

Selecting immunohistochemical cut-off scores for novel biomarkers of progression and survival in colorectal cancer

Inti Zlobec, Russell Steele, Luigi Terracciano, Jeremy R Jass, Alessandro Lugli

J Clin Pathol 2007;60:1112–1116. doi: 10.1136/jcp.2006.044537

See end of article for authors' affiliations

Correspondence to:
Ms Inti Zlobec, Department
of Pathology, McGill
University, Duff Medical
Building, Room B22, 3775
University Street, H3A 2B4
Montreal, Québec, Canada;
inti.zlobec@elf.mcgill.ca

Accepted 6 December 2006
Published Online First
20 December 2006

Background: Cut-off scores for determining positivity of biomarkers detected by immunohistochemistry are often set arbitrarily and vary between reports.

Aims: To evaluate the performance of receiver operating characteristic (ROC) curve analysis in determining clinically important cut-off scores for a novel tumour marker, the receptor for hyaluronic acid mediated motility (RHAMM), and show the reproducibility of the selected cut-off scores in 1197 mismatch-repair (MMR) proficient colorectal cancers (CRC).

Methods: Immunohistochemistry for RHAMM was performed using a tissue microarray of 1197 MMR-proficient CRC. Immunoreactivity was scored using a semi-quantitative scoring method by evaluating the percentage of positive tumour cells. ROC curve analysis was performed for T stage, N stage, tumour grade, vascular invasion and survival. The score with the shortest distance from the curve to the point with both maximum sensitivity and specificity, i.e. the point (0.0, 1.0), was selected as the cut-off score leading to the greatest number of tumours correctly classified as having or not having the clinical outcome. In order to determine the reliability of the selected cut-off scores, 100 bootstrapped replications were performed to resample the data.

Results: The cut-off score for T stage, N stage, tumour grade and vascular invasion was 100% and that for survival 90%. The most frequently selected cut-off score from the 100 resamples was also 100% for T stage, N stage, tumour grade, and vascular invasion and 90% for survival.

Conclusions: ROC curve analysis can be used as an alternative method in the selection and validation of cut-off scores for determining the clinically relevant threshold for immunohistochemical tumour positivity.

Immunohistochemistry (IHC) is an indispensable research tool frequently used to study tumour progression and prognosis in colorectal cancer (CRC). However, the clinical utility of its findings is largely dependent on the methods used to evaluate immunoreactivity. A large number of studies in CRC define positive protein expression using a predetermined and often arbitrarily set cut-off score, frequently 10%.^{1–11} In addition, staining intensity is often assessed despite concerns of subjectivity, reproducibility and the effect of storage time on tissue samples.^{12–16} The choice of scoring method, in particular the selection of cut-off scores for positivity is rarely addressed. The lack of standardised scoring systems has led to a wide range of methods, many unvalidated, for evaluating IHC in CRC. This factor may largely be responsible for the contradictory results of similar studies evaluating the same protein and the difficulty in ascertaining the prognostic value of potential tumour markers.¹⁷

ROC curves are commonly used in clinical oncology to evaluate and compare the sensitivity and specificity of diagnostic tests.^{18–23} In addition, they allow one to identify the threshold value above which a test result should be considered positive for some outcome.¹⁸ Established applications of ROC curve analysis in clinical oncology include the performance of standard and novel multi-marker models for the prediction of response in tamoxifen-treated breast cancer patients,²⁴ the accuracy of carcinoembryonic antigen to correctly diagnose recurrence of CRC compared to other serum markers²⁵ and the efficiency of MRI, CT and endoluminal ultrasonography to identify local invasion in patients with rectal cancer.²⁶

ROC curve analysis could be applied similarly to evaluate IHC protein expression and to select biologically or clinically relevant cut-off scores for tumour positivity. We have recently

shown that the receptor for hyaluronic acid mediated motility (RHAMM) is an independent prognostic factor and appears to play a role in tumour progression in CRC.²⁷ However, RHAMM is a novel tumour marker and an established cut-off score for this protein has not previously been reported. Therefore, in the present study we evaluate the performance of ROC curve analysis in determining clinically important cut-off scores for RHAMM and demonstrate the reproducibility of the selected cut-off scores in 1197 mismatch-repair (MMR) proficient CRCs.

MATERIALS AND METHODS

Tissue microarray construction

A tissue microarray (TMA) of 1420 unselected, non-consecutive CRCs was constructed.²⁸ Briefly, formalin-fixed, paraffin-embedded tissue blocks of CRC resections were obtained. One tissue cylinder with a diameter of 0.6 mm was punched from morphologically representative tissue areas of each donor tissue block and brought into one recipient paraffin block (3×2.5 cm) using a homemade semiautomated tissue arrayer.

