Human Pathology (2005) **36**, 1302–1308

ELSEVIER

Human PATHOLOGY

www.elsevier.com/locate/humpath

Correlation of spectral imaging and visual grading for the quantification of thymidylate synthase protein expression in rectal cancer☆

Gary Atkin[a,*], Paul R. Barber[a], Boris Vojnovic[a], Frances M. Daley[a], Rob Glynne-Jones[b], George D. Wilson[c]

[a]*Gray Cancer Institute, Mount Vernon Hospital, Northwood, HA6 2RN Middlesex, UK*
[b]*Department of Radiotherapy, Mount Vernon Hospital, Northwood, HA6 2RN Middlesex, UK*
[c]*Karmanos Cancer Institute, Wayne State University, Detroit, MI 48201-2013, USA*

Received 29 March 2005; accepted 30 August 2005

Keywords:
Immunohistochemistry;
Spectral imaging;
Thymidylate synthase

Summary Quantification of protein expression in tissue sections stained by immunohistochemistry has traditionally involved visual grading techniques. However, if these results are to be used to predict tumor behavior and permit targeted therapy, there is a need for more accurate, objective, and reproducible methods. This study investigated the utility of spectral imaging as a method of quantifying thymidylate synthase protein expression in immunohistochemically stained sections of primary rectal cancer and normal rectal mucosa by comparing it with the current gold standard of manual visual grading. There was good correlation between estimates of thymidylate synthase stain intensity and area derived by spectral imaging and visual grading in both tumor and normal mucosal sections, suggesting that spectral imaging is a valid way of quantifying biologic sections stained by immunohistochemistry.

1. Introduction

Immunohistochemistry (IHC) is widely used for assessing the distribution and expression of proteins within biologic sections to provide prognostic information and predict the likelihood of success of adjuvant therapy [1-3]. Accurate quantification of protein expression in tissue sections stained by IHC has been difficult because of lack of objective, reproducible, and widely available methods. Traditional techniques of visual grading have been the mainstay for many years and have included estimates of stain intensity and area, as well as visual scoring systems such as counting positively stained blood vessels to derive microvessel density estimates. However, these methods are subjective and may lack the accuracy needed to generate data suitable to guide patient therapy. Consequently, computer-assisted image analysis methods have been investigated as strategies to provide quantitative and reproducible data [4]. Spectral imaging is a form of image analysis that permits the separation of transmitted light

☆ This work was funded by The Freemasons' 250th Anniversary Fund, a Royal College of Surgeons of England Research Fellowship.
* Corresponding author. Muswell Hill, London N10 2LA, UK.
E-mail address: gkatkin@blueyonder.co.uk (G. Atkin).

0046-8177/$ – see front matter © 2005 Published by Elsevier Inc.
doi:10.1016/j.humpath.2005.08.016

into an arbitrarily large number of wavelength bands and produces a high-resolution spectrum of light intensity as a function of wavelength for each image pixel [5]. It is possible with spectral imaging to measure light in the UV and infrared regions of the visible spectrum, both of which are invisible to the naked eye [5], and images of the object under investigation can be generated depicting a range of qualitative and quantitative parameters [6].

Spectral imaging has been applied to various areas of biomedicine, such as improving the detection of nuclear atypia on cervical smears [7], determining gene expression profiles, and detecting early-stage disease in patients with melanoma [5]. Multiple dyes staining different markers within the same tissue section can be separated and segmented based on their spectral characteristics. They can then be analyzed qualitatively, for example, determining the distribution of hypoxic protein markers around blood vessels [6] or quantitatively, by measuring the estrogen receptor status in breast cancer [8].

The aim of this study was to investigate the utility of spectral imaging as a method of quantification of immunohistochemically stained samples of primary rectal cancer by comparing it with the current gold standard of manual visual grading. Thymidylate synthase (TS) is an enzyme involved in thymidine biosynthesis [9]. It is also the main site of action of the chemotherapy agent 5-fluorouracil [10] and has been investigated as a marker of prognosis and response to chemotherapy in a range of tumor types, including colorectal cancer [11,12]. It is preferentially expressed in colorectal tumor tissue compared with normal colonic mucosa [13] and was used in this study to investigate the validity of spectral imaging as a method of immunostain quantification.

