Pilot study comparing technologies to test for substandard drugs in field settings

Roger Bate1,2*, Richard Tren2,3, Kimberly Hess2, Lorraine Mooney4 and Karen Porter1

1American Enterprise Institute, 1150 Seventeenth Street, NW, Washington, DC 20036, USA.
2Africa Fighting Malaria, 1050 Seventeenth Street, NW, Suite 590, Washington, DC 20036, USA.
3Africa Fighting Malaria, P. O. Box 17156 Congella, 4013, South Africa.
4Africa Fighting Malaria, 4 Church Lane, Barton, Cambridge, CB3 7BE, UK.

Accepted 13 March, 2009

Researchers procured a range of antimalarial, antibiotic and antimycobacterial drugs from cities in six countries: Ghana, India, Kenya, Nigeria, Tanzania, and Uganda. Semi-quantitative thin-layer chromatography (TLC) and disintegration tests, Raman spectrometry, and near-infrared (NIR) spectrometry were used to measure the concentration of active ingredients and excipients (spectrometry only) to determine whether the tested samples were of good quality. Overall, 15% of tested samples failed TLC, 13% of tested samples failed disintegration tests, 41% of tested samples failed NIR spectrometry, and 47% of tested samples failed Raman spectrometry. The drug testing technologies were qualitatively compared in terms of time, cost, and reliability for identifying substandard drugs in the field. NIR and Raman spectrometry compared favorably to TLC in most respects except cost. If the indirect costs of TLC—including requirements for a climate controlled location and trained laboratory staff—are considered, the cost advantage of TLC may disappear in developing countries.

Key words: Raman and near-infrared spectrometry, thin-layer chromatography, counterfeit and substandard drug production, regulation of drug quality.

INTRODUCTION

The World Health Organization estimates that up to 200,000 of the one million deaths that occur from malaria each year could be avoided if antimalarial drugs were “effective, of good quality and used correctly” (World Health Organization, 2003). In May 2008, some of the authors published a study that found 35% of antimalarial drugs sold in private shops and pharmacies in six major African cities failed basic quality control tests (Bate et al., 2008). Additionally, tuberculosis and other bacterial infections cause millions of deaths a year; drugs to combat these diseases are also routinely counterfeited (World Health Organization, 2008).

Portable labs that perform thin-layer chromatography (TLC) provide a relatively inexpensive, versatile, and robust means of identifying substandard drugs at a fraction of the resources required for modern laboratory testing. Over 300 Global Pharma Health Fund e.V. Mini-labs (GPHF-Minilab®) are being used in 70 countries to help public authorities and private companies identify counterfeit and substandard drugs (Global Pharma Health Fund). TLC however, requires trained staff and may be time consuming.

New technologies are making it easier to test the authenticity of drugs in field settings. This paper compares two instruments that use the technologies of Raman spectrometry and near-infrared (NIR) spectrometry against TLC and disintegration testing to identify substandard drugs in the field.

MATERIALS AND METHODS

Phazir RX produced by Polychromix (Wilmington, Massachusetts, USA) utilizes NIR spectrometry to excite molecules in a material and then captures the unique pattern of vibrations emitted. The pattern, also referred to as the “fingerprint,” can be compared to a predetermined reference standard based on quantitative and qualitative attributes such as optical resolution, wavelength accuracy, wavelength range, signal to noise ratio, and linearity of the NIR platform, typically in less than five seconds (Polychromix). The reference standard can be created by a user with a sample (or
preferably multiple samples to allow for minor variations) of authenti-
cific product or by comparing the “fingerprint” to an electronic data-
based of excipients provided by Polychromix (based on samples
received from individual manufacturers).

