J Pathol Inform

Technical Note

NDER: A novel web application using annotated whole slide images for rapid improvements in human pattern recognition

Nicholas P. Reder[1], Daniel Glasser[1], Suzanne M. Dintzis[1], Mara H. Rendi[1], Rochelle L. Garcia[1], Jonathan C. Henriksen[1], Mark R. Kilgore[1]

[1]Department of Pathology, University of Washington Medical Center, Seattle, WA 98195, USA

E-mail: *Dr. Nicholas Patrick Reder - nreder@uw.edu
*Corresponding Author

Received: 17 January 2016 Accepted: 30 May 2016 Published: 26 Jul 2016

Abstract

Context: Whole-slide images (WSIs) present a rich source of information for education, training, and quality assurance. However, they are often used in a fashion similar to glass slides rather than in novel ways that leverage the advantages of WSI. We have created a pipeline to transform annotated WSI into pattern recognition training, and quality assurance web application called novel diagnostic electronic resource (NDER). **Aims:** Create an efficient workflow for extracting annotated WSI for use by NDER, an attractive web application that provides high-throughput training. **Materials and Methods:** WSI were annotated by a resident and classified into five categories. Two methods of extracting images and creating image databases were compared. Extraction Method 1: Manual extraction of still images and validation of each image by four breast pathologists. Extraction Method 2: Validation of annotated regions on the WSI by a single experienced breast pathologist and automated extraction of still images tagged by diagnosis. The extracted still images were used by NDER. NDER briefly displays an image, requires users to classify the image after time has expired, then gives users immediate feedback. **Results:** The NDER workflow is efficient: annotation of a WSI requires 5 min and validation by an expert pathologist requires An additional one to 2 min. The pipeline is highly automated, with only annotation and validation requiring human input. NDER effectively displays hundreds of high-quality, high-resolution images and provides immediate feedback to users during a 30 min session. **Conclusions:** NDER efficiently uses annotated WSI to rapidly increase pattern recognition and evaluate for diagnostic proficiency.

Key words: Adaptive learning, pathology education, smartphone, web application, whole-slide images

Access this article online
Website: www.jpathinformatics.org
DOI: 10.4103/2153-3539.186913
Quick Response Code:

INTRODUCTION

Our institution has a database of whole-slide images (WSIs) used at interdepartmental clinical conferences (i.e., tumor boards) and educational seminars. These WSI have great potential for education[1,2] and quality assurance;[3,4] however, they must be further annotated and classified

This article may be cited as: Reder NP, Glasser D, Dintzis SM, Rendi MH, Garcia RL, Henriksen JC, *et al.* NDER: A novel web application using annotated whole slide images for rapid improvements in human pattern recognition. J Pathol Inform 2016;7:31.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2016/7/1/31/186913

to be of greatest benefit. One potential use of these WSI is to improve rapid pattern recognition (RPR). RPR or so-called "fast thinking" is a key attribute of experienced pathologists.[5] The most established method to improve RPR among trainees is to encounter a large number of cases during residency.[6] However, there are additional methods for improving RPR that can be used in parallel to clinical duties during residency.

One promising method for improving histopathology pattern recognition is perceptual and adaptive learning techniques.[7] These techniques take advantage of preattentive processing, the phenomenon of forming a quick (<200 ms) impression of an image.[7] Herein, the term RPR is used to refer to this phenomenon, although many terms are used across multiple disciplines in the literature.[5,7-9]

We envisioned a system that would take advantage of our large WSI database to improve novice trainees' RPR skills. We developed a web application called novel diagnostic electronic resource (NDER) that uses still images extracted from annotated WSI, and engages users with an adaptive learning algorithm to improve trainees' RPR or "fast thinking" skills.

Our goals for this project:

- Create an efficient workflow for annotating 2000 × 2000 pixel regions of interest (ROI) from a large database of WSI
- Preserve annotation labels and create a pipeline to route images for use by NDER
- Create NDER, a user-friendly and aesthetically pleasing web application that provides high-throughput training for RPR.

