a false positive (FP), and a new test negative (–) and reference test positive (+) is a false negative (FN). (B) A 2x2 table that organizes the comparison of results data based off of the number of TP, FP, FN, and TN. The number of results from table A are indicated in parentheses. From this data, the diagnostic sensitivity and specificity of the new test can be calculated according to equations 5 and 6. The diagnostic sensitivity of the new test would be 67%[†] and the diagnostic specificity would be 75%[‡].

**A**

| **New test** | **+** | **–** | **+** | **+** | **–** | **–** | **–** |
|---|---|---|---|---|---|---|---|
| **Reference test** | **+** | **–** | **–** | **+** | **–** | **+** | **–** |
| **Result** | TP | TN | FP | TP | TN | FN | TN |

**B**

| | **Reference test** | |
|---|---|---|
| **New test** | Reference (+) | Reference (–) |
| New (+) | TP (2) | FP (1) |
| New (–) | FN (1) | TN (3) |

Using the 2x2 tables, the diagnostic sensitivity and diagnostic specificity can be directly calculated as:[1,3]

$$Sensitivity = \frac{TP}{TP + FN} \ , \ Specificity = \frac{TN}{TN + FP} \ . \quad (5), (6)$$

Diagnostic sensitivity can be thought of as the ability of the test to give positive results when the target condition is present and is calculated as the number of true positives detected by the test over the number of actual positives as determined by the reference test. Diagnostic sensitivity can often suffer when the test does not have the ability to detect low levels of target due to the target analyte level being below the LoD of the test or when interferents in a clinical sample inhibit detection of analytes. Diagnostic specificity can

[†] TP/(TP+FN)=2/(2+1)=67%

[‡] TN/(TN+FP)=3/(3+1)=75%

be thought of as the ability of the test to give negative results when the target condition is absent, and is calculated as the number of true negatives detected by the test over the number of actual negatives as determined by the gold standard. Examples where diagnostic specificity may suffer include non-specific binding for immunoassays, non-specific priming in nucleic acid assays, tests that have low signal-to-noise ratio, or tests that contain interferents that increase the test response (*e.g.* hemoglobin with fluorescence-based readouts).

### 2.3.2 Determining clinically relevant cutoffs and receiver operating curves (ROC)

When determining whether a particular test result indicates disease positive or negative for a given sample, a clinically relevant decision cutoff or threshold must be established, which plays a role in determining the diagnostic sensitivity and specificity of the test.[1,149–151] Diagnostic sensitivity and specificity are inversely related, meaning that as diagnostic sensitivity increases, diagnostic specificity decreases, and vice versa. Figure 5 shows how establishing different clinical cutoffs at certain test readouts affects the diagnostic sensitivity and specificity of the test.

In Figure 5A, a clinical cutoff value of 0.25 gives a diagnostic sensitivity of 100%, but a diagnostic specificity of only 60%. Conversely, a clinical cutoff value of 0.50 gives a diagnostic sensitivity of 50%, but a diagnostic specificity of 100%. A clinical cutoff in the middle at 0.35 gives a diagnostic sensitivity and diagnostic specificity of 80%. For tests where positive and negative readouts overlap to some extent, clinical cutoffs chosen to increase diagnostic sensitivity will have the effect of decreasing diagnostic specificity, with the reverse also true. Ideally the disease positive and disease negative test readings (y-value) are separated by some distance in the vertical direction allowing for a cutoff that would allow 100% diagnostic sensitivity and specificity.

The established LoD plays an important role because it provides the lower bound for the clinical cutoff choice for a test. For example, if the LoD of the test from Figure 5 is only 0.40, the ability to choose clinically relevant cutoffs is restricted by the LoD. In this case, the best diagnostic sensitivity that the test could achieve would be 60% if the clinical cutoff were established at the LoD of the test. For clinical situations where higher diagnostic sensitivity is required, the test would not be able to give the needed diagnostic

sensitivity. However, if the LoD of the test were 0.20, then the test developer would have greater freedom to choose which clinical cutoff gives the most clinical impact, including one with high diagnostic sensitivity.

### 2.3.3 2x2 tables for determining sensitivity and specificity

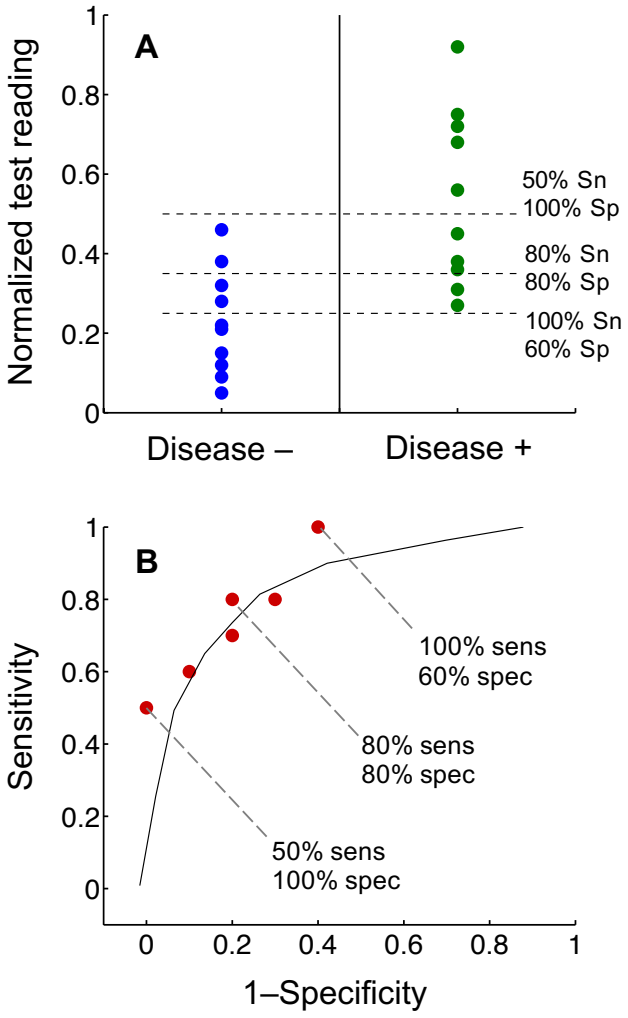



**Figure 2-5. Setting clinical thresholds and receiver operating curves**

(A) Plot showing the normalized measurement of a test for disease positive and disease negative patients. Depending on the chosen clinical cutoff, the diagnostic sensitivity and specificity of the test varies. For low cutoff values, the diagnostic sensitivity (Sn) of the test increases, while the diagnostic specificity (Sp) decreases. For high cutoff values, the diagnostic sensitivity of the test decreases, while the diagnostic specificity increases. (B) The diagnostic sensitivities and specificities associated with varying cutoffs can be visualized using receiver operating curves (ROC), which plots diagnostic sensitivity (true positive rate) versus 1–diagnostic specificity (false positive rate) for each different cutoff. Each point on the plot corresponds to a different cutoff, with the three points corresponding to the cutoffs from plot (A) labeled. In general, points closer to the upper left-hand corner indicate a more accurate test, but the clinical implications of false positive and false negatives should be taken into account when determining a cutoff.

The diagnostic sensitivity and specificity of a particular test changes depending on clinical cutoff selection and the cutoff for each test should be chosen with the particular clinical use in mind. For instance, tests for deadly illnesses without a cure would want a clinical cutoff with high diagnostic specificity to prevent the anguish associated with being falsely diagnosed with such a serious condition. On the other hand, tests for life threatening but treatable diseases, would require a lower clinical cutoff for high diagnostic sensitivity so that even low levels of disease could be detected and treated. In this case, the clinical cutoff may simply be set at the LoD of the test. These examples are extreme ends of the spectrum, but show that the chosen clinical cutoff for a given diagnostic needs to account for clinical relevance. In general, cutoffs chosen for screening tests should maximize diagnostic sensitivity, cutoffs for confirmatory testing should maximize diagnostic specificity, and cutoffs for general diagnosis should balance the two, but each choice will depend on the clinical application .[147]

The receiver operating curve (ROC) is one common method of examining how different cutoff values affect diagnostic sensitivity and specificity.[149,150] True positive rate (diagnostic sensitivity) on the y-axis is plotted versus false positive rate (1–diagnostic specificity) on the x-axis for different possible cutoffs, as shown in Figure 5B. A test with no diagnostic value would be a straight $y=x$ line, while in general, data points that lie in the upper left hand corner indicate better tests with higher diagnostic sensitivities and specificities. However, the cutoff used for a particular application should take into account the clinical implications of false positives and false negatives. Net benefit measures, such as the weighted comparison, provide the overall impact of test results by taking into account diagnostic sensitivity, diagnostic specificity, misclassification costs (clinical effect on the patient of false positives versus false negatives), and in some cases, disease prevalence.[151–153] Details of these measures are outside the scope of this review, but they can be used to determine cutoffs that provide the highest clinical impact and we encourage the reader to explore these ideas in further reading.[151–153]

Diagnostic sensitivity and specificity are likely the first values that scientists and engineers will determine when transitioning their LoC devices from the laboratory to clinical use. A small number of certified reference samples or clinical samples can be measured with the new device and compared to the gold standard test in order to determine preliminary diagnostic sensitivity and specificity, as well as evaluate potential clinical cutoffs. In addition to (or in place of) reporting diagnostic sensitivity and specificity for a

specific cutoff, we recommend generating plots similar to Figures 5A and 5B when reporting preliminary clinical sample results. These plots provide more details about the accuracy of the device being evaluated and allows potential users to fully evaluate clinical usefulness at different cutoffs as opposed to a single diagnostic sensitivity and specificity for one cutoff.

While diagnostic sensitivity and diagnostic specificity are important values for establishing the accuracy of a test compared to a reference test, they do not provide the probability that a certain test result will give a correct diagnosis. In other words, a test with 90% diagnostic specificity does not mean that a person who tests positive has a 90% chance of having the condition being tested for. In order for the test to give the probability that someone does or does not have a disease requires knowledge of diagnostic sensitivity, diagnostic specificity, and the probability that the person has the disease prior to testing is required, as will be explained in section 3.[1,112]

## 2.4 Clinical statistics: prevalence, predictive values, likelihood ratios

### 2.4.1 Prevalence and pre-test probability

The ultimate goal of a diagnostic test is to provide the physician with enough evidence about the patient's condition that they can make clinical decisions (*e.g.* treat, do not treat, perform further testing). Diagnostic sensitivity and specificity are important metrics for evaluating diagnostic tests because they provide a comparison between the index test and gold standard, but these metrics by themselves do not provide probability information for whether a patient does or does not have a disease for a given test result.[112] In order for a test to give the probability that a patient does or does not have a disease of interest following testing, knowledge of the prevalence or pre-test probability is required in addition to the diagnostic sensitivity and specificity.[112]

For a given population, prevalence can be thought of as the probability of the patient having the disease before the test is conducted.[1,112] The diagnostic test result will then change this probability with a positive diagnostic test increasing the probability of disease and a negative diagnostic test decreasing probability of disease. The extent that the diagnostic test changes the probability is determined by the diagnostic sensitivity and diagnostic specificity, but the test result only makes up one part of the decision making process, with the prevalence or pre-test probability of disease making up the other part. This idea can

generally (but not mathematically) be thought of as: **prevalence + test result → predictive value.**[3]

While prevalence and pre-test probability may seem to be analogous, prevalence has a very specific meaning and a few key differences between prevalence and pre-test probability (see section 3.3) should be noted. Prevalence is typically established through cross-sectional clinical studies and is characteristic of the studied population, meaning that it applies only to the studied population and cannot be transferred to dissimilar populations.[113,154] The population may be very broad or very specific. For example, a broad population may be any person presenting to a clinic complaining of a sore throat in any region at any time, while a specific population may be male children under the age of 18 from Seattle, WA presenting to a health clinic with complaints and symptoms of a sore throat between the months of November and April. In either case, the prevalence established by that study is defined only for that specific population. The key for researchers conducting a clinical trial that estimates prevalence is to sample the appropriate number and type of patients that will broadly represent the population that they are interested in, while still accurately portraying that specific population.[155,156]

Pre-test probability can be considered more general and while it may use prevalence as a starting point, it also takes into account other relevant patient information such as patient's signs and symptoms not included in the clinical study, patient history, as well as additional test results.[113,154] As a hypothetical example, we might know that the prevalence of Myocardial infarction (MI), *i.e.* heart attack, in US males over the age of 60 is 5%, but the pre-test probability of MI in a 65 year old male smoker with high cholesterol, high blood pressure, and a family history of MI, presenting to an emergency department with crushing chest pain (symptoms typical of an MI) would be far higher than 5%. For certain cases such as screening tests—asymptomatic people presenting without clinical features—prevalence and pre-test probability will likely be equal.

### 2.4.2 Predictive values and the impact of prevalence

Predictive values use both prevalence, as well as diagnostic sensitivity and specificity in order to determine the probability that a patient does or does not have a disease following testing. The probability that a patient is disease positive or disease negative following a diagnostic test is described by positive predictive values (PPV) and negative predictive values (NPV) respectively. PPV is the probability that someone who tests

positive truly has the condition of interest, while NPV is the probability that someone who tests negative truly does not have the condition of interest. PPV and NPV can be determined using 2x2 tables or mathematically with prevalence, diagnostic sensitivity, and diagnostic specificity as,[112]

$$PPV = \frac{TP}{TP+FP} = \frac{sensitivity \cdot prevalence}{sensitivity \cdot prevalence + (1-specificity) \cdot (1-prevalence)} \quad (7)$$

$$NPV = \frac{TN}{TN+FN} = \frac{specificity \cdot (1-prevalence)}{specificity \cdot (1-prevalence) + (1-sensitivity) \cdot prevalence} \quad (8)$$

It can be seen from the first equality in equations 7 and 8 that the PPV is the number of true positives over the total number of positive test results regardless of whether they are correct or not, and that NPV is the number of true negatives over the total number of negative test results regardless of whether they are correct or not. The second equality shows that prevalence plays a role in determining PPV and NPV, along with diagnostic sensitivity and specificity.

Figure 6 shows the impact of prevalence on PPV and NPV using natural frequencies.[157,158] Consider a scenario where 500 people are known to have a given disease out of every 10,000 people in the population, *i.e.* it has a prevalence of 5%, and a test with 95% diagnostic sensitivity and 95% diagnostic specificity exists for the disease. Figure 6 shows that this test would result in 475 of the diseased persons being correctly identified, but also 475 healthy persons incorrectly diagnosed. In other words, this translates to an NPV of 99.7%[§], but a PPV of only 50%[**]. The reason why the PPV is so much lower than the NPV despite high and equal diagnostic sensitivity and specificity is because the 5% prevalence means that there is low probability that the patient has the disease before the test, *i.e.* there is a 95% chance the person does **not** have the disease before testing. While a change from 5% to 50% probability is a relatively significant change, it means the patient still only has a 50/50 chance of being disease positive, simply because the prevalence was low before the test was run. Alternatively, a negative result helps to confirm what is most

[§] TN/(TN+FN)= 9025/(9025+25)=99.7%

[**] TP/(TP+FP)=475/(475+475)=50%

likely prior to testing—that the patient does not have the disease being tested for—and gives very high confidence that the patient does not have the target disease.

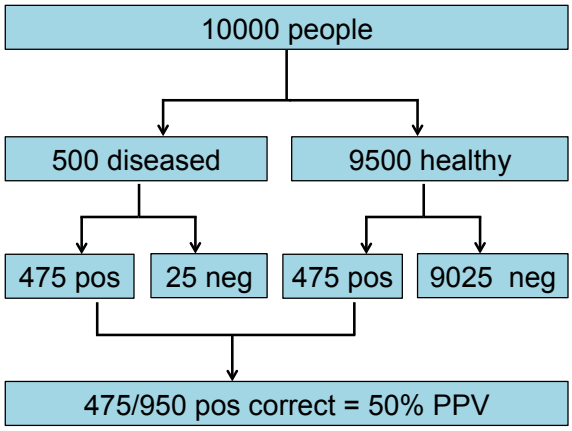



**Figure 2-6. Natural frequency diagram for predictive values**

A natural frequency diagram showing how prevalence affects the predictive value of a diagnostic test. In this case, there is 5% prevalence (500 diseased out of 10000 patients) and a test exists with 95% diagnostic sensitivity and specificity. Even with relatively high diagnostic sensitivity and specificity, the test only gives 50% PPV, which means that despite a positive test, the patient still only has a 50% probability of having the disease being tested for. The NPV in this case is 99.7%, which is much higher due to the low probability of the patient having the disease prior to testing.

Figure 7 shows how diagnostic sensitivity, diagnostic specificity, and prevalence affect the probability of a patient being disease positive (D+) for a positive or negative test result. Specifically, the PPV and 1–NPV are plotted as a function of prevalence for tests with different diagnostic sensitivities and specificities. For a given prevalence on the x-axis, the corresponding y-value on the PPV curve shows the probability that the patient has the disease for a positive test result, while the y-value on the 1–NPV curve shows the probability that the patient has the disease for a negative test result. There are a few key points that can be highlighted in Figure 7:

- Figure A shows a test with relatively high diagnostic sensitivity and specificity (96% and 98%). The test provides significant changes in post-test probabilities and gives relatively high confidence even for low and high prevalence diseases. For instance, at 5% prevalence, a positive test gives a PPV of approximately 80%. At 90% prevalence, a negative test will change the probability of being disease positive to approximately 20%. These changes represent large changes in probability from the initial prevalence, though it should be noted that even a relatively accurate test does not provide 100% certainty for positive results at low prevalence and negative results at high prevalence.

- Figure B shows that a poor diagnostic sensitivity and specificity test does very little to change the probability of disease regardless of the prevalence. For comparison, a test with no diagnostic benefit would a straight diagonal line from the bottom left corner to top right corner equivalent to $y=x$.
- Figure C shows that a high diagnostic specificity test with poor diagnostic sensitivity can still have value for ruling-in patients when the test result is positive, but that a negative test result gives very little new information.
- Figure D shows that a high diagnostic sensitivity, but low diagnostic specificity test can have value for ruling out patients with a negative test result, but offers very little new information for a positive test result.

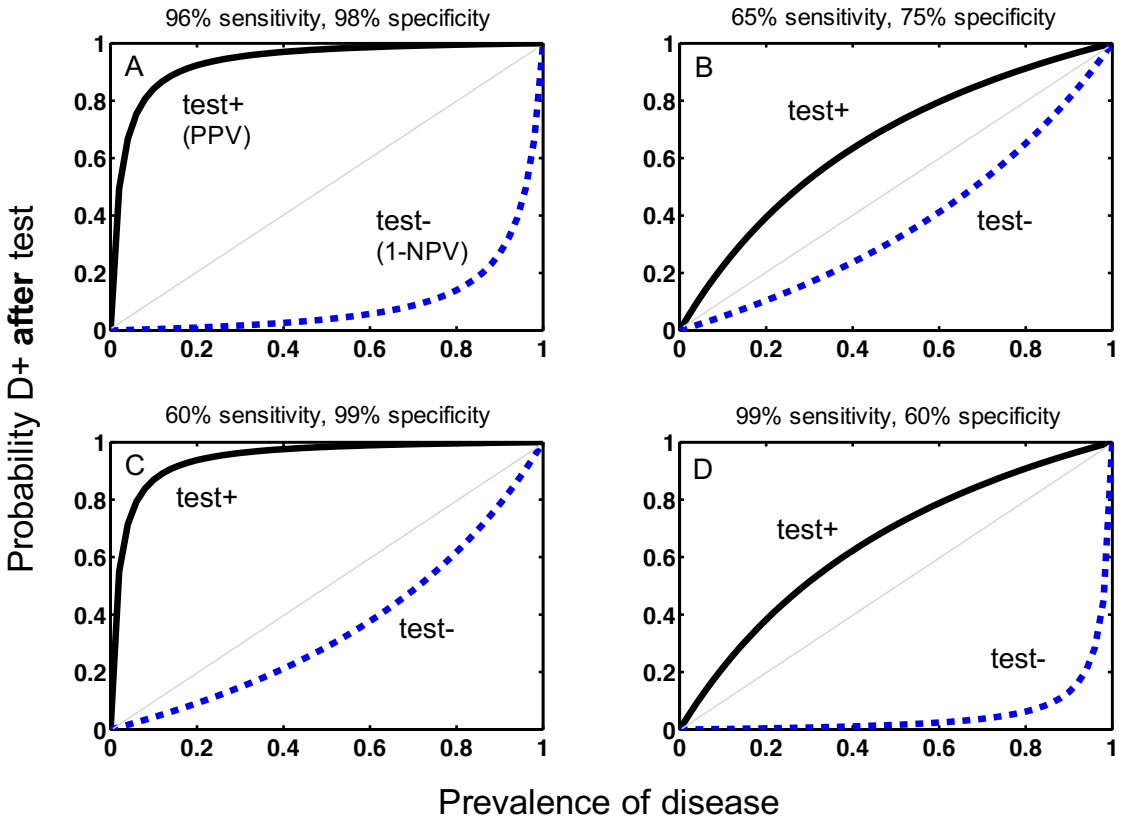



**Figure 2-7. Affects of prevalence and diagnostic sensitivity / specificity**

Series of plots showing how diagnostic sensitivity, diagnostic specificity, and prevalence affect the probability of disease status. For each plot, the probability of being disease positive (D+) after testing is plotted as a function of the probability of being D+ before testing, i.e. prevalence. The solid black line shows how the probability of being D+ changes when the test is positive (PPV) and the dashed blue line shows how the probability of being D+ changes when the test is negative (1–NPV). The y=x grey line shows no

change in pre- to post-test probability. Each plot A–D shows a test with a different diagnostic sensitivity and specificity. Tests with high diagnostic specificity have greater affects when the test is positive and tests with high diagnostic sensitivity have greater affects when the test is negative. However, the plots show that the probability of being D+ is also highly dependent on disease prevalence.

Plots C and D show the general idea that high diagnostic specificity allows for ruling-in a disease and that high diagnostic sensitivity allows for ruling out a disease. If a test is very good at detecting the target of interest with high analytical selectivity (resulting in high diagnostic specificity), then a positive test means that the correct target was indeed detected. Alternatively, if a test is very good at detecting any presence of target (quality LoD giving high diagnostic sensitivity), then a negative test likely means the target is truly not present or else the test would have detected it. Plots C and D also show that ruling-in a disease is more difficult at low prevalence, and ruling out a disease is more difficult at high prevalence. An example of a rule-out test would be a screening test that has high diagnostic sensitivity and tries to ensure that a negative test result truly means that no disease is present. If the screening test happened to give a positive result, then a subsequent rule-in test could be used to confirm whether or not the patient actually has the disease of interest.

In the SI, we have included a spreadsheet where the user can input various sensitivities and specificities, and plots similar to Figure 7 will be generated. We encourage the reader to enter different sensitivities and specificities and examine how disease probabilities change with not only diagnostic sensitivity and specificity, but also prevalence. Prevalence and predictive values are important to scientists and engineers because they dictate how useful a test will be to physicians and give guidelines for developing tests. For example, if a physician requires a rule-in test for a certain disease, the engineer and scientist would want to focus on maximizing the diagnostic specificity of the test. However, the required diagnostic sensitivity and specificity will ultimately be dependent upon clinical factors including prevalence, false positive and false negative consequences, treatment thresholds, and other factors.[159] Ideally the test result provides enough certainty that the clinician is able to decide treatment for the patient based off the test result.[1,3,160]

### 2.4.3 Positive and negative likelihood ratios

The final paired measures commonly used to determine diagnostic accuracy are likelihood ratios. Likelihood ratios use Bayesian reasoning and mathematics to calculate the probability that a patient has a disease

before and after testing.[113,154] Essentially, the pre-test probability is multiplied by a constant (the likelihood ratio) in order to determine the post-test probability, which can generally (but not in a strict mathematical sense) be thought of as **pre-test probability x likelihood ratio = post-test probability**. Likelihood ratios above 1 indicate an increase in the probability of disease presence, known as positive likelihood ratios (LR+), and likelihood ratios below 1 indicate a decrease in the probability of disease presence, known as negative likelihood ratios (LR–). LR+ and LR– can be defined as:[1]

$$LR+ = \frac{\frac{TP}{TP+FN}}{\frac{FP}{FP+TN}} = \frac{sensitivity}{1-specificity} \quad (9)$$

$$LR- = \frac{\frac{FN}{TP+FN}}{\frac{TN}{TN+FP}} = \frac{1-sensitivity}{specificity}, \quad (10)$$

where larger LR+ values (approaching infinity) indicate a more useful test for identifying presence of disease and smaller LR– values (approaching 0) indicate a more useful test for identifying absence of disease. General rules of thumb for likelihood ratios are that >10 (LR+) and <0.1 (LR–) indicate large and conclusive changes between pre- and post-test probabilities, while likelihood ratios of 5–10 and 0.1–0.2 indicate moderate changes, and likelihood ratios near 1 offer very little or no clinical value.[1]

When using likelihood ratios for calculations, it's important to note that probabilities cannot exceed one or 100%, *e.g.* if the pre-test probability is 10% and a positive result makes it 20 times more likely the patient has the disease, the post-test probability cannot be 200%. As a result, odds are used to facilitate the calculation between pre- and post-test probabilities to ensure the pre- and post-test probabilities range from 0 to 1, as shown in Figure 8. Understanding the meaning of odds is not required to use likelihood ratios because they are simply an intermediate math step, but the next two paragraphs will explain the meaning of odds and the details of the calculation.

Odds are a probability, *P*, expressed as a ratio, *P*:(1–*P*), which is the same as *P*/(1–*P*).[3] One way to visualize the difference between probability and odds is an example of splitting a pie into fractions. If a pie is split into equal thirds and person A takes 1/3 and person B takes 2/3, then person A owns 33% of the

total pie, which is equivalent to probability. In terms of odds, person A would own 1 part, while person B would own 2 parts, so the ratio or odds of person A to B is 1:2 or 1/2=0.5. Thus, probabilities express a fractional sample in terms of the total population, while odds express the fractional sample A versus fractional sample B.

To use likelihood ratios, the pre-test probability is converted to odds, the odds are multiplied by the likelihood ratio, and then the new odds are converted into a post-test probability to ensure calculated probabilities range from 0 to 1.[113] Figure 8 shows an example calculation for a test with pre-test probability of 0.20 and a LR+ of 15. In this case, the pre-test probability of 0.20 is equivalent to 1 out of 5 total patients having the disease of interest and after converting to pre-test odds, is equal to 1:4 healthy versus non-healthy patients for odds of 0.25. The odds are then multiplied by the LR+, which would be equal to 15:4 or 3.75, which is equivalent to 15 non-healthy patients to every 4 healthy patients. Converting these odds back into probabilities means 15 non-healthy patients out of 19 total patients and gives a post-test probability of 0.79. To automate the calculations required for using likelihood ratios, Fagan's nomogram has traditionally been used, with websites and mobile applications also available.[161–164] Our spreadsheet included in the SI also facilitates these calculations.

**Calculating post-test probability for a pre-test probability=0.20 and LR+=15**

Pre-test probability = $p_1$
Pre-test odds = $o_1 = p_1 / (1 - p_1)$
Post-test odds = $o_2 = o_1 \times LR$
Post-test probability = $p2 = o_2 / (1 + o_2)$

Pre-test probability = 0.2
Pre-test odds = 0.2 / (1 – 0.2) = 0.25
Post-test odds = 0.25 x 15 = 3.75
Post-test probability = 3.75 / (1 + 3.75) = 0.79

**Figure 2-8. Converting probabilities to odds**

Calculations for converting between probabilities and odds, which is necessary for using likelihood ratios to calculate disease probabilities.

Figure 9A shows how likelihood ratios change the pre- to post-test probability for a test with relatively high and low likelihood ratios. The larger LR+ (23 versus 2.8) and smaller LR– (0.07 versus 0.4) have more of an effect on the change between pre- and post-test probabilities. Only one pre-test probability is shown,

but it is important to note that the post-test probabilities from these likelihood ratios are dependent on the pre-test probability as well.

Figure 9B shows how the post-test probability for one test can become the pre-test probability for a second test, which together serve to create the final post-test probability of disease.[1] In this example, Test 1 has a diagnostic sensitivity of 95% and diagnostic specificity of 75% (LR+ = 3.8, LR– = 0.067), and the pre-test probability of disease is 5%. This situation may be similar to an inexpensive screening test where the diagnostic sensitivity is high to rule-out patients (LR– near zero), but at the expense of diagnostic specificity, which leads to a relatively low LR+. If the initial screening happens to give a positive diagnosis, the probability of disease is still only 17% because the screening test has low diagnostic specificity and as a result, low LR+. In this case, a second confirmatory test, often more expensive and time consuming, would likely need to be run. The second test has a diagnostic sensitivity of 93% and a diagnostic specificity of 96% (LR+ = 20), which means the high diagnostic specificity of a positive result would lead to a more confident diagnosis. A positive result using Test 2 leads to 80% post-test probability—up from 17% pre-test—of the patient having the disease. In order to have a single test give the same pre- to post-test probability change as both tests together, a test with LR+ of 76 would be required (*e.g.* 98.7% diagnostic sensitivity and 98.7% diagnostic specificity).

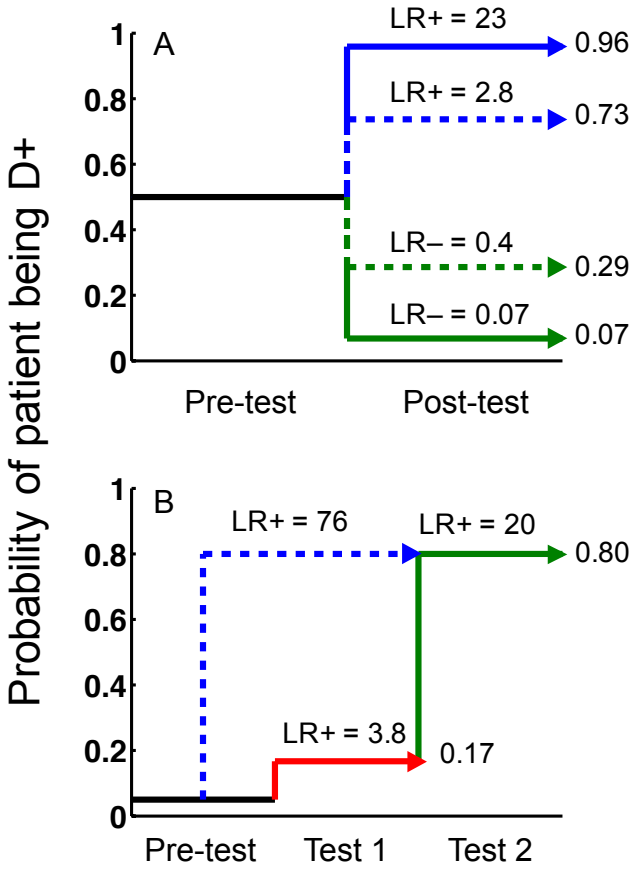



**Figure 2-9. Likelihood ratios**

Series of plots showing how the probability of being disease positive (D+) changes from pre-test to post-test based on diagnostic devices with different likelihood ratios. Figure (A) shows how tests with high and low likelihood ratios affect the post-test probability for both positive and negative tests. Figure (B) shows how the post-test probability from one test can be used as the pre-test probability for a second test, along with the required likelihood ratio for a single test to be equal to both tests together.

Likelihood ratios are useful for scientists and engineers who would like to determine how their devices may impact clinical care without knowing the exact prevalence or pre-test probability of their patient population. For example, the prevalence may not be known for a situation such as the 2014–2015 Ebola outbreak where new tests were being rapidly developed to respond to the crisis, but there was high variability in the number of cases being reported.[165,166] A benefit of calculating LR's is that if a researcher is creating a rule-in test and determines it has a high LR+, such as a LR+=50, they would know that their test has the potential to create the clinical impact they are striving for. However, while likelihood ratios give a general idea of how impactful the test can be, test developers will still want to determine the likely clinical use and estimated pre-test probabilities to ensure their test meets clinical needs. For instance, if the treatment threshold for their particular disease indication is 0.90 (treatment is only given when the post-test probability

is above 0.90) and the mean pre-test probability is 0.05, then the developed test with LR+=50 would only give a post-test probability of 0.72, which is still below the treatment threshold. In this case, the test wouldn't rule-in the disease enough to provide actionable data for the clinician.