2. Materials and methods

2.1. Study patients

Multiple biopsies of histologically proven untreated primary rectal carcinoma were obtained during standard surgical resections from 3 individual patients. In addition, from each patient, 2 samples of adjacent normal rectal mucosa were obtained and used as an internal control. Local ethics committee approval was obtained, and all patients gave informed consent. All biopsies were fixed in formalin solution (10% neutral buffered; Sigma, Poole, UK) for 24 hours. They were then dehydrated in graded alcohols (70%, 90%, and 100%), washed with a clearing agent, embedded in paraffin, and stored at room temperature until subsequent immunohistochemical analysis of TS protein expression.

2.2. Immunohistochemistry

Heat-mediated antigen retrieval was performed by boiling the sections in 250 mL 10 mmol/L citric acid, pH 6, for 3 cycles of 4 minutes each. The sections were then left to stand in citric acid for 10 minutes before washing in water for a further 5 minutes. Sections were then transferred to the Dako Autostaining machine (Dako, Ely, UK) containing peroxidase block (S2023; Dako), the detection reagents (ChemMate HRP, K5001; Dako), and anti-TS primary antibody (gift of Simon Joel, St Bartholomew's Hospital, London) diluted 1:300 in antibody diluent. The Autostainer program included 5 minutes in peroxidase block, 1-hour incubation in primary antibody, 30-minute incubation in ChemMate secondary and tertiary reagents, and 5 minutes in diaminobenzidine (DAB) substrate. When the program was complete, stained slides were removed from the machine and counterstained in Gills Haematoxylin (01500E; Surgipath Europe Ltd, Bretton, UK) for 5 seconds. Slides were then washed in tap water, dehydrated in graded alcohols (70%, 90%, and 100%), cleared in xylene, and mounted in DPX (08600E; Surgipath Europe Ltd). Each staining run incorporated a control slide that had previously demonstrated positive for the antibody of interest. A negative control was also incorporated and involved the substitution of the anti-TS primary antibody for an isotypic control antibody at the same protein concentration.

2.3. Visual grading of TS protein expression

The expression of TS within all sections was quantified by a visual grading score and by spectral imaging. The visual grading system outlined in Table 1 was used to score manually the intensity and area of TS staining for all sections. All sections were graded on 3 separate occasions by a single investigator, independent of the results of spectral imaging. If the grade varied between scoring runs, the results of the later analysis were chosen because this was considered to be more accurate (because it was derived after consideration of all other sections). An overall estimate of TS intensity was used rather than quantifying the regions of highest expression. In addition, the total positively stained area of the image was chosen rather than the ratio of positive to negative tumor cells because it was seen that almost all tumor cells were positive in all biopsies studied.

2.4. Quantification of TS protein expression by spectral imaging

Thymidylate synthase expression was quantified by spectral imaging using a system developed and constructed in our Institute, as reported previously [6]. The system uses a

Table 1 The visual grading system used for manual scoring of stain intensity and area for all sections

Visual grade	Stain intensity	Area of staining (%)
1	Negative	<20
2	Weak	20-50
3	Moderate	50-75
4	Strong	>75

standard monochrome charge-coupled device (CCD) camera (Brian Reece Scientific, Newbury, UK) with the spectrally selective element placed between the camera and the microscope output port using standard C-mount couplers. The element is based on a linearly variable dielectric bandpass filter, which has a resolution of 15 nm and covers the 400- to 700-nm band with a transmission of more than 40%. For this study, the spectrally resolved device was used with an upright microscope (Optiphot; Nikon, Kingston-Upon-Thames, UK). The spectral imaging process is outlined in Fig. 1.

Images were captured into a personal computer using a 1-GHz processor (Precision 220, 256 MB RAM; Dell, Bracknell, UK) via a frame grabber (type PCI-1409; National Instruments Ltd, Newbury, UK). Software was written in "C" programming language under the Lab Windows/CVI development environment (National Instruments Ltd) and Windows 2000 operating system (Microsoft Corp, San Francisco, CA). The optical density (OD) of the sample was determined from the intensity of light, at a given pixel and wavelength (λ), using Eq. 1:

$$\mathrm{OD}(\lambda) = -\log_{10}\frac{I(\lambda) - I_{\mathrm{black}}}{I_{\mathrm{blank}}(\lambda) - I_{\mathrm{black}}}, \quad (1)$$

where I_{blank} is the intensity through a blank part of the section (of equivalent optical thickness) and I_{black} is the intensity with no illumination. Optical density is proportional to the concentration (C) of immunostain present within the section, as given by the Beer-Lambert law in Eq. 2:

$$\mathrm{OD}(\lambda) = E(\lambda) \times C \times d, \quad (2)$$

where E is the wavelength-dependent extinction coefficient and d is the thickness of the sample (which was constant throughout the experiment).