An alternative spectrometer, the TruScan by Ahura Scientific
(also based in Wilmington, Massachusetts, USA), collects Raman
spectra to characterize the individual chemical components of a
material. Raman utilizes laser photons to excite molecules, and studi-
es low frequency modes. Based on the Raman effect, the Tru-
Scan measures the interaction of light and molecular bonds. Dif-
ferent bonds create peaks of varying intensity resulting in a spec-
trum that is a unique “fingerprint” (Witkowski, 2005). After creating
the spectral fingerprint, the Raman spectrometer automatically com-
pares it to the spectral reading from the reference method (ass-
essed from one or preferably more samples). Raman spectrometry
provides specific qualitative information on the identities of analy-
tes, characterization of sample matrices, and the molecular spectro-
metry used to analyze unknowns in a solid-state analysis (Witkow-
ski, 2005). “Point and shoot” testing against the reference method
takes approximately 30 s (Sherma, 2007). If the material assessed
against the method fails the first test, the TruScan provides a “Dis-
coverty” mode that accesses TruScan’s database of drugs and
chemical substances to determine the material’s identity (Sherma,
2007).

Both types of spectrometers can test through container liners and
glass vials in situ, allowing materials to be analyzed without chemi-
ical preparation or destruction of product. Manufacturers of spectro-
meters provide databases of drugs and chemical substances, which
would be updated via the Internet to ensure the availability of new and
authentic references. However, methods may not be available yet
for specific classes of drugs that are vital to developing countries,
such as antimalarial drugs, or for specific drugs produced by non-
Western companies, which may use different excipients and/or
coatings. Manufacturers report that many companies rely on their
own databases and reference standards. NIR and Raman spectrom-
etry are particularly useful for identifying what a drug is not—does
it match the standard of a given manufacturer and brand?—rather
than what it is.

The GPHF-Minilab® can be used to run semi-quantitative TLC
disintegration tests on samples to determine the presence and
relative concentration of active ingredients. This technology is well-
established in the literature for field assessments (Bate et al.,
2008). Each test is duplicated, with the generous assumption that the
result more consistent with the reference is recorded.

For Africa Fighting Malaria’s analytical drug quality work and for
this pilot study, a range of antimalarial, antibiotic and anticyclobac-
terial drugs were collected from cities in six countries, namely:
Ghana, India, Kenya, Nigeria, Tanzania, and Uganda. The simple
sampling protocol was developed in line with previously published
research (Lon et al., 2006; Bate et al., 2008). Treatment packs were
obtained by local nationals from randomly selected private pharma-
cies in major cities. Local nationals posed as customers and pur-
chased treatment packs on behalf of the project. Before assessing
antimalarial drugs, all of which are commonly available without a
prescription and regarded by the WHO as essential drugs. Treat-
ment packs were maintained in their original packaging as sold:
ether the manufacturer’s original packaging or loose.

Once the drugs were transported to the final field testing location
(outside a laboratory setting) in the United Kingdom, the drugs were
kept in ideal conditions: stored at ambient temperature, in low
humidity and away from sunlight.

Primary screening of samples was conducted at the United
Kingdom location in July 2008 using the GPHF-Minilab® protocol as
described above. Secondary screening of samples was conducted
in July and August 2008 using Raman spectrometry (TruScan) and
NIR spectrometry (Phazir) by adhering to protocols established both
by the manufacturers and by previously published research discern-
ing counterfeit from legitimate antimalarials in situ (Ricci et al.,
2007; Frosch et al., 2007; Ricci et al., 2008). Because blister pack-
ging material ranges in thickness and transparency, two tablets
from each treatment pack were removed from the packaging before
being subjected to spectrometer analysis.

TLC screening was based on the GPHF-Minilab® protocol, which
awards products a “pass” if 80% or more of the labeled active ingre-
dient(s) is present. In the United States, once the Food and Drug
Administration approves a drug formulation it allows for a 5% varia-
tion in contents post approval (US Food and Drug Administration,
1995). A more significant change is deemed unacceptable and the
product is not assumed to be bioequivalent for the patient. Conse-
quently, since both the NIR and the Raman spectrometers used in
this pilot study were set to United States pharmacopeial stand-
ards, product ranges had to be within 95-105% to be awarded a
“pass”. Due to this lower variation in standards, the NIR and Raman
spectrometers are likely to reject more samples than TLC.