Likelihood ratios may appear to be the same as predictive values, but a few key differences should be noted. The first difference is more theoretical in that LR's use general pre-test probability versus prevalence for predictive values according to their respective definitions. The second, more practical difference is that for tests that give continuous readouts, likelihood ratios can give more than a single yes or no answer to diagnose diseases, such as multiple levels of disease or intermediate results.[154,167] So far we have considered disease as either being present or not being present. In many cases, this assumption is a drastic over-simplification of actual medical conditions, which are more likely to be a spectrum of disease with varying severity.[3] Many tests give continuous or ordinal results and collapsing this data into a 2x2 table with only positive and negative results causes a loss of information.[154] For example, there isn't a single blood pressure cutoff that determines a healthy individual from a non-healthy individual, but rather a series of ranges that indicate hypotension, healthy, pre-hypertension, and different stages of hypertension.[154]

Likelihood ratios can be generated for multiple levels of measurements of continuous or ordinal readouts, as shown in Table 2 for a continuous measurement. In this example, a test gives continuous readings from 0 ng/mL to 200 ng/mL. Instead of establishing a single cutoff at 50 ng/mL to simply indicate yes or no, ranges of measurements can give different likelihood ratios so that the physician has more information in making the diagnosis. In this case, sensitivities and specificities would need to be generated for each measurement level in order to create the different likelihood ratios for each data range. Moderate levels of target in Table 2 (30–50 and 50–70 ng/mL) give only small changes in disease probability, but the low and high ranges (< 30 ng/mL and > 100 ng/mL) give more significant changes in pre- to post-test probability. Clinical studies have shown that seemingly non-useful diagnostic tests using a single clinical threshold may become useful when broken down into multiple levels.[154,168]

**Table 2-2. Multiple levels of likelihood ratios**

An example of creating multiple levels of likelihood ratios for a continuous test reading. Instead of using a single cutoff value that gives a single yes/no diagnosis, the continuous reading is broken down into ranges, with each range having an associated LR. Readings between 30–70 ng/mL give small changes in

probability because the LR is near 1, but low measurements (<30 ng/mL) and high measurements (>100 ng/mL) give moderate changes in probability due to LRs further from 1.

| Test measurement | Likelihood ratio |
|---|---|
| < 30 ng/mL | 0.17 |
| 30–50 ng/mL | 0.8 |
| 50–70 ng/mL | 3.2 |
| >100 ng/mL | 8.6 |

## 2.5 Diagnostic case studies using presented statistics

In this section, we will provide several examples of the importance of some of the above concepts for diagnostic tests relevant to the lab-on-a-chip community. The first example highlights the importance of prevalence when comparing the accuracy of a commercially available, point-of-care lateral flow test using blood versus oral samples to detect HIV antibodies in adults (Oraquick advance rapid HIV-1/2, OraSure Technologies Inc., PA, USA).[15] The second example presents accuracy data for rapid lateral flow tests and a CLIA waived molecular diagnostic test (Alere i, Waltham, MA, USA) used to diagnose group A *Streptococcal* (strep throat), and discusses test accuracy with regards to recommended treatment guidelines.[14,169] The final example will present accuracy data from rapid Chlamydia lateral flow tests and discusses the implications of poor LoD and diagnostic sensitivity.[17,170]

### 2.5.1 HIV lateral flow

Pant Pai *et al.* performed a systematic review and meta-analysis that compares the worldwide accuracy of the Oraquick advance rapid HIV-1/2 lateral flow test when performing the test using finger-stick blood samples versus oral swab samples for adult testing.[15] In addition to providing a direct comparison of the oral versus blood accuracy using diagnostic sensitivity, diagnostic specificity, and likelihood ratios, the authors also provide positive predictive values for higher-risk populations (>1% prevalence), e.g. intravenous drug users, sex workers, those who attended clinics for sexually transmitted diseases; and lower-risk (<=1% prevalence) populations, e.g. general outpatients.

Table 3 shows the diagnostic sensitivity, diagnostic specificity, likelihood ratios, and PPV for low prevalence

(<=1%) and high prevalence (>1%) populations.[15] Both sample types exhibit very high diagnostic sensitivity and specificity, but the blood-based sample has marginally higher diagnostic sensitivity (1.65%) and diagnostic specificity (0.17%) than the oral swab samples. While the diagnostic specificity difference seems tiny (0.17%), it has a noticeable affect on the PPV at low prevalence.[171] The PPV for both sample types is similar (98.50% and 97.65%) for populations with prevalence greater than 1%, but the oral swab test has a PPV of only 88.55% in low prevalence settings (<=1%) versus a PPV of 97.65% for the blood based test.[15]

**Table 2-3. HIV-1 lateral flow meta analysis results for Pant Pai et al.**

Adapted meta-analysis results from Pant Pai et al. of the diagnostic accuracy of the rapid HIV-antibody based point-of-care test (Oraquick advance rapid HIV-1/2, OraSure Technologies Inc., PA, USA) using blood or oral based samples.[15] The diagnostic sensitivity and specificity, positive and negative likelihood ratios, as well as the PPV of each test in low prevalence (< 1%) and high prevalence (> 1%) populations are shown. The blood test is more sensitive and specific than the oral test, resulting in higher LR+ and smaller LR–. The small difference in diagnostic specificity (99.91 versus 99.74 for the blood versus oral test) results in a significant PPV difference between the two sample types at low prevalence. While the PPV for the oral based sample is higher at high prevalence, the 95% confidence interval is larger—see Pant Pai et al. for details.

| | Sensitivity | Specificity | LR+ | LR– | PPV (>1%) | PPV (<=1%) |
|---|---|---|---|---|---|---|
| **Blood** | 99.68% | 99.91% | 1105.2 | 0.003 | 98.50% | 97.65% |
| **Oral** | 98.03% | 99.74% | 383.4 | 0.019 | 98.65% | 88.55% |

It is tempting to attribute the difference in PPV to the larger difference in diagnostic sensitivity, but as we noted earlier, a rule-in test is highly dependent on diagnostic specificity, especially at low prevalence.[171] In this case, we can see that a seemingly insignificant difference in diagnostic specificity (0.17%) results in a nearly 9% difference in PPV at low prevalence.[171] A PPV of 88.55% is still relatively high, especially for prevalence less than 1%, but this PPV means that approximately 1 out of every 10 positive tests will not be correct and possibly lead to considerable patient distress. This test is available over-the-counter in the United States using oral swab samples, meaning patients can use this test themselves at home.[172] The FDA recommends that anyone who tests positive contacts a doctor for confirmatory testing, which is likely needed given the PPV in low prevalence populations and the clinical implications of being HIV positive.[172]

This example demonstrates the importance of prevalence in determining disease probabilities, even for tests with seemingly near perfect accuracy. It also demonstrates the tradeoffs that can occur when

balancing accuracy versus other factors such as convenience or enabling more comprehensive testing. The oral swabs are less accurate, but are also more convenient for patients to gather on their own. It is possible that the diagnostic developers and/or public health officials decided that the reduction in PPV for oral swabs was worth the more convenient samples and/or increased uptake in testing. The test accuracy is obviously very important, but many other factors must also be weighed when determining its value to public health (*e.g.* cost, availability, patient convenience).

### 2.5.2 Group A strep lateral flow and molecular test

Over 13 million patients visit physicians' offices each year with pharyngitis, *i.e.* sore throat, with many of these cases caused by group A Streptococcal (GAS).[14] GAS can lead to acute rheumatic fever in rare cases, so antibiotics are often prescribed for patients with likely GAS pharyngitis.[14] The traditional gold standard for GAS testing is bacterial culture, but it requires 48 hours to provide diagnosis, so rapid GAS lateral flow immunoassays are often used to give faster results for immediate treatment decisions. Clinical guidelines recommend treating both adults and children with positive rapid tests.[18] For negative rapid tests, guidelines such as the Center for Disease Control (CDC) and Infectious Disease Society of America (IDSA) recommend only doing a culture backup for children, but not for adults.[18,14,173–175] This recommendation is due to adults having lower Strep A prevalence (less likely to have Strep A before testing) and less chance of developing complications compared to children.[18,14] However, no rapid test has been FDA approved for use without a backup confirmation for negative results and in some studies, the risk of complication for GAS has been shown to be as high in adults as in children.[176] In either case, the decision to use a backup affects the cost, time, and quality of care. These guidelines pertain to use in the United States and may varpy for other countries.

More recently, nucleic acid tests have been developed, in part, to offer improved diagnostic sensitivity compared to the rapid antigen detection tests because improved diagnostic sensitivity may lead to more confidence in negative results and possibly eliminate the need for culture backups. We will examine the diagnostic accuracy of rapid antigen tests, and a recently developed CLIA waived nucleic acid amplification test (Alere i) for adults and children, and discuss implications with regards to recommended guidelines.[169]

Table 4 shows data from a meta analysis from Stewart *et al.*[14] of trials of rapid strep A testing in children

(*n*=10,325) and adults (*n*=1216) as well as data from a single clinical study from Cohen *et al.*[169] (*n*=481) for the Alere i. The populations in both the rapid strep meta-analysis and nucleic acid clinical study included children and adult patients from emergency rooms and outpatient settings presenting with sore throat and signs of suspected pharyngitis.[14,169] In both cases, children exhibited higher prevalence of disease than adults. For the rapid test, the NPV for children was 94.2% and 97.5% for adults. The lower NPV for children is a combination of lower diagnostic sensitivity, as well as higher prevalence, both of which contribute to a lower probability of not having the disease for a negative test result. In other words, children have a 5.8% probability of having Strep A when the rapid test is negative, which when combined with the risk of complication, likely justifies the culture backup for children. In contrast, adults who test negative with the rapid strep test have a 2.5% probability of having Strep A. The combination of this lower probability with the lower risk of complication in adults is used by some to justify not using backups in adults.

Interestingly, the PPV of the rapid test in both adults and children is lower than the NPV, meaning that a positive result is less conclusive than a negative result, yet backups are not required for any positive test. In particular, adults with positive rapid tests have a 22.4%[††] probability of not actually having Strep A, meaning there are likely a number of unnecessary antibiotic prescriptions for positive rapid test results. This example demonstrates the importance of clinical context in interpreting the consequences of false positive or false negative tests. In this case, professional organizations consider missing a child with GAS to be worse than unnecessarily treating a person who does not have GAS. For other tests and other situations, the opposite might apply.

The nucleic acid test has higher diagnostic sensitivity and specificity than rapid tests, resulting in more drastic changes for pre- to post-test probabilities (likelihood ratios further from 1) and higher predictive values. The NPV in both children and adults are above 99%, meaning a backup culture is likely not justified in either population. Further, the PPV is higher, which may help reduce unnecessary antibiotic prescription

[††] 1–PPV = 1 – 0.776 (Table 4) = 0.224

and give more confidence in the diagnosis and treatment. It should be noted when comparing the predictive values between the Alere and rapid test studies that the prevalence's for the Alere study are different than the pooled rapid test prevalence's for the rapid test. However, applying the rapid test prevalence to the Alere test only changes the PPV by between 1–1.5%, and essentially does not change the NPV.

While the nucleic acid test reduces both false positive and false negative results, and hence seems an obvious choice to implement, in reality decisions at the clinic or policy level about which type of test to adopt involve many additional factors beyond accuracy, and even beyond the clinical consequences of false positive/negative results. These include those related to not only the test itself, but to multiple additional factors such as speed, ease-of-use, testing cost, etc. Discussion of these important factors are outside the scope of this review, but can be found elsewhere.[159,177]

**Table 2-4. Strep A results for rapid test and nucleic acid amplification**

Summary of data from a meta-analysis of rapid strep A tests[14] (n=11,541) and a clinical trial of the Alere i nucleic acid test[169] (n=481) for patients presenting with sore throat. The data is given for both children and adults because some guidelines recommend culture backup for negative rapid tests in children, but do not require them for adults.[18] The rapid test NPV of children is indeed lower than adults due to lower diagnostic sensitivity and higher prevalence. The nucleic acid test gives higher diagnostic sensitivity and specificity than the rapid test, resulting in more conclusive likelihood ratios and higher predictive values. The data presented for the Alere i is compared to culture with PCR adjudication of discrepant results.[169]

| | Patients | Sensitivity | Specificity | Prevalence | LR+ | LR– | PPV | NPV |
|---|---|---|---|---|---|---|---|---|
| **Rapid IA** | <18 years | 86% | 96% | 29.7% | 21.5 | 0.146 | 90.1% | 94.2% |
| **Rapid IA** | Adult | 91% | 93% | 21.3% | 13.0 | 0.097 | 77.6% | 97.5% |
| **Alere i** | <18 years | 98.5% | 98.2% | 36% | 65.8 | 0.013 | 97.1% | 99.1% |
| **Alere i** | Adult | 100% | 99.1% | 14% | 111.1 | 0 | 95.2% | 100% |

### 2.5.3 Chlamydia rapid tests

*Chlamydia Trachomatis* (CT) is the most prevalent sexually transmitted infection worldwide. It is curable by antibiotics, but often does not present symptoms and can lead to adverse outcomes such as infertility and pelvic inflammatory disease in women. As a result, the CDC recommends regular screening for all sexually active women <25 years of age, as well as screening of older women who may have increased risk.[178] The standard diagnostics for CT are nucleic acid amplification tests, but the turn-around time for laboratory testing can result in significant patient loss-to-follow-up in high-risk areas. To provide immediate treatment,

rapid Chlamydia tests have been developed, but the poor sensitivities of these tests present issues for their use as a screening tool.

Table 5 shows diagnostic sensitivity, diagnostic specificity, prevalence, and predictive values of four different Chlamydia rapid tests from two clinical studies by third parties.[17,170] The first study by van Dommelen *et al.* shows that the highest diagnostic sensitivity of the three studied tests is 25%, while specificities ranged from 88.9–99.7%.[17] The very low diagnostic sensitivity limits use of these tests as screening tools because diagnostic sensitivity is the most important factor in ruling-out patients. The NPVs from these tests may appear relatively high (89.0–91.5%), but they essentially do not change from the prevalence of 11%. In other words, the patient has an 11% probability of being disease positive before the test and after a negative test still has a probability of 8.5–10%. This minimal difference in pre- to post-test probability means that a negative test result essentially gives clinicians no new information for diagnosing the patient.

Despite poor diagnostic sensitivity and non-useful negative results, the high diagnostic specificity (99.7%) and large LR+ (83.3) of the QuickVue Chlamydia test may suggest that it could possibly be useful as a *rule-in* test. This idea can be evaluated by constructing a natural frequencies diagram for 10,000 theoretical patients with 11% prevalence for the QuickVue test with 25% diagnostic sensitivity and 99.7% diagnostic specificity, as shown in Supplementary Figure S3A. The high diagnostic specificity means that out of 8900 healthy individuals, only 27 would falsely test positive. However, out of the 1100 patients with the disease, only 275 would be correctly identified as positive using this test. It is true that the false positive rate is low, as 275/302 of positive tests would correctly identify patients that have the disease, meaning that a positive test would likely indicate the patient had the disease. However, when evaluating clinical usefulness, it must be considered whether running 10,000 tests to detect 275/1100 infected patients is worth the time and cost because 9,725 nucleic acid backup tests would still be required for negative QuickVue tests. Whether this test is clinically useful requires considering many additional factors beyond accuracy including the speed of the test, the follow-up rate for treatment after lab testing, the cost of initial and backup testing, the cost of treatment, and complications from infection. Gift et al. give examples of these “rapid test paradox” issues in detail and we will summarize their results at the end of this section.[16] The Biorapid CHLAMYDIA Ag test and Handilab-C both have LR+ and LR– near 1, meaning the test results do not result in useful

diagnostic changes for either positive or negative results.

**Table 2-5. Chlamydia rapid test meta analysis results**

Summary of rapid Chlamydia tests accuracy from van Dommelen[a] et al.[17] (n=763) and van der Helm[b] et al.[170] (n=912). The low diagnostic sensitivity from van Dommelen[a] et al. results in LR– near 1 and NPVs that do not significantly differ from the prevalence. The high diagnostic specificity of QuickVue may offer value as a rule-in test. The Chlamydia rapid test (CRT) test from van der Helm[b] et al. shows improved diagnostic sensitivity, but still does not offer significant LR+ values. The diagnostic sensitivity issue may be related to poor LoD for the tests.

| | Sensitivity | Specificity | Prevalence | LR+ | LR– | PPV | NPV |
|---|---|---|---|---|---|---|---|
| **Biorapid CHLAMYDIA Ag test[a]** | 17.1% | 93.7% | 11.0% | 2.7 | 0.89 | 24.6% | 90.4% |
| **QuickVue Chlamydia test[a]** | 25.0% | 99.7% | 11.0% | 83.3 | 0.75 | 91.3% | 91.5% |
| **Handilab-C[a]** | 22.5% | 88.9% | 11.0% | 2.0 | 0.89 | 19.8% | 90.4% |
| **Chlamydia Rapid test[b]** | 39.4% | 94.4% | 20.8% | 7.0 | 0.64 | 65.0% | 85.6% |
| **Chlamydia Rapid test[b]** | 42.0% | 96.8% | 9.2% | 13.1 | 0.60 | 56.9% | 94.3% |

The second study by van der Helm *et al.* studied the Chlamydia Rapid Test (CRT) in high-risk (20.8% prevalence) and low-risk (9.2% prevalence) populations.[170] The diagnostic sensitivity for this CRT test (39.4–42.0%) is higher than the tests from van Dommelen *et al.,* but still offers limited value for negative results (LR– of 0.60–0.64). The NPV of the high prevalence test is 85.6%, meaning that the probability of having the disease changes from 20.8% to 14.4% if the test is indeed negative. In the low prevalence areas, the probability of having the disease changes from 9.2% pre-test to 5.7% if the test is negative. These pre- to post-test changes likely do not give enough useful information for clinicians to make actionable decisions on whether to treat or not treat patients. For comparison, CDC recommended nucleic acid amplification tests would go from 9.2% pre-test probability to approximately 0.5% post-test probability for a negative result, which gives clinicians the confidence to not treat the patient.[178,179]

Further, while the CRT diagnostic sensitivity is higher than the QuickVue test, the CRT diagnostic specificities are lower than the QuickVue test, which reduces the confidence in the positive results (LR+=7.1, 13.1 versus LR+=83.3). The diagnostic specificity difference between the CRT and QuickVue test may not appear to be very significant (94.4%, 96.8% versus 99.7%), but LR+ values have a highly non-

linear dependence on specificity as specificities approach 100%. A natural frequencies diagram of the CRT (10,000 patients, Supplementary Figure S3B) shows that out of 671 positive tests, 285 would be falsely positive, a much higher false positive rate than the QuickVue test due to the lower diagnostic specificity. Supplementary information section 5 discusses this non-linear dependence in more detail with plots of diagnostic specificity versus LR+.

van der Helm *et al.* note that the bacterial load is 73 times higher in cases where the gold standard nucleic acid test and the CRT are both positive compared to when the nucleic acid test is positive, but the CRT is negative. This finding suggests that the low diagnostic sensitivity and poor NPV is likely due to the LoD of the CRT not being low enough to detect Chlamydia samples with low bacterial load.[170]

Despite the poor performance of these tests, relatively modest increases in diagnostic sensitivity may lead them to be useful in a clinical setting compared to highly accurate nucleic acid amplification tests. Gift et al. estimated that if the patient follow-up rate for treatment following testing was less than 65%, then a rapid point-of-care test for Chlamydia with diagnostic sensitivity of 63% and diagnostic specificity of 100%, combined with immediate treatment may lead to more cost-effective treatment of infected individuals than waiting for highly accurate nucleic acid amplification results.[16] This analysis takes into account the prevalence, follow-up rate, cost of testing and treatment, plus complications from cases that develop pelvic inflammatory disease.

The authors also note that a two-test algorithm of a rapid test with 63% sensitivity and 100% specificity followed by PCR testing for those identified as negative would likely treat the greatest number of cases and be the most cost-effective at prevalence above 9%.[16] While the recommendation for a two-test algorithm is given for an assumed set of numbers (prevalence, follow-up rate, etc.), it highlights the need to account for additional factors beyond diagnostic sensitivity and specificity. In this particular case, the cost of the large number of backups required (~95% of tests), may be offset by the earlier treatment for the cases that were detected by the rapid test. This situation is also analogous to using the rapid test as a rule-in test due to its high diagnostic specificity, but requiring backup for negative tests due to poor diagnostic sensitivity, similar to the Strep A rapid test with culture backup previously discussed.

# Chapter 3: NAIL—Nucleic Acid detection using Isotachophoresis and LAMP

## 3.1 Introduction

Nucleic acid amplification tests (NAAT) are the gold standard for many infectious disease diagnoses due to high sensitivity and specificity, rapid operation, and low limits of detection. Despite the advantages of nucleic acid amplification tests, they currently offer limited point-of-care (POC) utility due to the need for complex instruments and laborious sample preparation. We report the development of the Nucleic Acid Isotachophoresis LAMP (NAIL) diagnostic device. NAIL uses isotachophoresis (ITP) and loop-mediated isothermal amplification (LAMP) to extract and amplify nucleic acids from complex matrices in less than one hour inside of an integrated chip. ITP is an electrokinetic separation technique that uses an electric field and two buffers to extract and purify nucleic acids in a single step. LAMP amplifies nucleic acids at constant temperature and produces large amounts of DNA that can be easily detected. A mobile phone images the amplification results to eliminate the need for laser fluorescent detection. The device requires minimal user intervention because capillary valves and heated air chambers act as passive valves and pumps for automated fluid actuation.

In this chapter, we describe NAIL device design and operation, and demonstrate the extraction and detection of pathogenic *E. coli* O157:H7 cells from whole milk samples. We use the Clinical and Laboratory Standards Institute (CLSI) limit of detection (LoD) definitions that take into account the variance from both positive and negative samples to determine the diagnostic LoD. According to the CLSI definition, the NAIL device has a limit of detection (LoD) of 1,000 CFU/mL for *E. coli* O157:H7 cells artificially inoculated into whole milk, which is two orders of magnitude improvement to standard tube-LAMP reactions with diluted milk samples. We focus on *E. coli* cells in whole milk for its application towards food safety monitoring, as well as the availability of the target and sample matrix. The NAIL device potentially offers significant reductions in the complexity and cost of traditional nucleic acid diagnostics for POC applications.

## 3.2 Experimental methods: NAIL fabrication, chemistry, and operation[105]

### 3.2.1 Device fabrication

We molded our NAIL devices by solvent casting styrene-ethylene/butylene-styrene (SEBS, 42 wt% PS, A1536H, Kraton) polymer onto a silicon wafer patterned with SU-8 photoresist as previously described with

minor modifications (Figure 3-1).[180,181] We fabricated the SU-8 silicon wafers using photolithography (Figure 3-1A). An SU-8 2050 (Microchem, SU-8 2000 series) photoresist layer ~170 μm thick was spun (500 RPM for 5s, followed by 1250 RPM for 30s) onto a 4-in silicon wafer and prebaked for 20 min at 95°C. The resist was exposed to UV-light through a printed photomask (Fineline Imaging, Colorado Springs, CO) using an ABM contact aligner, baked again at 95 °C for 15 min, developed using SU-8 developer for 10 min (Microchem, Y020100, CAS: 108-65-6), and then hard-baked overnight at 150°C. An alpha step profilometer measured the height of the SU-8 microstructures, which correspond to the channel height. The master molds were silane coated by placing them in a desiccator under low vacuum for 30 min next to a Petri dish containing 50 μL of trichloro(1H,1H,2H,2H-perfluorooctyl)silane to create a nonstick surface for molding.

For SEBS fabrication, approximately 10 w/v% of SEBS was dissolved in toluene and mixed on a rotary mixer over night until the solution was transparent.[180] The solution was then sonicated for 2 h to further dissolve the polymer and placed under vacuum in a desiccator for two minutes to ensure there were no trapped air bubbles in the solution. If bubbles continuously formed during desiccation, the mixture was sonicated again for 1 h. We slowly poured the de-gassed solution onto the patterned wafer inside a PTFE coated ring (Norpro 666, Everett, WA) and baked for 3h at 65°C and 10h at 90°C on a hotplate inside a fume hood, as shown in Figure 3-1B. If bubble formation causes an issue, the solution may be poured through a funnel onto the wafer. A piece of aluminum foil with small holes was placed over of the ring to slow the rate of evaporation and reduce wrinkles from forming during annealing (Figure 3-1C). We heated the wafer to 40°C on a hotplate before de-molding to promote easier removal of the SEBS and protect the SU-8 structures (Figure 3-1D).

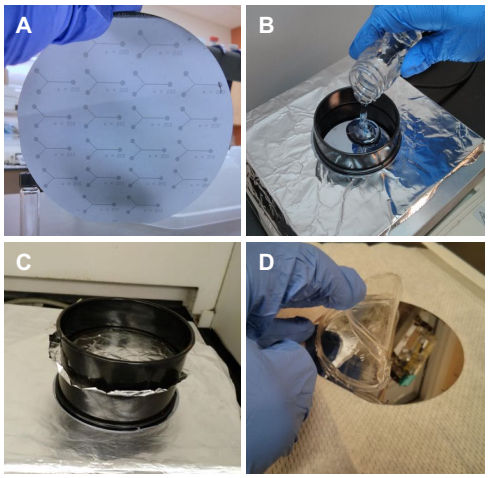



**Figure 3-1. SEBS fabrication on patterned silicon master molds**

(A) Image showing an SU-8 patterned wafer coated with silane. (B) The degassed mixture is slowly poured inside a Teflon ring onto the silicon master mold. (C) Aluminum foil with small holes punched into the top placed on top of the Teflon ring and weighted down with a second Teflon ring to slow evaporation and reduce wrinkles in the polymer film. (D) After heating, the SEBS devices are slowly peeled from the wafer on top of a hotplate set to 40ºC.

After fabrication of the molded substrate, ITP inlet ports and an air valve at the edge of the reaction chamber were punched using a hand punch (Fiskars 1/16" Circle Hand Punch) and 20-gauge needle respectively. Molded and flat substrates were air plasma treated at 200 mTorr and 30 W for 15 s (Harrick Plasma Cleaner, Ithaca, NY). After air plasma treatment, LAMP reagent mixture was pipetted onto the reaction chamber surface and the substrate was placed under vacuum with desiccant for 1-2 hours to dry the reagents on the surface before being manually pressed together with the flat substrate to seal. To ensure good room temperature bonding strength, the substrates were handled inside of a decontaminated biosafety cabinet with laminar flow as much as possible to keep the substrate surfaces clean. After bonding, we attached 5 mm high wells cut from polyvinyl chloride (PVC) tubing (3.175 mm ID, 6.35 mm OD) with UV curable epoxy and a UV transilluminator to hold LE and TE solutions during ITP. Aluminum foil strips were placed under the reaction chamber during UV exposure to protect the dried LAMP reagents.

### 3.2.2 ITP chemistry and operation

The leading electrolyte (LE) consisted of 600 mM Tris and 200 mM HCl buffer with 0.001% (w/v) Tween-20, 0.075% (w/v) polyvinylpyrridine (PVP), and 1x SYBR green (Life Technologies, Grand Island, NY). The trailing electrolyte (TE) was 90 mM Tris buffer and 37.5 mM HEPES buffer with 0.5 mg/mL proteinase K

(EO0491, Thermo Scientific, Pittsburgh, PA), 0.1% (w/v) Tween-20, 0.1% (w/v) PVP, 1x SYBR green, 5000 U Ready-Lyse™ Lysozyme (lysis buffer, Epibio, Madison, WI), 50% (v/v) whole milk (local grocery store), and varying amounts of E. coli O157:H7 genomic DNA (0801622DNA-10UG, Zeptometrix, Buffalo, NY) or E. coli O157:H7 cells (NATECO(933)-ERCM, Zeptometrix, Buffalo, NY). Chemicals were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise noted.

To initiate ITP, 45 µL of LE buffer was pipetted into the LE inlet and allowed to fill the device, followed by pipetting 30 µL of the TE buffer mixed with sample into the TE inlet. The increased LE inlet volume applies slight pressure against the migration of the ITP plug, resulting in increased separation and extraction of DNA.[74,78,182] We used platinum electrodes placed into the LE and TE inlets to apply 450 µA of constant current using a sourcemeter (Keithley 2410, Cleveland, OH). ITP was monitored using DNA-intercalating SYBR green dye and an inverted epifluorescence microscope (Nikon TE2000, Melville, NY), a fluorescent optical filter set (XF108-2, Omega, Brattleboro, VT), 2x or 4x objective, and cooled digital camera (Cascade IIb, Photometrics, Tucson, AZ).

After viewing the DNA plug enter the extraction chamber, the current was stopped and electrodes were removed before setting the chip onto a hotplate. The current or voltage can also be monitored to track the location of the plug without optics or labeling.[183] We quantified the ITP extraction efficiency by mixing known amounts of genomic DNA with the TE and imaging the fluorescence of DNA 3 mm away from the extraction chamber during ITP. This fluorescent intensity was compared to a calibration curve of fluorescent intensity versus DNA concentration to determine the extraction efficiency. Between each test, the platinum electrodes were soaked in DNA Erase (Sigma Aldrich) solution for 2 min, followed by ethanol for 5 min and DI water for 5 min to reduce contamination.

### 3.2.3 LAMP reactions

We conduct LAMP amplification experiments with three different methods. First, we conduct LAMP in the NAIL device with a range of target E. coli concentrations in whole milk diluted 2x. Second, we perform LAMP experiments in centrifuge tubes with a fixed concentration of target *E. coli* genomic DNA in a range of dilutions (10-10000x) of whole milk, as well as with no milk. Third, we complete LAMP experiments in

centrifuge tubes with a range of E. coli target concentrations ($10^2$-$10^6$ CFU/mL) in a fixed milk dilution of 1000x.

For the experiments conducted in the NAIL devices, we prepared 10 µL of LAMP reagents in a 1-mL microtube for drying on the reaction chamber surface. The six LAMP primers (F3, B3, FIP, BIP, LF, LB) are provided in supplementary table S1 and were developed by Wang et al. to target the *eae* gene of *E. coli* O157:H7.[98] The primers used in this study were shown to be specific for seven shiga toxin producing E. coli strains, including O157:H7, among 90 different bacteria tested to evaluate assay specificity.[98] A 10 µL stock mixture consisted of 8mM $MgSO_4$, 0.5 mM $MnCl_2$, 25 µM calcein disodium salt, 1.6 µM FIP and BIP, 0.8 µM LF and LB, 0.2 µM F3 and B3, (1.4 mM of each nucleotide, 2% (w/v) sucrose, and 0.32 U/µL Bst 2.0 Warmstart polymerase (M0538M, New England Biolabs, Ipswich, MA). The sucrose helps to stabilize the dried reagents and we observed quality amplification for up to two weeks after drying, the longest period tested in this study.[184] The LE and TE serve as the reaction buffer, so the 1x Thermopol buffer (New England Biolabs) is omitted in the drying mixture. 500 nL of this reaction mixture (the reaction chamber has a 500 nL volume) was pipetted and dried under vacuum on the reaction chamber surface before sealing the device. We placed the NAIL device onto a hotplate set at 65°C for 45 mins after DNA entered the extraction chamber from ITP. Silicone oil was pipetted over the air valve and into the inlets to prevent evaporation and equalize hydrostatic pressures.

For the tube-LAMP milk dilution tests, 10 µL reactions were prepared consisting of 1x Thermopol buffer (New England Biolabs), 6 mM $MgSO_4$, 0.5 mM $MnCl_2$, 25 µM calcein, 1.6 µM FIP and BIP, 0.8 µM LF and LB, 0.2 µM F3 and B3, 1.4 mM of each nucleotide, 0.32 U/µL Bst 2.0 Warmstart polymerase, 100 pg of *E. coli* O157:H7 genomic DNA, and 10-fold dilutions of whole milk. Positive and negative (*E. coli* DNA omitted) controls consisting of the same reaction mixture without the milk were run with each test. The reaction tubes were placed into a water bath heated to 65°C for 45 minutes. We determined the results by placing the tubes flat inside our mobile imaging unit and imaging the calcein fluorescence.

For the tube-LAMP reactions at 1000x milk dilution with varying cell concentrations, a 10 µL reaction consisted of the same reagents as the milk dilution tests except with 0.001 (v/v) whole milk, and 10-fold dilutions of *E. coli* O157:H7 cells (CFU/mL). Negative and positive controls with no milk and no *E. coli* (negative) or 100 pg of *E. coli* genomic DNA (positive) were tested with each experiment. The reaction

tubes were placed into a water bath heated to 65°C for 45 minutes. We determined the results by placing the tubes flat inside our mobile imaging unit and imaging the calcein fluorescence.