The characteristic spectra of the dyes used in the immunohistochemical process (DAB, hematoxylin-eosin) were generated by analyzing sections stained only with the single dye. These spectra were then used to determine the contribution of each dye to the spectra at each pixel of the sample image by applying a nonnegative least squares unmixing algorithm [14]. This allowed separation of DAB staining from hematoxylin and eosin staining.

There are several techniques that may be used to analyze spectral imaging data. Principal components analysis can be used to reduce the dimensionality and identify the major components that comprise the measured spectra. Because, in this case, the components are known, more straightforward

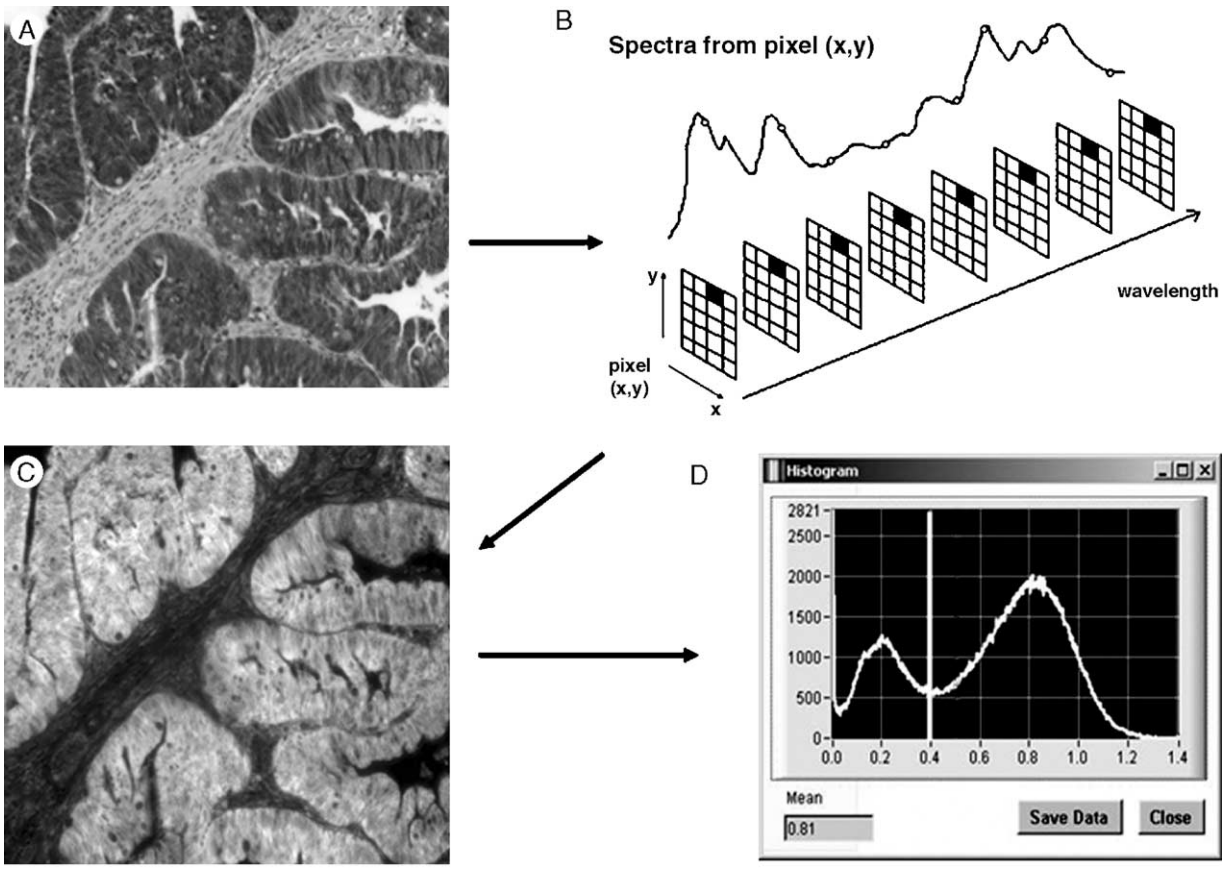

Fig. 1 The spectral imaging process. The tissue section (A) (DAB-labeled, original magnification ×100) is spectrally resolved (B) to generate a map of TS immunostaining (C). An arbitrary threshold is applied to the frequency histogram (D) (x-axis, OD; y-axis, pixel frequency), allowing segmentation of tumor (higher OD peak) and stromal (lower OD peak) TS expression.

