Clinical Chemistry 60:3
463–470 (2014)

Invention and Validation of an Automated Camera System That Uses Optical Character Recognition to Identify Patient Name Mislabeled Samples

Charles D. Hawker,[1,2*] William McCarthy,[3] David Cleveland,[4] and Bonnie L. Messinger[1]

BACKGROUND: Mislabeled samples are a serious problem in most clinical laboratories. Published error rates range from 0.39/1000 to as high as 1.12%. Standardization of bar codes and label formats has not yet achieved the needed improvement. The mislabel rate in our laboratory, although low compared with published rates, prompted us to seek a solution to achieve zero errors.

METHODS: To reduce or eliminate our mislabeled samples, we invented an automated device using 4 cameras to photograph the outside of a sample tube. The system uses optical character recognition (OCR) to look for discrepancies between the patient name in our laboratory information system (LIS) vs the patient name on the customer label. All discrepancies detected by the system's software then require human inspection. The system was installed on our automated track and validated with production samples.

RESULTS: We obtained 1 009 830 images during the validation period, and every image was reviewed. OCR passed approximately 75% of the samples, and no mislabeled samples were passed. The 25% failed by the system included 121 samples actually mislabeled by patient name and 148 samples with spelling discrepancies between the patient name on the customer label and the patient name in our LIS. Only 71 of the 121 mislabeled samples detected by OCR were found through our normal quality assurance process.

CONCLUSIONS: We have invented an automated camera system that uses OCR technology to identify potential mislabeled samples. We have validated this system using samples transported on our automated track. Full implementation of this technology offers the possibility of zero mislabeled samples in the preanalytic stage.

Mislabeled samples and patient identification errors are a serious problem in most, if not all, clinical laboratories *(1–8)*. In the preanalytic sequence, which was the focus of this study, mislabeling can occur at multiple points. A few of these points include patient misidentification at the time of collection, use of handwritten labels, labeling mix-ups immediately before or after collection, mislabeling during laboratory accessioning and aliquoting, and relabeling samples that already have an existing label from another system, such as may occur in a core laboratory or a reference laboratory.

In a 2002 review, Bonini et al. *(2)* reported that sample misidentification accounted for more laboratory errors than any other source. In 3 Q-Probe studies conducted by the College of American Pathologists (CAP),[5] the reported rates of mislabeled samples were 0.39/1000 in 120 institutions *(4)*, 0.92/1000 in 147 clinical laboratories *(7)*, and 1.12% in 122 clinical laboratories for blood bank samples *(8)*.

A discussion on the CAP website *(9)* entitled "When a Rose Is Not a Rose: the Problem of Mislabeled Specimens" has several suggestions to guide laboratories with procedures that minimize mislabel risk. This discussion also cites a poster (Kahn S, et al. "Improving Process Quality and Reducing Total Expense Associated with Specimen Mislabeling in an Academic Medical Center." Institute for Quality in Laboratory Medicine Conference, 2005) in which the authors determined a hypothetically incurred charge of a mislabeled sample at $712, if the patient payers had been rebilled for the additional resources used because of the mislabeling of the sample. This estimate did not include immeasurable costs such as patient anxiety and discomfort and delays or errors in diagnosis and treatment. The CAP discussion used the lowest identification rate cited above *(4)* of 0.39/1000 to estimate that mislabeled

[1] ARUP Laboratories, Salt Lake City, UT; [2] Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT; [3] Cognex Corporation, Natick, MA; [4] Custom Engineering Solutions, Fraser, CO.
* Address correspondence to this author at: ARUP Laboratories, 500 Chipeta Way, Salt Lake City, UT 84108-1221. Fax 801-584-5207; e-mail hawkercd@aruplab.com.

Received September 9, 2013; accepted November 21, 2013.
Previously published online at DOI: 10.1373/clinchem.2013.215434
[5] Nonstandard abbreviations: CAP, College of American Pathologists; OCR, optical character recognition; LIS, Laboratory Information System; FOV, field of view; QA, quality assurance.

samples may add as much as $280 000 in healthcare costs for each 1 million samples tested.