MATERIALS AND METHODS

Annotation of Whole-Slide Image

We chose intraductal proliferative breast lesions (IPLs) for our first NDER module because they represent a challenging biological spectrum with important ramifications for patient care.[10] Examples of IPLs were selected from our database, which contains over 4,400 WSI. The WSI were scanned at ×20 magnification using an Aperio ScanScope CS (Leica Biosystems Imaging, Inc., Vista, CA, USA) digital whole-slide scanner. A 2nd-year resident, who had completed a subspecialty breast pathology rotation, annotated (defined herein as outlining and labeling ROI) the WSI using Aperio ImageScope software (Leica Biosystems Imaging, Inc., Vista, CA, USA). Each ROI was labeled with one of five categories: No intraductal proliferation (NIP), usual ductal hyperplasia (UDH), atypical ductal hyperplasia (ADH), and a binary classification of carcinoma *in-situ* defined as low to intermediate grade ductal carcinoma *in-situ* (LGDCIS) and intermediate to high grade ductal carcinoma *in-situ* (HGDCIS).

Extraction of Still Images from Annotated Whole-Slide Image

Two methods for extracting still images and creating a large database of labeled images were compared:

- Extraction Method 1: Manual extraction of still images from WSI and subsequent validation of each still image by four breast pathologists
- Extraction Method 2: Validation of annotated regions on the WSI by a single experienced breast pathologist and automated extraction of still images tagged by diagnosis.

We determined that Extraction Method 1 would require validation by multiple pathologists because the still images are removed from the context of the slide. We used the literature as a guide to the level of pathologist agreement needed to qualify a still image as correctly classified. A study[11] of still images of ulcerative-colitis associated dysplasia, an entity of similar diagnostic challenge to IPLs, showed interobserver agreement ranging from poor to good (kappa coefficient 0.18–0.54). Assuming a kappa at the lower end (0.2), we estimated that four pathologists would agree on an "incorrect" classification only 2% of the time, an acceptable level of error.

NDER must be highly automated to be a practical solution for RPR training. Thus for Extraction Method 2, we built an automated pipeline to extract still images from annotated WSI [Figure 1]. Custom software was written in the CodeIgniter PHP framework (Utilizing Aperio ImageServer APIs) to authenticate, read annotations stored in the WSI database, and extract ROI as jpeg images from the WSI. The jpeg image files were labeled using a "uniqueid-annotation label" format when extracted. This file labeling scheme allowed us to sort the jpeg image files into a well-organized database to be used in our web application.

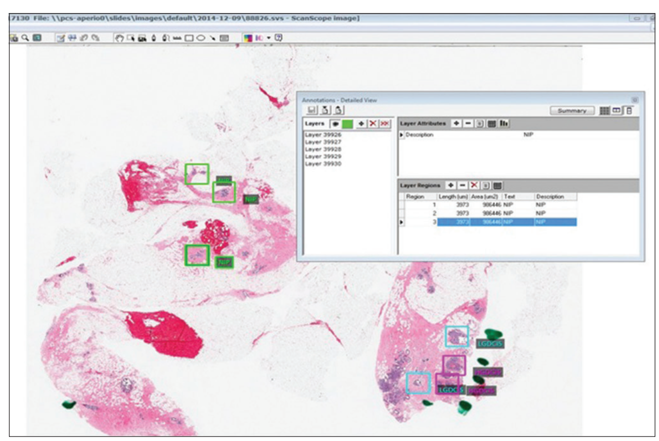

Figure 1: Slide annotation, whole-slide image are annotated with the following disease labels: No intraductal proliferation, usual ductal hyperplasia, atypical ductal hyperplasia, low to intermediate grade ductal carcinoma *in-situ*, or intermediate to high grade ductal carcinoma *in-situ*. An automated pipeline extracted still images from the annotated whole-slide image into a well-organized database to be used in our web application

Novel Diagnostic Electronic Resource, the Web Application

NDER is a web-based application written in PHP and JavaScript with an SQL database and utilizes components of the Bootstrap HTML (http://getbootstrap.com), CSS, and JS framework for responsive display on mobile devices. The welcome page [Figure 2] shows the active module, an about section, and contact information. A toolbar at the top of the page allows the user to choose the module they wish to complete. After choosing the module, the user selects "begin test" and selects his/her level of training (student, resident, fellow, etc.). The module begins with an untimed "pretest" consisting of twenty images with multiple-choice options corresponding to the disease categories, in this example IPLs of the breast.