### 3.2.4 Mobile phone imaging

The mobile imaging unit was made with 0.6 cm thick (0.23 inch) black acrylic and consisted of a box with detachable lid and embedded optical filters, as well as a mobile phone aligner. The box had a length of 16.5 cm (6.5 inches), width of 9.5 cm (3.75 inches), and height of 6.9 cm (2.72 inches). Acrylic cement was used to bond the walls of the box together. We allowed the lid and bottom of the box to be detachable from the walls by imbedding magnets (1/8 x 1/8 inch cylinder neodymium) into the plastic. A mobile phone alignment piece was also held in place with magnets on top of the lid to ensure proper alignment of the mobile phone. Two holes were cut in the lid for the excitation filter (Omega XF1073, Stamford, CT) and the emission filter (Omega XF3084, Stamford. CT). The holes were cut specifically to align the excitation and emission filters with a Samsung Galaxy III (Samsung, South Korea) camera and flash. The filters were held in place with a plastic piece attached to the bottom side of the lid with magnets to create a lip for the filters to rest on. A 0.5 mm (0.02 inch) thick silicon rubber sheet, cut with holes specifically aligned to the camera and flash, was placed over the lid to seal out contaminating light.

We imaged the reactions by placing the chip or tube inside of the mobile imaging unit and taking a picture using a Samsung Galaxy III mobile phone set on top of the box. The free Android application “Camera FV-5 Lite” was used for imaging so that camera settings could be controlled and better transferred between Android phones. We used the close-up focus mode, auto white balance, 1600 ISO, and +1 exposure for imaging. The images were analyzed in ImageJ. Each image was split into RGB images, the blue image was subtracted from the green image for background correction, and the reaction chamber or reaction volume in-tube was manually selected to calculate the average intensity. Images were also taken using an epifluorescence microscope to ensure quality of mobile phone results.

### 3.2.5 Limit of detection

We determined the LoD of the NAIL device with 2x diluted milk samples and compared it to tube-LAMP with unprocessed, 1000x diluted milk samples. We used 1000x diluted milk samples because it is the least amount of milk dilution that gave quality results. For the NAIL LoD, we tested 10-fold dilutions of cell

samples ranging from 1-$10^5$ CFU/mL in the original milk samples. We also ran identical NAIL tests without cells for negative controls. Devices were disposed into 10% bleach after each experiment to reduce the risk for contamination. For tube-LAMP, we added lysed cells ranging from 0.1-$10^3$ CFU/mL directly to the LAMP reaction and then calculated the corresponding concentration in an original milk sample based off of the 1000x dilution.

The threshold and LoD for these qualitative tests was determined according to the Clinical and Laboratory Standards Institute's EP12 "User protocol for Evaluation of Qualitative Test Performance; Approved Guideline—Second Edition" and EP17 "Protocols for Determination of Limits of Detection and Limits of Quantification; Approved Guideline" documents.[109,147] The limit of blank (LoB) was used as the threshold and calculated by, $LoB = \mu_B + 1.645\sigma_B$, where $\mu_B$ is the mean of the negative control samples and $\sigma_B$ is the standard deviation of the negative control samples.[109] The LoD was defined as the upper endpoint of the 95% interval, $C_{95}$. The $C_{95}$ is the lowest analyte concentration where at least 95% of the tested samples are determined to be positive according to the established threshold. The LoD was determined to be the sample concentration that has a mean response that is 1.645 sample standard deviations above the LoB, $LoD = LoB + 1.645\sigma_S$, where $\sigma_S$ is pooled the sample standard deviation from low level samples ranging from the LoB to 4xLoB.[109] Using these definitions, we have a 5% probability of committing either type I (false positive) or type II (false negative) error at the LoD. The confidence that our LoD is at or above the $C_{95}$ concentration is based on the number of correct tests run at the LoD and calculated according to Table A2 in the EP12 document.[147]

### 3.3 NAIL device overview

The NAIL device consists of six main components, shown in Figure 3-2: LE and TE reservoirs, LE channel and ITP separation channel, extraction chamber, capillary valve, enclosed air chamber, and reaction chamber. The LE and TE inlets allow filling of the device with buffer and sample, and hold electrodes for applying electric fields. The ITP separation channel is where extraction and purification of DNA from impurities in the sample occurs, and the LE channel connects the LE inlet to the extraction chamber. Purified DNA arrives in the extraction chamber before being pumped into the reaction chamber where amplification

and detection occurs. The capillary valve between the extraction and reaction chambers prevents buffer from wetting the reaction chamber with dried LAMP reagents before purified DNA has reached the extraction chamber. The enclosed air chamber generates pressure when heated to 65°C that forces the fluid in the extraction chamber to break the capillary valve, pump purified DNA from the extraction chamber to the reaction chamber, and isolate the reaction chamber from the rest of the device. The solution containing target DNA dissolves the LAMP and indicator reagents and undergoes isothermal amplification at 65°C. The amplified DNA is imaged in the reaction chamber using a mobile phone camera using qualitative fluorescent calcein detection to determine positive and negative samples.[66]

NAIL testing requires one hour (15 minutes for lysis, 4 minutes for ITP extraction, 45 minutes for LAMP amplification, and 1 minute for imaging) and five basic steps to operate. First, the sample containing target cells is added to TE buffer with lysing agents in a microtube that effectively dilutes the sample by two, and allowed to incubate for 15 minutes. Second, the NAIL device is filled with LE and the TE-sample mixture is added to the TE inlet. Third, applying an electric field extracts the DNA from whole milk and transports it to the extraction chamber. Fourth, the chip is placed onto a hot plate at 65°C to activate the air channel pump and the LAMP reaction occurs for 45 minutes. Finally, the chip is placed into the mobile imaging unit for fluorescent detection using a mobile phone.

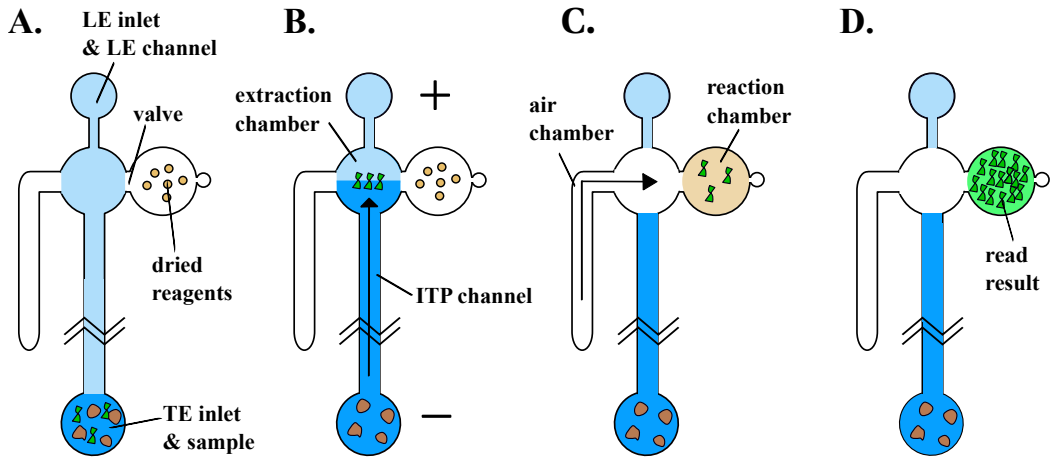



**Figure 3-2. Schematic of NAIL device components and operation**

Panel A shows the LE inlet and chip filled with LE buffer (light blue), and TE inlet filled with sample mixed with TE buffer (dark blue). (B) Applying an electric field isolates the DNA by ITP in the extraction chamber, purifying it from inhibitors in the sample. (C) Heating the chip to 65°C increases the pressure in the air chamber that breaks the capillary valve and drives flow of purified DNA solution in the extraction chamber to the reaction chamber. The purified DNA solution rehydrates the dry stored LAMP reagents and the air expansion isolates the reaction chamber from the rest of the chip. (D) LAMP specifically amplifies the target

DNA sequence to produce a fluorescent signal that can be imaged with a mobile phone to detect the presence (or absence) of target nucleic acid.

## 3.4 NAIL device design and operation

The geometry and surface properties of the microfluidic channels control the majority of device operation. Native styrene-ethylene/butylene-styrene (SEBS), our substrate material, is a moderately hydrophobic polymer, but offers a relatively stable hydrophilic surface following plasma treatment.[180] We plasma treat the flat and molded substrates prior to assembling the device to create hydrophilic surfaces for capillary flow. Following plasma treatment, adding LE buffer to the LE inlet causes the entire channel to fill almost instantaneously as shown in Figure 3-3 A. Liquid does not flow into the air channel because it is sealed, resulting in a dead-end channel.

### 3.4.1 Capillary valve

The capillary valve is designed to prevent liquid flow into the reaction chamber during filling and ITP DNA extraction, and then burst when the air pressure is increased by heating the chip to 65°C. The capillary valve, shown in Figure 3-3 B, works by forming a liquid-gas interface at a rapidly diverging channel wall that effectively increases the apparent contact angle of the flowing solution and creates an energetic barrier to flow into the reaction chamber.[185] An external pressure greater than the pressure drop across the bulging interface must be applied to force liquid to wet the diverging wall and allow fluid to continue advancing into the reaction chamber.

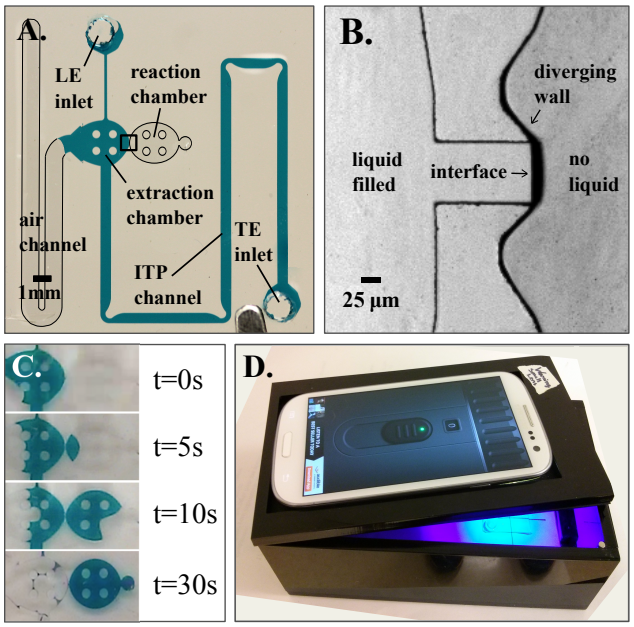



**Figure 3-3. Images of the NAIL device, capillary valve, heating, and mobile imaging unit**

Image A shows a fabricated device filled with colored LE via capillary flow. The black square is the location of the capillary valve between the extraction and reaction chambers. Each device was single-use and disposed after one experiment (B) A 4x micrograph of the capillary valve that prevents LE from wetting the reaction chamber. The liquid has fully filled the extraction chamber on the left and the valve up to the diverging wall, with no liquid present in the reaction chamber on the right. (C) Sequential images of fluid flow from the extraction chamber to the reaction chamber after the NAIL chip is placed onto a hotplate at 65°C. The liquid in the reaction chamber is completely isolated after 30 seconds of heating. (D) The mobile imaging unit consisting of a dark box with components that snap into place with magnets, an excitation and emission filter, and mobile phone.

The capillary valve pressure given by the modified Young-Laplace equation from Cho et al.

$$P_A - P_0 = -2\sigma\left(\frac{\cos(\theta_A + \beta)}{b} + \frac{\cos\theta_A}{h}\right) \tag{10}$$

is dependent upon the valve width, $b$, diverging wall angle, $\beta$, height of the channel, $h$, surface tension of the liquid, $\sigma$, and the contact angle of the liquid on the substrate surface, $\theta_A$.[185] The capillary valve uses a diverging wall characterized by an angle $\beta$ to effectively increase the contact angle of the flowing solution and create an energetic barrier to further flow. The maximum contact angle that a liquid meniscus can attain is 180°, so for $\theta_A+\beta > 180°$, the valve will burst when the liquid advances to a contact angle of 180° on the original wall. Figure 3-4 shows a graphic of a capillary valve with a diverging angle of $\beta$ =90°. The diverging wall must be a sharp angle in order for the capillary valve to properly function. Otherwise, the liquid will continue to flow along the smooth wall and not experience the increase in effective contact angle.

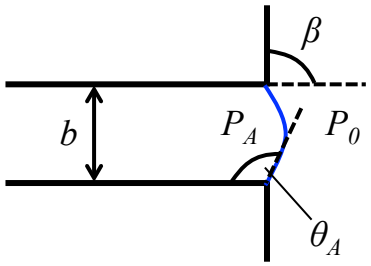



**Figure 3-4. Schematic of the capillary burst valve at a diverging wall**

The pressure inside the liquid, $P_A$, and outside the liquid $P_0$, are shown. The valve width, b, the advancing contact angle of the liquid on the substrate, $\theta_A$, and the diverging wall angle, β, are also shown.

We include tween-20 in our buffers to solubilize proteins and promote faster capillary flow, but the tween-20 affects both the surface tension and the contact angle of the liquid on the substrate, reducing the pressure required to burst the valve (*i.e.* tween-20 weakens the valve). Thus, we calculated various valve bursting pressures with different concentrations of tween-20 using equation 10. Niño and Patino give the surface tension of tween-20 in water at various concentrations,[186] while the contact angle of the LE solution on air plasma treated SEBS, $\theta_A$, was measured using contact angle goniometry. Figure 3-5 shows the effect of tween-20 concentration on the valve pressure at constant valve width (75 μm), height (170 μm), and diverging angle (120°). The valve pressure decreases with increasing tween-20 concentration due to decreased surface tension and more hydrophilic contact angle. Figure 3-5 also shows the effect of valve width and diverging angle on the valve pressure at constant tween-20 concentration (0.001%) and height (170 μm). The valve pressure decreases with increasing width and decreasing diverging angle. We observed that a valve pressure of approximately 800 Pa was necessary to prevent bursting prior to the end of ITP experiments. We used a valve width of 75 μm, height of 170 μm, diverging angle of 120°, and tween-20 concentration of 0.001% to give a valve pressure of approximately 900 Pa.

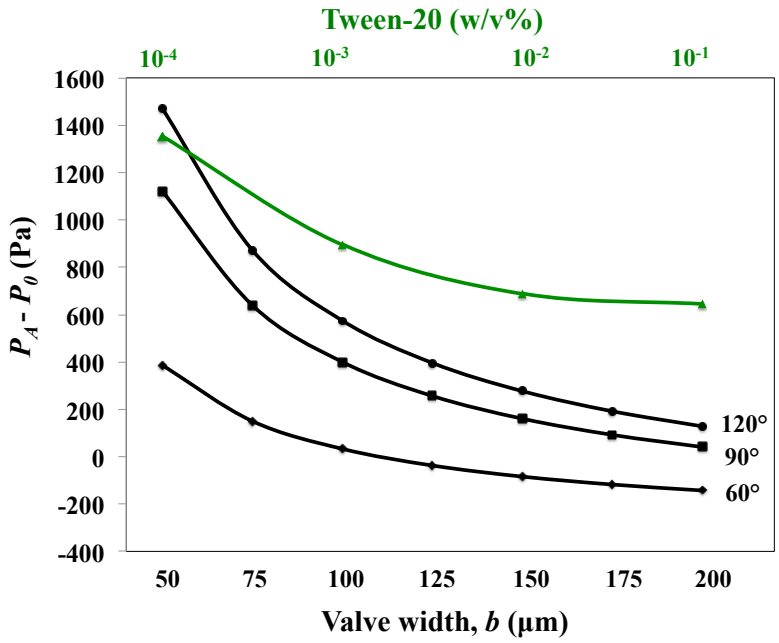



**Figure 3-5. Effect of tween-20, valve width, and diverging angle on valve pressure**

The green line (triangles) shows the effect of tween-20 concentration (top abscissa) on the valve pressure at constant width (b=75 µm), height (h=170 µm), and diverging angle (β=120°). Increasing tween-20 concentration gives smaller valve pressures due to smaller contact angles on SEBS and decreasing surface tension. The black lines show the effect of valve width (bottom abscissa) and diverging angle, denoted by the numbers to the right of the curves, on valve pressure at constant tween-20 concentration (0.001%) and height (170 µm). Increasing width decreases valve pressure and increasing diverging angle increases valve pressure. We experimentally determined that approximately 800 Pa of valve pressure drop was required to hold during ITP operation. We use a concentration of 0.001%, valve width of 75 µm, height of 170 µm, and β of 120° in our design for a valve pressure of approximately 900 Pa.

### 3.4.2 Air channel pump

Heating the microchannel to the 65°C required for LAMP activates the air channel pump. The pressure generated by the heated air channel achieves three goals: (1) it overcomes the energetic barrier created by the capillary valve; (2) assists in driving flow towards the reaction chamber along with capillary forces; and (3) displaces liquid from the extraction chamber so that the reaction chamber is fluidically isolated during amplification. The pressure generated by the air channel is dependent on the temperature after heating, $T_2$, and the temperature during device filling, $T_1$, and is given by the ideal gas or Gay-Lussac's law, $P_2 / P_1 = T_2 / T_1$. This pressure breaks the capillary valve and, along with capillary flow, drives liquid to fill the reaction chamber. The volume change caused by the enclosed air chamber is given by,

$$\Delta V = V_1\left[\left(\frac{T_2}{T_1}\right)-1\right] \quad , \tag{11}$$

where $\Delta V$ is the change in air volume, $V_1$ is the initial air channel volume, $T_1$ is the initial temperature before heating, and $T_2$ is the final temperature after heating. We designed the enclosed air chamber to have a volume change that is 90% of the extraction chamber volume when heated from room temperature, 25°C, to 65°C. This volume change causes liquid to evacuate the extraction chamber and isolate the filled reaction chamber during amplification of the target DNA, as seen after 30 seconds in Figure 3-3C. Isolation is necessary to prevent dilution and diffusion of LAMP reagents and purified nucleic acids during amplification.

The balance between the hydraulic resistances of the LE channel, ITP channel, as well as the valve and reaction chamber control the thermopneumatic flow driven by the air channel. First, the hydraulic resistances in the LE and ITP channels should be large relative to the valve burst pressure such that the pressure generated by heated air causes the valve to burst with only marginal flow through the LE and ITP channels. Second, the hydraulic resistance of the valve channel should be less than the hydraulic resistance of the ITP channel and LE channel to promote pressure driven flow, along with capillary flow, towards the reaction chamber after the valve bursts. Finally, the hydraulic resistances of the ITP channel and LE channel should be approximately equal to avoid preferential flow through either channel. After filling the reaction chamber, the remaining liquid will begin to evacuate the extraction chamber, thereby isolating the reaction chamber, by flowing through the LE and/or ITP channels towards the LE and/or TE inlets. Preferential flow through the LE or ITP channel may result in the reaction chamber retaining a fluid connection to one of the channels and prevent isolation of the reaction chamber. For example, Figure 3-3 at t=10s would have the possibility to retain fluid connection if there was preferential flow through either the ITP or LE channel.

We performed hydraulic circuit calculations to understand the flow direction based on channel geometry. Figure 3-6 shows a diagram of each channel considered as a resistor, $R_i$, that has a pressure applied to it, $P_i$, and corresponding volumetric flow $Q_i$. Channel 1 models the heated air chamber, channel 2 the LE channel, channel 3 the ITP channel, and channel 4 the reaction chamber. The resistance of each channel was calculated using $R_H \cong 8\eta L / r_H^2 A$, where $\eta$ is the fluid viscosity, $L$ is the channel length, $r_H = 2A/C$

is the hydraulic radius of the channel, *A* is the cross-sectional area of the channel, and *C* is the perimeter of the channel. Table 3-1 gives the pressures and dimensions for each channel of the optimized design. The flow rates for the given parameters were determined using a series of equations, $P_i = Q_i R_{Hi}$ , where *i* denotes each channel.

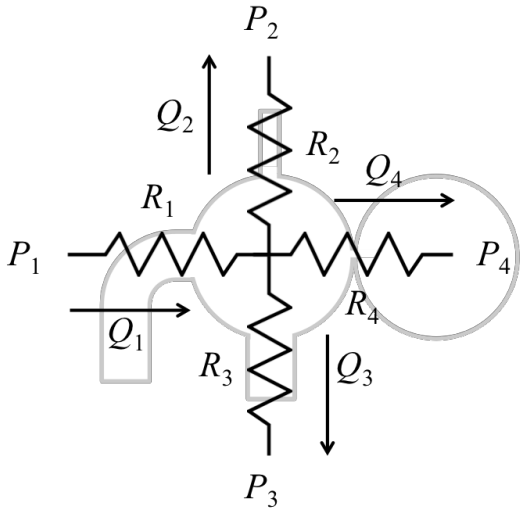



**Figure 3-6. Schematic of hydraulic resistance model**

Schematic showing the resistor models for each channel with pressure, P, hydraulic resistance, $R_H$, and volumetric flow, Q. Channel 1 corresponds to the air chamber, channel 2 to the LE channel, channel 3 to the ITP channel, and channel 4 to the capillary valve and reaction chamber.

**Table 3-1. Parameters used for each channel in the hydraulic circuit calculations**

| | Channel 1 | Channel 2 | Channel 3 | Valve (Channel 4) | Reaction chamber (channel 4) |
|---|---|---|---|---|---|
| **P (Pa)** | 108730 | 101365 | 101365 | N/A | 101325 |
| **Height (µm)** | 170 | 170 | 170 | 170 | 170 |
| **Width (µm)** | 2000 | 125 | 600 | 75 | 1500 |
| **Length (mm)** | 0.1 | 4 | 50 | 0.1 | 2.5 |

The pressure for channel 1 is the mean of the air channel pressure immediately after heating to 65°C from room temperature and atmospheric pressure according to the ideal gas law. The pressure for channel 2 and 3 is atmospheric pressure plus the hydrostatic pressure from liquid in the inlets. Channel 4 considers the valve and reaction chamber as resistors in series because the air valve (atmospheric pressure) for channel 4 is located at the end of the reaction chamber. Based off the calculations using the parameters in Table 3-1, the hydraulic resistances for channel 2 and 3 are $2.6 \times 10^{11}$ $Ns^{-1}m^{-5}$ and $2.5 \times 10^{11}$ $Ns^{-1}m^{-5}$

respectively, and the volumetric flow through channel 4 should be 9.39 times greater than the LE channel, *i.e.* $Q_4/Q_2$=9.39, and 9.06 times greater than the ITP channel, *i.e.* $Q_4/Q_3$=9.06.

### 3.4.3 Separation capacity

As described in Section 1.3.2, the ITP channel must contain sufficient ITP separation capacity in order to fully separate and purify DNA from contaminants.[88] Separation capacity is the amount of charge that passes through an ITP system before the target analyte reaches its final destination, in this case DNA reaching the extraction chamber, and is approximately dependent on channel volume and leading ion concentration.[42,88] The separation capacity should be greater than the separation parameter, the amount of charge required to separate two groups of ions with different mobility, to isolate one group of ions from the other during ITP.[88,187] If the separation capacity is greater than the separation parameter, the two ions groups should be completely separated by isotachophoresis. In this case, we were interested in designing HEPES adjusted TE to separate from the fatty acids in the milk. The separation parameter is given by equation 6 and the separation capacity is given by equation 8 in Section 1.3.2.[88] We used the values given in Table 3-2 to calculate $Q_s$ according to equation 6. The $Q_L$ value was determined using experiments that typically required 4 minutes at 450 µA to complete, *i.e.* $Q_L$=240s*0.45 mA = 108 mC. The calculated separation capacity value for our system was 98.0 mC using equation 8.

**Table 3-2. Inputs and results for separation parameter and capacity calculations**

| Component | $N_i$ ($x10^7$ mols) | $u_i$ ($x10^8$ $m^2V^{-1}s^{-1}$) | $Q_s$ (mC) | $Q_L$ (mC) |
|---|---|---|---|---|
| **A: HEPES** | 2.8 | -1.6 | 65.4 | 108.0 (experimental) |
| **B: Fatty acids** | 7.0 | -0.3 | | 98.0 (calculated) |
| **R: Tris ions** | – | +0.2 | | |

We assumed the fatty acids to be the fastest moving contaminant due to the high charge of triglycerides and the relatively low mobility of proteins, so the separation parameter is determined using the fatty acids in the milk. The moles of fatty acids were calculated based off of 8 g of fat per serving of milk (240 mL), with 15 µL input and an assumed average molecular weight of 711 g/mol using the size of common triglyceride chains in cow's milk.[188] The fatty acid mobility was measured using a Malvern Zetasizer with 10% milk in TE buffer because 50% milk was too turbid to recover results. The mobility was corrected for ionic strength differences between 10% and 50% milk TE using SPRESSO open source simulation

software.[189] The mobility of HEPES and Tris in the adjusted TE zone was determined using SPRESSO.[189]

### 3.5 Isotachophoresis extraction and purification

We have demonstrated the extraction and purification of nucleic acids from *E. coli* O157:H7 cells inoculated in whole milk samples using ITP. Whole milk samples contain an abundance of calcium ions, proteins, and fatty acids that can inhibit amplification reactions. ITP extracts and purifies nucleic acid for downstream amplification with minimal sample preparation and reduces manual steps compared to standard bead or column based purification techniques. ITP requires two pipetting steps (adding sample to TE and adding TE/sample mixture to chip) and pushing a button to apply an electric field, while standard purification techniques require over 10 pipetting steps and vacuum or centrifuging procedures. To prepare the samples, we add whole milk inoculated with target cells or genomic DNA to TE buffer in a microtube at a 1:1 ratio, resulting in two times dilution of the original milk sample. The TE buffer contains enzymatic lysing agent to lyse the cells at room temperature and release the DNA, as well as proteinase K to digest proteins that may bind DNA and alter its mobility.[42] The Ready-Lyse Lysozyme is added in excess to efficiently lyse the cells at room temperature and release the DNA into solution, while the tween-20 assists in lysis and solubilizing proteins following lysis.[62,65–67] After 15 minutes of incubation, the TE and sample mixture can be added directly to a device previously filled with LE buffer in order to extract and purify the DNA for amplification using ITP.

After the NAIL device is filled with LE and sample has been added, applying an electric field extracts the DNA from amplification inhibitors present in the milk. Cations, such as calcium present in milk, electromigrate in the opposite direction of the anionic DNA and do not exit the TE well. Anions, such as proteins and fatty acids electromigrate in the same direction as DNA, but have slower electrophoretic mobilities and are not able to overcome the TE buffer to enter the ITP plug. The adjusted TE mobility of the HEPES anions is approximately $1.6 \times 10^{-8}$ $Vm^{-2}s^{-1}$ after conductivity corrections.[82,83,189,190] The mobility of the proteins and fatty acids from milk in TE buffer are approximately $0.3 \times 10^{-8}$ $Vm^{-2}s^{-1}$, while the DNA should have an electrophoretic mobility of approximately $2.7 \times 10^{-8}$ $V/m^2s$ in the TE buffer.[85] Fully separating the DNA from the fatty acids and proteins requires 65.4 mC of separation capacity, which is less than the 108 mC separation capacity of the system, meaning DNA should separate from the fatty acids and

proteins in the milk sample.[88] The ITP separation process only requires four minutes after applying the electric field.

Figure 3-7 A shows a fluorescent image of *E. coli* O157:H7 DNA extracted and purified from cells inoculated in whole milk using ITP. The image is taken approximately 3 mm below the extraction chamber. Figure 3-7 B shows an image of the chip after DNA has entered the extraction chamber. The fatty acids and proteins, shown in Figure 3-7 C, only migrate about 60% of the channel length before the current is stopped, indicating that the DNA has been purified from the milk.

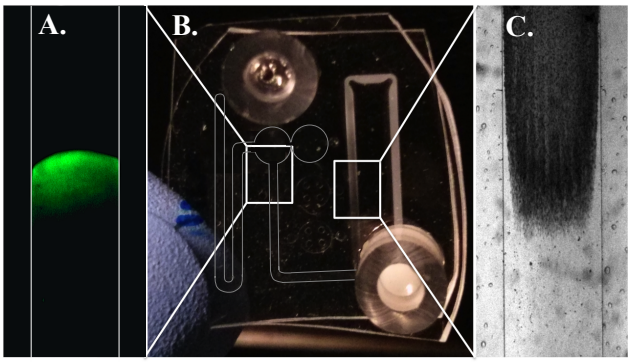



**Figure 3-7. ITP separation of DNA from whole milk**
(A) Fluorescent image of an ITP plug containing purified DNA from E. coli O157:H7 cells bound to 1x SYBR green dye 3 mm below the extraction chamber. (B) An image of the NAIL device after DNA has reached the extraction chamber, showing milk in the TE reservoir and filling approximately 60% of the channel. Approximate channel dimensions are drawn to show the location of the plug. (C) Transmission micrograph of milk in the channel (dark) at the termination of ITP. ITP sufficiently purifies the DNA from inhibitors present in the milk sample and allows for amplification downstream from only a 2x diluted sample.

LAMP is known to be robust to a number of perturbations and more tolerant to inhibitors than PCR reactions.[191] However, we found that significant dilution of unprocessed whole milk samples is required for the tube-LAMP experiments before quality amplification occurs. Figure 3-8 shows the normalized fluorescent intensity of LAMP experiments conducted in centrifuge tubes with 100 pg of pure genomic DNA in milk with various dilutions ranging from 10 to 10,000x. We normalize the data using a positive control with 100 pg of genomic DNA and a negative control with no DNA in water, denoted by the infinite (inf) data point. The data shows that the LAMP is unable to properly amplify the DNA in the presence of high concentrations of milk, *i.e.* at or below ~500x dilution. 1000x dilution was the smallest dilution amount that gave consistent, expected results according to the established LoB cutoff. At the dilutions above 1000x, the fluorescence intensity is relatively stable, indicating proper amplification. Table 3-3 gives a summary of all

milk dilution tests, indicating positive and negative results at each dilution. Using samples purified by ITP, we are able to consistently amplify nucleic acids from milk using only two times dilution of the original milk sample, indicating that ITP is sufficiently purifying the DNA for amplification. These results are supported in Figure 3-7 and Figure 3-9, by the separation capacity calculations, and by previous work that has shown ITP-purified DNA is compatible with PCR, which is known to be more sensitive to inhibitors than LAMP.[42,77,79,191–193]

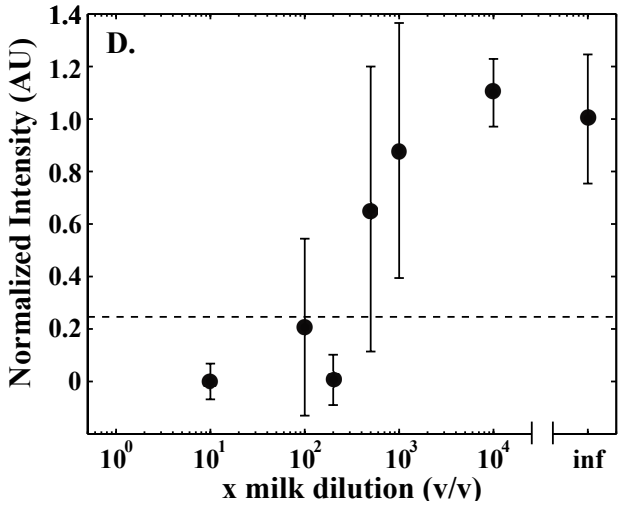



**Figure 3-8. Tube-LAMP amplification with various dilution of milk**

Average normalized fluorescent intensity of tube-LAMP reactions containing various dilutions of milk. The data is normalized by $(I_P - I_N)/(I_{P,\mathrm{inf}} - I_{N,\mathrm{inf}})$, where I is the fluorescent intensity; subscript P is the positive reaction; N is the negative reaction; P,inf is a positive control with no milk (infinite dilution); and N,inf is the negative control with no milk. The dashed horizontal line shows the cutoff used to determine positive versus negative reactions and is given by the LoB using the standard deviation of negative control reactions. Error bars represent $1.645\sigma_s$ where $\sigma_s$ is the standard deviation of the samples. 1000x milk dilution was the lowest dilution amount that gave reproducible LAMP results above the cutoff according to the LoD definition.