There has been a focus on barcodes as a method for reducing sample labeling errors. CLSI published a standard, AUTO02-A2 *(11)*, specifying that clinical laboratories are to use barcodes of the symbology Code 128. However, not all do, as codes 3 of 9 and Interleaved 2 of 5 are also used *(12)*. Inconsistent barcode quality has also been cited as a factor with various point-of-care devices in wrist band error *(13)*. On the other hand, a Laboratory Medicine Best Practices consensus committee recently published a systematic review and meta-analysis which concluded that barcoding is effective for reducing patient sample identification errors in diverse hospital settings and is recommended as an evidence-based "best practice" *(14)*.

Another approach focused on facilitating the actions of human effectors in the system through standardization. In 2011, CLSI published standard AUTO12-A to address label formats to standardize the human-readable content *(15)*. Laboratories and other providers were given 3 years to implement the standard, but it is too early to see any positive impact.

In our laboratory, the baseline measure of samples mislabeled in the preanalytic stages of testing was approximately 1/10 000, lower than those noted above *(4, 7, 8)*. Using documented corrections to final reports as a measure of efficacy, we believed that as many as 95% of these preanalytic mislabels were detected and corrected before analysis. Mislabeled samples that escape detection could have 1 of 2 possible outcomes: the test results received by the clients are not sufficiently different from the expected results to elicit an inquiry from the provider, or the erroneous result is not questioned. The last outcome, for which we have no baseline measure, represents the greatest risk to patients.

In 2006, we initiated development of an automated high-speed camera system using optical character recognition (OCR) technology to detect all mislabeled samples. We envisioned that this system might pass those samples in which the patient name printed on the customer label exactly matched the patient name in our Laboratory Information System (LIS), whereas all samples with labels that were not passed by the OCR system would route to a manual inspection location on the automation system. We knew this concept would be a significant challenge because of the large variety of tubes received and because the labels from our clients varied in fonts and in the location of the patient names on the labels.

The concept we initially considered and tested was a line-scan camera system, as we already had devices on our automated track that determined sample routing by rotating the tube past a slit in the automation transport carrier to read the barcode on our LIS label. In this concept, we replaced the barcode reader with the line-scan camera, which created a 2-dimensional photograph of the tube's exterior, read the barcode, and used OCR analysis to read all visible text on the client label. However, the images were not of adequate quality to complete the OCR analysis of the client label text on a sufficiently high percentage of tubes.

To resolve this issue, we tested a system using a robotic vacuum lifting device to lift the tube out of the transport carrier in front of 4 equidistant and simultaneously triggered cameras to photograph the outside of the sample tube. The vacuum lifter used suction on the smooth top of the tube cap to lift the tube, thereby not blocking the view of the tube exterior. Specially designed software stitched the 4 photographs together, creating an unwrapped 2-dimensional photograph of the entire exterior of the tube of sufficient quality to enable OCR analysis. This concept worked well as a prototype and demonstrated the concept feasibility. However, the vacuum lifting device was not consistently reliable and could not handle a wide range of tube sizes and caps.

The next prototype used a 6-axis robot, which firmly grasped the cap with a 3-jaw gripper and could lift a wide variety of tubes. After substantial testing in the engineering shop, the system was transferred to our laboratory and installed on our production track in May 2012. Additional months of problem solving and performance improvements followed, and on October 3, 2012, we initiated a validation study of 1 million samples.

On the basis of the estimated error rate of 1/10 000 discussed above and our estimated preanalytic correction of 95% of those errors, the examination of 1 million images was expected to yield 100 mislabeled samples and 5 corrected reports. These numbers were large enough to determine if those prior estimates were accurate. In addition to assessing these estimates, the goals of the validation study were to show that (*a*) the system would perform at the speed of our automation; (*b*) the system could handle the variety of tubes encountered; (*c*) the images of all mislabeled samples would be "failed" by the OCR analysis; and (*d*) the OCR pass rate would be high enough to minimize the labor required to inspect all failed images in a production setting.