Additionally, the spectral fingerprint for each spectrometer is of
the entire tablet (including excipients), whereas TLC measures only
the active ingredient(s). This makes method establishment more
complex with the spectrometers since excipients can be different for
two products that are bioequivalent. For a spectrometer to be use-
ful, one must establish a good quality product from each individual
manufacturer (since different manufacturers could use slightly differ-
ent excipients) before assessing products from the field. If refer-
ence samples from the manufacturer are not easily available,
methods may be created from samples collected in the field that
pass TLC (and preferably high performance liquid chromatography
or another more precise lab-based method such as mass spectro-
metry). For this pilot study, samples were collected either directly
from the manufacturers or from the GPHF, which collected them
from the manufacturers. Because few companies willingly provide
samples to researchers, the sample size in this study is small.
Method creation using field samples will be essential for field use
of spectrometers in developing countries, where there are hundreds of
manufacturers and brands.

RESULTS

78 treatment packs were tested comprising antimalarial
(amodiaquine, fixed-dose combination artemether-lume-
nfantrine, artesunate, dihydroartemisinin, mef-
loquine, sulfadoxine-pyrimethamine (SP), and chloroqui-
ne), antibiotic (erythromycin and ciprofloxacin), and anti-
mucobacterial drugs (isoniazid and rifampicin). Overall,
15% (12/78) of tested samples failed TLC, 13% (10/78) of
tested samples failed disintegration tests. 41% (32/78) of
tested samples failed NIR spectrometry (Phazir), and
47% (37/78) of tested samples failed Raman spectro-
metry (TruScan) (Table 2 and Figure 1). Overall, 10%
(8/78) of samples failed all four tests and 49% (38/78) of
samples passed all four tests. Nine samples had differ-
ent results for the NIR (Phazir) and Raman (TruScan)
spectrometry.

DISCUSSION

As expected, the spectrometers, which operate to more
exacting standards, failed more samples than the less
exacting methods of TLC and disintegration testing; nev-
ertheless, even these methods failed a substantial minori-
ty of sampled drugs. Although the sample size is too
small to draw definitive conclusions, these results along
with numerous other studies (Minzi et al., 2003; Amin et
Table 1. Comparison of NIR (Phazir) and Raman (TruScan) Spectrometry.

<table>
<thead>
<tr>
<th>Description</th>
<th>Near-infrared (NIR) spectrometry</th>
<th>Raman spectrometry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Phazir model produced by Polychromix</td>
<td>TruScan model produced by Ahura Scientific</td>
</tr>
<tr>
<td>Description</td>
<td>Stimulates sample molecules with near-infrared light and measures vibrations to obtain a unique fingerprint for the compound</td>
<td>Stimulates sample molecules with a laser and measures vibrations to obtain a unique fingerprint for the compound</td>
</tr>
<tr>
<td></td>
<td>Requires a dipole moment change</td>
<td>Requires a polarizable change</td>
</tr>
<tr>
<td></td>
<td>No laser related safety concerns or regulatory restrictions</td>
<td>Due to monochromatic light produced by a powerful laser the item must be cleared with Customs upon entering some countries</td>
</tr>
<tr>
<td>Weighs 4 pounds (1.8 kg)</td>
<td></td>
<td>Weighs less than 4 pounds (1.8 kg)</td>
</tr>
<tr>
<td>Operates in 5 to 45°C</td>
<td>Operates in -20 to 40°C</td>
<td></td>
</tr>
<tr>
<td>10-hour quick change battery</td>
<td>Minimum 5-hour battery life at 25°C</td>
<td></td>
</tr>
<tr>
<td>Application</td>
<td>Absorption with an NIR is based mainly on the overtones of C-H bonds, making this instrument less reliable for information-rich identification</td>
<td>Well-suited for symmetric vibrations in aromatic molecules, such as -S-S- bonds and C double bonds, seen in many pharmaceutical drugs</td>
</tr>
<tr>
<td></td>
<td>Limited ability to penetrate through packaging (co-blisters must be removed)</td>
<td>Laser can penetrate most translucent surfaces</td>
</tr>
<tr>
<td></td>
<td>Sensitive to changes in ambient light</td>
<td>Laser strength makes Raman less sensitive to external factors</td>
</tr>
<tr>
<td></td>
<td>Must control for humidity changes, sample position, and sample face (for tablets) or perform multiple tests</td>
<td>Controls for humidity, sample position and sample face</td>
</tr>
<tr>
<td></td>
<td>Produces results in approximately 5 seconds</td>
<td>Produces results in approximately 30 seconds</td>
</tr>
<tr>
<td>Both</td>
<td>Allow drug identification to be carried out rapidly in the field with minimal training and no sample preparation.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Fingerprints” materials without having to use external substances and assesses active pharmaceutical ingredients, excipients, fillers, dyes, and coatings.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moisture content can be ascertained.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Allows method creation with authentic sample not located in the drug database.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>21 CFR Part 11 Compliance Documentation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Provide qualitative, as well as quantitative information about the material in question and are non-invasive.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>With proper methods established, for most drugs, more accurate than the field standard GPHF-Minilab®.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No on-going cost for consumables or service.</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Testing results by formulation for TLC, disintegration, NIR and Raman.