techniques can be used to find the proportions of these components. Because absorptive components add linearly according to the Beer-Lambert law, linear least squares can be used to find a solution to this overdetermined fitting problem. Difficulties can arise because linear least squares may introduce nonphysical negative proportion values; this can be avoided with the use of a nonnegative least squares algorithm, which was the method we used. Better fits may be achieved with noisy data through the use of an iterative algorithm (such as Levenberg-Marquart), but this will be slower and is unnecessary because our data are relatively free from noise.

The results of the OD spectral unmixing represent proportions of the reference spectra and, as such, have arbitrary units (au) of OD normalized to the references. Frequency histograms of normalized OD were generated for DAB and presented as grayscale intensity maps for the whole image, which allowed spatial correlation of TS expression with the histocytological architecture (see Fig. 1).

From the frequency histogram, the mean normalized OD, representing TS intensity for that image, was determined. The area under the frequency histogram curve was also calculated and represented the number of pixels and, therefore, the area of the image, demonstrating TS expression. Because TS is differentially expressed in tumor tissue compared with stroma, the frequency histogram demonstrated 2 peaks (see Fig. 1), a higher OD peak representing TS within tumor tissue and a lower peak corresponding to TS stromal staining. An arbitrary threshold (0.4 au) was chosen to segment the image and minimize involvement of the stromal staining in the final analysis. This allowed selective and objective measurement of tumor-specific TS protein expression. The same threshold was applied to all images.

2.5. Statistics

A t test with unequal variance was used to compare visual grading and spectral imaging estimates. Spearman rank correlation coefficient (r_s) was used to determine the correlation between visual grading and spectral imaging estimates; the statistical significance was the 2-tailed P value for rejecting the hypothesis of zero correlation.

3. Results

3.1. Image capture and examples of staining

Twenty-eight sections of rectal carcinoma from 3 individual patients were stained for TS and used in this study. For 24 sections, an image was captured from 3 different tumor areas. Because of the small size of tumor tissue within 3 sections, 2 areas were used to capture images and 1 section had only 1 area imaged. Overall, there were 79 individual captured images that were quantified by both visual grading and spectral imaging (61 tumor and 18 control sections). Representative tumor and control sections showing the variability in stain intensity and area are shown in Fig. 2.