Materials and Methods

ROBOTIC SYSTEM

An Epson® #C3-A6018ST6 table-top 6-axis robot, with RC620+ control and RC+6.0 software, was used to lift the sample tubes into a central position at the common focal point of the 4 cameras (see below). Pneumatically

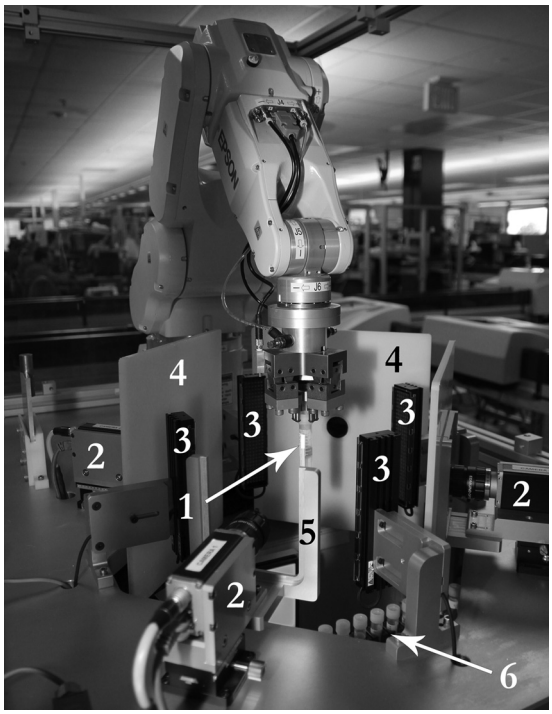

Fig. 1. Robotic and optical components for photographing sample tube exteriors for vision processing and OCR analysis.

Shown are sample tube (1) held by robot gripper, 3 of 4 cameras (2), 4 LED bar lights (3), 3 of 4 diffusing background panels (4), panel mounting frames (5), and sample tubes (6) in transport carriers on the automation conveyor.

operated mechanical stops positioned the transport carrier at precise locations for tube pickup and replacement in the transport carrier. Fig. 1 is a photograph of the robot and the optical system as described in the next section.

Each sample was lifted by its cap to allow a 360° unobscured view of the tube. To accommodate a wide variety of cap shapes and dimensions, the robot end effector was fitted with a 3-jaw parallel linkage gripper, each jaw containing 2 parallel posts spanned by a rubber o-ring. This gripper design was found sufficient to securely grip all tested caps.

A separate camera (Cognex® In-Sight® Vision System, 1100 Micro series, 640 by 480 pixels, 60 frames per second) was used with a backlight (Cognex ICRB-100100 red LED backlight, 635 nm, 100 mm square) to cast a profile of each arriving tube. From this the tube height was measured and sent as coordinates to the robot such that it would grip each cap approximately 4 mm below its top. In the event that a tube might not be lifted perfectly vertical, the actual angle as measured by the 4-camera array was sent to the robot controller with which a correcting motion was made to vertically align the tube for image capture and submission to the label unwrapping process. This vertical correction was also used to ensure reinsertion in the automation transport carrier. Fig. 2 is a photograph of the smart camera and mechanical system described in this paragraph and the robot gripper described in the preceding paragraph.

VISION AND OCR SYSTEMS

The vision (optical) system (Cognex Omniview®) used 4 high-resolution color cameras (Basler GigE 5 megapixel) fitted with 5-megapixel 12-mm lenses (Edmund Optics) and polarizing filters (Midwest Optics) to capture simultaneous images from all sides of the sample tube. The cameras were offset 90° to one another and mounted horizontally. Strobe illumination was provided through 4 bar lights fitted with polarizers (CCS America) and a 4-channel lighting controller (Gardasoft). A custom PC-based application was written in C# (Microsoft) to handle messaging, data archive, process inspection, and pass/fail logic.