<table>
<thead>
<tr>
<th></th>
<th>TLC (17%)</th>
<th>Disintegration (25%)</th>
<th>NIR (33%)</th>
<th>Raman (50%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amodiaquine</td>
<td>1/6</td>
<td>1/6</td>
<td>2/6</td>
<td>3/6</td>
</tr>
<tr>
<td>Artemether-lumefantrine fixed-dose combination</td>
<td>3/8 (38%)</td>
<td>2/8 (25%)</td>
<td>4/8 (50%)</td>
<td>5/8 (63%)</td>
</tr>
<tr>
<td>Artemether</td>
<td>1/3 (33%)</td>
<td>0/3 (0%)</td>
<td>2/3 (67%)</td>
<td>2/3 (67%)</td>
</tr>
<tr>
<td>Artesunate</td>
<td>1/12 (8%)</td>
<td>2/12 (17%)</td>
<td>4/12 (33%)</td>
<td>5/12 (42%)</td>
</tr>
<tr>
<td>Chloroquine</td>
<td>1/7 (14%)</td>
<td>1/7 (14%)</td>
<td>3/7 (43%)</td>
<td>3/7 (43%)</td>
</tr>
<tr>
<td>Dihydroartemisin</td>
<td>1/5 (20%)</td>
<td>1/5 (20%)</td>
<td>3/5 (60%)</td>
<td>3/5 (60%)</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>0/6 (0%)</td>
<td>0/6 (0%)</td>
<td>2/6 (33%)</td>
<td>2/6 (33%)</td>
</tr>
<tr>
<td>Isoniazid</td>
<td>0/7 (0%)</td>
<td>0/7 (0%)</td>
<td>2/7 (29%)</td>
<td>3/7 (43%)</td>
</tr>
<tr>
<td>Mefloquine</td>
<td>0/4 (0%)</td>
<td>0/4 (0%)</td>
<td>1/4 (25%)</td>
<td>2/4 (50%)</td>
</tr>
<tr>
<td>Rifampicin</td>
<td>1/6 (17%)</td>
<td>0/6 (0%)</td>
<td>3/6 (50%)</td>
<td>4/6 (67%)</td>
</tr>
<tr>
<td>Sulfadoxine-pyrimethamine</td>
<td>3/10 (30%)</td>
<td>3/10 (30%)</td>
<td>5/10 (50%)</td>
<td>5/10 (50%)</td>
</tr>
<tr>
<td>Ciprofloxacin</td>
<td>0/4 (0%)</td>
<td>0/4 (0%)</td>
<td>1/4 (25%)</td>
<td>0/4 (0%)</td>
</tr>
<tr>
<td>Total</td>
<td>12/78 (15%)</td>
<td>10/78 (13%)</td>
<td>32/78 (41%)</td>
<td>37/78 (47%)</td>
</tr>
</tbody>
</table>

1 Percentages are supported by total that failed testing/total treatments tested.
al., 2005; Bate et al., 2008) indicate a substandard drug problem for some developing countries.