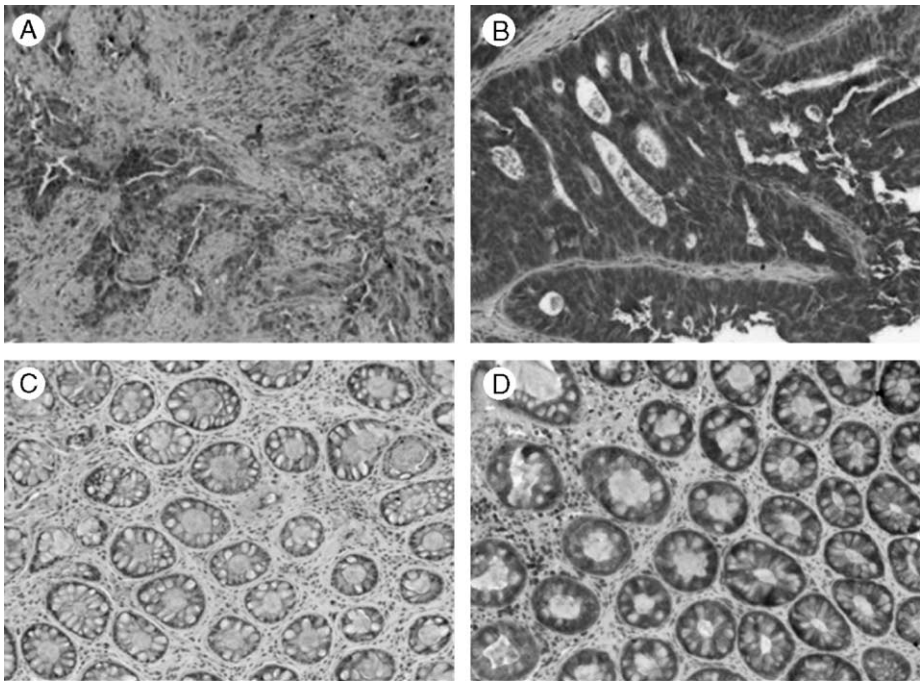

Fig. 2 Representative images of tumor (A and B) and control (C and D) sections showing variability in TS stain intensity and area (DAB-labeled, original magnification ×100).

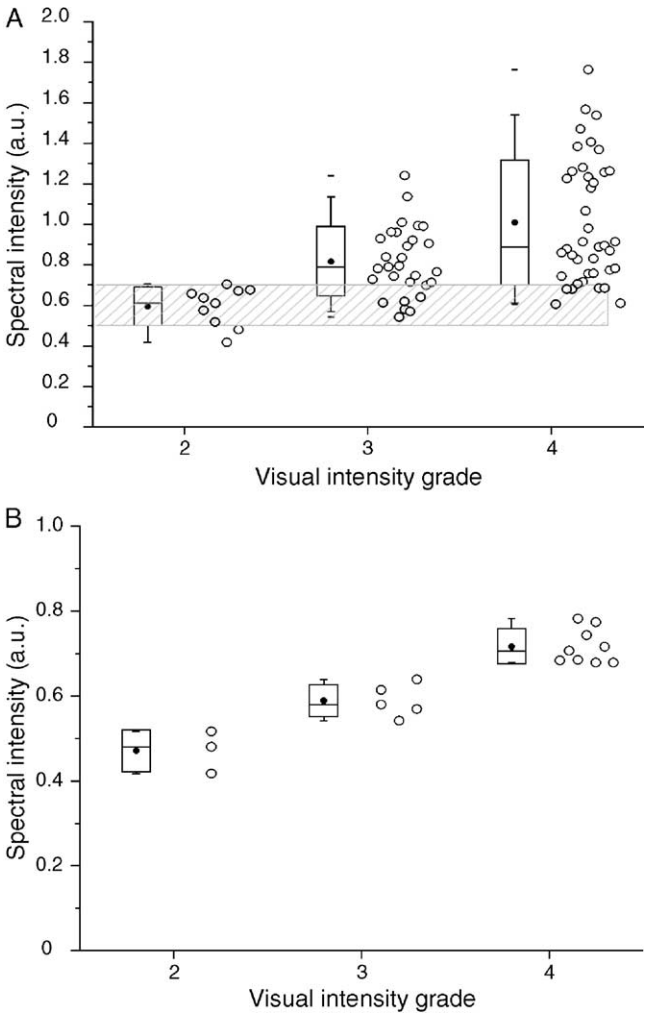

Fig. 3 Comparison of stain intensity measured by spectral imaging (au) and visual intensity grade for all tumor (A) and control (B) sections. The shaded area of A represents the zone of uncertainty for spectral stain intensities between 0.5 and 0.7 au. Box and whisker plots are shown with the mean (dot) and median (line) values; the box represents the SD, and the whiskers, the 95% confidence limits.

3.2. Comparison of visual grading and spectral imaging stain intensity

All tumor and control sections demonstrated at least partial TS expression; hence, no sections were visually graded as 1 for stain intensity. Fig. 3A shows the spectral imaging data plotted as a function of the visual grading scores. The mean OD and SD for visual grades 2, 3, and 4 were 0.60 ± 0.10, 0.82 ± 0.17, and 1.01 ± 0.31 au, respectively. Linear regression analysis revealed a significant correlation between the 2 quantification methods ($r_s = 0.51$, $P < .0001$). The mean spectral TS stain intensity for images with a visual grade of 2 was significantly less than for images with a visual grade of 3 or 4 (both $P < .0001$). Similarly, the mean spectral TS stain intensity for images with a visual grade of 3 was significantly less than for images with a visual grade of 4 ($P = .0014$). It is evident that there is a "zone of uncertainty" for spectral imaging values between 0.5 and 0.7 au (see Fig. 3A). This represents 20 of the 79 images, and within this zone, the probabilities of being visually graded as 2, 3, or 4 were similar (.35, .35, and .3, respectively). The variability of spectral imaging scores increased with each visual grade. The coefficients of variation were 16.0%, 21.0%, and 30.4% in grades 2, 3, and 4, respectively.