The robotic system and the vision system communicate via Windows Communication Foundation. The robotic system sends a message to the vision system that a sample is in the field of view (FOV), and the vision system then acquires all 4 images simultaneously. Each image is analyzed to ensure the sample is oriented vertically in the FOV. If the sample is not perfectly vertical (or within some degree of acceptable error), the vision system will compute and send correction coordinates back to the robot system. After the vision system has accepted a vertical alignment, all 4 images are blended together to create a single unwrapped image containing the entire 360° circumference of the sample. A barcode on the unwrapped image is used to query an LIS database to obtain the patient name. The unwrapped image is analyzed to determine if there is a single label or multiple labels contained on the sample. An automatic pass is determined when a single label is recognized and no further analysis is needed. These are called singles. When multiple labels are found, the customer label portion of the image is cropped out and analyzed by the OCR engine (ABBYY FineReader®). The OCR engine returns character data to the vision system. The vision system then searches the character data for the patient name. If the patient name is not found in the character data, the sample is classified as a fail. If the patient name is found, the sample is classified as a pass.

A video of the OCR system in operation accompanies the online version of this article at http://www.clinchem.org/content/vol60/issue3.

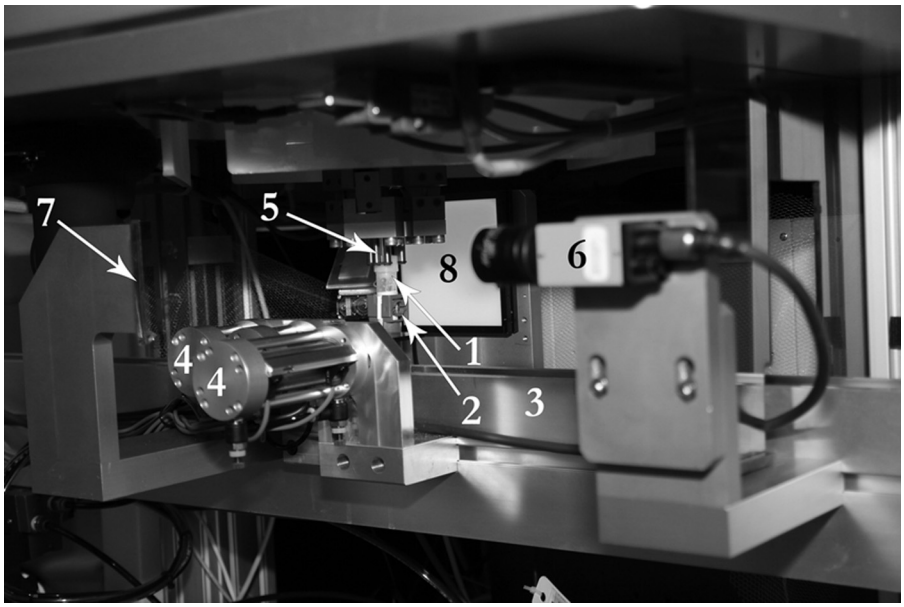

Fig. 2. Robotic components to facilitate correct pick-and-place operations for the OCR system.
Shown are a sample tube (1) in an automation transport carrier (2) on the automation conveyor (3), pneumatically operated mechanical stops (4), robot gripper (5), along with the smart camera (6), 45° mirror (7), and backlight (8) which together analyze cap height and diameter.

Results

The validation study of 1 million images was terminated on May 31, 2013, with 1 009 830 total images collected by the OCR system. This was only a fraction of our laboratory's workload during that time, since there was just the single prototype OCR system on 1 of 4 automated tracks. Moreover, not all samples on the track had images collected, since the robotic system was undergoing continual improvement. Nevertheless, the total of 1 009 830 images constituted a valid data set for this study. The speed of the system, measured from pickup to replacement of the tube in the automation transport carrier, was approximately 3 s, not quite at the desired speed of 2 s. The reliability of the robotic system was acceptably similar to other robotic systems in our laboratory, with only a minimal number of jams or service interruptions per week, despite the range of tube sizes handled (from 12 by 75 mm to 16 by 110 mm) and the use of a transport carrier not designed for robotic pick-and-place operations.

