ELSEVIER

Digital pathology in drug discovery and development: multisite integration

Steven J. Potts

Aperio Technologies, Inc., 1360 Park Center Drive, Vista, CA 92081, United States

Digital pathology is an emerging technology that provides an image-based environment for managing and interpreting the information generated from a digitized glass slide, offering substantial improvements in pharmaceutical drug development across discovery, preclinical GLP pathology and oncology clinical trials. Digital pathology is transforming global pharmaceutical research by enabling data sharing to integrate dispersed pharma pathology labs around the world. This article reviews the stages of multisite digital pathology integration in large pharmaceutical companies, offering suggestions for success and highlighting challenges.

Introduction

Tissue-based research is a key component to drug development. The local tissue response contains valuable efficacy and therapeutic information on toxicity that can only be deciphered by expert, credentialed pathologists. Tissue-based research occurs across the entire drug development process as shown in Table 1, dominated by biologists in discovery, veterinary pathologists in preclinical toxicological and medical anatomic pathologists in oncology clinical trials. Unlike biochemical or molecular assays, pathology testing requires manual interpretation by an expert.

Because veterinary and medical anatomic pathologists are central contributors to tissue assessment, pathologist's access and communication is vital. Yet there are many barriers in global pharmaceutical drug development preventing effective tissue assessment, including physical geographic location, accessibility within a corporation and differences between D.V.M. and M.D. pathologists. In the past decades, most veterinary pathologists were employed inside pharmaceutical companies and preclinical operations were fully integrated. A study director had direct access to the pathologist, frequently located in the same building.

The growth of the large contract research organization (CRO) in recent years has, however, concentrated veterinary and human anatomic pathologists outside the pharmaceutical corporations they serve. Glass slides produced in one location are frequently shipped to a pathologist in another location. When pathologists at a CRO are involved, the communication of slide results must be shared between organizations. International preclinical operations have globalized and clinical trials often involve remote, developing countries for patient access. Communication is made even more difficult by the slow process of mailing slides between international locations.

The cross-discipline contributions of veterinary pathologists in preclinical studies and medical anatomic pathologists in clinical trials also create challenges. Clear communication between these two pathologist specialties is needed regularly [1]. For example, to characterize mouse and complex xenograft models, such as human primary xenografts, which preserve human tumor-like morphology but have a mouse stromal environment, both veterinary and human oncology knowledge is required [2–4].

Digital pathology improves efficiency

Digital pathology has the potential to eliminate the pathology barriers of geography, subjectivity and cross-discipline communication in tissue-based research. Digital pathology allows diagnosis via whole slide images from a computer screen rather than glass slides from a microscope. The entire tissue section on a slide is scanned with 20×, 40× or 100× magnification, and then viewed on a monitor. Digital pathology software allows multiple pathologists to view and annotate the same image concurrently, in a digital slide conference or simultaneously to view or align multiple slides on a computer screen. Pathologists can access remote slides via web-based secure portals. Because these digital slides are simply

E-mail address: spotts@aperio.com.

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2009.06.013

TABLE 1

Tissue research is common across drug development, and offers many areas where digital slides are utilized in the industry

Discovery	Preclinical	Clinical trials
Oncology	Predictive toxicology	Phosphomarker expression in phase I, II and III trials as surrogate endpoints
Xenografts and matrigels		
Nuclear, membrane and cytoplasm protein expression	Quantitative measurements of lesions or abnormalities in toxicology studies	
Phosphomarkers		
Angiogenesis and microvessel density		Protein biomarker expression for patient stratification
Metabolism	Digital slide conferencing in preclinical studies	Confirmatory diagnosis during patient enrollment
Alpha and beta cell mass		
Neuroscience	Digital pathology working Groups	
Amyloid plaque		Secondary consults
Microfibular tangles		
Tau expression		Tumor boards
Grey versus white matter	GLP toxicologic pathology assessment in animal models	Tissue microarrays for biomarker analysis
Inflammation and autoimmune		
Inflammatory cells in bone or joint tissue	Peer reviews	
Vascular measurements		
Ophthalmology		
Retinal ganglion cell counting	Standardization of toxicology terminology and scoring	
Angiogenesis in AMD		

large image files, image analysis programs run quantitative, computer-generated measurements that a pathologist then reviews. Owing to its complexity, Pathology has historically lagged behind other disciplines like Radiology in the use of digital media, but this is gradually changing in the clinic [5] and in pathology education [6].