Fig. 3B shows the correlation between the 2 methods in control sections of normal rectal mucosa. The expression of TS measured by spectral imaging was lower overall for control sections compared with tumor sections (0.64 versus 0.96 au; $P < .0001$). There was a significant linear relationship between spectral imaging and visual grading for the control sections ($r_s = 0.9$, $P < .0001$). The mean OD and SD for visual grades 2, 3, and 4 were 0.47 ± 0.05, 0.59 ± 0.04, and 0.72 ± 0.05 au, respectively. Although fewer images were assessed, both the variation and the

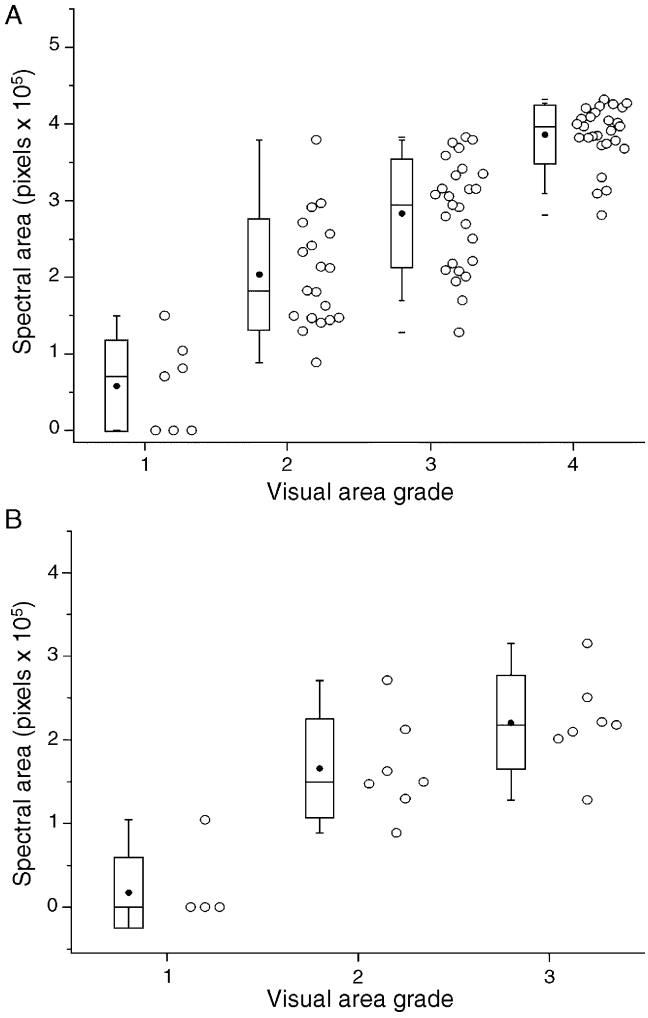

Fig. 4 Comparison of stain area measured by spectral imaging and visual area grade for all tumor (A) and control (B) sections. Box and whisker plots are shown.

overlap in spectral image values within and between visual grades were less than what was seen in the tumor specimens.

3.3. Comparison of visual area grading and spectral imaging pixel number

Fig. 4A shows the spectral pixel data plotted as a function of the visual area grading system. The mean pixel number and SD for visual area grades 1, 2, 3, and 4 were 0.6 ± 0.6, 2.0 ± 0.7, 2.8 ± 0.7, and 3.8 ± 0.3 pixels $\times 10^5$, respectively. Linear regression showed a significant correlation between mean pixel number and visual area grading ($r_s = 0.72$, $P < .0001$). The mean number of pixels for area grade 1 was significantly less than for grade 2 ($P = .0002$). Similarly, pixel number for grade 2 was less than for grade 3 ($P = .0007$), and grade 3 was less than for grade 4 ($P < .0001$).

Fig. 4B shows the correlation between visual area grade and pixel number in control sections. There was a significant linear relationship between spectral imaging and visual grading ($r_s = 0.72$, $P < .001$). The mean pixel number and SD for area grades 1, 2, and 3 were 0.26 ± 0.52, 1.66 ± 0.59, and 2.21 ± 0.56 pixels $\times 10^5$, respectively.