After completing the "pretest," the user proceeds to the "training module." The training module consists of images that are displayed for a specified amount of time, and then disappear [Figure 3]. The user must choose from the same multiple-choice options as in the pretest (NIP, UDH, ADH, LGDCIS, HGDCIS). After choosing the disease category, the image re-appears and the user are given instant feedback [Figure 4]. The user's selection and the correct answer are displayed next to the image. The user then moves onto the next image. The amount of time the image displays is variable (range: 1.5–10 s) depending on the user's performance. An adaptive algorithm decreases the image display time of each subsequent image with increasing user accuracy to maintain user engagement.

The user takes an untimed twenty image "posttest" after finishing the training module, and a feedback page is

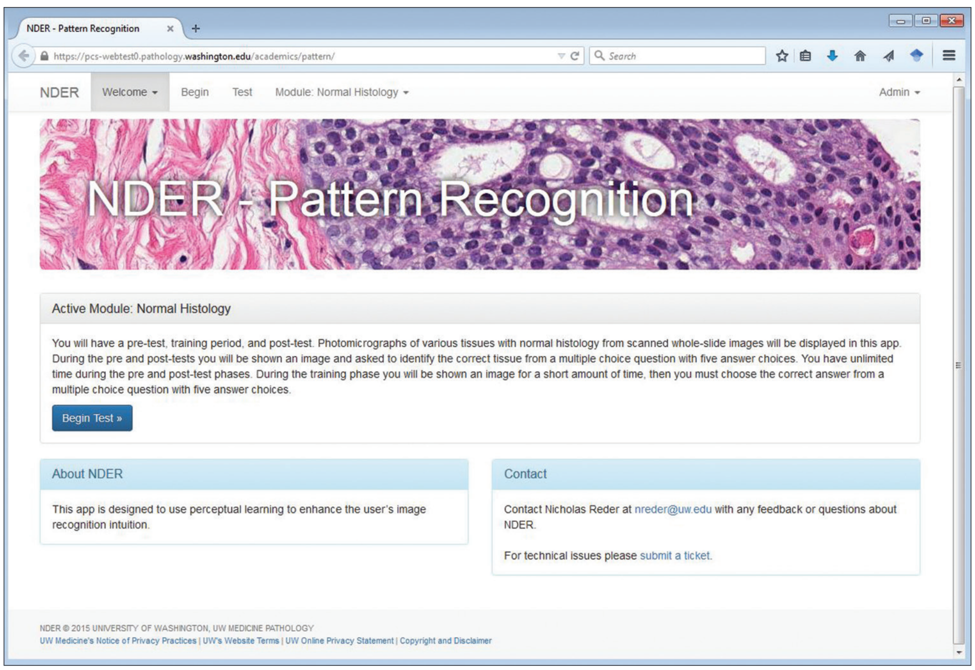

Figure 2: Welcome page, the welcome page shows information regarding the active module, an about section, and contact information. The toolbar at the top of the page allows the user to choose the module he/she wishes to complete

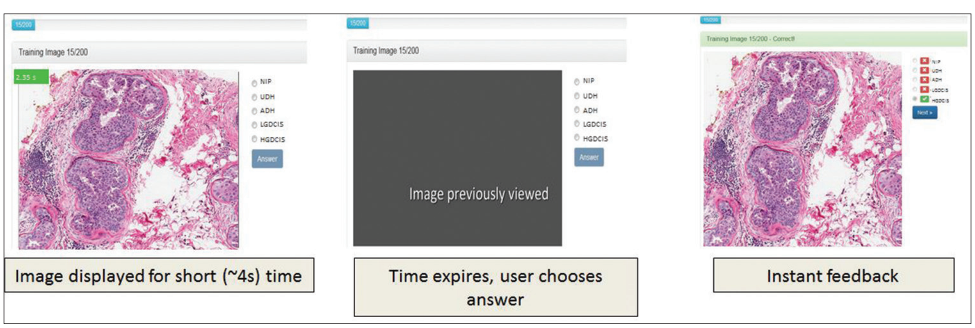

Figure 3: Training module, example screenshots from the novel diagnostic electronic resource intraductal proliferative breast lesion module. An image is displayed on the screen for a short time (range 1.5–10 s). The image disappears after time has expired and the user must choose from the available multiple-choice options. The user is given instant feedback on their selection and can move onto the next image. A progress bar is displayed at the top of the page