**Table 3-3. Summary of milk dilution results**

| Dilution | Results |
|---|---|
| 10x | 0/3 |
| 100x | 1/5 |

| | |
|---|---|
| 200x | 0/3 |
| 500x | 2/3 |
| 1000x | 5/5 |
| 10000x | 3/3 |
| Pure water | 15/15 |

### 3.6 Limit of detection

To determine the limit of detection of the NAIL device, we tested 10-fold dilutions of cells (CFU/mL) inoculated in whole milk with the device. We compare these results to tube-LAMP amplification using 10-fold dilutions of cells (CFU/mL) inoculated in whole milk and then diluted 1000x, and report results from previous studies using immunological or molecular methods to detect bacteria from milk. The purpose of these experiments is to show that ITP sufficiently purifies DNA for amplification from a minimally diluted sample, that the limit of detection for an initial sample concentration may be improved by using NAIL with ITP purification as opposed to simply diluting the sample, and that the NAIL LoD is comparable to other molecular methods such as lab-based PCR that require user intensive sample extraction preparation protocols. We use the Clinical and Laboratory Standards Institute (CLSI) definitions of limit of blank (LoB) and limit of detection (LoD) for determining our LoD.[109,147] Many publications currently report LoDs based only on measurements from blank samples, if any statistical measure is used at all. The rigorous CLSI statistical definitions take into account the variance in both positive and negative (blank) results when determining the LoD, and ensure that test concentrations at or above the LoD (up to the upper measuring limit) are truly positive at least 95% of the time according to the cutoff established by the LoB.

Figure 3-9 A and B show unprocessed negative and positive mobile phone images of the reaction chamber following 45 min of LAMP amplification of *E. coli* O157:H7 DNA purified from cells inoculated in whole milk. The positive reaction chamber is clearly visible in the mobile phone image, while the negative reaction chamber is only faintly visible. Figure 3-9 C shows the average normalized fluorescent intensities of at least 5 measurements at each cell concentration for the NAIL device (black circles) and at least 3 measurements for tube-LAMP with 1000x diluted milk (blue crosses). Table 3-4 shows a summary of all experimental

results. Both the NAIL device and tube-LAMP exhibit roughly sigmoidal responses due to the qualitative response of the calcein dye.

The solid horizontal line represents the cutoff for determining positive and negative samples for the NAIL device and the dashed horizontal line represents the cutoff for the tube-LAMP tests. The cutoffs are given by the LoB, or 1.645 times the standard deviation above the mean blank (negative reactions) response. Repeated tests at a concentration directly at the cutoff should yield positive results 50% of the time and negative results 50% of the time, meaning there is equal probability of committing type I and type II error.[147] The errors bars in Figure 3-9 may appear to show large variance, but note that we have plotted 1.645 times the standard deviation in order to provide clarity for the CLSI LoD definitions because the error bars, along with the cutoff, are used to determine the LoD. The lower bound of $1.645\sigma_S$ should be equal to or greater than the cutoff value to be considered at or above the LoD, which is visually shown in Figure 3-9. Plotting the same data with the 95% confidence intervals, as shown in Figure 3-10, results in much more narrow intervals for the fluorescence signal at each tested concentration. The 95% confidence intervals indicate our confidence in creating an interval that contains the true mean and show that the mean fluorescence for the tests conducted above the LoD are unambiguously larger than those conducted below the LoD.

We determined the LoD of our NAIL device to be $10^3$ CFU/mL because it is the lowest concentration for the given data that has a mean that is $1.645\sigma_S$ above the cutoff. Due to the qualitative response of the dye, we only considered actual test concentrations as possible LoD concentrations, as opposed to directly calculating the LoD concentration using the pooled sample standard deviation protocol from section 4.3 of the EP17-A document, which gives a lower LoD concentration. At this concentration, 10/10 samples were determined to be positive, giving us 76.6% confidence that $10^3$ CFU/mL is at or above the $C_{95}$ concentration.[109] Testing more samples would likely increase this confidence according to Table A2 in the EP17 document. For the 1000x dilution tube-LAMP experiments, the LoD was determined to be $10^5$ CFU/mL, or two orders of magnitude higher than the NAIL device using ITP purification. Table 3-4 shows a summary of LoD determination results for the NAIL tests, as well as the tube-LAMP results with 1000x diluted milk. The differences in initial dilution and ITP extraction efficiency approximately account for the LoD difference between NAIL and tube-LAMP. The tube-LAMP requires 1000x dilution, or 500x more

dilution than NAIL, but the NAIL device has 10% extraction efficiency, which means that NAIL should be 50x better by a first approximation. The LoD of the NAIL device also has potential to be further improved with larger sample volumes and higher extraction efficiency, which are not feasible with diluted samples.

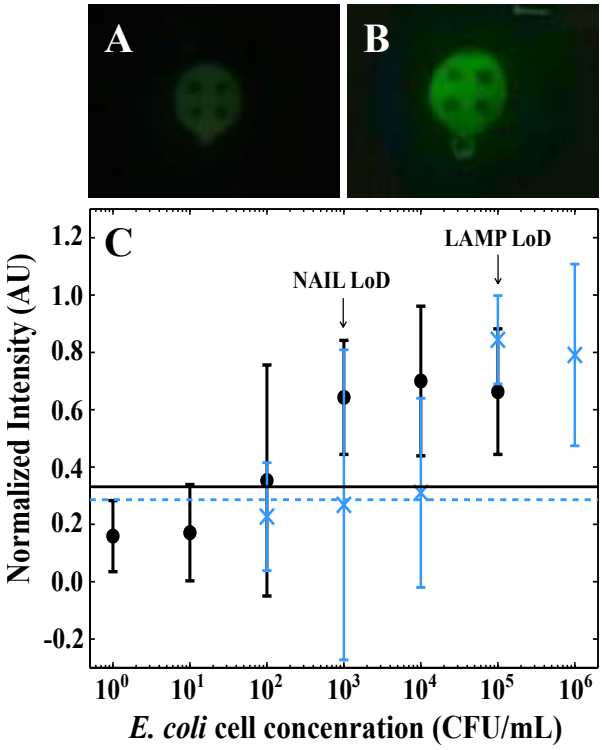



**Figure 3-9. Limit of detection results for NAIL and tube-LAMP**

(A) Unprocessed image of the NAIL reaction chamber following a negative reaction. (B) Unprocessed image of the NAIL reaction chamber following a positive reaction. The fluorescent intensity of the positive reaction is noticeably greater than the intensity of the negative reaction. (C) Average normalized intensity of the NAIL device reaction chamber (black circles) and tube-LAMP reactions with 1000x diluted milk (blue cross) as a function of cell concentration (CFU/mL) in the original milk sample. Error bars represent 1.645 times the sample standard deviation, $1.645\sigma_S$, for each data point according to the CLSI definition of LoD. The two separate data sets were normalized between 0 and 1 using $(I-I_{min})/(I_{max}-I_{min})$, where I is fluorescent intensity, $I_{min}$ is the lowest intensity from negative responses, and $I_{max}$ is the highest intensity from positive responses for each data set. The black solid line represents the NAIL device cutoff, while the blue dashed line represents the tube-LAMP cutoff as determined by the means and 1.645 times the standard deviations of negative control experiments. The data shows that the NAIL device can reliably detect $10^3$ CFU/mL in the original milk sample, while the 1000x diluted milk samples can detect $10^5$ CFU/mL. Using the NAIL device with ITP purification improves upon the limit of detection by two orders of magnitude.

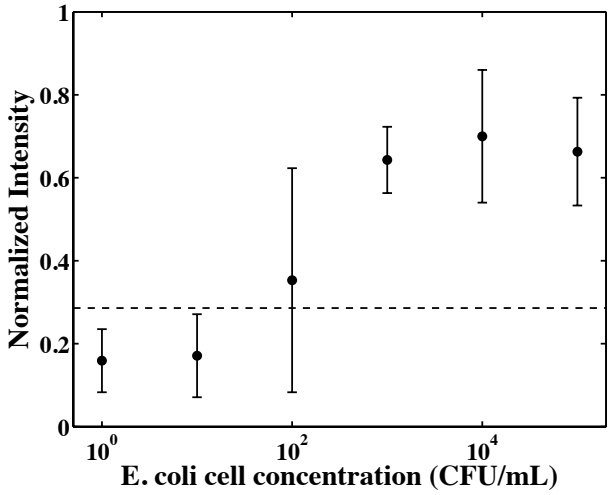



**Figure 3-10. LoD results with 95% confidence intervals.**

Normalized fluorescent intensity versus E. coli cell concentration (CFU/mL) for NAIL limit of detection tests with 95% confidence intervals as the error bars. The dotted line shows the threshold for determining positive (above the threshold) or negative (below the threshold) tests. Concentrations above 1000 CFU/mL show consistently positive positive results.

**Table 3-4. Summary of LoD results for the NAIL and tube-LAMP tests.**

The LoD results are for varying initial cell concentration levels in whole milk. The results are given as the number correct out of the total number of valid runs.

| Initial cell concentration (CFU/mL) | NAIL | tube-LAMP |
|---|---|---|
| 0 | 0/22 | 0/8 |
| 1 | 0/5 | – |
| $10^1$ | 0/5 | – |
| $10^2$ | 1/5 | 0/3 |
| $10^3$ | 10/10 | 0/3 |
| $10^4$ | 5/5 | 0/3 |
| $10^5$ | 5/5 | 5/5 |
| $10^6$ | – | 3/3 |

Culture based methods are generally considered the gold standard for pathogen detection in food, but typically consist of multiple incubation steps (pre-enrichment, selective enrichment, selective and differential plating) before a biological, serological, or molecular confirmation test. More rapid tests including immunological and molecular tests have been developed to speed the detection process. Our LoD compares favorably to LoDs reported for other immunological and molecular based tests. Nogva et al. and

Ibekwe et al. reported qPCR detection limits on the order of $10^3$ (between 2E2-2E3 and 6.4E3 CFU/mL respectively) for detecting *L. monocytogenes* from milk and *E. coli* from milk wastewater, respectively.[34,194] Both of these studies used manual purification techniques such as bead based assays or other extraction kits. Hoehl et al. reported an LoD of 1,000 inserted *E. coli* cells from milk for qPCR and LAMP in their semi-automated LabSystem that corresponds to 10,000 CFU/mL for their 100 µL sample input. Daly et al. combined PCR amplification with ELISA detection to achieve an LOD of 100 *E. coli* CFU/mL from milk.[195] They state that manual purification and PCR without ELISA results in an LoD of 1,000 CFU/mL. Immunological based tests with surface plasmon resonance or electrochemical detection have reported detection limits of 100-1,000 and 10,000 *E. coli* CFU/mL from milk and lettuce respectively.[196,197] NAIL shows a comparable LoD, but has the advantage of being a rapid, integrated device that does not rely on considerable user interaction and significant ancillary, lab equipment.

# Chapter 4: Simultaneous nucleic acid extraction and amplification using electrokinetic paper substrates

## 4.1 Motivation and background

The key aspect of the NAIL device presented in Chapter 3 is that it shows the potential of combining isotachophoresis (ITP) sample preparation with isothermal amplification, but further innovations are required for point-of-care (POC) diagnostic use, including faster time-to-result and increased automation. The first needed improvement is to remove any form of pumps or valves. While the NAIL device uses passive valves and pumps that are relatively simple to fabricate and operate, removing these fluid actuation mechanisms will decrease complexity and reduce failure. Second, using loop-mediated isothermal amplification (LAMP) eliminates the need for a thermocyler, but still uses significant energy and a hot plate to be heated to 65°C for 45–60 minutes. An improved system would require little to no external heating for amplification. Finally, the time-to-result for the NAIL device is approximately 1 hour, which is faster than many diagnostic technologies, but should be fast enough for patients to be tested and treated in a single visit. For these applications, the time-to-result should be less than 30 minutes, and ideally less than 15–20 minutes for rapid diagnosis and immediate care.

This chapter discusses the development of a porous glass fiber “paper-based” device that combines ITP and recombinase polymerase amplification (RPA) to simultaneously extract and amplify nucleic acids in a single step. This platform alleviates the issues encountered during NAIL development by providing: (1) single step extraction and amplification with no complex fluid actuation, (2) minimal or no external heating requirements, and (3) time-to-result between 10–20 minutes. We will discuss the integration of RPA reaction chemistries with ITP operation on paper-based substrates to demonstrate the detection of synthetic HIV-1 nucleic acids from serum and whole blood samples.

The HIV/AIDS epidemic is a major global health challenge and a significant cause of mortality over the last decade.[198] The World Health Organization (WHO) estimates that there were 35 million people living with HIV (PLHIV) as of 2013.[199] Of those diagnosed HIV+, nearly 15 million PLHIV receive antiretroviral therapy (ART), a figure that falls short of the WHO/UNAIDS targets to provide lifelong ART for all 35 million PLHIV. Initiation of ART can maintain durable viral suppression, slowing progression from HIV infection to acquired immune deficiency syndrome (AIDS).[200] Viral load (VL) tests are needed to monitor the ongoing effectiveness of ART therapy, so that treatment regimens can be adjusted to maintain viral suppression if needed.[200] The WHO defines ART failure as two consecutive VL tests above 1000 cp/mL after receiving

ART for >6 months.[201]

By 2020, UNAIDS aims to have 90% of PLHIV aware of their status, 90% to receive sustained ART, and 90% of those on ART to have sustained VL suppression (90-90-90).[202] To achieve these goals, there is a need for 35–70 million global HIV VL tests per year. However, performing routine laboratory-based VL testing for the 15 million people receiving ART has been a challenge, particularly in resource limited settings, where lack of VL testing leads to undiagnosed virological failure, late treatment switches, and potential spread of HIV drug resistance.[12]

Current tests for measuring HIV VL rely on molecular amplification methods that detect HIV nucleic acids from plasma samples. Nucleic acid amplification tests (NAATs) are the gold standard for HIV-1 viral load testing because they provide high diagnostic sensitivity and specificity, as well as quantitative information.[12] Despite the clear advantages of NAATs, most systems are restricted to central laboratories due to assay, instrumentation, and/or protocol complexity that requires skilled personnel to operate. The logistics around specimen collection, transport, and returning of results for laboratory NAATs typically delay obtaining diagnoses and often negate patient management benefits. For example, a study in South Africa showed that only 69% of newly diagnosed HIV-infected adults received lab-based results within 90 days, and less than half of those began ART within 12 months, *i.e.* ~35% of HIV+ patients received ART within 12 months.[9,203]

The motivation for this work began through a collaboration with David Boyle from the Program for Appropriate Technologies in Health (PATH, Seattle, WA), whose work focuses on developing a POC diagnostic for HIV-1 VL in low resource settings. They have developed the biochemistry to detect HIV-1 using RPA, but require an integrated device that includes sample preparation from whole blood.[99,101,103,204] Thus, our goal is to develop a low cost and rapid NAAT platform that can perform accurate diagnosis and VL for HIV-1 from a finger-prick whole blood sample to improve patient care in peripheral settings. Clinic-based HIV VL monitoring would allow same day adherence counseling, simplify patient management, and allow for earlier detection of ART failure, all at reduced cost.[9,12] Current tests do not meet the needs for austere settings due to a combination of long test times, complex instrumentation, high-cost, operation requirements <30°C, and large plasma volumes (~1 mL obtained through intravenous draw and centrifugation of whole blood to generate plasma). Specific requirements for a viable HIV VL

diagnostic include: (1) results in under 30 minutes, (2) robust hardware and operation, and (3) high performance at low-cost.

Figure 4-1 summarizes an example of the envisaged technology that the current work could enable in future prototypes. An ideal operation workflow would be the user adding finger-prick whole blood to a collection filter using a disposable constant volume sampler such as the Safe-Tec Microsafe (Figure 4-1A).[205] After sealing the device to prevent contamination, the cartridge could be inserted into a reader for testing where a blister pack would be burst for buffer release and rehydration of dried reagents on the paper substrate. The mobile phone based reader would consist of basic electronics and optics, and perform all required steps including extraction, amplification, and detection (Figure 4-1B,C). Besides following on-screen prompts, the user would simply be required to add sample and insert the cartridge into the reader (Figure 4-1D,E), providing a test that could be operated by minimally trained users.

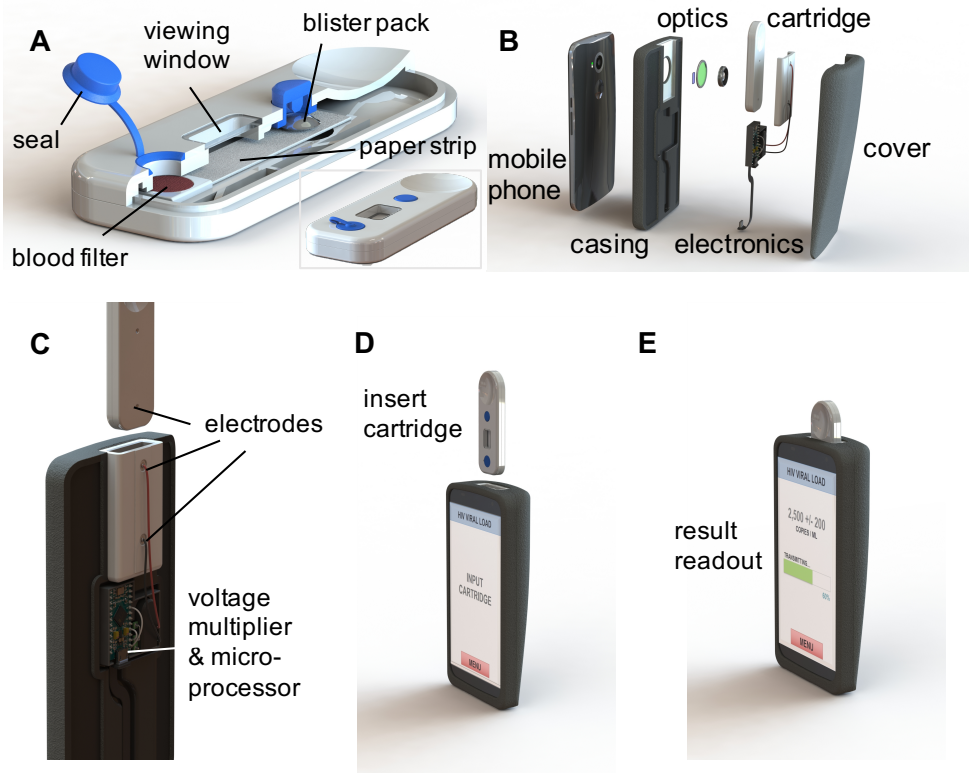



**Figure 4-1. HIV-1 nucleic acid detection concept**

The concept device consists of a paper-based cartridge and mobile phone reader for automated, 15–20 minute detection of viral RNA. (A) Cartridge containing paper strip, whole blood filter, electrodes, and laminate heater. The disposable cartridge accepts unprocessed whole blood and can be sealed to prevent contamination (A inset). (B) Components of mobile phone reader, consisting of basic plastics, electronics, and optics. (C) Detailed view of electronics that provide current for ITP-RPA extraction and amplification of

HIV-1 target (D) User inserting the sealed cartridge with sample for test operation. (E) Unambiguous readout of quantitative VL for HIV-1 diagnosis and treatment guidance.

The remainder of this chapter will focus on the paper-based device design and the associated chemistry that is required to enable the envisaged diagnostic device in Figure 4-1. The described work will show the development of an ITP-compatible chemistry that is able amplify RPA reactions on porous glass fiber substrates within an ITP plug using a buffer sample. We will then show the adaption of the system to detect nucleic acids from blood serum samples, as well as a filtering method to passively generate plasma from whole blood in order to provide detection from unprocessed finger-prick samples. Finally, we will show analytical evaluation of the current device, including LoD determination and demonstration of result linearity.

## 4.2 Experimental methodology: simultaneous ITP-RPA fabrication, chemistry, and operation

### 4.2.1 Substrate and device fabrication

Our ITP-RPA device (ITP extraction and concentration combined with RPA amplification) consists of a porous glass fiber strip placed between two liquid buffer reservoirs and housed within a sealed petri dish (60 mm diameter, Fisher Scientific). Figure 4-2A shows an image of the petri dish with the glass fiber strip creating a fluidic connection between reservoirs for the leading electrolyte (LE) and trailing electrolyte (TE). The 30 x 3.5 mm (l x w) glass fiber (GFCP203000**,** EMD Millipore) strip was fabricated with a $CO_2$ laser cutter set to 7 power and 4 speed. The reservoirs were fabricated from 1/8" thick acrylic using a $CO_2$ laser cutter (Universal Laser Systems, Scottsdale, AZ) at 100 power/1 speed for vectorized cutting, and 70 power/10 speed for rastering. The rastered portion of the acrylic reservoir creates a notch that enables correct positioning of the glass fiber strip within the reservoir, as well as promotes better fluidic contact between the glass fiber strip and the liquid in the reservoir.

If whole blood is being used as the sample, the left side of the glass fiber strip also contains a Vivid© plasma separation membrane (GR, Pall Corporation) attached to acrylic using double-sided adhesive (3M). The Vivid membrane filters whole blood samples to create plasma containing target within the glass fiber substrate by capturing red and white blood cells in pores of the filter. The hydrophilic glass fiber in contact with the Vivid membrane creates the driving force for plasma separation via capillary action. If serum or buffer is the sample liquid, the Vivid filter is not required and sample is added directly to the left side of the glass fiber strip. The petri dish lid contains disposable electrodes (Titanium wire, 22 gauge,

UnkamenSupplies) placed through the lid to enable external electrical connections between the buffer reservoirs and electrical power source. By connecting to the power source outside the sealed dish, the amplification reaction inside the dish is not exposed to ambient air, which helps prevent amplicon contamination (Figure 4-2B). After placing the electrodes through laser cut holes in the top of the dish, they are sealed into place using UV curable adhesive to prevent leaks.

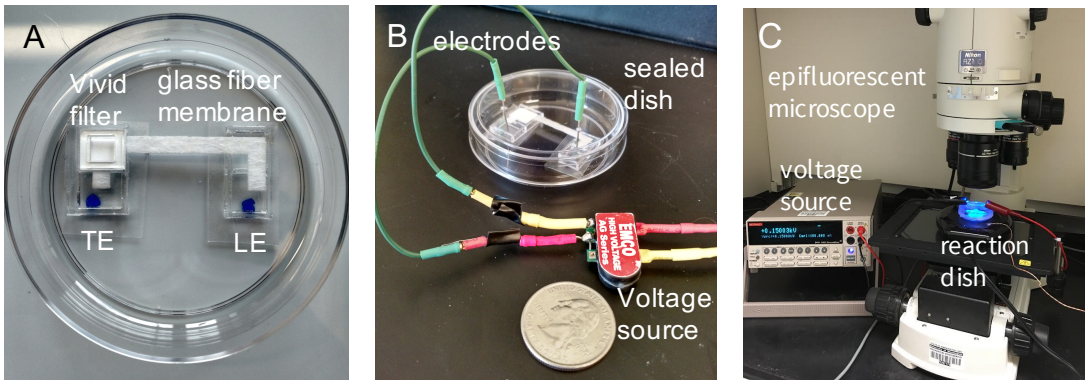



**Figure 4-2. ITP-RPA reaction setup**

(A) The ITP-RPA reactions consists of a glass fiber strip and two buffer reservoirs inside of a petri dish. If whole blood is being processed, a Vivid membrane filter is included on the glass fiber sample pad to create plasma with target on the glass fiber strip, while removing red and white blood cells. (B) Electrodes embedded through the lid of the petri dish provide external electrical connections for ITP operation. (C) Experimental setup used for the work presented in Chapter 4. The sealed reaction dish is monitored using an epifluorescent microscope and powered by a Keithly 2410 voltage source.

### 4.2.2 Reaction chemistry and operation

TwistDx has developed an RPA kit that consists of fives components: a dried pellet containing proteins, nucleotides, ATP, polyethylene glycol (PEG) and other necessary molecules for the reaction; a separate vial containing target primers and probes; a separate vial containing magnesium acetate; a separate vial containing rehydration buffer; and a separate vial containing target DNA.[104] The RPA amplification is initiated by mixing together the five different components at specified concentrations before incubation at 38–41°C. For example, the standard commercial reaction in a tube consists of 29.5 µL of rehydration buffer, 5 µL of primer/probes (120–500 nM final concentration), 1 µL of target DNA; 10.5 µL of water; and 4 µL of magnesium acetate (12–14 mM final concentration) for a total of 50 µL. This reaction mixture is used to rehydrate the dried pellet within the microtube and vortexed to mix.[104] The reaction is then placed within the Twista© unit that heats the mixture to 38–41°C and performs real-time monitoring of the fluorescent

output of the reaction to generate intensity versus time amplification curves. Figure 4-3 shows images of the reaction buffers, dried enzyme pellets, and Twista© unit.

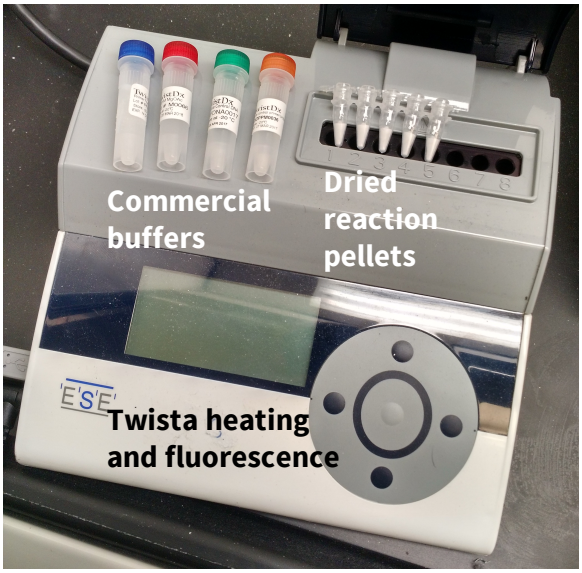



**Figure 4-3. Images of TwistDx reaction buffer, enzyme pellet, and Twista© unit**

For a standard tube reaction, the commercial reaction buffers are mixed with primers, probe, and target before addition to the dried pellet that contains all other reaction components (enzyme, nucleotides, etc.). This tube reaction can then be heated and monitored using laser fluorescence detection inside the Twista© heating and fluorescence unit. For the ITP-RPA reactions described in Chapter 4, only the dried reaction pellets are used. In these reactions, we used a custom ITP compatible buffer and the reaction setup shown in Figure 4-2.

As described in Chapter 4, we adapted the TwistDx RPA reaction chemistry to be suitable for use in an ITP system. The general protocol for running an ITP-RPA reaction consists of:

(1) Create TE solution

(2) Create LE solution

(3) Rehydrate commercial RPA pellet with LE solution, primers, and probe

(4) Add sample (buffer, serum or whole blood) containing target nucleic acids directly to left side of the glass fiber pad (buffer, serum) or to the Vivid membrane filter attached to the glass fiber (whole blood).

(5) Add rehydrated pellet solution to the center of the glass fiber strip

(6) Add non-pellet LE solution to LE well

(7) Add TE (without target) to TE well

(8) Seal petri dish lid using Parafilm with electrodes inserted into TE and LE reservoirs (cathode in TE reservoir, anode in LE reservoir)

(9) Apply 100–150 V across membrane and monitor reaction reaction fluorescence (with or without heat at 34ºC)

The TE buffer consists of glycine or serine weak acid, tris counterion, polyvinylpyrrolidone (PVP) and Triton-X100 at pH 8.5–9.2. The LE buffer consists of HCl, tris counterion, $MgCl_2$, polyethylene glycol (PEG), PVP, Triton-X100, and tetramethylammonium chloride (TMAC) at pH 8.1. We used the Curtipot Excel spreadsheet to calculate the ionic strength of the buffers that we used for ITP-RPA reactions. This spreadsheet allows input of buffer concentrations and calculates the resulting solution pH, buffering capacity, and ionic strength using the Davies equation. After creating TE and LE buffers, 45 µL of the LE solution is combined with primers and probe that specifically detect the target nucleic acid sequence of interest. This LE/primers/probe solution is then used to rehydrate the TwistDx RPA pellet. The rehydrated pellet (50 µL total) consists of:

- approximately 1x LE buffer
- 500–1000 nM of forward and reverse primers for nucleic acid amplification,
- 120–240 nM of sequence-specific fluorescence indicator probe, and
- 1x rehydrated RPA pellet containing all of the enzymes, nucleotides, and other biomolecules required for amplification using RPA.

The first step of operation is to add the sample to the left side of the glass fiber pad near the TE reservoir. If blood serum is being using as the sample, proteinase K (0.5 µg) and T-X (0.1%) should be dried on the strip left side of the strip where sample will be added before adding the serum. After adding sample, the rehydrated pellet is added to the center of the glass fiber strip and allowed to wet the glass fiber strip by capillary action. The LE reservoir is then filled with 200 µL of standard LE solution that does not contain primers, probe, or pellet. 180 µL of TE is then added to the TE reservoir. At this point, the glass fiber strip is fully wetted, both TE and LE reservoirs are full, and the glass fiber strip has established a fluidic connection between the two reservoirs. The lid to the dish is then placed on top to seal the reaction, and embedded electrodes in the lid are inserted into the TE and LE reservoir, as shown in Figure 4-2.

The dish containing the glass fiber with LE, TE, and RPA reaction is sealed using Parafilm and set on the microscope stage. The external electrode leads are connected to a voltage supply (Keithley 2410) and 100–150 V is applied to the paper strip with a current compliance of 1.5–3.5 mA. The reaction creates amplicons determined by the forward and reverse primers. These amplicons are detected by a sequence specific probe that creates a fluorescent signal upon binding to the center of the amplicon.[58] We performed quantitative fluorescence imaging to visualize the fluorescent signal generated by the amplicons using a Nikon AZ100 microscope equipped with a 0.5x (NA 0.05) magnification objectives (Nikon Corporation, Tokyo, Japan), an epifluorescence excitation and emission filter cube set (488 nm excitation, 518 nm emission, Omega Optics, Brattleboro, VT), and a 16-bit, cooled CCD camera (Cascade 512B, Photometrics, Tucson, AZ). Images are captured at one second intervals using MicroManager software and saved as image stacks using ImageJ software for later analysis.

### 4.2.3 Target and primer set

We began our ITP-RPA work aiming to detect HIV-1 nucleic acids using the RPA primer set developed by our collaborators from PATH.[99,204] However, in order to simply the initial experiments for proof-of-concept (*e.g.* not using RNA with reverse transcriptase), we used a synthetic DNA strand synthesized using IDT's gBlock gene sequence product line. Our synthetic gBlock consisted of a 200 base pair double stranded nucleic acid that nearly matched the HIV-1 pol gene that PATH's RPA primers targeted. However, one issue that we encountered was that the pol gene of interest contained too high of A-T content and too many repeating A or T sequences in order to be synthesized by IDT. As a result, we were required to make minor modifications to the sequence. We used these modifications as an opportunity to maximize primer specificity through IDT's oligo analyzer tool. In all, we changed approximately 20% of the nucleotides in each primer compared to the primer set from PATH and Boyle et al, but we did use the exact fluorescent probe sequence.[99] Thus, the ITP-RPA results in Chapter 4 use a primer set similar to Boyle et al for HIV-1 detection, but with minor modifications and a synthetic nucleic acid target. The exact target, primer, and probe sequences are given below.

Target:

5'AGGCTGAACATCTTAGGACAGCAGTACAAATGGCAGTATTCATTCACAATTTTAAAAGAAAAGGGG

GGATTGGGGGGTACAGAGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAGAG

AACTAGGTTGAGAAATTAGAAAAGTTGAATATGTTAGGGTTTATTACAGGGACAGCAGAGATCCACTT

3’

Green=forward primer, grey=probe, teal=reverse primer

Forward primer: 5’-TGGCAGTATTCATTCACAATTTTAAAAGAAAAGG-3’

Reverse primer: 5’CCCTAACATATTCAACTTTTCTAATTTCTCAACC3’

Highlighted sequences represent locations of modifications from Boyle et al. primers.

### 4.2.4 Data analysis of ITP-RPA reactions

The fluorescent intensity generated by each reaction indicates the success (or lack thereof) of target amplification. We captured a fluorescent grayscale image of the paper strip for each second during the course of the reaction (typically 900–1200 seconds) and created a MATLAB algorithm in order to efficiently process the intensity for each experiment. The full MATLAB code is included in the Appendix. After loading the image stack into MATLAB and running the code, the user is presented with a static image from the experiment, as shown in Figure 4-4A. The user should select the region of interest for which the analysis will be performed for all images. The migration of the plug during the experiment was always within the viewable range of a static microscope setup for these experiments, so we simply selected the entire paper strip outline for the region of interest, as indicated by the white dotted line in Figure 4-4.