In pharmaceutical companies, the transition to digital pathology will not happen overnight. Because of the proprietary nature of drug development, little has been published in the literature on digital pathology in pharmaceutical research, although nearly all of the large pharmaceutical companies have implemented it. The technology utilized for digital pathology has, however, been well reviewed [7,8]. Waves of technology improvements rapidly overflow the published reviews, as barriers erode in ease-of-use of image analysis, slide scanning speeds and storage and network costs. In pharmaceutical development using digital pathology, there have been few published reviews [9] and even fewer publications [10], despite widespread and increasing adoption of the technology.

Digital pathology changes workflow

Digital pathology begins with whole slide imaging, the complete capture of the tissue on a glass slide as a single large image. Current scanning speeds are approximately 2 min to scan a 15 mm × 15 mm slide section at 20× magnification 40× scanning taking 4 times longer and 100× oil scanning approximately 12 times longer. Partial or complete automation allows several hundred slides to be scanned at 20× magnification in an eight-hour shift.

The software infrastructure has become a key component to digital pathology. Figure 1 shows one potential layout based on a global implementation at a large pharmaceutical company. Because compressed whole slide image files are generally large (160 MB for a 20× magnification 15 mm × 15 mm section), the images are best stored near the scanner with a large local bandwidth connection.

As the glass slides are scanned, a message via XML over secure Internet is sent to a central and, frequently, remote database that contains all metadata, annotations and security information for the pathology system. A placeholder is made for the slide and an electronic pointer to this slide is stored in the database. An important part of the scalable architecture is that the large image files are never stored inside the database, but instead generally stay near where the glass slides are scanned.

A user from any location logs into the database via the Internet, searches for their slides, and is able to view thumbnails. Once a slide is found, a secure connection is made between the user's computer and the slide location, and the pathologist is able to view the slide as he would do with a microscope. The slide is viewed as pixels on demand, so that only the portion of interest at any given time is transferred to the pathologist's desktop. This makes viewing scalable and does not require any additional Internet bandwidth beyond what is commonly already in place at most modern pharmaceutical companies or CROs.

There are many time and economic drivers for multisite integration of pathology. It is important when embarking on integration that attempts be made to develop metrics for success, since ultimately any new technology should be valued upon whether it saves time or cost in pharmaceutical development. Some of the major drivers [11] are depicted in Table 2.

Multisite adoption in stages

Multisite integration of digital pathology in a pharmaceutical environment can be roughly broken into several stages as shown in Table 3. As would be expected, there are many exceptions, and the transition from one stage to another is gradual and overlapping.