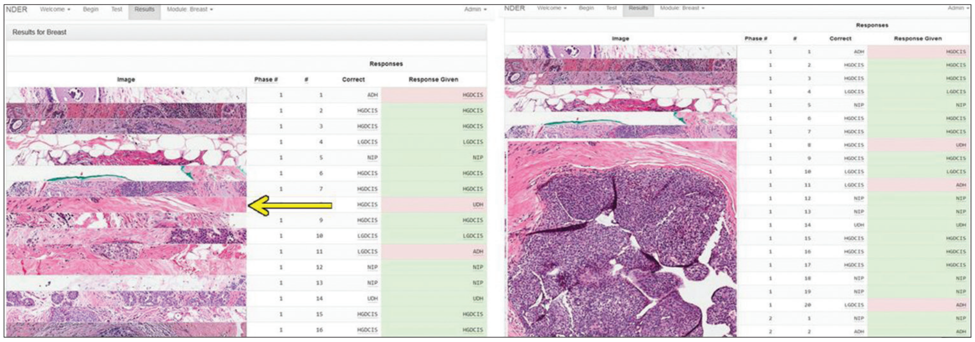

Figure 4: Feedback page, after completing the pretest, training module, and posttest, a feedback page is displayed. There is a thumbnail of each image, the correct answer, and the response given. Users can click each thumbnail to expand the image for further study. The yellow arrow indicates the thumbnail image that has been expanded

displayed. The feedback page contains a thumbnail of each image, the correct answer, and the response given. Users can click each thumbnail to expand the image for further study.

Novel diagnostic electronic resource can be used on multiple devices

RPR training lends itself well to short training sessions and is thus a good fit for mobile phones. We enabled users to take advantage of NDER on their mobile phones using a responsive image CSS class in the Bootstrap HTML, CSS, and JS framework. This feature allows for seamless integration with mobile phones and tablet devices of all sizes, allowing users to train with NDER anytime in any place with an internet connection.

RESULTS

Extraction Method 1: The resident performing the annotation required 2 min on average to annotate and extract each image (200 images in total). Four experienced breast pathologists categorized each of the 200 images, taking on average 30 min to categorize all 200 images. Although the agreement was "substantial,[12]" (Fleiss' kappa = 0.63), only 101 images remained after removal of discordant images.

Extraction Method 2: The resident required 5 min to annotate one WSI with an average of 6.4 ROI. Validation by an experienced pathologist took 1 min per WSI. All other steps were automated.

Table 1 compares the average time to annotate, extract, and validate 200 images using each method. Extraction Method 2 was almost an order of magnitude more efficient than Extraction Method 1 for creating our database of labeled jpeg images. Absolute time saving was greater for the resident portion of the workflow using Extraction Method 2. Relative efficiency gains were similar for both attendings and residents. The resulting

Table 1: Comparison of efficiency of extraction methods

	Time to extract and validate 200 images		
	Total resident time	Total attending time	Total overall time
Extraction method 1*	792 minutes	238 minutes	1030 minutes
Extraction method 2**	156 minutes	31 minutes	187 minutes

* Individual still image capture from WSI by resident followed by validation by four pathologists. **WSI annotation by resident and validation of WSI by a single pathologist

NDER breast module takes users 25–30 min to complete and displays 240 images. Qualitative feedback from initial users regarding website design, usability, and interactivity has been very positive.