Both NIR (Phazir) and Raman (TruScan) spectrometers allow drug identification to be carried out rapidly in the field with minimal training and no physical sample preparation. Both spectrometers “fingerprint” materials without using external substances. The spectra generated will reflect all contents of the sample: active pharmaceutical ingredients, excipients, fillers, dyes, and coatings. The spectra will change when any of these contents is changed or is inherently different due to different manufacturers producing drugs with different concentrations of excipients, and perhaps entirely different excipients. Furthermore, temperature degradation or moisture degradation of a sample will affect the spectra, which is critical when assessing the viability of compounds such as artemisinin, whose effectiveness is lowered by moisture. NIR and Raman spectrometry provide qualitative as well as quantitative information about the material in question and are non-invasive. A comparison of NIR and Raman spectrometry can be found in Table 1.

It was simple to establish a standard for each drug type using the samples provided by the manufacturers and calibration of both instruments was fast and easy. Early experimental work of the spectrometers revealed that the chosen NIR spectrometer (Phazir) was ergonomically more pleasing and slightly easier to use than the chosen Raman spectrometer (TruScan). It took less than one hour to produce 250 images of 50 samples with the NIR (Phazir) (on average five runs were used to confirm a result). The NIR (Phazir) allows users to change the relative importance of various aspects of the spectral measurements in the creation of methods to make it more precise in recording a “pass” or a “fail”.

Prior to initiating drug testing, the authors experimented with the spectrometers to determine each one’s ability to test through packaging. It was discovered that the Phazir’s ability to penetrate through various packaging surfaces was variable and some tablets had to be removed from their blister packaging in order to produce consistent readings. This undermines a key attribute of the Phazir spectrometer—that it can assess through many packaging materials with no obvious change to the sample. Therefore during drug testing, all samples were removed from the packaging for quality control purposes. The Phazir was also sensitive to surrounding light and results were therefore altered if the ambient light changed significantly, so testing was done in a light controlled environment. Another drawback that has been noted with NIR spectrometry is that humidity changes, sample position, and sample face (for tablets) need to be controlled for in order for the results to be acceptable (Deisingh, 2005). With the exception of the need to remove samples from the packaging, early experimental work of the spectrometers found that most of these issues could be overcome with repeated testing—that is, aiming the instrument at several sides of the tablet and conducting tests over a period of time to allow for changes in humidity—but they did weaken one of the greatest advantages of NIR: its speed.

The Raman spectrometer (TruScan) took approximately 1 h to run 50 samples during field testing. It took up to 30 s to produce results, but in situ and field readings were less susceptible to environmental interference than those of the NIR (Phazir) due to the intensity of the monochromatic light produced by a laser within the Raman instrumentation. The high-powered laser component in the instrumentation requires the Raman (TruScan) to be registered with the Customs authority in some foreign countries, while the NIR (Phazir) has no laser-related safety concerns or regulatory restrictions (Ahura Scientific, Inc., 2007; Australian Government, 2008). In some circumstances Customs may delay the release of the Raman spectrometer, and delays could be more severe or problematic in countries with less sophisticated customs procedures. The authors did not encounter significant delays in transporting the Raman spectrometer.

Ahura states that the rechargeable Lithium battery in the TruScan lasts for approximately 5 h and can operate at a wide temperature range. According to Polychromix, the Phazir will operate in almost as wide a temperature range and its battery life is 10 h. The authors’ own observations broadly confirmed these assertions, although the batteries were not run to exhaustion. Both instruments weigh around four pounds (1.8 kg) and are user-friendly, requiring very little time for the authors to become acquainted with their features. Data transfer to the computer was easier with the Phazir than the TruScan, with the former having a simple Universal Serial Bus (USB) attachment and fast data transfer. The TruScan requires a non-standard adaptor for data transfer, which took several minutes to work during testing but is designed to ensure data are never lost in transfer.