Each of the 1 009 830 images was reviewed by the lead author or by 1 of 4 technicians temporarily assigned to the project. This employee review was normally conducted 1–3 days after the image was obtained, based on availability of staff. Its purpose was to determine that the OCR system had classified each image according to the laboratory's policies. It was not designed to prevent errors or change the flow of production. The study objective was to validate system performance, and time to inspection was not a study parameter. Once the OCR system has been validated and placed in production, all review of potential mislabeled samples would occur before analysis.

As described in Materials and Methods, the system classified images into 1 of 3 categories on the basis of the OCR analysis: passes, singles, and fails. Although placed in a separate category by the system, singles were automatically considered passes. Fig. 3, A and B, shows the operator's screen on the OCR system for fictitious examples of images that are, respectively, a pass and a fail. Fig. 4 shows the images of the first actual mislabeled samples detected by the OCR system with protected health information redacted. These mislabeled samples were not detected by existing quality assurance (QA) procedures. After detection by the OCR system, we investigated and then issued corrected reports to the client according to procedure.

In addition to mislabeled samples or spelling discrepancies, there were many reasons labels could be classified as a fail by OCR. These included nonstandard fonts not trained in the OCR system, poor-quality client labels, solid colored or striped labels instead of white labels, name truncations, marks on the label that