Table 2 displays the result of a pilot study involving ten users including four novice trainees (medical students and junior residents), four senior trainees (senior residents and fellows), and two attending pathologists. Users of all experience levels showed increased accuracy in the posttest compared to the pretest, with novice trainees showing the largest gains. The overall accuracy improved from 53% pretest to 83% posttest, a statistically significant result ($P < 0.0001$, paired t-test). The effect size (Cohen's d = 2.11) was very large, where 0.2 is a small effect, 0.5 moderate, and 0.8 large.

DISCUSSION

NDER is a novel web application for improving RPR. A 25–30 min high-yield NDER training module can be created in roughly 3 h (resident: Two hours, attending: One hour). The training module can be accessed via desktop computer, tablet device, or mobile phone web browsers and responds to device screen size. A pilot study of ten users demonstrated large increases in accuracy from pre- to post-test. Qualitative feedback from initial users has been positive, focusing on intuitive design,

Table 2: Pilot study results

	Novice trainee (*N*=4) (%)	Senior trainee (*N*=4) (%)	Attending (*N*=2) (%)	Overall (*N*=10) (%)
Pre-test	33.8	61.3	72.5	53.0
Post-test	77.5	85.0	90.0	83.0

usability, and responsive display features. One issue highlighted by users was image display latency caused by web connection bottlenecks.

Training in pathology consists of training in both "slow" and "fast" thinking. Discussion at the double-headed microscope, didactic lectures, and textbooks are common methods for teaching pathology. All these methods target deliberate or "slow" thinking. The only current strategy for teaching "fast" thinking or RPR is seeing a large volume of cases during residency.[13] NDER is a novel tool for RPR to be used in parallel to traditional residency training. Previous studies using electrocardiograms have shown that a combination of pattern-based and feature-based instruction performs better than either method alone.[9] We believe NDER, in combination with traditional methods, will enhance residency education through improvement of residents' histopathologic intuition.

There are many high-quality digital tools developed for pathology training including examples in medical education,[14] histology,[7,15] and cytopathology.[15] However, none of the currently available tools specifically target RPR. The closest analog to NDER is the excellent introductory histopathology training module developed by Krasne *et al*.[7] This tool shows images of basic pathologic processes, forces users to choose an answer within 24 s, and gives immediate, succinct feedback. In contrast, NDER forces the user to choose an answer in as little as 1.5 s or at most 10 s, depending on performance. The brief time period forces users to focus solely on improving the accuracy of their initial impression of the image. Improved accuracy of this initial impression, or intuition, has never been targeted for pathology training by a digital tool (to our knowledge). An additional advantage of the brief period is an increased volume of images viewed by the user.

NDER has applications to pathology beyond resident and medical student education. Quality assurance and quality improvement are particularly intriguing areas for future modules. For example, we are currently creating a module for breast prognostic/predictive immunohistochemical markers in hopes of decreasing both inter- and intra-observer variability in the interpretation of 2013 CAP/ASCO human epidermal growth receptor 2 (HER2) immunohistochemistry (IHC) guidelines.[16] Current evidence suggests an increase in the number of "equivocal 2+" interpretations since the release of the revised guidelines in 2013,[17] some of which may be related to overinterpretation. We hope NDER will help train pathologists to accurately separate IHC HER2 protein expression into the appropriate category (0, 1+, 2+, 3+) as defined by the guidelines.[16] Such a module would decrease costs for pathology departments by decreasing the number of unnecessary reflex HER2 fluorescence *in situ* hybridization tests in cases better interpreted as either negative or positive. Additional quality assurance and improvement applications are being explored, of which the possibilities are essentially endless.

One current limitation of NDER includes the use of still images rather than WSI for the pre- and post-test phases of the module. We have begun work on integrating WSI into the pre- and post-test phases for a more accurate assessment of actual pathology practice. Incorporation of WSI will investigate how context affects interpretation when users can evaluate an entire slide rather than an isolated static image. The addition of WSI will be of particular interest in borderline lesions such as the distinction of ADH and LGDCIS, in which size and span may be the limiting factor in differentiating these two entities.[18,19]

In the future, we aim to create an open-access version of NDER available to anyone with an internet connection. NDER relies on a well-curated and annotated database of WSI, which is available to users at only a select group of institutions. We envision a system where collaborators from across the world can contribute annotated WSI for the creation of a broad range and depth of modules.