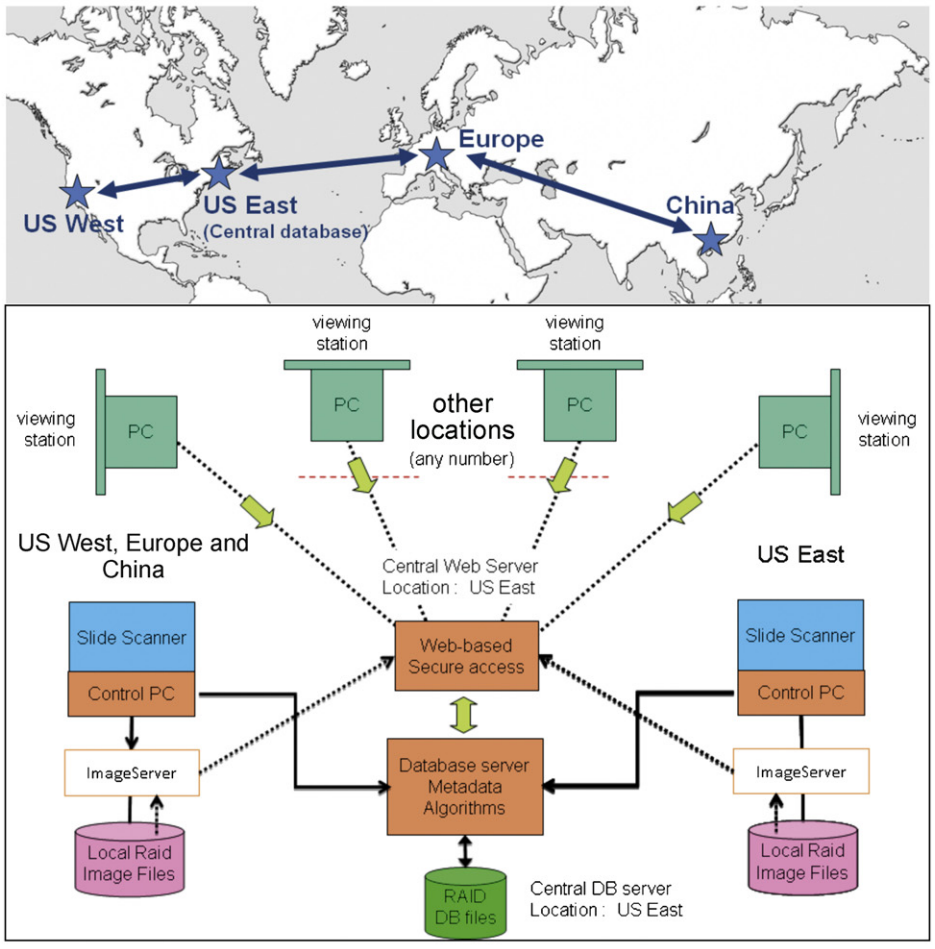

FIGURE 1

Multisite integration architecture. With pathologists at multiple locations globally (top), the architecture allows the images to be stored where they are scanned but accessed via secure Internet from a database at a single location (bottom).

Stage I: single-site digital pathology

Initially, one group in a location begins to implement digital pathology as a single site, and starts to have some of the readings conducted from a monitor rather than a digital slide. There are substantial advantages to utilizing digital slides at a single site; the greatest is typically the ability to make more quantitative, reproducible computer-based measurements of efficacy in tissues. The local group can discuss pathology findings in a conference room on a computer screen, or with individuals at other locations in *ad hoc* digital slide conferences. Interesting cases are cataloged for future comparisons, helping with standardized scoring.

The greatest two challenges in this first stage are pathologist adoption and demonstrating return on investment. It also takes time for pathologists to realize fully the advantages and increased functionality of digital slides, such as digital slide conferencing with remote peers, alignment and simultaneous viewing of multiple consecutively cut tissues with different biomarkers, or computer tracking of locations on a slide that have been previously viewed. The second challenge is that at the preliminary stage of adoption of any new technology, the financial returns in efficiency are typically minimal. Groups will be using their prior technology in parallel with the new technology, which initially adds rather than cuts expense. In the activities where it does eliminate expense (e.g. travel for pathology working groups or informal peer reviews), it is not always easy to quantitate these savings early on in adoption.

Stage II: multisite pathology in discovery and non-GLP preclinical within a single organization

In the first stage, users at other sites may participate with a site that has adopted whole slide imaging, with *ad hoc* digital slide conferences, or using digital slides that have been mailed, ftp downloaded or accessed remotely with pixels on demand image viewing. Other sites with histology may, however, also want to scan their own slides and conduct their own workflow. This leads to a transition to a second stage, where slides are scanned and used at multiple sites within the same organization.

If these organizations have adopted different vendor's technologies, or have had independent adoption, as in the case of a merger between companies, more time is needed for integration. A key architecture design question that must be answered at this stage is

TABLE 2

Advantages of multisite integration of digital pathology

Eliminate geography	Slides can be made in one place and read in another Standardized pathology procedures at all sites Provide cross coverage globally (particularly pathologist support) Rapid recall for internal conferencing Image recovery on demand for presentations Portability of slides and work Reduced pathology travel time and costs Reduced slide shipping time and cots
Decrease subjectivity	Quantitative assessment of efficacy or toxicity Higher consistency of interpretation and quality control Searchable archive of scoring standards Remote conferencing for calibration Quantitate biomarker assessment in tissue microarrays Use of the whole slide context to verify subjective interpretation
Improve pathology	Maximize value of work already done and reduce duplication Expanded knowledgebase of digital slides No long-term deterioration of glass slides Improve collaboration between preclinical biomarker teams and clinical researchers with whole slide images Drug development project team members become more familiar and comfortable with pathology data by having access to the whole slides Increase communication between DVM and MD pathologists Improve training

whether the workflows, types of users, security roles and so on are sufficiently similar between the different sites that all users should work from the same database (as shown in Fig. 1). The answer depends on the degree that the multiple sites share common activities. For example, a pharmaceutical company with tissue biomarker works in oncology that is conducted at three sites globally, but under the same department head, will have substantial benefits to having all users access the same database. Electronic roles, access privileges, naming conventions, reagent and antibody links, image analysis algorithm settings and so on can then be shared across this group and reused by all.