After selecting the region of interest, the algorithm will perform automated analysis of all of the images, including (1) y-average the intensity of the image, (2) calculate and subtract the background, (3) set a threshold, and (4) integrate the intensity over the threshold for each image / time point. Figure 4-4B-E shows an example of these steps for the 1200$^{th}$ image of a positive reaction, *i.e.* after the reaction has occurred for 1200 seconds. The initial y-average intensity of the plug is shown in B, which has a background intensity near 5000 arbritary fluorescence units. The y-averaged background for the strip is calculated over the first 100 images before the ITP plug forms and subtracted from each subsequent image to give the background subtracted profile, as shown in Figure 4-4C. We then set a threshold, above which all signal intensity is integrated to create a single fluorescence value for each image, as shown in Figure 4-4D. The single fluorescence value for each image / time point is then plotted as a function of time to create

amplification intensity curves that closely resemble curves commonly seen from exponential reactions, such as qPCR, as demonstrated in Figure 4-4E.

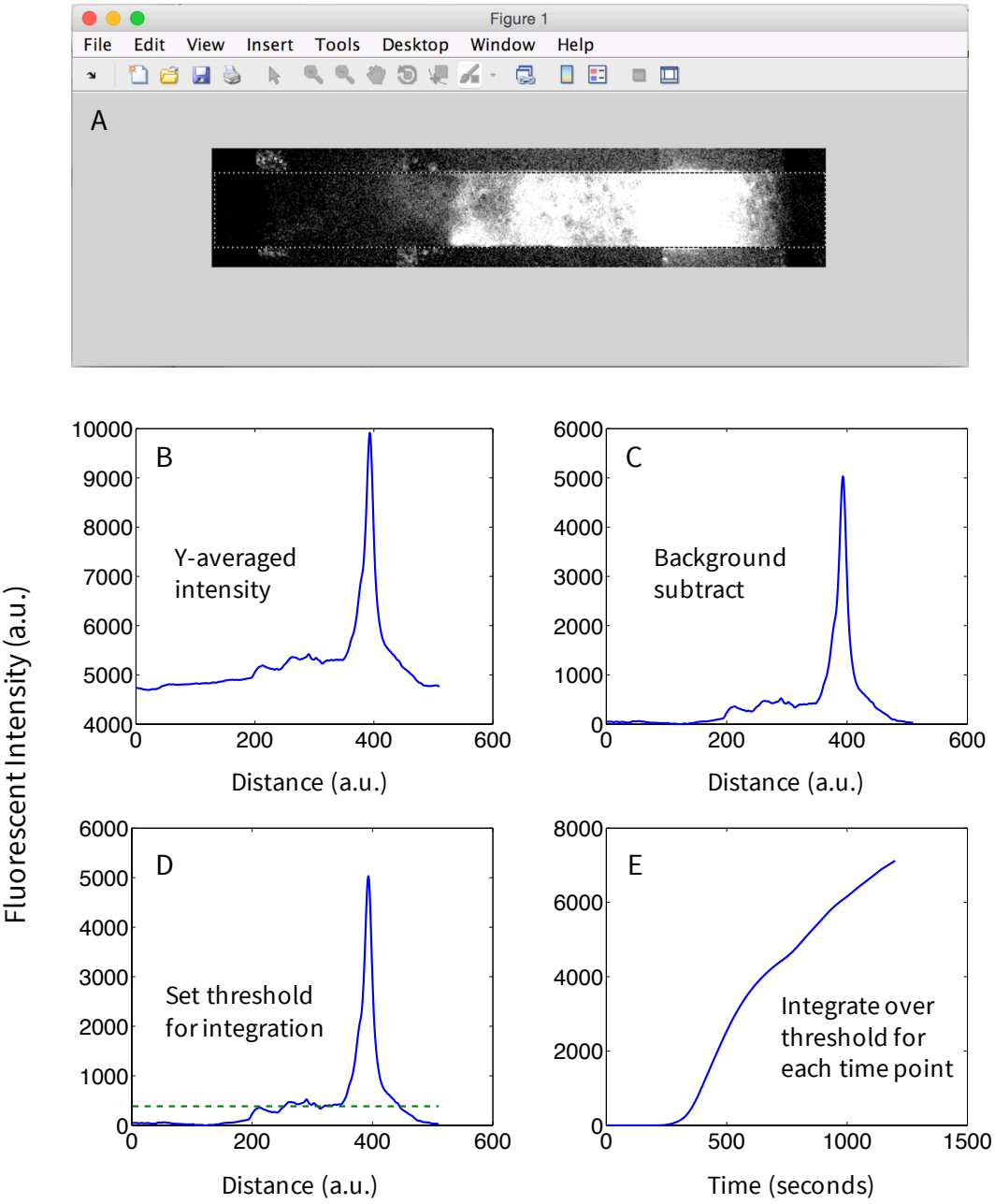



**Figure 4-4. Example of ITP-RPA data analysis**

We chose the threshold for the majority of data presented in Chapter to be 108% of the average background signal (the value used for background subtraction correction). This value was selected based on experience

with the system and stayed constant for every experiment, including for positive and negative controls. Using the threshold allowed us to eliminate non-specific background signal that appeared to result from adsorption of molecules on the surface of the substrate. By eliminating this non-specific signal, we were able to achieve increased signal to noise ratio between positive and negative experiments. The threshold value that we used resulted in many negative experiments exhibiting zero or near zero fluorescence because the entire signal is below the threshold. However, samples with low concentration will also approach this threshold, at which point using an LoD procedure to differentiate between blank and low level samples becomes necessary.

### 4.3 Glass fiber substrate use

We chose to use porous glass fiber substrates as opposed to standard microchannels in this work due to three key considerations: (1) higher separation capacities can more easily be achieved on glass fiber compared to microchannels; (2) glass fiber can offer higher repeatability without practical issues often experienced in microchannels including bubble formation, leaking, and loading difficulty; and (3) ease-of-fabrication and operation compared to microchannels.

High efficiency extraction of the target analyte from large sample volumes is necessary to provide low limit of detections (LoD) and high diagnostic sensitivity for clinical applications. As discussed in section 1.4.3, the separation capacity determines the amount of sample that an ITP system can process before the experiment ends, *i.e.* the efficiency of extracting the target analyte into the ITP interface, and the separation distance between the ITP interface and non-focusing inhibitors. Larger separation capacities can be achieved by increasing the concentration of the LE and/or increasing the volume of LE that a substrate holds. For this particular application, the LE concentration is restricted by the ionic strength dependency of the amplification reaction, meaning large substrate volumes are required for high separation capacities. Glass fiber substrates can easily hold one or two orders of magnitude higher volumes of LE (10–100+ µL) compared to a standard microchannel (<1–10 µL). For example, the glass fiber substrate used in this work holds approximately 100 uL of LE (approximately 30 mm long x 3.5 mm wide, with an absorbency of 88 $\mu L/cm^2$). In order for a microchannel with identical spatial footprint (30 mm long x 3.5 mm wide) to hold a similar volume of LE, it would require 1 mm height. Common photolithography and soft lithography fabrication processes could likely not be used to fabricate a chip of these dimensions, and operation

procedures such as loading the channel by capillary flow would become considerably more difficult (or impossible) due to the low capillary pressures that result from the large channel dimensions. In contrast, the glass fiber can easily be fabricated using laser fabrication and can be filled in less then 5 seconds by capillary forces with extremely high repeatability (the glass fiber is hydrophilic and can be visualized as many parallel microchannels, which individually have high capillary forces). The importance of the large cross-sectional area offered by glass fiber can also be highlighted when considering the time for separation. Passing large electric charge through a separation channel in an acceptable time for operation requires a high driving current. For a given applied voltage, channel length, and solution conductivity, applying a high driving current requires increasing cross-sectional area, which can readily be achieved using glass fiber substrates.

The glass fiber substrate used in this work has a separation capacity of approximately 1000–1800 mC, which is close to 20–40x higher than the microchannel used during the NAIL work if equal LE concentrations are assumed (50 mC). This separation capacity allows the use of 20–40 µL of undiluted serum sample in our experiments, with the ability to process larger sample volumes in future experiments. A specialized microchannel for large separation capacity ITP was recently published, showing a 159 mC separation capacity and 80% extraction of nucleic acids from a 25 µL sample. However, this separation capacity is still an order of magnitude smaller than the glass fiber substrate, and the blood sample was diluted 20x so that only 1.25 uL of blood sample was processed. Further, the chip required a specialized valve system to load the sample and a solid phase change material in the reservoirs to prevent high pressure driven flow velocities (due to reservoir liquid height differences and large channel dimensions). In contrast, sample can be added directly to the open surface of the glass fiber substrate (see section 4.5) and we have observed no operational issues from height differences in the liquid reservoirs. Overall, glass fiber substrates offer advantageous ITP properties for extraction and separation, while providing considerable practical benefits including ease of fabrication and operation.

### 4.4 Developing ITP-compatible RPA reaction chemistry

We intended to create a device that integrates ITP and RPA to rapidly extract, amplify, and detect target nucleic acids on a strip of porous glass fiber, *i.e.* “paper”. Figure 4-5 illustrates the application of ITP to separate nucleic acids from blood, while simultaneously focusing with reagents required for RPA. The glass

fiber strip and right reservoir are initially filled with LE buffer containing RPA reagents, while the left reservoir is filled with TE sample mixed with whole blood. Applying an electric field begins extraction of target nucleic acids using ITP, while focusing the nucleic acids with reagents required for RPA reactions. As the ITP plug migrates across the glass fiber substrate, amplicons are continuously produced, resulting in a fluorescent signal that can be detected to confirm amplification.

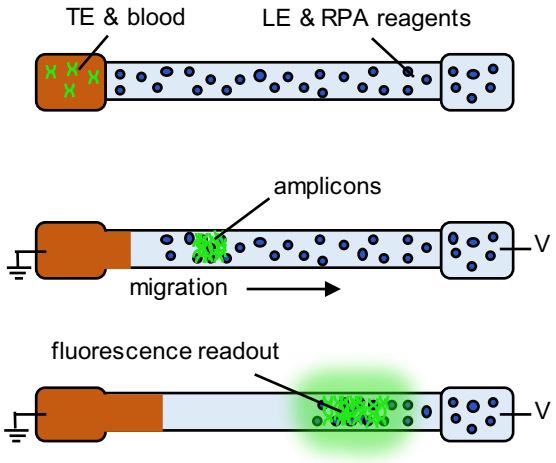



**Figure 4-5. Schematic of ITP-RPA operation**

The porous substrate is initially filled with LE buffer mixed with RPA reagents. The TE in the left reservoir contains nucleic acids mixed with the blood sample. Applying an electric field begins extraction and focusing of the nucleic acids within the ITP interface. As the ITP interface migrates across the substrates, the RPA reagents begin to focus with the nucleic acids, resulting in amplification of the target sequence via RPA. As the amplification reaction proceeds, fluorescent signal is produced that can be imaged for target detection.

The RPA reaction chemistry is produced and sold by TwistDx, a company formed in 2006 to manufacture and distribute this amplification chemistry. TwistDx provides reaction kits that contain all the reagents necessary to perform RPA reactions within a standard microtube format.[104] The reaction kits from TwistDx arrive with two main components—the dried enzyme pellet and a reaction buffer. To initiate an RPA reaction, the user adds their custom primers, probe, and target to the commercial reaction buffer, and this reaction mixture (50 µL) is used to rehydrate the enzyme pellet. After rehydration and mixing, the reaction mixture is heated to 38–41ºC and monitored using laser fluorescence within a Twista© heating and detection unit from TwistDx.

Our initial goal with this project was to create an ITP-compatible chemistry that matched the conditions of an RPA reaction inside of a tube. This step was required because the commercial reaction buffer supplied with the kit was not ITP compatible, and even if it was, its use would have restricted design and optimization

of the ITP-RPA reaction chemistry. The TwistDx reaction pellet and buffer contain proprietary components that include multiple enzymes, nucleotides, crowding agents, electrolytes, and other biomolecules (*e.g.* ATP, phosphocreatine).[58,104] We were able to deduce which components were likely in the pellet from patents detailing their RPA kits.[104] The main components contained within the reaction buffers were electrolytes (*e.g.* potassium acetate), crowding agents (*e.g.* Carbowax), and magnesium acetate. In order to recreate these conditions with a custom ITP-compatible buffer, we sought to determine the ionic strength of the RPA reaction pellet and full reaction (reaction pellet rehydrated with commercial buffer) so that our custom buffer could be used to rehydrate the pellet and give similar amplification results.

We measured the conductivity of known ITP-compatible Tris-HCl LE buffers by measuring their output current in a microchannel under an applied voltage to create a calibration curve that relates calculated ionic strength (using Curtipot spreadsheet) to measured conductivity. We then measured the conductivity of the RPA pellet rehydrated in DI water and the full reaction mixture (commercial reaction buffer rehydrating RPA pellet) in order to determine the calculated ionic strength of the reaction mixtures. The calibration curve, ionic strength of the pellet, and ionic strength of full reaction are shown in Figure 4-6A.

After determining the ionic strength of the pellet and full reaction, we created an ITP compatible Tris-HCl LE buffer that would give ionic strengths near 172 mM (the determined full RPA reaction ionic strength) when mixed with the RPA pellet and performed amplification reactions in the standard tube format to compare our custom reactions to the commercial reaction. Figure 4-6B shows that ITP-compatible reactions between 147–172 mM ionic strength give near-identical efficiency compared to the commercial positive control. All ITP-compatible reaction mixtures contained 5% PEG as an added crowding agent because reactions with no crowding agent showed no amplification (data not shown).

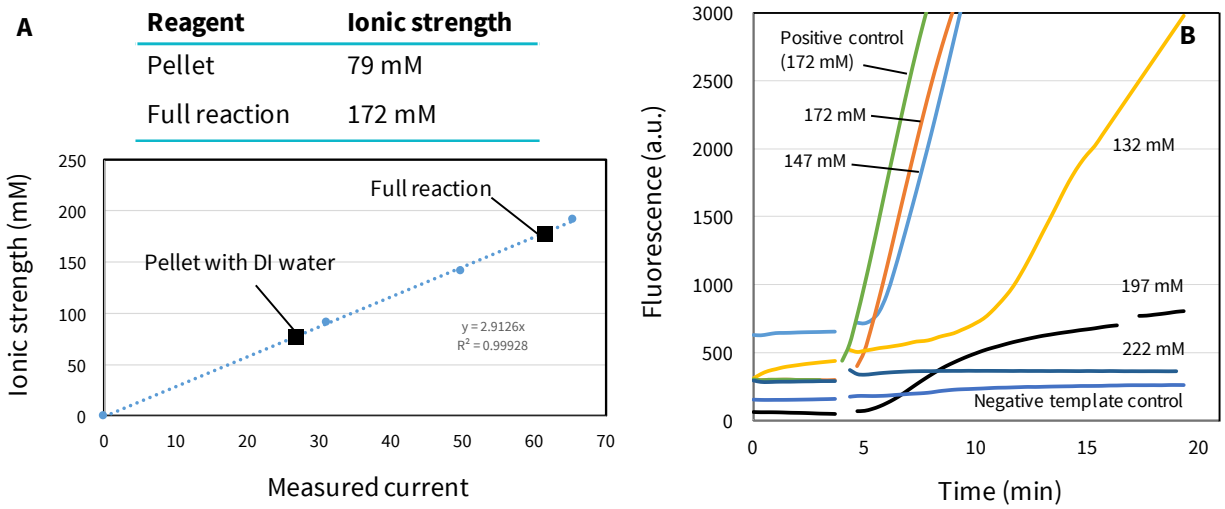



**Figure 4-6. Determining the RPA reaction conditions with ITP buffer**

(A) We sought to determine the ionic strength of the RPA reaction pellet and full reaction so that a custom ITP-compatible buffer could be used to rehydrate the pellet and give similar amplification results. The determined ionic strength of the RPA pellet rehydrated with DI water, and full reaction rehydrated with the commercial amplification buffer are shown. (B) A plot of amplification curves for standard tube reactions using various Tris-HCl LE buffers with total reaction ionic strengths near 172 mM, as well as positive and negative commercial controls. Reactions with ITP-compatible buffers between 147–172 mM ionic strength show similar efficiency compared to the commercial positive control. All ITP-compatible reaction mixtures contained 5% PEG. Ionic strength calculations were performed using the Curtipot Excel spreadsheet.

After observing the ability to perform RPA reactions in the standard tube format using our custom ITP-compatible chemistry, we sought to amplify this reaction mixture within porous substrates. We initially tested nitrocellulose and polyethersulfone (PES) porous substrates for this purpose. Nitrocellulose substrates exhibited high adsorption levels of oligonucleotides, enzymes, and target nucleic acids that partially or completely inhibited amplification reactions. PES substrates demonstrated the ability to amplify, but still exhibited adsorption of biomolecules that proved to be problematic for ITP operation. We also experienced difficulty fabricating PES substrates using our preferred laser ablation technique due to burning. Glass fiber substrates proved to be the best choice for this application due to their low adsorption, resistance to burning during ITP operation, and ease of fabrication. Figure 4-7 shows the amplification of our target nucleic acids on glass fiber. In these experiments, the custom reaction buffer (172 mM ionic strength) was used to rehydrate the RPA pellet and the reaction was pipetted directly onto the glass fiber. The reaction was then sealed in a petri dish and set on a hotplate at 41ºC. The fluorescence was monitored every second using an epifluorescence microscope. The positive reaction shows amplification after approximately 10 minutes

and continues to amplify up to 30 minutes. The negative control (no target nucleic acids) shows minimal increase in fluorescence intensity

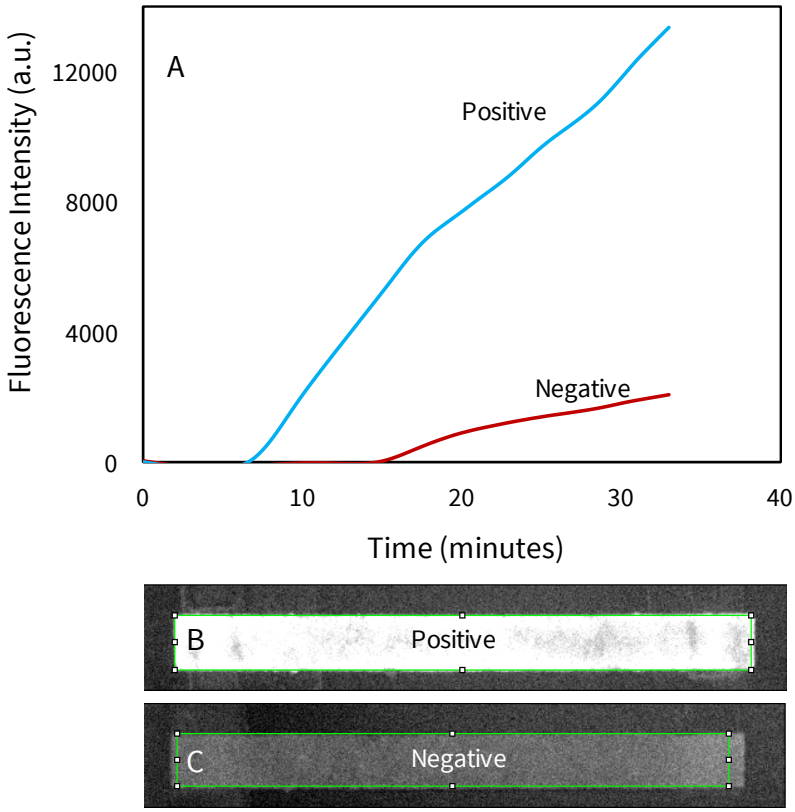



**Figure 4-7. RPA amplification on glass fiber using custom buffer**

(A) Integrated fluorescent intensity curves of positive (contains target DNA) and negative (no target DNA) amplification reactions on glass fiber substrates as a function of time. The positive reaction begins amplification around approximately 10 minutes and increases in intensity until the end of the experiment. The negative control shows a minimal increase in fluorescence. These reactions were conducted using the ITP-compatible custom buffer.

After demonstrating the ability to perform RPA on glass fiber using a custom ITP-compatible LE buffer, we designed a trailing anion that could focus all reagents required for amplification within an ITP plug. We chose to introduce the RPA reagents into the LE, as shown in Figure 4-5, because there are a number of enzymes and low mobility molecules in the RPA mixture that must be focused within the ITP plug for amplification. As shown in Chapter 3, ITP is ideal for separating high mobility nucleic acids from low mobility molecules (*e.g.* proteins) when both are present in the TE. By including the RPA reagents in the LE, we ensure some degree of interaction between the target nucleic acids and enzymes, primers, nucleotides, etc. within the ITP interface, while also increasing the potential stacking of RPA reagents.

To design a trailing anion that maximized focusing of RPA reagents within the ITP plug, we used SPRESSO simulations to quickly screen multiple TE chemistries and their affects on the focusing of nucleic acids with RPA reagents.[86] In these simulations, we placed proteins with various mobilities (5, 8, 12, and 20 x $10^{-9}$ $m^2$ $V^{-1}$ $s^{-1}$) into the LE in order to take into account molecules in the RPA pellet that may have a range of mobilities. We also added primers (40 x $10^{-9}$ $m^2$ $V^{-1}$ $s^{-1}$) to the LE and nucleic acids (40 x $10^{-9}$ $m^2$ $V^{-1}$ $s^{-1}$) into the TE. Some results from these simulations are shown in Figure 4-8 for a HEPES, serine, and glycine TE system—all aspects of the simulation are identical except for the trailing anion species. The adjusted TE zone mobility (Section 1.3.1) of the HEPES, serine, and glycine were 18.5, 9.3, and 5.3 x $10^{-9}$ $m^2$ $V^{-1}$ $s^{-1}$ respectively.

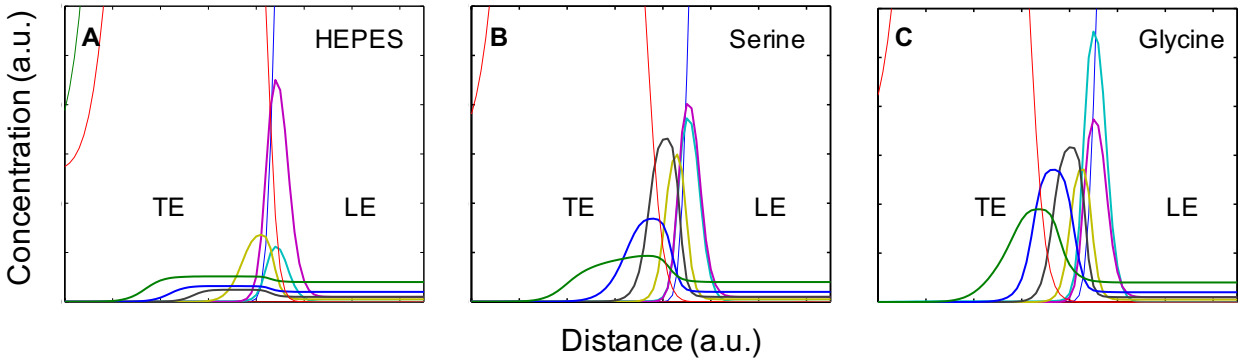



**Figure 4-8. SPRESSO simulations of TE anions for ITP-RPA focusing**

In order for RPA to amplify target nucleic acids within the ITP interface, all RPA reagents (proteins, primers, nucleotides, and other biomolecules) must focus together. These simulations show results for three different anions with differing mobility (HEPES>serine>glycine) for focusing of proteins of various mobilities with nucleic acids and primers. (A) HEPES only focused nucleic acids, primers, and the fastest protein due to its high mobility, while proteins with low mobility are not focused in the plug. (B) Serine focuses all reagents, including the slowest proteins to some extent. (C) Glycine focuses almost all molecules in the reaction mixture. We will use serine and glycine TE anions for this ITP-RPA work.

The HEPES anion used in the NAIL work has higher mobility than all species in the system except for the nucleic acids, primers, and fastest protein (20 x $10^{-9}$ $m^2$ $V^{-1}$ $s^{-1}$), resulting in minimal stacking of nucleic acids with required proteins. The serum anion system provides more stacking, but the two slowest proteins do not have as high of concentration as the other ions in the ITP zone due their lower mobility than the ATE zone mobility (5 and 8 vs. 9.3 x $10^{-9}$ $m^2$ $V^{-1}$ $s^{-1}$). However, depending on the actual mobility of the proteins in the system (these are just a range of possible values), the majority of reagents from the RPA pellet will likely focus to some extent. Finally, the slowest TE anion, glycine, shows high stacking of all of the proteins in the ITP interface. We were able to confirm the ability to focus proteins placed into the LE and nucleic

acids placed into the TE using a simple model system, as shown in Figure 4-9. In this experiment, we placed fluorescent labeled IgG proteins into the LE, Alexafluor 488 dye (similar mobility to DNA) in the TE, and used serine as the trailing anion to provide a conservative estimate of focusing DNA with proteins in the ITP interface.

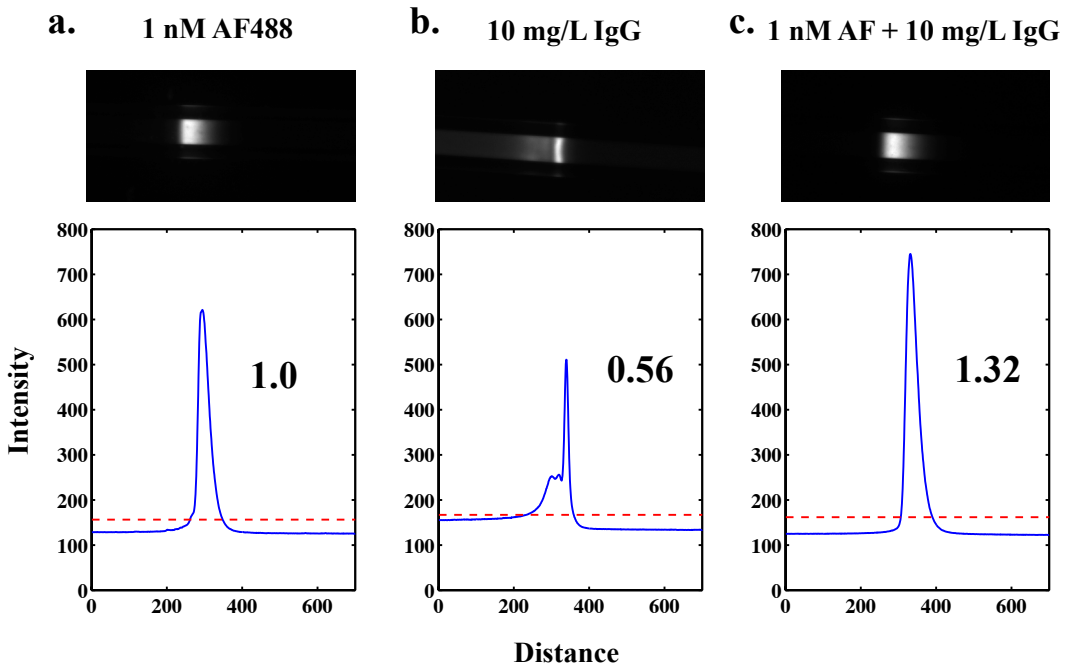



**Figure 4-9. Model experiments showing co-focusing of AF488 and IgG protein**

Model experiments showing co-focusing of (a) AlexaFluor 488 (AF488) dye, (b) fluorescent IgG proteins, and (c) both AF488 and IgG proteins into an ITP plug. AF488 simulates DNA in these experiments, while the IgG simulates the RPA mixture, so the AF488 was mixed with the TE and the IgG was mixed with the LE. The top images show analyte fluorescence within the ITP plug, with the y-averaged intensity profile plotted beneath the images (blue solid line). The dashed red line shows a baseline for integrating the area under the curve. The values next to the fluorescence curves are the total intensities under the curve normalized by the total intensity of the AF488 curve from (a). The plug in Figure 2c with both AF488 and IgG shows greater intensity than either the AF488 (Figure 2a) or IgG (Figure 2b) alone, indicating that the two analytes are both present in the plug. This result demonstrates that despite the low mobility of the IgG protein, that the AF488 and IgG can be focused together into an ITP plug. Co-focusing is crucial because all reactants must be present in the reaction volume for amplification to occur.

In the following sections, we will introduce two different ITP-RPA systems. The first system utilizes the glycine anion from Figure 4-8C due to its ability to provide high focusing for all of the reagents in the amplification reaction. In the second portion of the work, we adapted the TE anion chemistry to be compatible with blood serum. We encountered reaction inhibition with the glycine TE anion when using blood serum samples because the low glycine mobility focused all of the serum proteins into the ITP

interface, leading to reduced or inhibited amplification. As a result, we used the serine TE anion with higher mobility from Figure 4-8B in order to provide better separation of serum proteins from the reaction zone at the ITP interface, while still providing relatively high stacking of the majority of reaction components.

## 4.5 Amplifying target nucleic acids using ITP-RPA in buffer

The ability to perform an RPA reaction within an ITP plug requires two main conditions to be met: (1) all reaction components are present in the ITP plug, at least momentarily, (2) the reactants have maximal ITP stacking, and (3) the reaction environment within the ITP plug recreates the commercial tube assay environment as closely as possible. Section **Error! Reference source not found.** discusses meeting the first and second conditions by including the RPA reagents in the LE so that as the ITP plug migrates, it will interact with a constant source of reagents, as well as designing TE anions that focus as much of the reagents as possible. This section will discuss optimization of reaction and ITP chemistry to meet the conditions for the second requirement.

The most critical factors for the reaction environment are the ionic strength of the solution, the ITP concentrating effect on the reagents, and the temperature of the reaction. The difficulty in designing and optimizing parameters for ITP-RPA is that the ITP buffer system is non-linear, resulting in non-uniform spatial and temporal distributions of parameters such as ionic strength, pH, temperature, etc. between the TE and LE zones as the ITP plug migrates.[67] Addtionally, many of the parameters are confounded with each other, meaning that adjusting a parameter to better account for one aspect of the reaction may have a negative effect on another portion of the reaction. Figure 4-10 summarizes the key operating principles of the ITP-RPA reaction with some of the critical factors and relationships between the different aspects of the system.

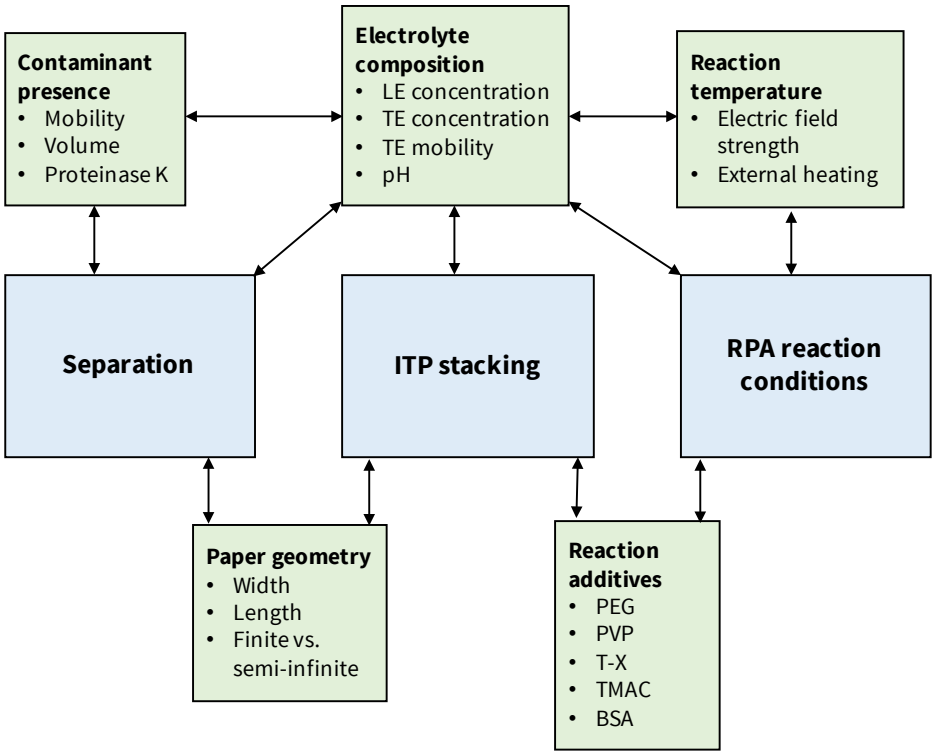



**Figure 4-10. Factors affecting the operation of ITP-RPA reactions**

The three main components that dictate the success of ITP-RPA reactions are the separation efficiency of the ITP plug from contaminants in the sample, the concentrating or stacking effect of ITP, and the RPA reaction conditions within the ITP interface. Summaries of the factors that control each of these components are shown, as well as the relationships between each factor. The design of the ITP electrolyte system has the most confounding affects with other factors in the system, but if any factor isn't sufficiently addressed, the reaction will not proceed efficiently or be completely inhibited.