CONCLUSIONS

NDER is a novel web app that uses annotated WSI to rapidly improve pattern recognition and evaluate for diagnostic proficiency. It has wide applicability to education, training, and quality assurance. In the near future, we envision an open-access platform where NDER modules can be created from annotated WSI from collaborators across the world, including experts in their fields.

Financial Support and Sponsorship

Nil.

Conflicts of Interest

There are no conflicts of interest.

REFERENCES

1. Dee FR. Virtual microscopy in pathology education. Hum Pathol 2009;40:1112-21.
2. Foster K. Medical education in the digital age: Digital whole slide imaging as

an e-learning tool. J Pathol Inform 2010;1. pii: 14.
3. Ho J, Parwani AV, Jukic DM, Yagi Y, Anthony L, Gilbertson JR. Use of whole slide imaging in surgical pathology quality assurance: Design and pilot validation studies. Hum Pathol 2006;37:322-31.
4. Wei BR, Simpson RM. Digital pathology and image analysis augment biospecimen annotation and biobank quality assurance harmonization. Clin Biochem 2014;47:274-9.
5. Evered A, Walker D, Watt AA, Perham N. To what extent does nonanalytic reasoning contribute to visual learning in cytopathology? Cancer Cytopathol 2013;121:329-38.
6. Talbert ML, Ashwood ER, Brownlee NA, Clark JR, Horowitz RE, Lepoff RB, *et al.* Resident preparation for practice: A white paper from the College of American Pathologists and Association of Pathology Chairs. Arch Pathol Lab Med 2009;133:1139-47.
7. Krasne S, Hillman JD, Kellman PJ, Drake TA. Applying perceptual and adaptive learning techniques for teaching introductory histopathology. J Pathol Inform 2013;4:34.
8. Healey CG, Booth KS, Enns JT. High-speed visual estimation using preattentive processing. ACM Trans Computer Hum Interact 1996;3:107.
9. Ark TK, Brooks LR, Eva KW. Giving learners the best of both worlds: Do clinical teachers need to guard against teaching pattern recognition to novices? Acad Med 2006;81:405-9.
10. Leonard GD, Swain SM. Ductal carcinoma *in situ*, complexities and challenges. J Natl Cancer Inst 2004;96:906-20.
11. Odze RD, Goldblum J, Noffsinger A, Alsaigh N, Rybicki LA, Fogt F. Interobserver variability in the diagnosis of ulcerative colitis-associated dysplasia by telepathology. Mod Pathol 2002;15:379-86.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
13. Cross SS, Betmouni S, Burton JL, Dubé AK, Feeley KM, Holbrook MR, *et al.* What levels of agreement can be expected between histopathologists assigning cases to discrete nominal categories? A study of the diagnosis of hyperplastic and adenomatous colorectal polyps. Mod Pathol 2000;13:941-4.
14. Kanthan R, Senger JL. The impact of specially designed digital games-based learning in undergraduate pathology and medical education. Arch Pathol Lab Med 2011;135:135-42.
15. Van Es SL, Pryor WM, Belinson Z, Salisbury EL, Velan GM. Cytopathology whole slide images and virtual microscopy adaptive tutorials: A software pilot. J Pathol Inform 2015;6:54.
16. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, *et al.* Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. J Clin Oncol 2013;31:3997-4013.
17. Long TH, Lawce H, Durum C, Moore SR, Olson SB, Gatter K, *et al.* The new equivocal: Changes to HER2 FISH results when applying the 2013 ASCO/CAP guidelines. Am J Clin Pathol 2015;144:253-62.
18. Page DL, Dupont WD, Rogers LW, Rados MS. Atypical hyperplastic lesions of the female breast. A long-term follow-up study. Cancer 1985;55:2698-708.
19. Tavassoli FA, Norris HJ. A comparison of the results of long-term follow-up for atypical intraductal hyperplasia and intraductal hyperplasia of the breast. Cancer 1990;65:518-29.