In this case, Fig. 1 shows one applicable approach – all security, metadata, access privileges and database customization will be in a single database location, with the slides sitting on remote servers at each individual site. There are, however, many other cases where two labs are not naturally sharing the same workflows, common terminologies or standard operating practices (SOPs), and doing so would not add any substantial economies of scale. For example, a cardiovascular group and an oncology biomarkers group may prefer to have one database permanently assigned to each laboratory that does its own work. *Ad hoc* conferencing and user access at other sites can still be granted to different remote users, but the sites will have databases that are specifically customized to their individual needs. While this will provide a lower level of standardization, it allows individual labs the freedom to build the user experience in ways most appropriate for their own needs.

TABLE 3

Multisite adoption occurs in stages

Stage	Major productivity drivers	Challenges
Single-site usage of digital slides	Quantitative scoring Standardization across a single site Improvements in turnaround time and workflow Improved communication between pathologists and others in the organization Introduction of the technology to pathologists	Pathologist aversion to digital slides ROI justification of technology new to the organization and only partially implemented
Multisite integration in discovery and non-GLP preclinical	Standardization of scoring across geographic sites Ability to shift work Savings on travel and slide shipping Greater communication between pathologists	In-house IT engagement and planning at the pharmaceutical company Migrating multisite users to a shared lexicon and pathology work environment
Non-GLP data sharing between pharmaceutical company and contract research organization (CRO) pathologists	Standardization of scoring at the beginning of studies Earlier notification of failure Greater communication with contracted pathologists Electronic record of lesions across preclinical studies Increased access to specialist pathologists	CRO and Pharma Information Technology buy-in and engagement Security and compliance Access to digital pathology technology by CROs
GLP assessment with digital slides	Full digital workflow in pathology – savings in time Complete electronic record Ability to peer review international studies Higher throughput (as the technology develops) Improved ergonomics	Confidence in the digital slide as equivalent to a glass slide Compliance by the pharmaceutical company, CRO and technology vendor Integration with existing toxicology GLP computer system vendors

Stage III: non-GLP preclinical data sharing between pharma, CROs and independent consultants

Because more users across a pharmaceutical company conduct multisite integration, a next logical step is the ability for external pathologists at CROs to work with internal pharmaceutical pathologists. While the early stages allow external consultants or contract pathologists access via virtual personal networks (VPN) to slides for peer review, or for digital slide conferencing, this stage becomes important when the slides are generated by an outside organization, or substantial and ongoing pathology support is required from outside the pharmaceutical company. It is more challenging architecturally, because the external CRO will have its own processes, will be working with multiple pharmaceutical clients and have several studies ongoing at any given time. Indeed, a majority of the largest preclinical and clinical trials CROs already have their own multisite integration infrastructure plans ongoing.

There are two times that are outside of GLP processes in most preclinical studies where substantial communication needs to occur via images. First, before the start of a study, it is helpful for pathologists at the two organizations to agree on terminology and scoring. This can be done with digital slide conferences or sharing of slides in a secure location in secure location outside of either company's firewall (often called a demilitarized zone, or DMZ). Second, during a study if unexpected abnormalities are encountered, the ability to do conferencing between the organizations saves delays. Neither of these scenarios requires fulltime IT integration between organizations and the majority of the largest preclinical CROs now offer this service. While there are many exceptions, there may be a general difference between a short-term study which might be conducted within a pharmaceutical company and a long-term study which would be conducted at a contract laboratory. In the short-term study, contract pathologists at a CRO may need to access slides at the pharmaceutical company and in a long-term study, pharmaceutical pathologists may need to access slides at the CRO partner.

Stage IV: GLP assessment with digital slides

The use of whole slide imaging is gradually expanding into GLP studies, where the full power of a digital environment will probably offer the strongest contribution to time and cost savings in preclinical toxicity studies. For the reader unfamiliar with GLP studies, thousands of microscope slides must be examined by board certified veterinary pathologists for signs of preclinical toxicity before filing an IND. This pathology assessment must be done in compliance with regulatory guidelines under 21 CFR 11 and 21 CFR 58. There is active dialog and discussion over some key concepts in this area, including the definition of raw data [12] as it relates to the digital slide and the requirements for validation [13]. Study validation approaches used in clinical studies can be applied to the preclinical reading of digital slides [14]. All databases used for GLP, however, should be kept separate from non-GLP, to preserve innovation and freedom in discovery groups, yet enforce rigid data control in regulated groups. Work is also required to integrate whole slide imaging into the GLP workflows of existing preclinical lab information software.