According to some scientific literature, Raman spectro-
metry is preferred for symmetric vibrations that are present in aromatic molecules, -S-S- bonds and C double bonds, as seen in many pharmaceutical drugs. Raman may also be more reliable than NIR under circumstances likely to be found in pharmaceutical analysis (McCreery et al., 1998). However, some drugs such as SP, an anti-malarial drug, have considerable fluorescence and using Raman spectrometry to test these types of drugs could be problematic and invalidate its assessment (this is a type 2 error – where SP is passed when its active ingredients are too low). This was not an obvious problem in the small number of SP samples that were analyzed.

When choosing a testing technology to be used in the field, time, cost, reliability, and usability must all be taken into account. The results of this study suggest that both NIR and Raman spectrometry compare favorably to the established standard set by TLC with the GPHF-Minilab® in most respects except cost. While TLC is relatively sensitive, specific, and accurate, the sample preparation and analysis may be time consuming and requires user patience and attention to detail or results can be biased. Further, TLC may have limited use in the field because of competency and training that is required to perform TLC and interpret the results. Additionally, TLC requires a dedicated climate controlled location with potable water and electricity. These barriers may make this technology more difficult to use in typical developing country settings, such as malaria-affected areas, and implies that TLC can only be used by organizations (e.g. Departments of Health) that have sufficient staff dedicated to drug testing, which undermines the cost advantage of TLC.

One critical advantage of TLC is that it is established within the academic literature, which means its results are more accepted by government agencies. In other words, both NIR and Raman spectrometry have to be established as accurately robust field technologies in the literature before their results will be accepted without question. Currently, the technologies are only being used by the military or private companies; the United States military, for instance, uses the technology to assess dangerous products, and drug manufacturers use it to assess fake copies of their products in the field.

With experience, each technology can be made to work well in the field. For example, the more exacting standards of the spectrometers can be lowered in order to “pass” samples with slightly more variation in spectra, which allows more borderline product approval (to mirror TLC). Of course, the simplicity of the spectrometers is lost if more discretionary interpretation is required from the user. If the market for handheld spectrometers increases in developing countries, then manufacturers could lower measurement standards (to more closely mirror TLC), thereby not requiring specialist adaptation in the field. This is not to say that developing countries should accept lower quality of drugs, just that primary field screening could mirror TLC to unearth drugs with very low or absent active ingredients rather than borderline substandard drugs.

Overall, choice of technology will come down to a variety of factors: how quickly results are required (spectrometry is generally quicker; however, given that different but bioequivalent products produce different spectra, methods must be established for all new brands, which means the initial setup time can be longer for spectrometers); cost (TLC is less expensive – at most $10,000 for a fully equipped lab and training costs for one person, compared with approximately $50,000 for a spectrometer and training); reliability of results to an uninitiated user (Raman spectrometry is generally more reliable); ease of transport (spectrometry is more easily transported); and transport across borders (NIR does not contain a laser and therefore involves the least bureaucracy). While spectrometers have advantages in terms of ease of transportation, since the GPHF-Minilab® includes potentially dangerous reagents, it should be noted that GPHF staff are adept at handling the paperwork necessary for such transportation.

For most resource-constrained developing countries, the GPHF-Minilab® is the product of choice based on cost. However, as aid agencies import and purchase more drugs for developing countries (notably those to combat HIV/AIDS, tuberculosis and malaria), and individuals and governments of the developing world become wealthier and purchase more pharmaceuticals, it might not be long before NIR and Raman spectrometers are deployed in even the poorest countries.

ACKNOWLEDGEMENTS

The authors would like to thank Richard Jaehnke of the German Pharma Health Fund for his advice on using the GPHF-Minilab®, as well as Robert Brush and Julien Bradley (and other scientific staff) of Ahura Scientific, Igor Nazarov of Polychromix, and Sharon Flank of InfraTrac. The authors would also like to thank Philip Coticelli and Rachel Mitrovich for their background research and editorial advice. This research was funded by the Legatum Institute and the Ohrstrom Foundation through grants made to Africa Fighting Malaria.

REFERENCES