Figure 4-10 shows that the important factors for designing the ITP-RPA system include the electrolyte composition, reaction temperature, reaction additives, contaminant presence, and paper geometry. We will discuss the electrolyte composition, temperature, and additives in the following subsections. Contaminant presence and paper geometry will be discussed in section 4.4 with the adaption of the system for serum and whole blood.

### 4.5.1 ITP LE chemistry design

The electrolyte or ITP buffer composition has the largest effect on the system due to its association between multiple aspects of ITP and RPA. For example, the ITP buffers set the ionic strength, pH, separation capacity, ITP stacking, and also indirectly affect heating through electric field strength. The electrolyte composition is broken down into the LE and TE, with each buffer made up of a weak acid and buffering counterion. In the LE, we aimed to achieve an ionic strength in the ITP plug that was the same as the 172

mM Tris-HCl system shown in Figure 4-6, which consisted of 50 mM HCl and 100 mM Tris (pH 8.1). We hypothesized that the reaction ionic strength in the ITP interface could be approximated as ½ the LE ionic strength because the center of the interface would be a 50:50 mixture of TE and LE (and the TE contributes negligible ionic strength). Figure 4-11 shows amplification curves for an ITP-RPA reaction with 80, 100, 120, and 140 mM HCl and a corresponding 1:2 ratio of HCl:Tris for each solution. The 100/200 mM HCl/Tris system showed the fastest and highest intensity amplification for a positive reaction containing target nucleic acid. The 140/280 and 120/240 mM systems showed less amplification, with the 140/280 mM LE having slightly higher intensity after 10 minutes. The 80/160 mM system demonstrate the lowest amplification intensity, which may have been due to poor stacking and/or too low of ionic strength in the ITP plug. All negative template controls (NTCs) showed essentially no amplification.

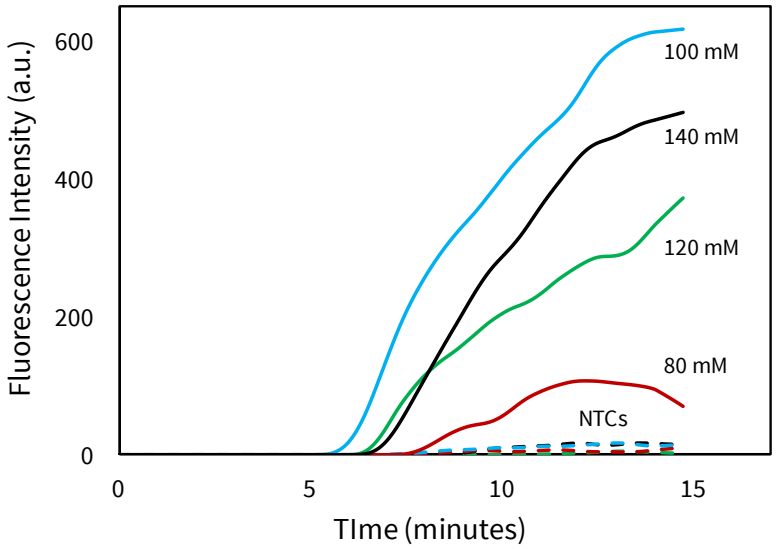



**Figure 4-11. Effect of LE concentration on ITP-RPA reactions**

The LE concentration plays an important role in determining the reaction ionic strength and the focusing of the ITP plug. Amplification curves showing various HCl concentrations (with a corresponding 1:2 HCl:Tris ratio) are shown. 100/200 mM HCl/Tris provides the fastest and highest intensity amplification. 80 mM showed the worst amplification, potentially due to poor stacking and/or low ionic strength. All negative template controls showed no amplification.

### 4.5.2 ITP TE chemistry design

In general, we found that lower TE concentrations led to better amplification. We believe this effect is due to the improved stacking that lower TE concentrations offer, as discussed in section 1.3.2, where the concentrating effect of ITP was shown to be inversely related to TE conductivity. Higher stacking likely

leads to improved localization and higher concentrations of all reaction components within the ITP interface that may lead to improved amplification reaction kinetics of primers and proteins, as other work has shown.[22,71,72] Alternatively, it may simply ensure all reaction components are present for the reaction within a specified region.

The limiting factor for decreasing the TE concentration was electroosmotic flow (EOF) that was induced due to large conductivity differences between the LE and TE, as discussed in section 1.3.3.[89] Optimizing the TE consisted of setting the LE to the 100/200 mM HCl/Tris concentration and then decreasing the glycine concentration until EOF severely dispersed the reaction. Figure 4-12A shows sequential experimental images of a positive reaction with the optimized glycine TE solution and Figure 4-12B shows a spatiotemporal plot of the entire reaction from (A). During the first 250 seconds, the plug rapidly migrates approximately half the distance of the glass fiber length. However, as the TE migrates onto the substrate, EOF increases and between 250 and 600 seconds, the ITP plug essentially does not migrate due to EOF. From 600 seconds until the end of the experiment, the plug begins to migrate again, which we believe may result from evaporation in the reservoirs that reduces the amount of bulk liquid that can be transported across the strip by EOF. Decreasing the TE concentration further causes severe drying and EOF that negatively impacted the reaction.

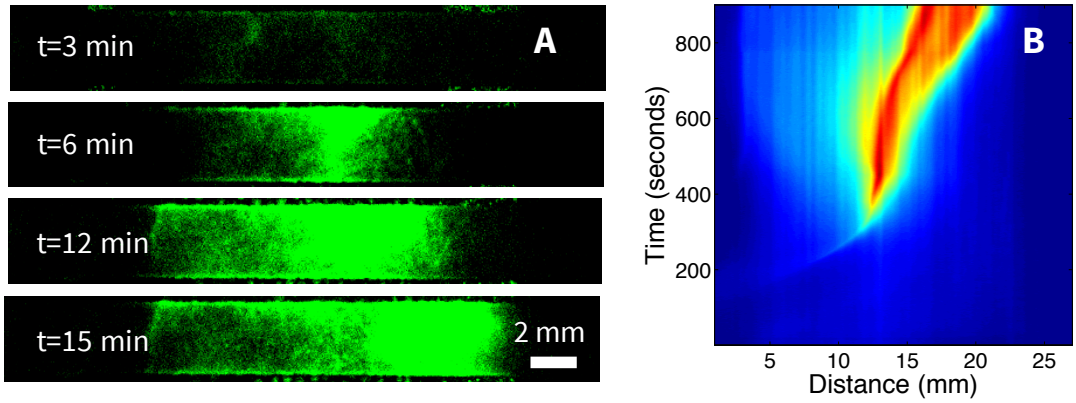



**Figure 4-12. Determining TE concentration with EOF constraint**

We set the LE concentration according to Figure 4-11. ***Effect of LE concentration on ITP-RPA reactions*** and then decreased the TE concentration until EOF severely impacted the reaction. (A) A series of fluorescent experimental images showing the amplification reaction with optimized TE chemistry. The plug rapidly migrates to the middle of the strip, but remains stationary for the majority of the experiment due to EOF. However, EOF does not significantly affect amplification, as fluorescent intensity increases while stationary, but would cause too great of dispersion with further TE concentration decreases. (B) Spatiotemporal map of the entire reaction that shows the migration of the plug in more detail.

### 4.5.3 ITP buffer additives

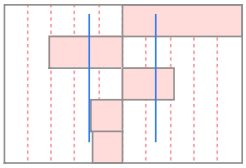
A number of additives significantly affect the operation of ITP. In order to determine the most critical additives and their relatively optimal levels, we performed fractional factorial screening experiments. Following the design of experiments and gathering of experimental data, we performed analysis of variance (ANOVA) analysis in order to determine which additives have a significant effect on reaction efficiency. Table 4-1 details results of additives that most strongly affect amplification efficiency, with PEG being the most significant additive. PEG is often used as a crowding agent and similar to the CarboWax crowding agent included in the commercial reaction buffer. This result was surprising because we believed ITP stacking would eliminate the need for a crowding agent to be included in the reaction. However, this result lends support towards the hypothesis that the ITP-RPA reaction should aim to mimic the tube reaction as much as possible. BSA blocking of the glass fiber surface also increases amplification, likely due to reduced adsorption of molecules on the surface. TMAC inclusion reduces non-specific primer hybridization, PVP helps to reduce EO flow by coating the surface of the glass fiber, and T-X is included to improve wetting.

**Table 4-1. Additive screening for ITP-RPA**

We performed screening for various additives that affect the amplification of the ITP-RPA reaction. The data below shows the ANOVA results for the additives that most strongly affect the reaction efficiency. The estimate gives the effect size for change fluorescence intensity (a.u.), while the prob>|t| gives the p-value for significance. All additives shown here are included in all experimental data.

**Sorted Parameter Estimates**

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| PEG(0,2.5) | 398.375 | 35.17202 | 11.33 | 0.0015* |
| BSA block[no] | -235.05 | 33.36711 | -7.04 | 0.0059* |
| TMAC(0,5) | 171.375 | 35.17202 | 4.87 | 0.0165* |
| BSA block[no]*PVP | -104.125 | 35.17202 | -2.96 | 0.0595 |
| T-X(0.01,0.1) | -98.125 | 35.17202 | -2.79 | 0.0684 |



Because PEG was shown to be the most critical additive, we studied its concentration level in more detail. We varied PEG concentrations from 0 to 7.5% w/v in order to determine the optimal level of PEG. Figure 4-13 shows that concentrations between 2.5–5% exhibit similar amplification efficiency, while higher and lower concentrations provide 2–6x less fluorescence due to non-ideal amplification conditions. Low concentrations may not mimic the tube reaction conditions closely enough due to the lack of crowding agent. At high levels, e.g. 7.5% PEG, the reaction mixture becomes too viscous and the ITP plug does not properly form and migrate, severely inhibiting the reaction. We found that 2.5% provided the best ITP

operation (plug shape, migration, etc.), while still enhancing the amplification. As a result, we chose 2.5% w/v as the PEG concentration for the remainder of the work.

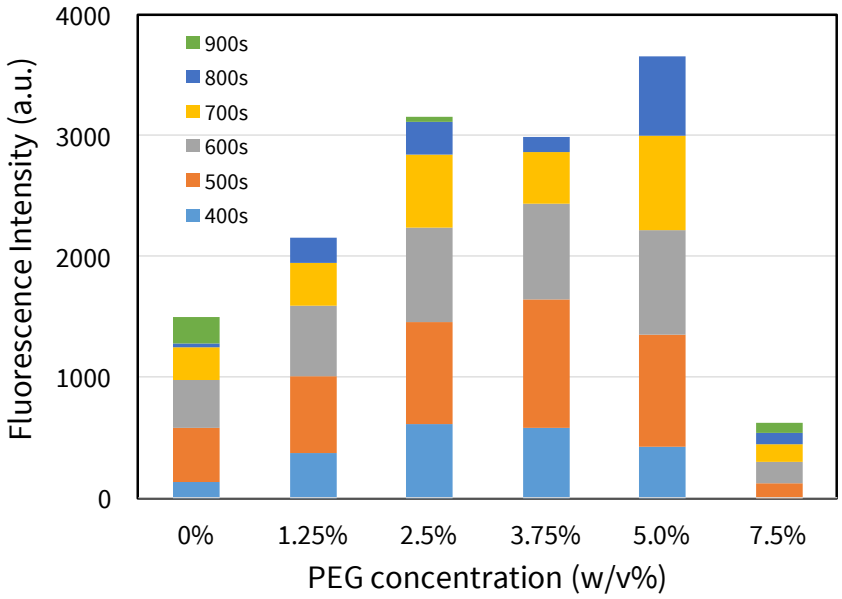



**Figure 4-13. Affect of PEG concentration on ITP-RPA**

The affect of variable levels of PEG on the ITP-RPA reaction intensity. The amount of intensity increase for 100 second increments between 400–900 seconds is shown for each concentration. High and low amounts of PEG (>5% and <1.25%) do not provide as high of intensity output as concentrations between 2.5–5%. We chose 2.5% due to its minimal effects on ITP, while still providing high intensity amplification.

### 4.5.4 Internal Joule heating and external heating

When a constant electric field is applied, a significant amount of current is passed through the glass fiber substrates during ITP-RPA reactions (approximately 1–3.5 mA depending on whether the strip is filled with TE or LE). The large current passed through the system can lead to a significant amount of Joule heating if the generated heat is not properly dissipated from the system. Previous work from the group using nitrocellulose found that currents above 1–1.5 mA resulting in the nitrocellulose substrates burning, indicating that a significant portion of the heat is not dissipated from the porous substrate, even in an open system. Glass fiber does not burn at high temperatures like nitrocellulose, so the internal temperature rise from Joule heating may be able to be leveraged to provide heat for the RPA reaction. We used infrared (IR) thermal imaging with the iPhone Flir© camera attachment in order to approximate the spatial and temporal variation of heating due to electrical current application during ITP-RPA.

Figure 4-14 shows experiments where the voltage is set to 100 V and 150 V, and the corresponding temperature at the ITP interface is measured and shown. Under these conditions, the average temperature of the ITP zone as it migrates across the paper strip is approximately 30–34ºC for 100 V and 40–42ºC for 200 V. Different temperature ranges can be achieved depending on the applied electric field strength, the composition of the ITP electrolytes, and the dimensions of the glass fiber strip.

This internal heating may be used as the heating source for an amplification reaction so that external heating is not required. The main difficulty with controlling this heating is the spatial non-uniformity across the substrate due to the differences in electric field between the TE and LE (the TE side is hotter than the LE side due to higher electric field). Further, the temperature varies with time because the current continually drops as the TE migrates onto the paper. Methods to improve these issues may be using a thermally conductive layer under the substrate that better distributes the generated heat and using a source control that maintains constant power through the system, as opposed to constant current or voltage. Applying these methods are outside the scope of this work, but we did demonstrate that we could amplify reactions using ITP-RPA starting from room temperature (~20ºC) using only internal heating (data not shown). However, to maintain better reaction consistency, we generally set the reaction on a hotplate set to 34ºC, with the additional heat supplied by Joule heating. Future work may be able to adapt the reaction to not require any external heating while maintaining high consistency reactions.

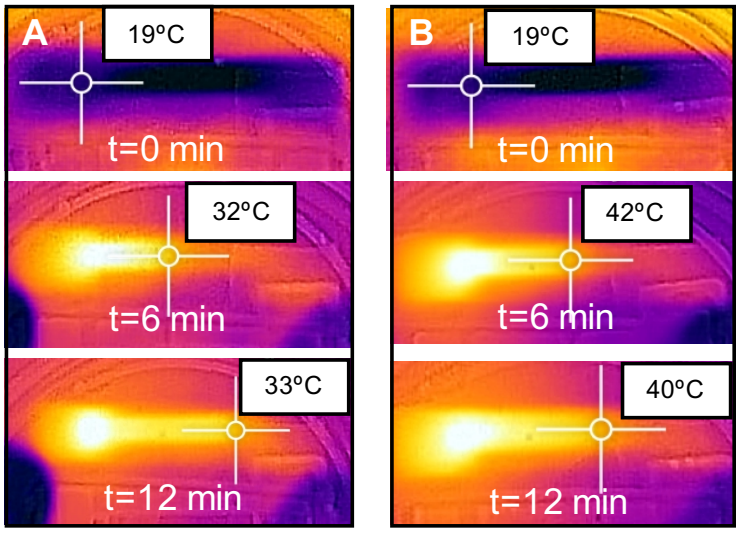



**Figure 4-14. Internal heat generation due to Joule heating**

The temperature increase within the substrate due to Joule heating is dependent upon the electric field strength, electrolyte solution composition, and the substrate geometry. Due to the conductivity difference between the TE and LE electrolytes, the substrate experiences non-uniform heating. The location of the ITP plug was tracked using a colorimetric dye and the approximate temperatures of the ITP zone on the corresponding IR thermal images are shown. The internal temperature generation may be able to be controlled enough to provide the required heating for RPA reactions (38–41ºC). The IR images from A show a reaction with 100 V applied, while the IR images from (B) show an identical reaction with 150 V applied. The higher electric field in (B) results in a larger temperature rise than in the experiment for (A).

## 4.6 Adapting the ITP-RPA reaction for serum and whole blood samples

Providing HIV-1 VL diagnostic information at the POC necessitates that the diagnostic device provides results from unprocessed whole blood samples. We aimed to develop an ITP chemistry that was capable of extracting and amplifying target nucleic acids from blood serum, with the intention of using a filtration membrane to passively generate plasma (serum with clotting factors) from whole blood. Previous work has shown that >80% of plasma can be generated from whole blood with near 100% recovery of virus within 5 minutes using a passive filtration membrane.[206,207] Even with filtering of the whole blood inputs, using plasma or serum as the sample required a re-design of the ITP TE anion chemistry in order to separate proteins and other biomolecules in the serum sample from the ITP-RPA reaction zone.

As mentioned in section **Error! Reference source not found.**, we used glycine for proof-of-concept in buffer, but had to adapt the chemistry to serine to separate serum proteins. This modification required precise adjustment of TE mobility because it needed to have higher mobility than serum proteins, yet remain slow enough to focus all required reagents for RPA reactions. Figure 4-15A shows the approximate

difference in mobility between the serine and glycine in the adjusted TE zone. The glycine mobility is lower than the approximate mobility of serum proteins, leading to contaminants focusing in the ITP reaction zone. Using serine as the TE allows for the majority or all of the serum proteins to be separated from the reaction zone during ITP-RPA operation. Figure 4-15B shows a simulation of the designed ITP system that focuses all reagents and nucleic acids required for RT-RPA in the ITP reaction zone between the LE and TE, while serum proteins, cations, and other biomolecules are separated out.[86] The success of the ITP system is determined by the separation distance between the reaction plug and the serum proteins, and the concentrating effect on reaction reagents.

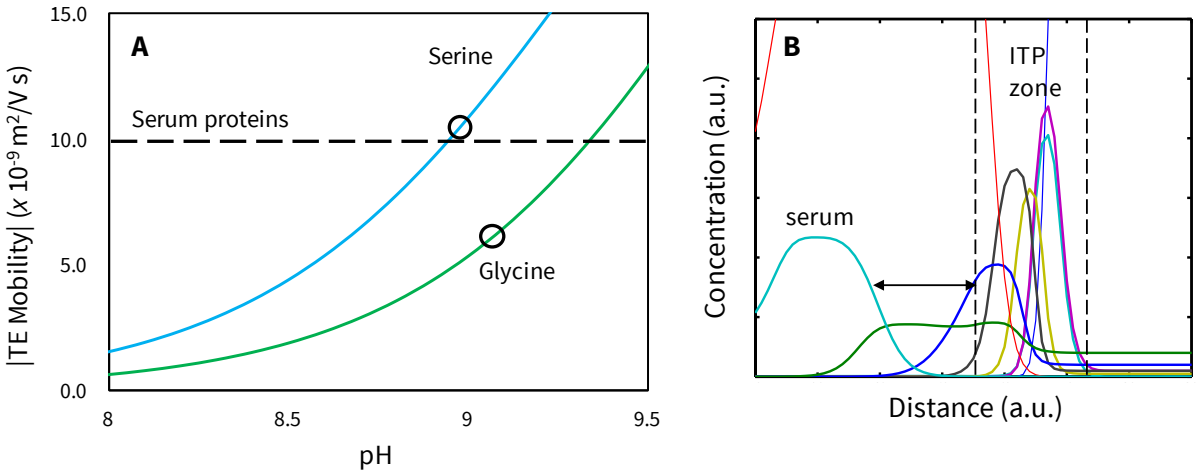



**Figure 4-15. Design of TE anion mobility for serum separation**

(A) The plot shows the absolute value of mobility of the serine and glycine TE anions in the adjusted TE zone. We adjusted the TE anion to separate serum proteins from the ITP-RPA reaction by using serine at an adjusted TE mobility that was slightly greater than the approximate mobility of serum proteins. The amplification reactions with buffer samples used glycine in order to provide high focusing of all reagents, but also led to focusing of contaminants from the sample. (B) A SPRESSO simulation showing separation from the approximate serum proteins using the serine chemistry, while maintaining focusing of other reagents in the ITP zone.

Figure 4-16A shows a micrograph of the separation of serum proteins from the reaction plug via ITP extraction. We also evaluated the inclusion of proteinase K (PK) protein digestion into the system to completely eliminate any potential confounders (*e.g.* proteins binding to target, RNAse presence). The result is shown in Figure 4-16B, where the protein band from the sample has been digested, while the amplification reaction still proceeds. Furthermore, the ITP system prevents PK from entering the ITP-RT-RPA reaction, ensuring the integrity of enzymes required for RT-RPA.

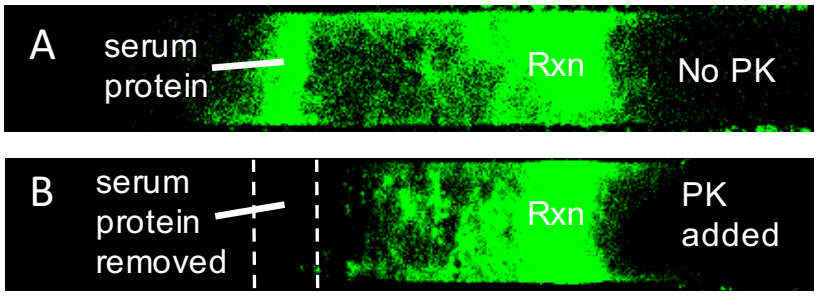



**Figure 4-16. Separation and removal of serum proteins**

(A) Experimental fluorescence image showing ITP separation of nucleic acids and RPA reagents from serum proteins using the modified serine TE anion chemistry (C) Despite the separation, we aimed to eliminate any confounding affects of serum proteins, so we added PK to the rsample pad to digest serum proteins, while not affecting the RPA reaction.

We also aimed to show proof-of-concept that this ITP-RPA reaction is compatible with passively filtered whole blood samples. Spiked whole blood was added directly to a commercial Vivid© blood separation filter and passively fractionated to create plasma on the glass fiber surface.[206] The glass fiber surface contains desiccated Triton-X100 to promote better wetting of plasma on the glass fiber surface. The proteinase K for protein digestion is also desiccated on the glass fiber surface to enable automated sample preparation after the plasma reached the glass fiber surface. Figure 4-17A shows an image of the blood separation filter saturated with 30 µL of human whole blood. After 3 minutes of passive microfiltration, the fractionated plasma was collected on the glass fiber (Figure 4-17B). ITP-RT-RPA can be run directly from this sample without further processing using the developed ITP chemistry. Figure 4-17C compares ITP-RT-RPA fluorescent outputs from serum and whole blood under identical experimental conditions. The similarity of the amplification curves for serum and fractionated whole blood demonstrate that the ITP-RT-RPA system can amplify target nucleic acid targets in unprocessed whole blood from a finger or heel prick within 15–20 minutes (3 minutes filtering + 10-minute separation and amplification) using an integrated filter.

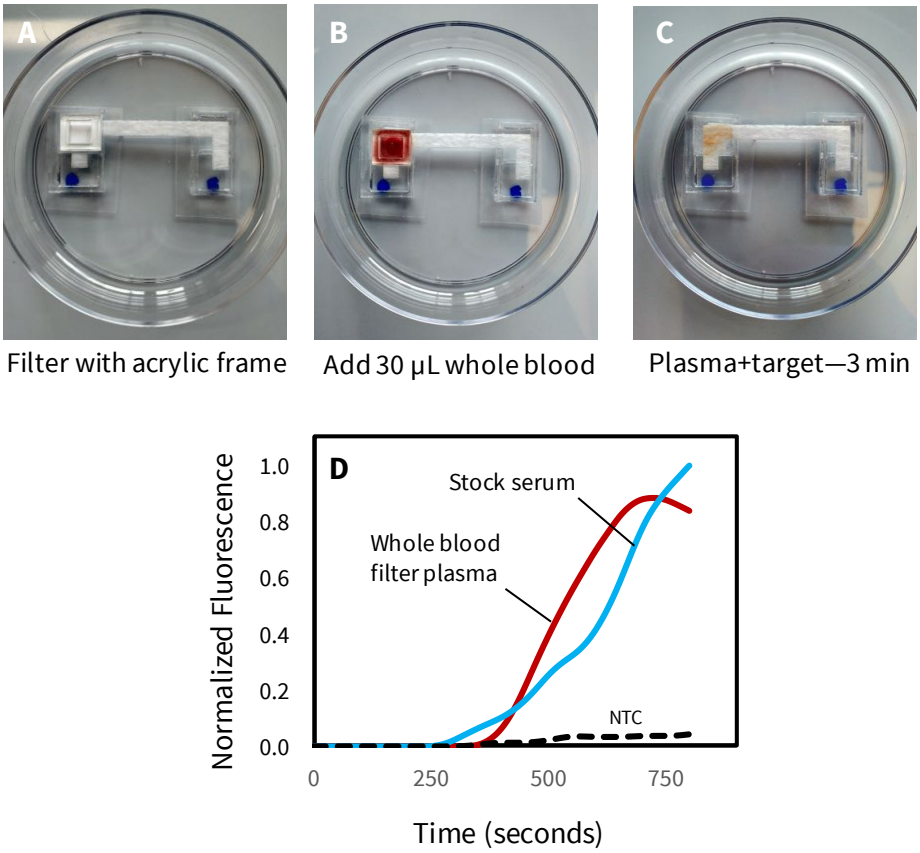



**Figure 4-17. Whole blood filtering for plasma sample generation**

(A) A plasma separation membrane is used to fractionate whole blood and create a target-laden plasma sample of the surface of the glass fiber substrate via capillary forces. (B,C) Application of whole blood to the integrated blood separation filter passively generates plasma from whole blood within 3 minutes. (D) ITP-RT-RPA of fractioned plasma from whole blood amplifies similarly to serum, showing proof-of-concept that whole blood may be used as the test sample.

## 4.7 Analytical assessment of the ITP-RPA reaction in serum samples

We obtained preliminary linearity and LoD data of the ITP-RPA system using spiked serum samples ranging from $10^3$–$10^8$ copies of DNA per milliliter of serum (cp/mL), as shown in Figure 4-18. In these experiments, nucleic acids were spiked into 20 µL of human serum, with the added copy numbers ranging from 200–2E6 copies in the 20 µL sample (multiplying the added copy number by 50 gives the amount in units of cp/mL serum). These serum samples were added directly the the glass fiber surface and run using the developed serine ITP-RPA chemistry. Figure 4-18A shows the average integrated intensity curves for each dilution. All dilutions exhibit greater fluorescent intensity than the negative controls, which produce essentially no fluorescence using the algorithm described in Chapter 2. Target concentrations above 1E7 cp/mL begin to asymptote, meaning these values likely cannot be differentiated and quantified. However, VL above 1E7

cp/mL are likely not clinically relevant in the majority of cases and the quantification of these values above this concentration is likely unnecessary.

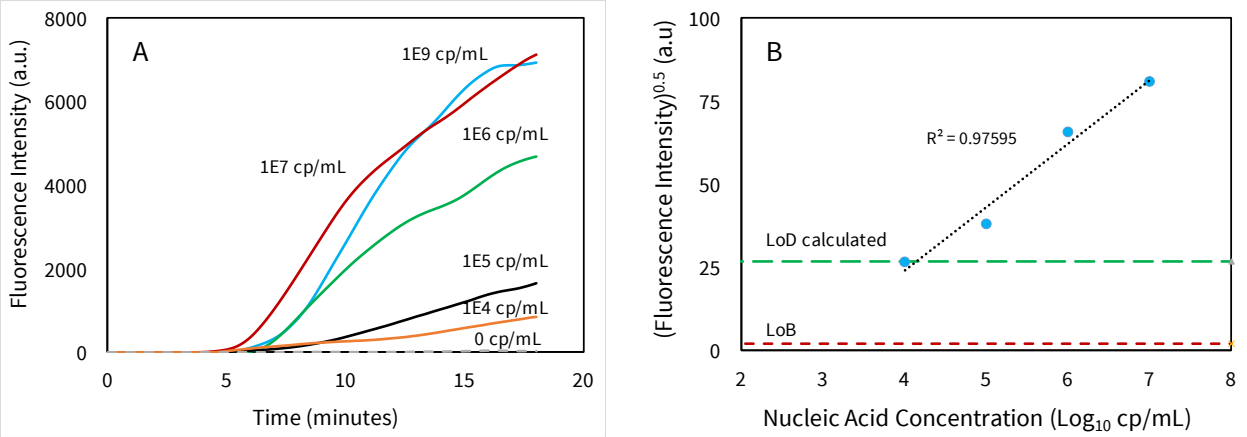



**Figure 4-18. Linearity and LoD of ITP-RPA dilution series**

(A) Integrated intensities of ITP-RPA experiments vs. time for a log dilution of DNA concentrations spiked into serum. The differences in endpoint intensity, slope, etc. allow for an algorithm to be developed for quantitative or semi-quantitative determination of initial copy number of viral load. (B) A plot of the square root of endpoint intensity at 15 minutes from (A) versus initial log nucleic acid concentration. The data shows relatively high linearity over four orders of magnitude and an LoD of 10,000 cp/mL of serum (green dashed line). The red dashed line shows the limit of blank (LoB).

The data in Figure 4-18A shows decreasing intensity with decreasing copy number, as well as differences in slope and time to amplification between dilutions. These differences can be used to create an algorithm to differentiate between initial copy numbers for potential VL monitoring applications. Figure 4-18B shows an example using a basic algorithm to create a fluorescent intensity vs. nucleic acid concentration linear calibration plot. For this example, we simply used the square root of the reaction endpoint intensity at 15 minutes. This method was chosen because it gave relatively linear results for the obtained data—future work will likely develop more sophisticated algorithms. The data in Figure 4-18B shows reasonable linearity ($R^2$=0.976) over four orders of magnitude that can be used to quantify the initial VL of the sample from a given intensity response using ITP-RPA. We also used the data from Figure 4-18B to show the current LoB and LoD of the test, which are labeled and shown by the red and green dashed lines respectively. The LoB of the test is extremely low because the thresholding method used to analyze the experimental images gives fluorescent intensities (and corresponding standard deviations) of essentially zero for negative samples. The LoD was calculated as $LoD = LoB + 1.645\sigma_S$, with the weighted sample standard deviations from 1E4 and 1E5 cp/mL (n=8).[109] The samples showed relatively high standard deviation ($\sigma_S = 20.7$),

resulting in an LoD that is significantly higher than the LoB. However, the calculated LoD corresponds almost exactly to 10,000 cp/mL, which can reasonably be said to deviate from blank samples according to the the data from Figure 4-18A,B.

As mentioned at the beginning of this chapter, the VL clinical threshold for adjusting ART treatment is 1,000 cp/mL (with a range of approximately 500–5,000 cp/mL).[200] Our determined LoD is currently one order of magnitude higher than the required clinical threshold. Chapter 6 conclusions and recommendations will discuss future work in more detail, but there are three methods that can be pursued to further decrease this LoD: increase sample volume, reduce number of required copies, and improve repeatability. The 10,000 cp/mL LoD used 20 µL of serum and 200 copies of target. Increasing the added sample volume to 50 or 100 µL (the upper limit for a finger prick) would reduce the LoD to between 2,000–4,000 cp/mL. Increasing volume will require adjusting the paper geometry and potentially electrolyte concentrations for great separation capacity. Decreasing the required number of copies to 50 or 100 will also reduce the LoD by 2 to 4-fold. For example, if 100 µL of serum is used as the sample volume and only 50 copies are required, the given LoD would be 500 cp/mL and would be able to differentiate samples around the 1000 cp/mL cutoff. The HIV-1 assay developed by Boyle and co-workers has shown as low as 5 copy/reaction LoD in a standard tube reaction, so requiring 50 copy/reaction in the ITP-RPA format is likely achievable, assuming >50% extraction efficiency.[99,204] Finally, improving repeatability of the assay at low copy numbers will decrease the calculated LoD through less sample standard deviation. Increasing repeatability will require perturbation analysis of the ITP chemistry (particularly LE, TE, and sample), as well as the heating due to electric field and/or external heat.