4. Discussion

Traditional methods of assessing immunohistochemical staining of tissue sections have involved visual estimation techniques to derive qualitative and quantitative data regarding the expression of marker proteins. These techniques have been applied to a wide range of antigens, generating abundant literature encompassing a broad field of interest. Their ubiquity stems from the fact that they do not require specialist equipment and are easy to perform. They are limited, however, by their subjectivity and lack of consensus regarding a positive result. In addition, they are generally considered to be only semiquantitative because they lack the discriminatory power to differentiate subtle differences in protein expression between different sections [15]. Furthermore, reproducibility is an issue, resulting in interobserver and interlaboratory discrepancies in interpretation [16,17]. These include which area of the tissue to assess (random areas versus peripheral/central areas), which staining pattern to consider (nuclear versus cytoplasmic, overall staining versus most intense regions, and stain intensity versus area staining positively), and which cutoff level dictates a positive result.

Computer-assisted image analysis has the potential to provide the required objectivity and accuracy, as well as deliver the automation needed for clinical practice. Spectral imaging is a promising form of image analysis that is suitable for both bright field and fluorescence microscopy [7]. Its ability to measure a large number of wavelengths at each pixel is an advantage over standard color imaging and enables the image to be spectrally resolved. This allows histological dyes with characteristic spectra to be separated and permits the simultaneous measurement of multiple analytes present within a section [7]. Spectral imaging has been applied to various areas of biomedicine, but this is the first report of the use of spectral imaging in the quantification of protein expression in colorectal cancer sections. The only previous study comparing spectral imaging and visual grading techniques was concerned with the assessment of the estrogen receptor status in patients with breast cancer [8]. Our study investigated the utility of spectral imaging in the quantification of rectal cancer sections stained for TS by IHC, but the concepts and results are applicable to many other potential prognostic markers in different tumor types. The results show a significant correlation between the 2 quantification methods for intensity and stained area, suggesting that spectral imaging is a valid way of quantifying biologic sections stained by IHC.

Although there was good correlation for stain intensity between the 2 methods, there was considerable overlap and variation of spectral imaging parameters within each visual grading score. This is exemplified by the zone of uncertainty in Fig. 3A, which shows that there is an almost equal probability that a section could be classified as visual grade 2, 3, or 4 if its spectral imaging value fell between 0.5 and 0.7 au. This suggests that there is considerable difficulty in accurately classifying immunostaining by visual estimation when the staining pattern is not at the extremes of stain intensity or area. It confirms the subjective nature of visual estimation and suggests that it lacks the discriminatory power needed to classify the staining pattern when the marker expression is at the median level. The major source of instrumental noise in the system will be due to the light intensity measurement of the camera, with an additional source in the stability of the tungsten illumination lamp. The stain reference spectra used in this study can be called upon to convert the stain intensity into real OD at the spectral peak. A stain intensity of 0.6 au at the center of this zone of uncertainty represents an OD of around 0.2. Considering the sources of noise, we expect the signal-to-noise ratio to be around 15 for measurements of this magnitude (0.200 ± 0.014 OD), with a ratio exceeding 20 at higher stain intensities. The scatter in measurements within the zone of uncertainty would represent a signal-to-noise ratio of less than 12 (0.200 ± 0.017 OD), and therefore, although there may be a contribution due to instrumental noise within the uncertainty zone, the scatter at higher intensities must be due to biologic variation or subjective errors on manual grading.

The lack of correlation noted in some sections may be related to a problem of sampling in that if only a very small number of pixels demonstrated stain intensity above the chosen threshold, an erroneously high spectral intensity could be produced if the selected pixels were strongly positive, even if overall the staining was weak, and quantified as such by manual grading. Analysis without a threshold or the application of an extremely low threshold may reduce these erroneously high means at the expense of poor tumor segmentation. Alternatively, it may be that an integrated intensity (product of average intensity and pixel number)

above the threshold would have been a more appropriate method of quantification.