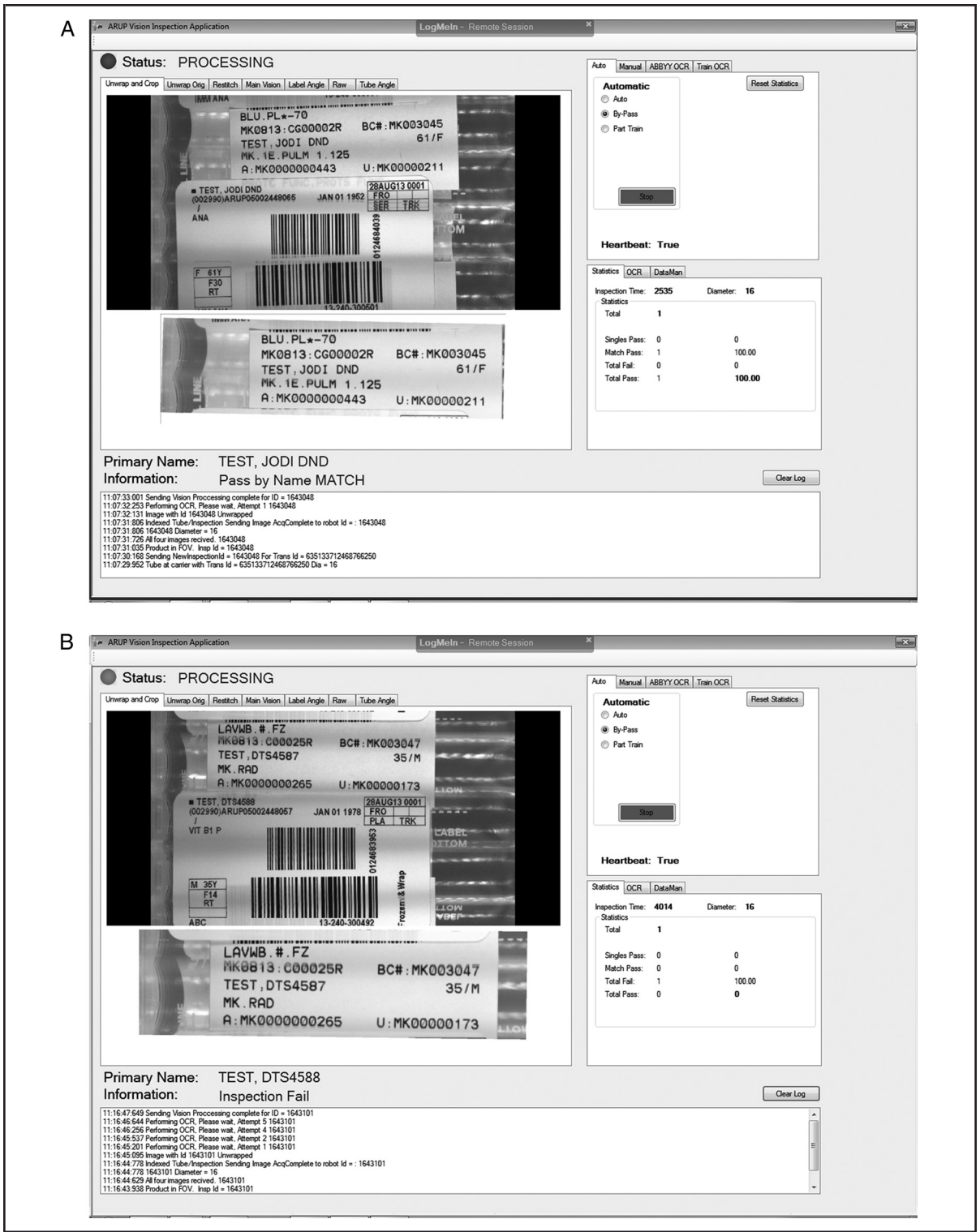

Fig. 3. Screen shots of the OCR system's operator's screen for examples of a pass image (A) and a fail image (B). In the latter example, the alphanumeric name on the client label was intentionally increased by one digit from . . .4587 to . . .4588 in our laboratory's LIS to demonstrate that a discrepancy of one character in the string will cause a fail.

Fig. 4. Photographic images of the labels on the first pair of mislabeled samples identified by the OCR system.
Brenda and Linda are twin sisters with identical surnames whose samples were drawn on the same date by the same physician for the same test. The upper names shown are on the client's label, the lower names are on our laboratory's label.

touched the patient name, handwritten labels, patient names turned 90° to the length of the tube, patient names split as two lines of text, and patient names on the client label partially covered by our laboratory's LIS label. However, human inspection can immediately see that these OCR fails are, in fact, not mislabeled samples.

Of the total of 1 009 830 images in the study, there were 562 286 passes and 180 691 singles for a total of 742 977 passed images (73.6%) and 266 853 fails (26.4%). In the first 3 weeks of the study, the pass rate was low until adjustments were made to the system. Likewise, in May, after some engineering changes, the system performed poorly for about 10 days, until adjustments were made. If those 31 days are excluded, the resulting pass rate on 833 934 total images was 75.3%.

Of the combined total of 742 977 images that were passes and singles, there were zero samples mislabeled by patient name and zero samples with patient name spelling discrepancies. This confirms that OCR is not classifying sample identification errors as passes.

Included in the 266 853 images classified as fails by OCR were 121 true patient name mislabels, resulting in a ratio of 1 mislabel per 8346 images, slightly higher than our prestudy estimate of 1/10 000. Of the 121 mislabeled samples identified, 71 had been found and corrected before analysis by our QA process and 4 were duplicate samples that were not tested; notifications for the remaining 46 were forwarded to the testing laboratory for investigation. Of these 46, technical staff determined that 25 did not require a correction to the final report (per laboratory policy, the clinical or statistical significance of the result did not change), whereas 21 led to corrected reports, higher than our prestudy estimate of 5 corrected reports in 1 million images. Thus,

for the series of samples inspected by OCR, the rate of error not detected by our QA protocols was approximately 4.6/100 000, of which 2.1/100 000 had the potential to affect patient care.

The 266 853 images that were fails also included 148 instances of spelling discrepancies between the patient name in our LIS record and the name printed on the client label. Often this was a single letter (e.g., Kristen vs Kristin). In the current QA system, if the identification is close and can be confirmed with a secondary identifier (i.e., medical record number), the sample is not rejected. When a secondary identifier is not present, the laboratory's policy requires that the correct name be verified with the client. Of the 148 spelling discrepancies found among the OCR fails, 46 were found by our QA process before analysis. The remaining samples were investigated and identification confirmed per policy after OCR found the discrepancies.