The LOD is correlated to the diagnostic sensitivity of the assay, but a diagnostic test must also be specific to the target of interest. Boyle and co-workers have performed extensive specificity studies of the assay that we intend to use for HIV-1 VL monitoring in future work.[204] They spiked 50,000 genome equivalents of non HIV-1 pathogens or commensal flora into their tube assay and found that none of the non-specific targets amplified.[204] Table 4-2 gives a list of the tested pathogens. The primers in nucleic acid amplification tests provide target specificity, but we wanted to ensure that the unique ITP-RPA environment did not lead to unusual false positives compared to the tube reaction. We spiked E. coli DNA and control DNA from the RPA reaction kit into and ITP-RPA reaction with the primers used throughout this work and

found no non-specific amplification. We also ran a test with very high copy numbers (1E9 copies) and the probe only (*i.e.* no primers) to ensure our results were simply not hybridization of the probe to the target. These experiments also exhibited no amplification, as shown in Figure 4-19.

**Table 4-2. RPA specificity testing in tube format for PATH assay**

List of organisms tested with the PATH RPA assay for HIV-1. All non-target showed no amplification, indicating the developed primers are specific for HIV-1.

| Pathogen | Reactive | Pathogen | Reactive |
|---|---|---|---|
| Adenovirus 1 | No | HSV type 2 | No |
| Adenovirus 2 | No | HPV type 16 | No |
| Epstein Barr virus | No | HPV type 18 | No |
| Hepatitis A virus | No | Varicella zoster virus | No |
| Hepatitis B virus | No | SIV | No |
| Hepatitis C virus | No | *Candida albicans* | No |
| HIV-2 | No | *Mycobacterium smegmatis* | No |
| Human cytomegalovirus | No | *Staphylococcus aureus* | No |
| Human herpes simplex virus 1 | No | *Staphylococcus epidermidis* | No |

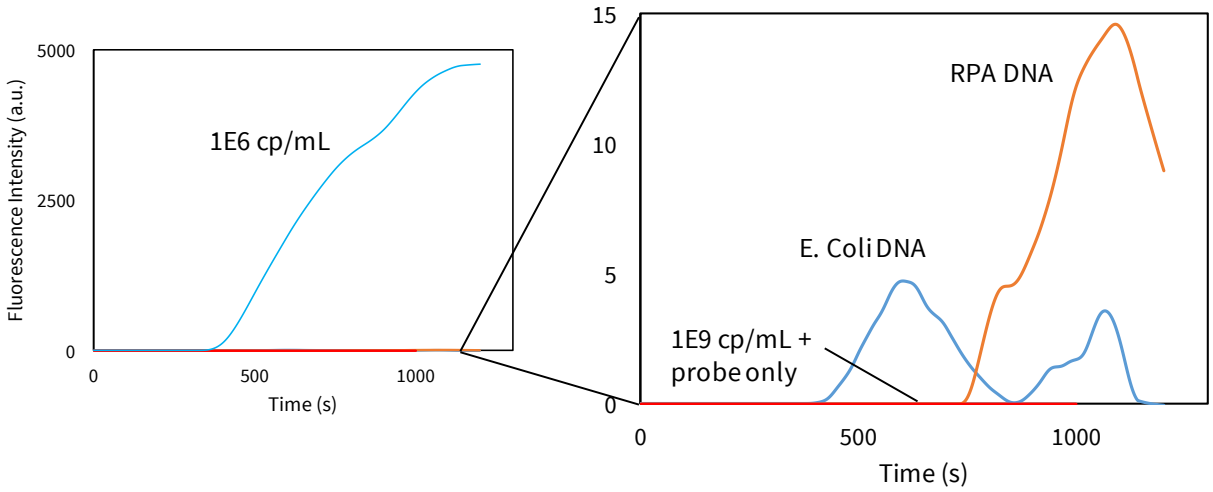



**Figure 4-19. Specificity testing for ITP-RPA reactions**

We added two non-target DNA's into the ITP-RPA reactions with the primers used throughout this work—neither showed any amplification. We also added high target copy numbers (1E9 copies) into the reaction with only the fluorescent probe (no primers) to demonstrate that the reaction requires all components to amplify. Further specificity analysis is required, but preliminary results show ITP-RPA does not exhibit different specificity than the tube reaction.

# Chapter 5: Conclusions and Recommendations

## 5.1 NAIL conclusions and recommendations

The NAIL device demonstrated the ability to use ITP to extract and concentrate nucleic acids from a complex sample within an integrated microchip. It had an LoD of 1000 CFU/mL, which was a two orders of magnitude improvement compared to tube-LAMP. The LoD was also comparable to many reported molecular and immunological based methods, but NAIL used integrated sample extraction and amplification within a compact, portable device. Altogether, the lysis, extraction, and amplification process required approximately one hour (15 minutes for lysis, 4 minutes for ITP, 45 minutes for LAMP, and 1 minute for detection) and five steps: adding sample to TE buffer, adding LE and TE-sample mixture to NAIL, applying an electric field for ITP, heating the chip, and taking an image of the chip with a mobile phone.

While the NAIL device showed some of the potential for using ITP and isothermal amplification as a general strategy to create low-cost nucleic acid diagnostic tools, it suffered from a few drawbacks that affected its performance in the laboratory that would preclude its use in clinical applications in its current form. The main drawback was that the NAIL device did not exhibit sufficient repeatability—approximately 40% of the experiments would fail due to practical operational issues such as poor channel filling (*e.g.* bubble formation), the capillary valve prematurely bursting, or the reaction chamber not filling uniformly during pumping. A lot of these issues (*e.g.* channel filling, valve bursting) were likely due to fabrication issues that resulted from wall roughness (bubble formation) or rounded fabrication patterns (valve bursting), but highlight the precision that's required to create a robust NAAT. Steps could have (and still could be) taken to improve aspects of this device, but ultimately, our method to improve the NAIL device was to eliminate pump, valve, and loading complexities by creating the ITP-RPA device presented in Chapter 4.

The NAIL device did represent some important contributions and learning experiences. It showed ITP working with complex whole milk samples, which is the first time ITP had been shown to work with a food-based samples. It also showed the ability to use a basic imaging setup to detect the efficacy of amplification reactions in a microchip with a metal-dye indicator. Finally, it was the first work to show the potential benefits of using ITP and isothermal amplification to create POC NAAT. Specifically, this setup allows for the creation of devices with small footprints and low equipment requirements because the extraction is performed without multiple wash steps, waste reservoirs, pumps, etc. and the amplification can be

performed using a simple resistive heater. Finally, the lessons learned during this work—device filling, pump, valve failure, etc.—led to the ideation and development of ITP-RPA.

## 5.2 ITP-RPA conclusions and recommendations

ITP-RPA offers an intriguing strategy to create point-of-care diagnostics. Simultaneously extracting and amplifying nucleic acids on a strip of glass fiber substrate achieves two of the most challenging steps of nucleic acid amplification in a manner that does not require mechanical parts. The potential to also eliminate the need for external heating by using internal Joule heating to achieve necessary reaction temperatures further reduces the equipment requirements of the test.

In Chapter 4, we described the design of the ITP electrolyte chemistry to focus the RPA reagents within the ITP interface. We also determined the important parameters that affect the operation of the reaction, such as LE concentration, TE concentration, and additives such as PEG. Our current target application for the ITP-RPA device is viral load monitoring for HIV-1 at the point-of-care using a drop of whole blood. The performed LoD analysis showed the ability to detect approximately 10,000 copies of nucleic acids in a blood serum sample. Further, we demonstrated the ability to integrate a Vivid plasma separation filter to passively create plasma on the glass fiber surface from an unprocessed whole blood sample.

The presented work is an important first step towards demonstrating proof-of-concept of ITP-RPA reactions in general, as well as applying this technology towards POC HIV-1 VL monitoring. Future work will require additional improvements in order to achieve this goal. For example, the LoD will need to be reduced by over an order of magnitude so that the device provides clinically relevant results in accordance with the 1,000 copy/mL threshold established by the World Health Organization for monitoring ART treatment. Achieving this LoD will require increasing the sample volume to between 50–100 µL of whole blood, as well as improving the reaction sensitivity to low copy numbers (e.g. 25–50 copies required per reaction). These objectives can likely be achieved by further optimization of ITP chemistry, particularly the TE separation and concentration, as well as using a finite injection scheme to maximize the extraction efficiency of ITP.

Once the clinically relevant LoD has been achieved, the reaction will need to be demonstrated using HIV-1 virions of various subtypes. This step will include using a lysis agent to lyse the virions and release target nucleic acids into the glass fiber substrate with ITP solutions. Our current design with dried Triton-X100 and proteinase K on the glass fiber surface that is wet by plasma containing target was designed with

virion lysis in mind, lowering the barrier towards achieving this objective. We have demonstrated linearity for potential quantification using a single type of nucleic acids, but sequence mismatches between subtypes present challenges with quantification due to changes in reaction efficiency. We have preliminary plans to use multiple metrics associated with the amplification curves to create a more sophisticated algorithm for inter sub-type quantification. Finally, the operation of all of these steps will require high repeatability in whole blood samples.

Another possibility for simultaneous ITP extraction and amplification is using alternative isothermal amplification strategies within the ITP interface. As mentioned in Section 1.5, many different isothermal amplification chemistries have been created with differences based on enzymes, primers, temperatures, etc. The majority of these reactions require temperatures around 65ºC, which would necessitate using external heating. This expansion to other chemistries would open up a wide diversity of targets and primer sets for diagnosis applications.

### 5.3 Analytical and clinical metrics conclusions and recommendations

The microfluidics community has leveraged microscale technologies to create a wide-array of unique healthcare assays and devices that have the potential to improve patient care. We believe that more stringent use of analytical statistics recommended by standards organizations (*e.g.* CLSI, FDA, Westgard) to characterize device performance, and recognition of clinical statistics will improve the overall quality and translation of microfluidic health research. Specific areas where we see potential for improvement include: (1) detailed use of analytical performance metrics, (2) moving towards representative samples, and (3) establishing the connection between laboratory and clinical performance.

**Detailed use of analytical performance metrics:** Many papers reporting microfluidic healthcare technologies do not use thorough statistical terminology and definitions to evaluate and describe device performance, such as the metrics presented in the analytical metrics section. Further, when these metrics are said to be determined, full statistical definitions and experimental protocols are often not given in detail.

For example, a search for “analytical sensitivity” in Lab on a Chip articles shows 33 articles that use this term (searched Feb. 2016). The majority of these articles use “analytical sensitivity” to describe detection limits that would be more suitably characterized as LoD or to describe analytical accuracy that would be

better described by trueness (a search for trueness in Lab on a Chip articles produced no results). Only a small handful use “analytical sensitivity” to describe the calibration slope with regards to quantification.

A search for “limit of detection” in Lab on a Chip articles shows 472 articles that use this term (searched Feb. 2016). We selected 25 of these articles related to biomolecule detection at random and found that over 50% of the papers that state an LoD for the device do not report the statistical and/or experimental methods used to determine the reported LoDs. None of the papers used the limit of blank / limit of detection method outlined by the CLSI.

While these examples are brief scans of the literature, and many articles related to microfluidic health technologies do use high experimental and statistical rigor, we believe that more recognition and use of the metrics laid out in section 1 will improve the overall quality and consistency of publications. A few examples include: clearly defining the analytical sensitivity, reportable range, and quantitative resolution for quantitative assays; using the combined statistical and empirical approach to determine LoD with reported answers given in concentration units where possible (*e.g.* copies/mL, CFU/mL, ng/mL); and providing detailed protocols used to study an assay’s precision and/or trueness (e.g. within day replicates, between day replicates, sample matrices used).[109,110,117,119,126,145] For all of these metrics, the experimental protocols (e.g. replicates, type of sample used, statistical methods and calculations) should be clearly reported, and ideally align with guidelines recommended by standards organization.[109,110,117,119,126,145]

**Moving towards representative samples:** Proof-of-concept studies are an important first-step in understanding and demonstrating the physics, chemistry, and overall feasibility of a new assay or device. Once feasibility has been demonstrated however, we recommend a clear focus on establishing performance of new technologies using complex matrices, reference materials, or patient samples. By focusing on more representative samples, researchers will more quickly discover areas for improvements when their devices are challenged with the complexities and heterogeneity of reference or patient samples.

Determination of analytical parameters (e.g. LoD, analytical selectivity, precision) with representative samples will provide a better characterization of actual test performance, as opposed to the possibly overstated performance when using more idealized samples. In either case, the strengths and weaknesses

of experimental protocols used should be acknowledged and discussed. Additionally, preliminary diagnostic sensitivity and specificity can be determined using a small amount of patient samples, which allows for researchers to better understand the potential clinical impact (predictive values and likelihood ratios) of their devices early in development. In addition to (or in place of) diagnostic sensitivity and specificity, we also recommend generating plots similar to Figure 5A and 5B to give more detailed results. To date, only 19 papers published in Lab on a Chip contain the terms “diagnostic sensitivity”, “diagnostic specificity”, “clinical sensitivity”, or “clinical specificity” (searched Feb. 2016).

**Establishing the connection between laboratory and clinical performance:** Analyzing the intended clinical use of microfluidic diagnostics during research and development will help researchers to create devices that are more likely to be beneficial to patients and clinicians. By focusing on the clinical use, researchers can focus their attention and resources on improving the most critical aspects of their devices. For instance, a screening test that requires high NPV or LR– would want to detect the smallest amount of analyte possible, and effort should be focused on improving LoD. Alternatively, tests that require high diagnostic specificity (high PPV and LR+) for rule-in would want to complete a thorough analytical selectivity study in the laboratory to account for a wide range of interferents at varying concentrations. Additionally, researchers should consider additional clinical factors such as the speed of result, false positive and false negative consequences, disease prevalence, patient convenience, and other factors discussed in the examples from section 4.

The importance of considering intended clinical use can further be highlighted by considering quality control for POC devices, which is an area of interest for microfluidic researchers. Establishing the trueness and precision for POC devices becomes much more challenging and critical due to the single-use nature of many of these tests and the variable environments they are used in. When considering potential sources of variability, the entire quality control process, including factors beyond actual testing must be taken into account. The three main quality control phases include: the pre-analytical phase (*e.g.* sample collection, patient identification), analytical phase (*e.g.* actual sample testing), and post-analytical phase (*e.g.* reporting of results).[208]

Early POC devices such as blood glucose meters suffered from design features that reduced the quality of results. Examples include (1) preanalytical: no mechanism to prevent untrained users from using the

machine and no safeguard to prevent use of expired strips, (2) analytical: operation required several steps including timing operation between dropping blood, wiping blood off the strip, and inserting strip in the correct orientation into reader, and (3) post-analytical: no electronic data management system, leaving manual documentation up to the operator.[209]

Due to the issues encountered with early POC devices, significant research and development has been focused on simplifying user operation, incorporating internal calibrations and controls, giving automated readouts that do not require timing, and using informatics and connectivity to transmit and store results.[209] For example, some devices read in barcodes to identify patients, the cartridges lock users out until an automated calibration is performed, and test results are automatically acquired and transmitted electronically. The balance between reducing the amount of resources (*e.g.* operator training, device cost, hardware requirements, time-to-result) needed for testing, while maintaining high analytical (*e.g.* precision, trueness) and clinical performance (*e.g.* PPV and NPV) is extremely important for improving the ability for clinicians to treat patients in all settings. Researchers must consider the entire testing process including pre- and post-analytical phases, as well as other previously discussed clinical factors during device development in order to provide the most accurate, precise, and accessible tools possible.

Overall, we believe that the metrics presented in Chapter 2 will help researchers to conduct consistent and rigorous laboratory assay evaluation, as well as guide them in designing clinically useful diagnostic tests. While all diagnostic tests aim to achieve perfect accuracy, the majority of tests will suffer from some combination of imperfect accuracy, technological barriers, usability issues, and/or high cost. Maximizing the benefit of imperfect tests can be achieved by taking into account the clinical metrics of device performance, and focusing on establishing and improving the analytical metrics that impact clinical use.

# Bibliography

1. Thompson M, Van den Bruel A. Diagnostic Tests Toolkit. Wiley-Blackwell; 2012.

2. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996 Jan 13;312(7023):71–72.

3. Newman TB, Kohn MA. Evidence-Based Diagnosis. Cambridge University Press; 2009.

4. Weigl BH, Gaydos CA, Kost G, Beyette FR, Sabourin S, Rompalo A, de los Santos T, McMullan JT, Haller J. The Value of Clinical Needs Assessments for Point-of-Care Diagnostics: Point Care J -Patient Test Technol. 2012 Jun;11(2):108–113.

5. Peeling R. Review of Candidate Technologies for Rapid Diagnostic Tests.

6. Drain PK, Hyle EP, Noubary F, Freedberg KA, Wilson D, Bishai WR, Rodriguez W, Bassett IV. Diagnostic point-of-care tests in resource-limited settings. Lancet Infect Dis. 2014;14(3):239–249.

7. Ciaranello AL, Park J-E, Ramirez-Avila L, Freedberg KA, Walensky RP, Leroy V. Early infant HIV-1 diagnosis programs in resource-limited settings: opportunities for improved outcomes and more cost-effective interventions. BMC Med. 2011;9:59. PMCID: PMC3129310

8. Bassett IV, Regan S, Chetty S, Giddy J, Uhler LM, Holst H, Ross D, Katz JN, Walensky RP, Freedberg KA, Losina E. Who starts antiretroviral therapy in Durban, South Africa?... not everyone who should. AIDS Lond Engl. 2010 Jan;24 Suppl 1:S37–44. PMCID: PMC3521614

9. Drain PK, Hyle EP, Noubary F, Freedberg KA, Wilson D, Bishai WR, Rodriguez W, Bassett IV. Diagnostic point-of-care tests in resource-limited settings. Lancet Infect Dis. 2014 Mar;14(3):239–249.

10. Alto, M.D. B. Indian Health Clinic Interview. 2015.

11. Kiechle FL, Holland CA. Point-of-care testing and molecular diagnostics: miniaturization required. Clin Lab Med. 2009 Sep;29(3):555–560. PMID: 19840687

12. Garrett NJ, Drain PK, Werner L, Samsunder N, Karim SSA. Diagnostic Accuracy of the Point-of-Care Xpert HIV-1 Viral Load Assay in a South African HIV Clinic. JAIDS J Acquir Immune Defic Syndr. 2016;72(2):e45–e48.

13. Posthuma-Trumpie GA, Korf J, Amerongen A. Lateral flow (immuno)assay: its strengths, weaknesses, opportunities and threats. A literature survey. Anal Bioanal Chem. 2008 Aug 13;393(2):569–582.

14. Stewart EH, Davis B, Clemans-Taylor BL, Littenberg B, Estrada CA, Centor RM. Rapid Antigen Group A Streptococcus Test to Diagnose Pharyngitis: A Systematic Review and Meta-Analysis. Reid SD, editor. PLoS ONE. 2014 Nov 4;9(11):e111727.

15. Pai NP, Balram B, Shivkumar S, Martinez-Cajas JL, Claessens C, Lambert G, Peeling RW, Joseph L. Head-to-head comparison of accuracy of a rapid point-of-care HIV test with oral versus whole-blood specimens: a systematic review and meta-analysis. Lancet Infect Dis. 2012;12(5):373–380.

16. Gift TL, Pate MS, Hook EW, Kassler WJ. The Rapid Test Paradox: When Fewer Cases Detected Lead to More Cases Treated. Sex Transm Dis. 1999;26(4):232–240.

17. Van Dommelen L, Van Tiel FH, Ouburg S, Brouwers EE, Terporten PH, Savelkoul PH, Morré SA, Bruggeman CA, Hoebe CJ. Alarmingly poor performance in Chlamydia trachomatis point-of-care testing. Sex Transm Infect. 2010;86(5):355–359.

18. Shulman ST, Bisno AL, Clegg HW, Gerber MA, Kaplan EL, Lee G, Martin JM, Beneden CV. Clinical Practice Guideline for the Diagnosis and Management of Group A Streptococcal Pharyngitis: 2012 Update by the Infectious Diseases Society of America. Clin Infect Dis. 2012 Sep 9;cis629. PMID: 22965026

19. Minion J, Leung E, Talbot E, Dheda K, Pai M, Menzies D. Diagnosing tuberculosis with urine lipoarabinomannan: systematic review and meta-analysis. Eur Respir J. 2011 Dec 1;38(6):1398–1405.

20. Life Technologies Animal Health. Strategic Use of Diagnostics. 2013.

21. ELISA [Internet]. Healthline. [cited 2016 Aug 11]. Available from: http://www.healthline.com/health/elisa

22. Moghadam BY, Connelly KT, Posner JD. Two orders of magnitude improvement in detection limit of lateral flow assays using isotachophoresis. Anal Chem. 2015;87(2):1009–1017.

23. Kerr KJ, Moghadam BY, Levenson A, Posner JD. Utilizing isotachophoresis and colorimetric quantification to lower Strep A limit of detection by two orders of magnitude. Anal Chem. 2015;Submitted.

24. Fu E, Liang T, Houghtaling J, Ramachandran S, Ramsey SA, Lutz B, Yager P. Enhanced Sensitivity of Lateral Flow Tests Using a Two-Dimensional Paper Network Format. Anal Chem. 2011 Oct 15;83(20):7941–7946.

25. Gordon J, Michel G. Analytical Sensitivity Limits for Lateral Flow Immunoassays. Clin Chem. 2008 Jul 1;54(7):1250–1251. PMID: 18593968

26. Ngom B, Guo Y, Wang X, Bi D. Development and application of lateral flow test strip technology for detection of infectious agents and chemical contaminants: a review. Anal Bioanal Chem. 2010 Apr 27;397(3):1113–1135.

27. U.S. Food and Drug Adminstration. CLIA - Clinical Laboratory Improvement Amendments. CLIA - Clin Lab Improv Amend [Internet]. 2014 Nov 10; Available from: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCLIA/search.cfm

28. Food and Drug Administration F. Recommendations for Clinical and Laboratory Improvement Amendments of 1988 (CLIA) Waiver Applications for Manufacturers of In Vitro Diagnostic Devices. Guid Ind FDA Staff. 2008;

29. Clinical and Laboratory Standards Institute (CLSI). Risk Management Techniques to Identify and Control Laboratory Error Sources; Approved Guideline—Second Edition. 2009.

30. World Health Organization. Mapping the landscape of diagnostics for sexually transmitted infections [Internet]. 2004. Available from: http://www.who.int/tdr/publications/documents/mapping-landscape-sti.pdf

31. Mothershed EA, Whitney AM. Nucleic acid-based methods for the detection of bacterial pathogens: Present and future considerations for the clinical laboratory. Clin Chim Acta. 2006 Jan;363(1-2):206–220.

32. Greco S, Girardi E, Navarra A, Saltini C. Current evidence on diagnostic accuracy of commercially based nucleic acid amplification tests for the diagnosis of pulmonary tuberculosis. Thorax. 2006 Sep 1;61(9):783–790.

33. Ieven M, Goossens H. Relevance of nucleic acid amplification techniques for diagnosis of respiratory tract infections in the clinical laboratory. Clin Microbiol Rev. 1997;10(2):242–256.

34. Ibekwe AM, Watt PM, Grieve CM, Sharma VK, Lyons SR. Multiplex Fluorogenic Real-Time PCR for Detection and Quantification of Escherichia coli O157:H7 in Dairy Wastewater Wetlands. Appl Environ Microbiol. 2002 Oct 1;68(10):4853–4862.

35. Pabinger S, Rödiger S, Kriegner A, Vierlinger K, Weinhäusel A. A survey of tools for the analysis of quantitative PCR (qPCR) data. Biomol Detect Quantif. 2014 Sep;1(1):23–33.

36. Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. Genome Res. 1996 Oct 1;6(10):986–994. PMID: 8908518

37. U.S. Food and Drug Adminstration. Nucleic acid based tests. Nucleic Acid Based Tests [Internet]. 2014 Jun 17; Available from: http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/ucm330711.htm

38. Promega. PCR Amplification [Internet]. [cited 2016 Aug 11]. Available from: https://www.promega.com/resources/product-guides-and-selectors/protocols-and-applications-guide/pcr-amplification/

39. Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science. 1988 Jan 29;239(4839):487–491. PMID: 2448875

40. Mackay IM. Real-time PCR in the microbiology laboratory. Clin Microbiol Infect. 2004 Mar;10(3):190–212.

41. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem. 1987 Apr;162(1):156–159. PMID: 2440339

42. Rogacs A, Marshall LA, Santiago JG. Purification of nucleic acids using isotachophoresis. J Chromatogr A. 2014 Mar 28;1335:105–120.

43. Niemz A, Ferguson TM, Boyle DS. Point-of-care nucleic acid testing for infectious diseases. Trends Biotechnol. 2011 May;29(5):240–250.

44. Sannes and Associates. The Market and Potential for Molecular Point of Care Diagnostics. Kalorama Information; 2014.

45. GenomeWeb. Cepheid Introduces True Point-of-Care Molecular Platform at AACC Analyst Event. 2015.

46. GenomeWeb. As Roche Grows its Rapid MDx Menu for Liat, Clinical Evaluations Hit the Presses. 2015.

47. GenomeWeb. With CLIA Waiver and Widespread Flu, Alere Ramps Up Molecular Test Production. 2015.

48. Cepheid Delays Launch of Omni Point-of-Care MDx System [Internet]. GenomeWeb. [cited 2016 Jun 14]. Available from: https://www.genomeweb.com/pcr/cepheid-delays-launch-omni-point-care-mdx-system

49. Niemz A, Ferguson TM, Boyle DS. Point-of-care nucleic acid testing for infectious diseases. Trends Biotechnol. 2011 May;29(5):240–250.

50. Byrnes S, Thiessen G, Fu E. Progress in the development of paper-based diagnostics for low-resource point-of-care settings. Bioanalysis. 2013 Nov;5(22):2821–2836.

51. Cui F, Rhee M, Singh A, Tripathi A. Microfluidic Sample Preparation for Medical Diagnostics. Annu Rev Biomed Eng. 2015 Dec 7;17(1):267–286.

52. Hoehl MM, Bocholt ES, Kloke A, Paust N, Stetten F von, Zengerle R, Steigert J, Slocum AH. A versatile-deployable bacterial detection system for food and environmental safety based on LabTube-automated DNA purification, LabReader-integrated amplification, readout and analysis. The Analyst. 2014;139(11):2788–2798.

53. Lutz S, Weber P, Focke M, Faltin B, Hoffmann J, Müller C, Mark D, Roth G, Munday P, Armes N, Piepenburg O, Zengerle R, von Stetten F. Microfluidic lab-on-a-foil for nucleic acid analysis based on isothermal recombinase polymerase amplification (RPA). Lab Chip. 2010;10(7):887.

54. Connelly JT, Rolland JP, Whitesides GM. "Paper Machine" for Molecular Diagnostics. Anal Chem. 2015 Aug 4;87(15):7595–7601.

55. Byrnes SA, Bishop JD, Lafleur L, Buser JR, Lutz B, Yager P. One-step purification and concentration of DNA in porous membranes for point-of-care applications. Lab Chip. 2015;15(12):2647–2659.

56. Rodriguez NM, Wong WS, Liu L, Dewar R, Klapperich CM. A fully integrated paperfluidic molecular diagnostic chip for the extraction, amplification, and detection of nucleic acids from clinical samples. Lab Chip. 2016;16(4):753–763.

57. Notomi T, Okayama H, Masubuchi H, Yonekawa T, Watanabe K, Amino N, Hase T. Loop-mediated isothermal amplification of DNA. Nucleic Acids Res. 2000;28(12):e63.

58. Piepenburg O, Williams CH, Stemple DL, Armes NA. DNA detection using recombination proteins. PLoS Biol. 2006 Jul;4(7):e204. PMCID: PMC1475771

59. Craw P, Balachandran W. Isothermal nucleic acid amplification technologies for point-of-care diagnostics: a critical review. Lab Chip. 2012;12(14):2469.

60. Myers FB, Henrikson RH, Bone J, Lee LP. A handheld point-of-care genomic diagnostic system. PloS One. 2013;8(8):e70266.

61. Stedtfeld RD, Tourlousse DM, Seyrig G, Stedtfeld TM, Kronlein M, Price S, Ahmad F, Gulari E, Tiedje JM, Hashsham SA. Gene-Z: a device for point of care genetic testing using a smartphone. Lab Chip. 2012 Apr 21;12(8):1454–1462.

62. Selck DA, Karymov MA, Sun B, Ismagilov RF. Increased Robustness of Single-Molecule Counting with Microfluidics, Digital Isothermal Amplification, and a Mobile Phone versus Real-Time Kinetic Measurements. Anal Chem. 2013 Nov 19;85(22):11129–11136.

63. Kersting S, Rausch V, Bier FF, Nickisch-Rosenegk M von. Rapid detection of Plasmodium falciparum with isothermal recombinase polymerase amplification and lateral flow analysis. Malar J. 2014 Mar 15;13(1):99. PMID: 24629133

64. Roskos K, Hickerson AI, Lu H-W, Ferguson TM, Shinde DN, Klaue Y, Niemz A. Simple System for Isothermal DNA Amplification Coupled to Lateral Flow Detection. PLOS ONE. 2013 Jul 26;8(7):e69355.

65. Goto M, Honda E, Ogura A, Nomoto A, Hanaki K-I. Colorimetric detection of loop-mediated isothermal amplification reaction by using hydroxy naphthol blue. BioTechniques. 2009 Mar;46(3):167–172.

66. Tomita N, Mori Y, Kanda H, Notomi T. Loop-mediated isothermal amplification (LAMP) of gene sequences and simple visual detection of products. Nat Protoc. 2008 Apr;3(5):877–882.

67. Bocek P, Deml M, Gebauer P, Dolnik V. Analytical Isotachophoresis. VCH, Weinheim; 1988.

68. Kaigala GV, Bercovici M, Behnam M, Elliott D, Santiago JG, Backhouse CJ. Miniaturized system for isotachophoresis assays. Lab Chip. 2010 Sep 7;10(17):2242–2250.

69. Alberty RA. Moving Boundary Systems Formed by Weak Electrolytes. Theory of Simple Systems Formed by Weak Acids and Bases. J Am Chem Soc. 1950 Jun 1;72(6):2361–2367.

70. Jung B, Bharadwaj R, Santiago JG. On-Chip Millionfold Sample Stacking Using Transient Isotachophoresis. Anal Chem. 2006 Apr;78(7):2319–2327.

71. Bercovici M, Han CM, Liao JC, Santiago JG. Rapid hybridization of nucleic acids using isotachophoresis. Proc Natl Acad Sci. 2012 Jul 10;109(28):11127–11132. PMID: 22733732

72. Han CM, Katilius E, Santiago JG. Increasing Hybridization Rate and Sensitivity of DNA Microarrays Using Isotachophoresis. Lab Chip [Internet]. 2014 May 20 [cited 2014 May 28]; Available from: http://pubs.rsc.org/en/content/articlelanding/2014/lc/c4lc00374h

73. Kondratova V, Serd'uk O 'ga, Shelepov V, Lichtenstein A. Concentration and isolation of DNA from biological fluids by agarose gel isotachophoresis. BioTechniques. 2005 Nov;39(5):695–699.

74. Marshall LA, Han CM, Santiago JG. Extraction of DNA from Malaria-Infected Erythrocytes Using Isotachophoresis. Anal Chem. 2011 Dec 15;83(24):9715–9718.

75. Bercovici M, Kaigala GV, Mach KE, Han CM, Liao JC, Santiago JG. Rapid Detection of Urinary Tract Infections Using Isotachophoresis and Molecular Beacons. Anal Chem. 2011 Jun;83(11):4110–4117.

76. Marshall LA, Wu LL, Babikian S, Bachman M, Santiago JG. Integrated Printed Circuit Board Device for Cell Lysis and Nucleic Acid Extraction. Anal Chem. 2012 Oct 9;84:9640–9645.

77. Rogacs A, Qu Y, Santiago JG. Bacterial RNA Extraction and Purification from Whole Human Blood Using Isotachophoresis. Anal Chem. 2012 Jul 17;84(14):5858–5863.

78. Strychalski EA, Konek C, Butts ELR, Vallone PM, Henry AC, Ross D. DNA purification from crude samples for human identification using gradient elution isotachophoresis. Electrophoresis. 2013;34(17):2522–2530.