Challenges in integration

The differences between laboratories in global integration cannot be overlooked and must not be underestimated. There may be differences in work requests from project teams, from organized web forms and formal regular team meetings to casual phone or email requests. SOPs and histology processes can be substantially different, including antibodies, staining equipment and approaches to image analysis methods. Labeling schemes for blocks and slides, as well as metadata annotations, will vary between labs. Multiple approaches are taken in sample data storage and tissue and specimen tracking, from web-enabled databases or laboratory information management systems (LIMS), to stand-alone spreadsheet or access files to paper notebooks. The final report format, and its delivery and archiving, whether freeform or formatted PowerPoints, web reports, electronic notebooks or formal pharmaceutical company document reports will vary.

These challenges are not unique to pathology and have been handled successfully in many other fields. Much of what has been learned in centralized compound and target database storage should be reusable and applied in pathology multisite integration.

Suggestions for success

Integration challenges can be addressed by some simple approaches, which allow laboratories to move toward a common goal. This does require buy-in from senior stakeholders in the pharmaceutical science and IT organizations, as well as commitment to long-term support from the vendors themselves. It is helpful to create common practices documents for each stage of tissue-based research, including sample collection, histology, pathology scoring, image analysis and archiving. These documents are a single set of common beliefs for best practices backed by peer reviewed research or internal experiments, including detailed SOPs for each laboratory's processes, which helps to consolidate the SOPs across sites over time.

Because the work is linked electronically through a single database, more standardization occurs, particularly as reagents, tissues, cell blocks or unique antibodies are shared across laboratories. Standardization can be encouraged across the organization as the database provides links to areas beyond tissue research, including pharmacology results, transcript, mutation and Western blot data on the same or similar samples. Computer user interface functions that limit free text typing in the database will encourage common terminology in projects, specimens and gene targets. For project coordination, a weekly multisite conference call can be useful. It is recommended to assign a person to the project team from each site, and clear goals and timelines agreed to by senior stakeholders. The goals should be set early in the project lifecycle, and preferably in measurable outcomes directly beneficial to drug discovery, like time or cost savings, or increased standardization of efficacy or toxicity measurements.

Staffing and pathology training considerations are key to successful multisite integration. Similar to how bioinformatician's and cheminformatician's careers expanded as chemical compound and searchable target database projects were completed at pharmaceutical companies in the 1990s, project success hinges on a few key personnel with fluency in both IT and biology. As there are no training programs producing pathology informatics scientists, suitable backgrounds include bioinformaticians, computer scientists with biology training, and also histologists with strong computer skills.

Emergence of innovative on-line pathology teaching programs [1] and local pathology workshops makes some of the required pathology knowledge more accessible to biologists or IT trained personnel (e.g. Armed Forces Institute of Pathology, CL Davis, Johns Hopkins, and The Jackson Laboratory all offer regional courses in pathology). As the slides go digital, there are more opportunities for the few pathologists employed within a company to discuss findings in context or tissue heterogeneity with pathology novices. Regular use of digital slide conferencing for reviewing slides and teaching purposes also expands the tissue knowledge across the organization. Although few pathology informaticians will have enough knowledge in any organ to do independent research, their ability to communicate with the pathologists and help with the infrastructure is greatly facilitated by these regular discussions.

It is important that the software database product be easy enough to customize without advanced software skills, as the people who usually understand the workflow and the best fields to associate with slides, specimens, or projects, are histotechnicians who usually are not programmers. Some pharmaceutical companies require that project team leaders must review and understand all primary data, which would include pathology. This requirement makes leadership more interested in the details of pathology and multisite integration, and makes the SOPs more transparent across the team.

Limitations and further technology requirements

As multisite usage increases, the need for cross-platform reading of images and third party scanner integration will also increase. Pharmaceutical companies should insist on open image file formats (e.g. TIFF images with JPEG2000 compression) which makes this much easier. Vendors should provide XML descriptors for the image annotation formats, allow customization of their user interface web pages, and offer open database access methods like SOAP (Simple Object Access Protocol) or web services to their underlying databases. There is much healthy dialog between pathology and radiology vendors regarding the adaptation of the medical imaging standard DICOM for digital pathology. While DICOM could allow pathology data to be integrated into the larger framework of medical images (e.g. radiology), this decades-old standard was designed for much smaller image file sizes and has its own set of limitations.