Partway through the validation period, we moved the OCR system from 1 conveyor on our automation system to another conveyor. The first conveyor served a high-volume testing area that required our laboratory's standardized transport tube (Sarstedt #62.612.016), whereas the laboratory area served by the second conveyor did not have that requirement. This enabled us to gain considerable experience with the entire range of tube sizes permitted on our automation. However, this necessitated some design changes. The conveyor caused vibrations in the automation transport carrier. Insertion of our standardized transport tube in a vibrating carrier after OCR analysis was not difficult for the robot due to the beveled bottom of that tube, but was difficult for some large-diameter tubes without beveled bottoms. Thus, we added a pneumatic pinching device to hold the carrier tight so that the robot could insert those tubes. Thereafter, the incidence of mechanical faults was no greater than that of any of our other automated systems.

This relocation of the system also enabled us to observe the impact of the OCR inspection on the 2 laboratories served by the system. The first OCR location served a high-volume laboratory where all tests are performed daily or more than once per day. The second location served a group of smaller laboratory sections in which most tests are performed less frequently, sometimes only 2–3 times per week. Because the review of the OCR images was usually 1–3 days after the images were collected, notification of OCR mislabel detection to the high-volume laboratory was generally after results had been reported, whereas notification of OCR mislabel detection to the second laboratory was often before analysis. Of the 21 corrected reports sent to clients, as noted above, 17 were issued after tests performed in the high-volume laboratory, whereas only 4 were issued following tests performed in the second laboratory location. Although part of the explanation may reside with the nature of the tests, the data gathered demonstrated that the OCR system was a factor in reducing corrected reports.

Discussion

We have invented an automated camera system that photographs the entire exterior of sample tubes and uses OCR to identify potential patient name mislabels. All 1 009 830 images obtained by the system were inspected by employees. Among 742 977 images classified as passes by the OCR system, there were zero samples mislabeled by patient name or with a spelling discrepancy in the name. Of the total of 266 853 images that were classified as fails by the system, subsequent inspection found 121 that were mislabeled and 148 that had spelling discrepancies. The rate of mislabeled samples among all OCR images (1/8346) was higher than the 1/10 000 we had expected from our current system, but lower than the rates in the published literature *(4, 7, 8)*. An unexpected finding was the number of mislabels (4.6/100 000) that escaped detection by our normal QA system.

We recognize that improvement to the overall pass rate of 75% is needed to reduce the labor required to inspect samples that are fails. Several activities to improve the pass rate, such as establishing rules to address name truncations, are underway. As more laboratories implement the CLSI standard AUTO12-A *(15)*, labels will become more uniform and both the OCR pass rate and the global incidence of mislabeled samples should improve. Additionally, widespread implementation of the standard may enable an OCR system to inspect for other parameters such as the date of birth or the medical record number.

As stated before, this project was not designed to prevent errors or to change the flow of our ongoing production, but to demonstrate acceptable performance characteristics for the future use of OCR in a production setting. The next step will be to integrate an improved OCR system into our automation system, which will eliminate the 1- to 3-day delay for manual inspection. All samples classified as fails will route on the automation to a station for human inspection and relabeling, if necessary, before routing for analysis. On the basis of our study, on average, we expect to find 1 name error and 1–2 spelling errors per 2000 samples classified as fails. The total of the inspection times (OCR and human review, if needed) plus track routing time might range from 3 s to 1–2 min. This has no appreciable impact on turnaround time in our reference laboratory, but might be significant in a hospital setting with stat testing. However, we think the positive impact on patient safety is worth that added time.