Automated image processing techniques are not without their limitations. Background staining is included in the analysis of stain intensity and area, contributing to a false-positive result. This can be reduced by the application of a threshold to the spectral data, but this depends on greater expression of the protein marker in tumor tissue compared with stroma and will inevitably lead to the exclusion of weakly stained tumor areas in the analysis of some sections. Artifacts such as staining of adhesives used in tissue section production or excessive staining around the periphery of a section, the so-called edge artifact [15], also create problems in interpretation. These can be ignored when analyzing the staining pattern manually but would be included with automatic assessment. This highlights the need for robust image segmentation such that defined regions or thresholds can be generated and applied independently.

5. Conclusions

As the biology of cancer and the mechanisms of antitumoral agents are further elucidated, there will be a greater need for accurate quantification of biologic constituents to predict tumor behavior and permit targeted therapy based on the results of tissue marker quantification. It is vital, therefore, that accurate, objective, and reproducible methods of quantifying IHC are devised, and this study suggests that spectral imaging is a valid method and would meet these requirements. Although visual estimation was considered to be the gold standard, it may be difficult to confirm the validation of spectral imaging because of the subjectivity of the visual estimation technique against which it is being compared. The ultimate validation will come in the ability of spectral imaging to improve the prognostic power of a molecular marker such as TS in clinical specimens.

References

[1] Aebersold DM, Burri P, Beer KT, et al. Expression of hypoxia-inducible factor-1 alpha: a novel predictive and prognostic parameter in the radiotherapy of oropharyngeal cancer. Cancer Res 2001;61:2911-6.

[2] Aschele C, Debernardis D, Casazza S, et al. Immunohistochemical quantitation of thymidylate synthase expression in colorectal cancer metastases predicts for clinical outcome to fluorouracil-based chemotherapy. J Clin Oncol 1999;17:1760-70.

[3] Boku N, Chin K, Hosokawa K, et al. Biological markers as a predictor for response and prognosis of unresectable gastric cancer patients treated with 5-fluorouracil and *cis*-platinum. Clin Cancer Res 1998;4:1469-74.

[4] Aziz DC, Barathur RB. Quantitation and morphometric analysis of tumors by image analysis. J Cell Biochem Suppl 1994;19:120-5.

[5] Farkas DL, Becker D. Applications of spectral imaging: detection and analysis of human melanoma and its precursors. Pigment Cell Res 2001;14:2-8.

[6] Barber P, Vojnovic B, Atkin G, et al. Applications of cost-effective spectral imaging microscopy in cancer research. J Phys D:Appl Phys 2003;36:1729-38.

[7] Farkas DL, Du C, Fisher GW, et al. Non-invasive image acquisition and advanced processing in optical bioimaging. Comput Med Imaging Graph 1998;22:89-102.

[8] Rothmann C, Barshack I, Gil A, et al. Potential use of spectral image analysis for the quantitative evaluation of estrogen receptors in breast cancer. Histol Histopathol 2000;15:1051-7.

[9] Aschele C, Lonardi S, Monfardini S. Thymidylate synthase expression as a predictor of clinical response to fluoropyrimidine-based chemotherapy in advanced colorectal cancer. Cancer Treat Rev 2002;28:27-47.

[10] van der Wilt CL, Peters GJ. New targets for pyrimidine antimetabolites in the treatment of solid tumours. 1: thymidylate synthase. Pharm World Sci 1994;16:84-103.

[11] Edler D, Hallstrom M, Johnston PG, et al. Thymidylate synthase expression: an independent prognostic factor for local recurrence, distant metastasis, disease-free and overall survival in rectal cancer. Clin Cancer Res 2000;6:1378-84.

[12] Okonkwo A, Musunuri S, Talamonti M, et al. Molecular markers and prediction of response to chemoradiation in rectal cancer. Oncol Rep 2001;8:497-500.

[13] Paradiso A, Simone G, Petroni S, et al. Thymidilate synthase and p53 primary tumour expression as predictive factors for advanced colorectal cancer patients. Br J Cancer 2000;82:560-7.

[14] Lawson C, Hanson R. Solving least squares problems. 1995 rev. ed. New Jersey: Prentice-Hall; 1974 [SIAM, Society for Industrial and Applied Mathematics].

[15] Heyderman E, Warren PJ, Haines AM. Immunocytochemistry today—problems and practice. Histopathology 1989;15:653-8.

[16] Levenson RM, Hoyt C. Spectral imaging and microscopy. Am Lab 2000:26-33.

[17] Seidal T, Balaton AJ, Battifora H. Interpretation and quantification of immunostains. Am J Surg Pathol 2001;25:1204-7.