Glass slides are used in many different situations, which can be extremely challenging to emulate digitally. A box of slides might be used in a controlled study, then sent to a pathologist for peer review, then years later used for teaching purposes, or a portion captured digitally and used in image analysis. With a digital slide, this multipurpose usage does present challenges to database design. A slide might need to be referenced under a given project, and from a certain specimen, but then shared as an image in a teaching course, and later linked to an image analysis program. Systems that will allow for a many purpose referencing of the slide, while still enforcing rigidity when necessary, still require further work. Currently this requires a lot of copying of slides, when multiple uses are required for each copy.

Workflows for pathologists working away from the Internet will require additional technology enhancements, because many pathologists find the least interruptions during travel or evenings. The digital pathology solution should work similarly to email, where it is standard practice to do a lot of work offline, then update the server when back online or in the office.

Scanning times will need to increase, particularly as all slides start to be scanned and the technology moves toward GLP. Sample data management and barcoding in histology operations is also a challenge. Scanning being the last link in the chain, the autostaining vendors upstream must allow open barcode formats so that a single barcode can be used for staining and reagent as well as scanning information.

Conclusion

Veterinary and human anatomic pathologists remain central contributors to drug development, even while spread globally across pharmaceutical companies and CROs. Pharmaceutical companies can no longer operate as they have done in the decades past, when a pathologist resided down the hall, and results could be discussed in an office by two people peering into a multiheaded microscope.

The adoption of digital pathology across multiple sites helps to eliminate geography barriers, reducing subjectivity and ultimately improving pathology. It requires time and planning to implement, and long-term organizational commitment by the corporation. The technology continues to improve and additional enhancements in this dynamic field are eagerly awaited. When regular project team meetings include participants across continents with veterinary and human anatomic pathology disciplines, and when these participants are not aware of the geographic location of a slide under discussion, a pharmaceutical company is well on the road to multisite pathology integration.

Acknowledgements

This article benefited from numerous suggestions and the pioneering vision of James Deeds and his team from Novartis Institutes of BioMedical Research. The author would like to thank John Banks from Aperio Technologies for his technical work on multisite integration at numerous large pharmaceutical organizations.

References

1 Cardiff, R.D. *et al.* (2008) 'One medicine–one pathology': are veterinary and human pathology prepared? *Lab. Invest.* 88, 18–26
2 Warren, M.V. *et al.* (2009) An impending crisis in the provision of histopathology expertise for mouse functional genomics. *J. Pathol.* 217, 4–13
3 Tan, A. (2008) Do-it-yourself (DIY) pathology. *Nat. Biotechnol.* 26, 978–979
4 Offerhausm, G.J. *et al.* (2008) A transgenic mouse model for ''lipid hang-up'', or why pathologists need to be involved in genetically engineered mouse modeling. *Gut* 57, 1739–1740
5 Kayser, K. *et al.* (2006) Towards an automated virtual slide screening: theoretical considerations and practical experiences of automated tissue-based virtual diagnosis to be implemented in the Internet. *Diagn. Pathol.* 1, 10
6 Della Mea, V. *et al.* (2006) User attitudes in analyzing digital slides in a quality control test bed: a preliminary study. *Comput. Methods Programs Biomed.* 82, 177–186
7 Mulrane, L. *et al.* (2008) Automated image analysis in histopathology: a valuable tool in medical diagnostics. *Expert Rev. Mol. Diagn.* 8, 707–725
8 Treanor, D. (2009) Virtual slides: an introduction. *Diagn. Histopathol.* 15, 99–103

9 Ying, X. and Monticello, T.M. (2006) Modern imaging technologies in toxicologic pathology. *Toxicol. Pathol.* 34, 815–826
10 Wiederschain, D. *et al.* (2007) Contribution of polycomb homologues Bmi-1 and Mel-18 to medulloblastoma pathogenesis. *Mol. Cell Biol.* 27, 4968–4979
11 Deeds, J. and Gardner, H. (2008) Multi-site integration with digital pathology: three countries, four time zones, one oncology. *Pathology Visions Conference*, San Diego, California
12 Tuomari, D. *et al.* (2007) Society of toxicologic pathology position paper on pathology image data: compliance with 21 CFR parts 58 and 11. *Tox. Pathol.* 450–455
13 Pyrah, I. et al. (2008) Bioresearch modernization – the quality path forward: validation of whole slide imaging. Presentation at the PhRMA Meeting, September 15th
14 Nassar, A. et al. (2007) Reading digital IHC slides on a computer monitor. USCAP Annual Meeting