We believe our invention of an automated camera system that can detect patient name mislabels will be significant for any high-volume laboratory. The undetected error rate during the OCR validation study of 2.1/100 000 represents a significant opportunity for improvement for our laboratory. We, along with our patient-care partners, consider the goal to be zero patient/sample identification errors escaping the system. We trust this goal will be realized when OCR systems are fully implemented on our automation.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:*

Employment or Leadership: C.D. Hawker, ARUP Laboratories; W. McCarthy, Cognex Corp.; B. Messinger, ARUP Laboratories.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.
Patents: None declared.
Other Remuneration: D. Cleveland, Custom Engineering Solutions.

Role of Sponsor: No sponsor was declared.

Acknowledgments: The authors acknowledge the contributions of 4 technicians in our laboratory, James Fuller, Amanda Leech, Maggie Redmond, and Kaylene Sawyer, who each reviewed tens of thousands of OCR images during the validation study. We also gratefully acknowledge the late Dr. William Roberts who suggested to the lead author the need to "find a solution for the problem of mislabeled specimens."

References

1. Plebani M, Carraro P. Mistakes in a stat laboratory: types and frequency. Clin Chem 1997; 43:1348–51.
2. Bonini P, Plebani M, Ceriotti F, Rubboli F. Errors in laboratory medicine. Clin Chem 2002;48: 691–8.
3. Howanitz PJ. Errors in laboratory medicine: practical lessons to improve patient safety. Arch Pathol Lab Med 2005;129:1252–61.
4. Valenstein PN, Raab SS, Walsh MK. Identification errors involving clinical laboratories: a College of American Pathologists Q-Probes study of patient and specimen identification errors at 120 institutions. Arch Pathol Lab Med 2006;130:1106–13.
5. Wagar EA, Tamashiro L, Yasin B, Hilborne L, Bruckner DA. Patient safety in the clinical laboratory: a longitudinal analysis of specimen identification errors. Arch Pathol Lab Med 2006; 130:1662–8.
6. Plebani M. Errors in clinical laboratories or errors in laboratory medicine? Clin Chem Lab Med 2006;44:750–9.
7. Wagar EA, Stankovic AK, Raab S, Nakhleh RE, Walsh MK. Specimen labeling errors: a Q-Probes analysis of 147 clinical laboratories. Arch Pathol Lab Med 2008;132:1617–22.
8. Grimm E, Friedberg RC, Wilkinson DS, AuBuchon JP, Souers RJ, Lehman CM. Blood bank safety practices: mislabeled samples and wrong blood in tube – a Q-Probes analysis of 122 clinical laboratories. Arch Pathol Lab Med 2010;134: 1108–15.
9. CAP. When a Rose Is Not a Rose: the problem of mislabeled specimens. http://www.cap.org/apps/portlets/contentViewer/show.do?printFriendly=true&contentReference=practice_management%2Fdirectips%2Fmislabeled_specimens.html. (Accessed January 2014).
10. Kahn SE, Jarosz C, Webster K, Czerlanis C, Barnish D, Robertazzi S, et al. Improving process quality and reducing total expense associated with sample mislabeling in an academic medical center. Poster session presented at: 2005 Institute for Quality in Laboratory Medicine Conference: *Recognizing Excellence in Practice*; 2005 Apr 28–30; Atlanta, GA.
11. CLSI. Laboratory automation: bar codes for specimen container identification; approved standard—second edition. Wayne (PA): CLSI; 2005. CLSI document AUTO02-A2.
12. Hawker CD. Bar codes may have poorer error rates than commonly believed. Clin Chem 2010; 56:1513–4.
13. Snyder ML, Carter A, Jenkins K, Fantz CR. Patient misidentifications caused by errors in standard bar code technology. Clin Chem 2010;56:1554–60.
14. Snyder SR, Favoretto AM, Derzon JH, Christenson RH, Kahn SE, Shaw CS, et al. Effectiveness of barcoding for reducing patient specimen and laboratory testing identification errors: a Laboratory Medicine Best Practices systematic review and meta-analysis. Clin Biochem 2012;45:988–98.
15. CLSI. Specimen labels: content and location, fonts, and label orientation; approved standard. Wayne (PA): CLSI; 2011. CLSI document AUTO12-A.