79. Marshall LA, Rogacs A, Meinhart CD, Santiago JG. An injection molded microchip for nucleic acid purification from 25 microliter samples using isotachophoresis. J Chromatogr A. 2014 Feb;1331:139–142.

80. Persat A, Marshall LA, Santiago JG. Purification of Nucleic Acids from Whole Blood Using Isotachophoresis. Anal Chem. 2009 Nov 15;81(22):9507–9511.

81. PURIGEN [Internet]. [cited 2016 Aug 11]. Available from: http://www.purigenbio.com/

82. Persat A, Chambers RD, Santiago JG. Basic principles of electrolyte chemistry for microfluidic electrokinetics. Part I: Acid–base equilibria and pH buffers. Lab Chip. 2009;9(17):2437–2453.

83. Milanova D, Chambers RD, Bahga SS, Santiago JG. Electrophoretic mobility measurements of fluorescent dyes using on-chip capillary electrophoresis. Electrophoresis. 2011;32(22):3286–3294.

84. Rogacs A, Marshall LA, Santiago JG. Purification of nucleic acids using isotachophoresis. J Chromatogr A. 2014 Mar 28;1335:105–120.

85. Stellwagen E, Stellwagen NC. The free solution mobility of DNA in Tris-acetate-EDTA buffers of different concentrations, with and without added NaCl. Electrophoresis. 2003;23(16):1935–1941.

86. Bercovici M, Lele SK, Santiago JG. Open source simulation tool for electrophoretic stacking, focusing, and separation. J Chromatogr A. 2009;1216(6):1008–1018.

87. Khurana TK, Santiago JG. Sample Zone Dynamics in Peak Mode Isotachophoresis. Anal Chem. 2008 Aug 1;80(16):6300–6307.

88. Bocek P, Deml M, Kaplanova B, Janak J. The concept of the separation capacity. J Chromatogr. 1978;160:1–9.

89. Garcia-Schwarz G, Bercovici M, Marshall LA, Santiago JG. Sample dispersion in isotachophoresis. J Fluid Mech. 2011 May 12;679:455–475.

90. Kirby BJ, Hasselbrink EF. Zeta potential of microfluidic substrates: 1. Theory, experimental techniques, and effects on separations. Electrophoresis. 2004 Jan;25(2):187–202.

91. Kirby BJ, Hasselbrink EF. Zeta potential of microfluidic substrates: 2. Data for polymers. Electrophoresis. 2004 Jan;25(2):203–213.

92. Milanova D, Chambers RD, Bahga SS, Santiago JG. Effect of PVP on the electroosmotic mobility of wet-etched glass microchannels. Electrophoresis. 2012;33(21):3259–3262.

93. Craw P, Balachandran W. Isothermal nucleic acid amplification technologies for point-of-care diagnostics: a critical review. Lab Chip. 2012;12(14):2469–2486.

94. Curtis KA, Rudolph DL, Nejad I, Singleton J, Beddoe A, Weigl B, LaBarre P, Owen SM. Isothermal Amplification Using a Chemical Heating Device for Point-of-Care Detection of HIV-1. Landay A, editor. PLoS ONE. 2012 Feb 23;7(2):e31432.

95. Huang S, Do J, Mahalanabis M, Fan A, Zhao L, Jepeal L, Singh SK, Klapperich CM. Low Cost Extraction and Isothermal Amplification of DNA for Infectious Diarrhea Diagnosis. Nübel U, editor. PLoS ONE. 2013 Mar 28;8(3):e60059.

96. Lillis L, Lehman D, Singhal MC, Cantera J, Singleton J, Labarre P, Toyama A, Piepenburg O, Parker M, Wood R, Overbaugh J, Boyle DS. Non-Instrumented Incubation of a Recombinase Polymerase Amplification Assay for the Rapid and Sensitive Detection of Proviral HIV-1 DNA. López-Galíndez C, editor. PLoS ONE. 2014 Sep 29;9(9):e108189.

97. Mori Y, Kitao M, Tomita N, Notomi T. Real-time turbidimetry of LAMP reaction for quantifying template DNA. J Biochem Biophys Methods. 2004 May 31;59(2):145–157. PMID: 15163526

98. Wang F, Jiang L, Ge B. Loop-Mediated Isothermal Amplification Assays for Detecting Shiga Toxin-Producing Escherichia coli in Ground Beef and Human Stools. J Clin Microbiol. 2011 Oct 26;50(1):91–97.

99. Boyle DS, Lehman DA, Lillis L, Peterson D, Singhal M, Armes N, Parker M, Piepenburg O, Overbaugh J. Rapid detection of HIV-1 proviral DNA for early infant diagnosis using recombinase polymerase amplification. mBio. 2013;4(2). PMCID: PMC3622927

100. Boyle DS, McNerney R, Teng Low H, Leader BT, Pérez-Osorio AC, Meyer JC, O'Sullivan DM, Brooks DG, Piepenburg O, Forrest MS. Rapid detection of Mycobacterium tuberculosis by recombinase polymerase amplification. PloS One. 2014;9(8):e103091. PMCID: PMC4138011

101. Lillis L, Lehman D, Singhal MC, Cantera J, Singleton J, Labarre P, Toyama A, Piepenburg O, Parker M, Wood R, Overbaugh J, Boyle DS. Non-instrumented incubation of a recombinase polymerase amplification assay for the rapid and sensitive detection of proviral HIV-1 DNA. PloS One. 2014;9(9):e108189. PMCID: PMC4180440

102. Urdea M, Penny LA, Olmsted SS, Giovanni MY, Kaspar P, Shepherd A, Wilson P, Dahl CA, Buchsbaum S, Moeller G, Hay Burgess DC. Requirements for high impact diagnostics in the developing world. Nature. 2006 Nov 23;444 Suppl 1:73–79. PMID: 17159896

103. Lillis L, Siverson J, Lee A, Cantera J, Parker M, Piepenburg O, Lehman DA, Boyle DS. Factors influencing Recombinase polymerase amplification (RPA) assay outcomes at point of care. Mol Cell Probes. 2016 Apr;30(2):74–78.

104. Piepenburg O, Armes NA. Recombinase Polymerase Amplification Reagents and Kits. US2012/0129173A1.

105. Borysiak MD, Kimura KW, Posner JD. NAIL: Nucleic Acid detection using Isotachophoresis and Loop-mediated isothermal amplification. Lab Chip [Internet]. 2015 Feb 10 [cited 2015 Feb 12]; Available from: http://pubs.rsc.org/en/content/articlelanding/2015/lc/c4lc01479k

106. Saah AJ, Hoover DR. “Sensitivity” and “specificity” reconsidered: the meaning of these terms in analytical and diagnostic settings. Ann Intern Med. 1997;126(1):91–94.

107. Currie LA. Nomenclature in evaluation of analytical methods including detection and quantification capabilities:(IUPAC Recommendations 1995). Anal Chim Acta. 1999;391(2):105–126.

108. Vessman J, Stefan RI, Van Staden JF, Danzer K, Lindner W, Burns DT, Fajgelf A, Helmut M. Selectivity in Analytical Chemistry. IUPAC Recomm Pure Appl Chem. 2001;73(8):1381–1386.

109. Clinical and Laboratory Standards Institute (CLSI). Protocols for Determination of Limits of Detection and Limits of Quantitation; Approved Guidline. EP17-A; 2004.

110. Clinical Chemistry. Information for Authors [Internet]. http://www.clinchem.org/site/info_ar/info_authors.xhtml; 2015 [cited 2015 Oct 20]. Available from: http://www.clinchem.org/site/info_ar/info_authors.xhtml

111. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. BMJ. 1994;308(6943):1552.

112. Altman DG, Bland JM. Diagnostic tests 2: predictive values. BMJ. 1994;309(102):102.

113. Deeks JJ, Altman DG, others. Diagnostic tests 4: likelihood ratios. Bmj. 2004;329(7458):168–169.

114. Jacobson RH. Principles of validation of diagnostic assays for infectious diseases. Diagn Epidemiol Anim Dis Lat Am [Internet]. International Atomic Energy Agency (IAEA) Technical Documents (TECDOC); 1998 [cited 2015 Oct 21]. p. 15–25. Available from: http://www.iaea.org/inis/collection/NCLCollectionStore/_Public/29/067/29067668.pdf#page=22

115. Linnet K. Partly Nonparametric Approach for Determining the Limit of Detection. Clin Chem. 2004 Apr 1;50(4):732–740.

116. Clinical and Laboratory Standards Institute (CLSI). User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition. EP15–A2; 2005.

117. Burd EM. Validation of Laboratory-Developed Molecular Assays for Infectious Diseases. Clin Microbiol Rev. 2010 Jul 1;23(3):550–576.

118. Clinical and Laboratory Standards Institute (CLSI). Interference Testing in Clinical Chemistry; Approved Guideline—Second Edition. EP7–A2; 2005.

119. Pardue HL. Counterpoint The inseparable triad: analytical sensitivity, measurement uncertainty, and quantitative resolution. Clin Chem. 1997 Oct 1;43(10):1831–1837. PMID: 9342000

120. Zhou M. Regulated Bioanalytical Laboratories: Technical and Regulatory Aspects from Global Perspectives [Internet]. John Wiley & Sons; 2011 [cited 2015 Oct 21]. Available from: http://books.google.com/books?hl=en&lr=&id=MMHi3ayYcbUC&oi=fnd&pg=PT7&dq=%22immunological+and+microbiological+procedures,+and+to+other+biological+matrices,+such+as+tissue%22+%22guidance+provides+general+recommendations+for+bioanalytical+method+validation.%22+%22This+guidance+has+been+prepared+by+the+Biopharmaceutics+Coordinating+Committe&ots=ERYPzOAfJR&sig=bAEShmpBUhUFRP6juym1Gk4CmUs

121. Bunk DM. Reference Materials and Reference Measurement Procedures: An Overview from a National Metrology Institute. Clin Biochem Rev. 2007 Nov;28(4):131–137. PMCID: PMC2282405

122. Ekins R, Edwards P. Point On the meaning of "sensitivity." Clin Chem. 1997 Oct 1;43(10):1824–1831. PMID: 9341999

123. Currie LA. Detection and quantification limits: origins and historical overview. Anal Chim Acta. 1999;391(2):127–134.

124. Laerd Statistics. QuickVue frequency diagram.png. "Linear Regression using Minitab", https://statistics.laerd.com/minitab-tutorials/linear-regression-using-minitab.php.

125. Frost J. Regression Analysis: How Do I Interpret R-squared and Assess the Goodness-of-Fit? [Internet]. "Regression Analysis: How Do I Interpret R-squared and Assess the Goodness-of-Fit?", http://blog.minitab.com [cited 2016 Mar 4]. Available from: http://blog.minitab.com/blog/adventures-in-statistics/regression-analysis-how-do-i-interpret-r-squared-and-assess-the-goodness-of-fit

126. National Association of Testing Authorities, Australia. NATA tech note, Asens and Aselect.pdf. Guidelines for the validation and verification of quantitative and qualitative test methods; 2013.

127. Frost J. Multiple Regression Analysis: Use Adjusted R-Squared and Predicted R-Squared to Include the Correct Number of Variables [Internet]. "Multiple Regression Analysis: Use Adjusted R-Squared and Predicted R-Squared to Include the Correct Number of Variables", http://blog.minitab.com/ [cited 2016 Mar 4]. Available from: http://blog.minitab.com/

128. Madej RM, Davis J, Holden MJ, Kwang S, Labourier E, Schneider GJ. International Standards and Reference Materials for Quantitative Molecular Infectious Disease Testing. J Mol Diagn JMD. 2010 Mar;12(2):133–143. PMCID: PMC2871718

129. ISO. Capability of detection — Part 1. Terms and definitions. Geneva: International Organization for Standardization: 1997; Report No.: ISO 11843-1.

130. Tholen DW, Linnet K, Kondratovich M, Armbruster DA, Garrett PE, Jones RL, Kroll MH, Lequin RM, Pankratz TJ, Scassellati GA, Schimmel H, Tsai J. Protocols for Determination of Limits of Detection and Limits of Quantitation; Approved Guidline. Clin Lab Stand Inst. 2004;24(EP17-A).

131. Needleman SB, Romberg RW. Limits of Linearity and Detection for Some Drugs of Abuse. J Anal Toxicol. 1990 Jan 1;14(1):34–38. PMID: 2314061

132. Armbruster DA, Pry T. Limit of blank, limit of detection and limit of quantitation. Clin Biochem Rev. 2008;29(Suppl 1):S49–52.

133. Christian GD. Analytical Chemistry. 5th ed. New York: Wiley; 1994.

134. Kaiser H. Z Anal Chem. 1972;260(252).

135. Danzer K. Selectivity and specificity in analytical chemistry. General considerations and attempt of a definition and quantification. Fresenius J Anal Chem. 2001;369(5):397–402.

136. Dimeski G. Interference Testing. Clin Biochem Rev. 2008;29(Suppl (i)).

137. Food and Drug Administration (FDA). Regulatory Information: Search for Guidance Documents [Internet]. http://www.fda.gov/RegulatoryInformation/Guidances/default.htm; [cited 2015 Dec 27]. Available from: http://www.fda.gov/RegulatoryInformation/Guidances/default.htm

138. Krouwer JS. Interference Testing: Why Following Standards Is Not Always the Right Thing to Do. J Diabetes Sci Technol. 2012;6(5):1182–1184.

139. Clinical and Laboratory Standards Institute (CLSI). Evaluation of Matrix Effects: Approved Guideline—Second Edition. EP14–A2; 2005.

140. Westgard J. Comparison [Internet]. "The Comparison of Methods Experiment", https://www.westgard.com/lesson22.htm. Available from: https://www.westgard.com/lesson23.htm

141. Chesher D. Evaluating assay precision. Clin Biochem Rev. 2008;29(Suppl 1):S23–S26.

142. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Int J Nurs Stud. 2010;47(8):931–936.

143. Clinical and Laboratory Standards Institute (CLSI). Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Second Edition. EP09; 2010.

144. Food and Drug Administration (FDA). Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests [Internet]. 2007. Available from: http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071148.htm

145. Clinical and Laboratory Standards Institute (CLSI). Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline—Second Edition. EP5; 2004.

146. Westgard J. Precision. "The Replication Experiment", https://www.westgard.com/lesson23.htm.

147. Clinical and Laboratory Standards Institute (CLSI). User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline--Second Edition. EP12; 2008.

148. Clinical and Laboratory Standards Institute (CLSI). Preliminary Evaluation of Quantitative Laboratory Measurement Procedures; Approved Guideline—Third Edition. EP10–A3; 2014.

149. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982 Apr 1;143(1):29–36.

150. Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. Acta Paediatr. 2007 May;96(5):644–647.

151. Mallett S, Halligan S, Thompson M, Collins GS, Altman DG. Interpreting diagnostic accuracy studies for patient care. BMJ. 2012 Jul 2;345(jul02 1):e3999–e3999.

152. Moons KG, Stijnen T, Michel BC, Büller HR, Van Es G-A, Grobbee DE, Habbema JDF. Application of Treatment Thresholds to Diagnostic-test Evaluation An Alternative to the Comparison of Areas under Receiver Operating Characteristic Curves. Med Decis Making. 1997;17(4):447–454.

153. Pencina MJ, D'Agostino RB, D'Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008 Jan 30;27(2):157–172; discussion 207–212. PMID: 17569110

154. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. The Lancet. 2005;365(9469):1500–1505.

155. Mann C. Observational research methods. Research design II: cohort, cross sectional, and case-control studies. Emerg Med J EMJ. 2003 Jan;20(1):54–60. PMCID: PMC1726024

156. Levin KA. Study design III: Cross-sectional studies. Evid Based Dent. 2006;7(1):24–25.

157. Gigerenzer G, Hoffrage U. How to improve Bayesian reasoning without instruction: frequency formats. Psychol Rev. 1995;102(4):684.

158. Gigerenzer G. HIV screening: helping clinicians make sense of test results to patients. BMJ. 2013 Aug 21;347(aug21 2):f5151–f5151.

159. Thompson M, Weigl B, Fitzpatrick A, Ide N. "More Than Just Accuracy": A Novel Method to Incorporate Multiple Test Attributes in Evaluating Point of Care Tests. Compr Eval Valid POC Technol. 2015.

160. Hozo I, Djulbegovic B. Using the Internet to calculate clinical action thresholds. Comput Biomed Res Int J. 1999 Apr;32(2):168–185. PMID: 10337498

161. Nomogram for Bayes's Theorem. N Engl J Med. 1975 Jul 31;293(5):257–257. PMID: 1143310

162. Caraguel CGB, Vanderstichel R. The two-step Fagan's nomogram: ad hoc interpretation of a diagnostic test result without calculation. Evid Based Med. 2013 Aug 1;18(4):125–128. PMID: 23468201

163. DocNomo, iPhone and iPad application. DocNomo good. 2014.

164. Alan Schwartz, Diagnostic Test Calculator, http://araw.mede.uic.edu/cgi-bin/testcalc.pl. Diagnostic Test Calculator [Internet]. 2002 [cited 2015 Oct 23]. Available from: http://araw.mede.uic.edu/cgi-bin/testcalc.pl

165. Chan M. WHO | WHO Director-General addresses UN Security Council on Ebola. http://www.who.int/dg/speeches/2014/security-council-ebola/en/;

166. World Health Organization (WHO). WHO | Ebola virus disease. http://www.who.int/csr/don/archive/disease/ebola/en/;

167. Shinkins B, Thompson M, Mallett S, Perera R. Diagnostic accuracy studies: how to report and analyse inconclusive test results. BMJ. 2013 May 16;346(may16 2):f2778–f2778.

168. Hoffman RM, Gilliland FD, Adams-Cameron M, Hunt WC, Key CR. Prostate-specific antigen testing accuracy in community practice. BMC Fam Pract. 2002 Oct 24;3(1):19. PMID: 12398793

169. Cohen DM, Russo ME, Jaggi P, Kline J, Gluckman W, Parekh A. Multicenter Clinical Evaluation of the Novel Alere i Strep A Isothermal Nucleic Acid Amplification Test. Bourbeau P, editor. J Clin Microbiol. 2015 Jul;53(7):2258–2261.

170. van der Helm JJ, Sabajo LOA, Grunberg AW, Morré SA, Speksnijder AGCL, de Vries HJC. Point-of-Care Test for Detection of Urogenital Chlamydia in Women Shows Low Sensitivity. A Performance Evaluation Study in Two Clinics in Suriname. Dean D, editor. PLoS ONE. 2012 Feb 29;7(2):e32122.

171. Leung CC, Lee SS. Rapid HIV tests: from meta-analysis to field application. Lancet Infect Dis. 2012;12(5):356–357.

172. Food and Drug Administration (FDA). Premarket Approvals (PMAs) - Information regarding the OraQuick In-Home HIV Test [Internet]. Premarket Approvals (PMAs) - Information regarding the OraQuick In-Home HIV Test [cited 2015 Oct 23]. Available from: http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMAs/ucm311895.htm

173. Centers for Disease Control and Prevention (CDC). Get Smart: Know When Antibiotics Work [Internet]. Adult Treatment Recommendations for Pharyngitis, http://www.cdc.gov/getsmart/community/for-hcp/outpatient-hcp/adult-treatment-rec.html#ref8 [cited 2015 Nov 23]. Available from: http://www.cdc.gov/getsmart/community/for-hcp/outpatient-hcp/adult-treatment-rec.html#ref8

174. Centers for Disease Control and Prevention (CDC). Get Smart: Know When Antibiotics Work [Internet]. Pediatric Treatment Recommendations for Pharyngitis, http://www.cdc.gov/getsmart/community/for-hcp/outpatient-hcp/pediatric-treatment-rec.html [cited 2015 Nov 23]. Available from: http://www.cdc.gov/getsmart/community/for-hcp/outpatient-hcp/pediatric-treatment-rec.html

175. Snow V, Mottur-Pilson C, Cooper RJ, Hoffman JR. Principles of Appropriate Antibiotic Use for Acute Pharyngitis in Adults. Ann Intern Med. 2001 Mar 20;134(6):506–508.

176. Food and Drug Administration (FDA). Tips and Articles on Device Safety - Use of Backup Testing for Negative Rapid Group A Strep Tests [Internet]. http://www.fda.gov/MedicalDevices/Safety/AlertsandNotices/TipsandArticlesonDeviceSafety/ucm109407.htm; 2015 [cited 2016 Jan 4]. Available from: http://www.fda.gov/MedicalDevices/Safety/AlertsandNotices/TipsandArticlesonDeviceSafety/ucm109407.htm

177. Weigl BH, Gaydos CA, Kost G, Beyette FR, Sabourin S, Rompalo A, de los Santos T, McMullan JT, Haller J. The Value of Clinical Needs Assessments for Point-of-Care Diagnostics: Point Care J -Patient Test Technol. 2012 Jun;11(2):108–113.

178. Papp JR, Schachter J, Gaydos CA, Van Der Pol B. Recommendations for the laboratory-based detection of Chlamydia trachomatis and Neisseria gonorrhoeae—2014. MMWR Recomm Rep Morb Mortal Wkly Rep Recomm ReportsCenters Dis Control. 2014;63:1.

179. Gaydos CA, Quinn TC, Willis D, Weissfeld A, Hook EW, Martin DH, Ferrero DV, Schachter J. Performance of the APTIMA Combo 2 assay for detection of Chlamydia trachomatis and Neisseria gonorrhoeae in female urine and endocervical swab specimens. J Clin Microbiol. 2003 Jan;41(1):304–309. PMCID: PMC149571

180. Borysiak MD, Bielawski KS, Sniadecki NJ, Jenkel CF, Vogt BD, Posner JD. Simple replica micromolding of biocompatible styrenic elastomers. Lab Chip. 2013;13(14):2773–2784.

181. Borysiak MD, Yuferova E, Posner JD. Simple, Low-Cost Styrene-Ethylene/Butylene-Styrene Microdevices for Electrokinetic Applications. Anal Chem. 2013 Nov 22;11700–11704.

182. Shackman JG, Ross D. Gradient Elution Isotachophoresis for Enrichment and Separation of Biomolecules. Anal Chem. 2007 Sep;79(17):6641–6649.

183. Karsenty M, Rosenfeld T, Gommed K, Bercovici M. Current monitoring in microchannel with repeated constrictions for accurate detection of sample location in isotachophoresis. Anal Chem [Internet]. 2014 [cited 2014 Dec 3];Just Accepted Manuscript. Available from: http://pubs.acs.org/doi/abs/10.1021/ac5036346

184. Klatser PR, Kuijper S, van Ingen CW, Kolk AH. Stabilized, freeze-dried PCR mix for detection of mycobacteria. J Clin Microbiol. 1998;36(6):1798–1800.

185. Cho H, Kim H-Y, Kang JY, Kim TS. How the capillary burst microvalve works. J Colloid Interface Sci. 2007 Feb;306(2):379–385.

186. Niño MRR, Patino JR. Surface tension of bovine serum albumin and tween 20 at the air-aqueous interface. J Am Oil Chem Soc. 1998;75(10):1241–1248.

187. Dolnik V, Deml M, Bocek P. Large sample volume preseparation for trace analysis in isotachophoresis. J Chromatogr. 1985;320:89–97.

188. Breckenridge WC, Kuksis A. Molecular weight distributions of milk fat triglycerides from seven species. J Lipid Res. 1967;8(5):473–478.

189. Bercovici M, Lele SK, Santiago JG. Open source simulation tool for electrophoretic stacking, focusing, and separation. J Chromatogr A. 2009 Feb;1216(6):1008–1018.

190. Persat A, Suss ME, Santiago JG. Basic principles of electrolyte chemistry for microfluidic electrokinetics. Part II: Coupling between ion mobility, electrolysis, and acid–base equilibria. Lab Chip. 2009;9(17):2454.

191. Francois P, Tangomo M, Hibbs J, Bonetti E-J, Boehme CC, Notomi T, Perkins MD, Schrenzel J. Robustness of a loop-mediated isothermal amplification reaction for diagnostic applications: LAMP for microbiological diagnosis. FEMS Immunol Med Microbiol. 2011 Jun;62(1):41–48.

192. Marshall LA, Han CM, Santiago JG. Extraction of DNA from Malaria-Infected Erythrocytes Using Isotachophoresis. Anal Chem. 2011 Dec 15;83(24):9715–9718.

193. Strychalski EA, Konek C, Butts ELR, Vallone PM, Henry AC, Ross D. DNA purification from crude samples for human identification using gradient elution isotachophoresis: Nucleic Acids. ELECTROPHORESIS. 2013 Sep;34(17):2522–2530.

194. Nogva HK, Rudi K, Naterstad K, Holck A, Lillehaug D. Application of 5'-Nuclease PCR for Quantitative Detection of Listeria monocytogenes in Pure Cultures, Water, Skim Milk, and Unpasteurized Whole Milk. Appl Environ Microbiol. 2000 Oct 1;66(10):4266–4271.

195. Daly P, Collier T, Doyle S. PCR-ELISA detection of Escherichia coli in milk. Lett Appl Microbiol. 2002;34(3):222–226.

196. Waswa J, Irudayaraj J, DebRoy C. Direct detection of E. Coli O157:H7 in selected food systems by a surface plasmon resonance biosensor. LWT - Food Sci Technol. 2007 Mar;40(2):187–192.

197. Radke SM, Alocilja EC. A high density microelectrode array biosensor for detection of E. coli O157:H7. Biosens Bioelectron. 2005 Feb;20(8):1662–1667.

198. Ortblad KF, Lozano R, Murray CJL. The burden of HIV: insights from the Global Burden of Disease Study 2010. AIDS Lond Engl. 2013 Aug 24;27(13):2003–2017. PMCID: PMC3748855

199. World Health Organization. WHO | Size of the epidemic [Internet]. WHO. [cited 2015 Sep 29]. Available from: http://www.who.int/gho/hiv/epidemic_status/en/

200. World Health Organization (WHO). The Use of Antiretroviral Drugs for Treating and Preventing HIV Infection: Recommendations for a Publich Health Approach [Internet]. 2016. Available from: http://www.who.int/hiv/pub/arv/arv-2016/en/

201. World Health Organization (WHO). Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV [Internet]. 2015. Available from: http://apps.who.int/iris/bitstream/10665/186275/1/9789241509565_eng.pdf

202. UNAIDS. 90-90-90: An ambitious treatment target to help end the AIDS epidemic [Internet]. 2014. Available from: http://www.unaids.org/sites/default/files/media_asset/90-90-90_en_0.pdf

203. Kieffer MP, Mattingly M, Giphart A, van de Ven R, Chouraya C, Walakira M, Boon A, Mikusova S, Simonds RJ. Lessons Learned From Early Implementation of Option B+: The Elizabeth Glaser Pediatric AIDS Foundation Experience in 11 African Countries. J Acquir Immune Defic Syndr 1999. 2014 Dec 1;67(Suppl 4):S188–S194. PMCID: PMC4251909

204. Lillis L, Lehman DA, Siverson JB, Weis J, Cantera J, Parker M, Piepenburg O, Overbaugh J, Boyle DS. Cross-subtype detection of HIV-1 using reverse transcription and recombinase polymerase amplification. J Virol Methods. 2016 Apr;230:28–35. PMCID: PMC4767662

205. Safe-Tec. Safe, Effective Blood Sampling Products [Internet]. Available from: http://www.safe-tecinc.com/microsafe

206. PALL Life Sciences. Vivid Plasma Separation Membrane [Internet]. Available from: https://www.pall.com/pdfs/OEM-Materials-and-Devices/09.2730_VividPlasma_DS_6pg.pdf

207. Nabatiyan A, Parpia ZA, Elghanian R, Kelso DM. Membrane-based plasma collection device for point-of-care diagnosis of HIV. J Virol Methods. 2011 Apr;173(1):37–42.

208. Gill JP, Shephard MD. The conduct of quality control and quality assurance testing for PoCT outside the laboratory. Clin Biochem Rev. 2010;31(3):85.

209. Lewandrowski K, Gregory K, Macmillan D. Assuring quality in point-of-care testing: evolution of technologies, informatics, and program management. Arch Pathol Lab Med. 2011;135(11):1405–1414.

# Appendix 1: MATLAB code

```
% This code reads a multipage tif image file,
% integrates fluorescent intensity above a threshold,
% and creates intensity versus time plots, as well as spatiotemporal maps

clear all;
close all;
clc;
clf;

set(0,'DefaultFigureColor','White',... %sets default
    'DefaultAxesFontSize',18,...
    'DefaultTextFontSize',18,...
    'DefaultAxesFontName','<Arial>',...
    'DefaultTextFontName','<Arial>',...
    'DefaultLineMarkerSize',12,...
    'DefaultLineLineWidth', 1);

%% inputs
namefile = 'Pos';   %Name of input file
n = 1200; % number of images to read
background_images = 50; % number of images to use as background
thresh_coeff=1.08;  % threshold value above background
exposure=1; % exposure time of camera (seconds)

% Cropping region of interest on the image
figure(1)
I=mimread('',namefile,n,1); %view last image in stack for cropping
imshow(I,[4800 5300]);  % set intensity values for when cropping image
rect = getrect();
newImage=imcrop(I,[rect(1) rect(2) rect(3) rect(4)]);   % extracting image
dimensions
xmin=int32(rect(1));    %minimum pixel x position to crop image
ymin=int32(rect(2));    %minimum pixel y position to crop image
xmax=int32(rect(3)+rect(1));    %maximum pixel x position to crop image
ymax=int32(rect(4)+rect(2));    %maximum pixel y position to crop image
height=double(rect(4));     %height of image
L=double(rect(3));   %width of image

% Calculates intensity verus time plots, spatio temporal map, and exports
% values

% blank matrices
numMoles = zeros(1,background_images-1);
numMoles2 = zeros(1,background_images-1);
Gaussian2=[];



tic

%% Reading in the images, y-averaging, subtracting background, integrating
above threshold
for j=1:n;

    % Reading image
    imcell=mimread('',namefile,j,1);  % read in multiple image tiff file
```

```
    newImage=imcrop(imcell,[xmin ymin L height]);   % crop image according to
selection
    NSGaussian(j,:)=mean(newImage);                 % y-average image
    Gaussian2(j,:)=smooth(NSGaussian(j,:),0.025,'loess');
%Gaussian2=[Gaussian2; Gaussian];

    if j>background_images-1                          % establishes average
background intensity for X number of images
       background = mean(Gaussian2(1:background_images,:)); %calculating y-
average background
       BG_subtract = (Gaussian2(j,:)-background);   % subtracts mean
background
       avg_background = mean(background);   %   calculating mean background
       thresh_value = thresh_coeff*avg_background;  %setting threshold value
       threshold=thresh_value-avg_background;
       Diff = (BG_subtract(BG_subtract>threshold)-threshold)/1000;  %
determining values above threshold
       Diff2 = BG_subtract/1000;    % determining values above background (no
threshold)
       numMoles=[numMoles sum(Diff)*(ymax-ymin)];      % calculates
integrated intensity over threshold
       numMoles2=[numMoles2 sum(Diff2)*(ymax-ymin)];      % calculates
integrated intensity over background
    end


end
%% Plots and exporting data
% Number of moles
numMoles=double(numMoles);
figure(4)
xtime=1:n;
time=xtime*exposure;
smoothnumMoles=smooth(numMoles,0.1,'lowess');   % smoothing filter
plot(time,smoothnumMoles)   %plotting time vs integrated moles
xlabel('time (seconds)'); ylabel('Number of Moles Thresh');

figure(5)
numMoles2=double(numMoles2);
smoothnumMoles2=smooth(numMoles2,0.1,'lowess');
plot(time,smoothnumMoles2)
xlabel('time (seconds)'); ylabel('Number of Moles No Thresh');


% Spatio temporal plot
xmmEnd = double(xmax*pixelDim*10^-3);%mm
maxTime=j*exposure;
spatioX = 1:xmmEnd;
spatioY = 1:maxTime;
 figure(6)
 imagesc(spatioX, spatioY, Gaussian2);
 axis xy
 xlabel('Distance (mm)')
 ylabel('Time (seconds)')


% excel data output
```

```
datawrite=[smoothnumMoles smoothnumMoles2]; %smoothnumMoles3];
xlswrite(namefile,datawrite)
%displaying values in workspace
disp('500 sec'),uint16(smoothnumMoles(500))
disp('750 sec'), uint16(smoothnumMoles(750))
disp('1000 sec'), uint16(smoothnumMoles(1000))
disp('1200 sec'), uint16(smoothnumMoles(1200))
```