Clinicopathological data

The clinicopathological data for all patients included T stage (T1, T2, T3 and T4), N stage (N0, N1 and N2), tumour grade (G1, G2 and G3), vascular invasion (presence or absence) and disease-specific survival. The distribution of these features is described elsewhere.²⁹

Abbreviations: AUC, area under the curve; CRC, colorectal cancer; IHC, immunohistochemistry; MMR, mismatch-repair; RHAMM, receptor for hyaluronic acid mediated motility; ROC, receiver operating characteristic; TMA, tissue microarray

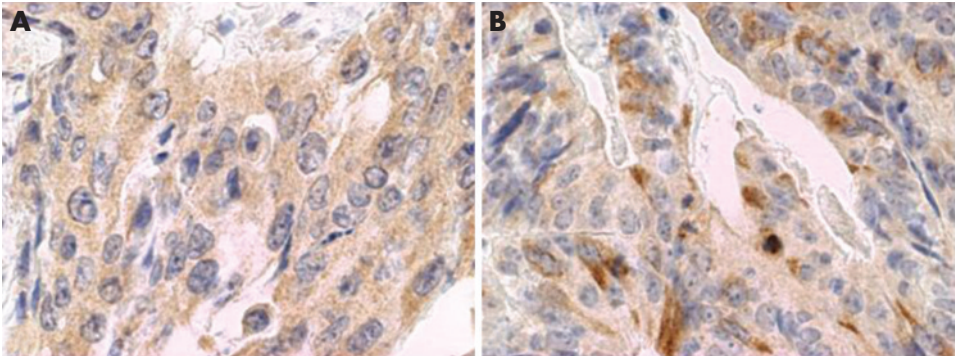


Figure 1 Cytoplasmic expression of receptor for hyaluronic acid mediated motility (RHAMM) in a moderately differentiated mismatch-repair-proficient colorectal cancer (40×) with (A) 100% tumour cell positivity, and (B) <100% tumour cell positivity.

Immunohistochemistry

Sections (4 µm) of TMA blocks were transferred to an adhesive-coated slide system (Instrumedics, Inc., Hackensack, NJ, USA). Briefly, 1420 CRC punches were dewaxed and rehydrated in dH₂O. Endogenous peroxidase activity was blocked using 0.5% H₂O₂. The sections were incubated with 10% normal goat serum (Dako Cytomation, Carpinteria, CA, USA) for 20 min and incubated with primary antibody at room temperature (MLH1 clone MLH-1, BD Biosciences Pharmingen, San Jose, CA, USA; MSH2 clone MSH-2, BD Biosciences Pharmingen; MSH6 clone 44, Transduction Laboratories, San Jose, CA, USA; RHAMM clone 2D6; Novocastra, UK). Subsequently, sections were incubated with peroxidase-labelled secondary antibody (DakoCytomation) for 30 min at room temperature. For visualisation of the antigen, the sections were immersed in 3-amino-9-ethylcarbazole + substrate-chromogen (DakoCytomation) for 30 min, and counterstained with Gill’s haematoxylin.

IHC evaluation

Cytoplasmic immunoreactivity was scored in a semi-quantitative manner by evaluating the proportion of positive tumour cells over total tumour cells in 5% increments (0%, 5%, 10%, ..., 100%). MLH1, MSH2 and MSH6 were scored in the nucleus as negative (0%) or as positive (>0%).

MMR status

The 1420 CRCs were stratified according to DNA MMR status: (1) MMR-proficient tumours expressing MLH1, MSH2 and MSH6; (2) MLH1-negative tumours; and (3) presumed hereditary nonpolyposis CRC cases showing loss of MSH2 and/or MSH6 at any age, or loss of MLH1 at <55 years.³⁰ Only MMR-proficient tumours were included in this study (n = 1197, 84.4%).

Statistical methods

Selection of cut-off scores

The selection of clinically important cut-off scores for RHAMM expression was based on ROC curve analysis.¹⁸ At each percentage score, the sensitivity and specificity for each outcome under study was plotted, thus generating a ROC curve. The score having the closest distance to the point with both maximum sensitivity and specificity, ie the point (0.0, 1.0) on the curve, was selected as the cut-off score leading to the greatest number of tumours which were correctly classified as having or not having the clinical outcome. In order to use ROC curve analysis, the clinicopathological features were dichotomised: T stage (early (T1+T2) or late (T3+T4)), N stage (N0 (no lymph node involvement) or >N0 (any lymph node involvement)), tumour grade (low (G1+G2) or high (G3)), vascular invasion (absent or present), and survival (death due to CRC or censored (lost to follow-up, alive or death from other causes)).

Reproducibility of cut-off scores

In order to determine the reliability of the selected cut-off scores, 100 bootstrapped replications were performed to resample the data.³¹ With bootstrapping, 100 resamples of equal size are created and ROC curve analysis is performed for each subgroup. The most frequently obtained cut-off score (mode) over the 100 resamples and the area under the ROC curve (AUC) and 95% CI were acquired for each analysis. The AUCs summarise the discriminatory power of RHAMM over the entire range of scores for each outcome with values of 0.5 indicating low power and those closer to 1.0 indicating higher power. All analyses were carried out using SAS V.9 (SAS Institute, Cary, NC, USA).

RESULTS
IHC

Immunoreactivity was evaluated in 967 of the 1197 MMR-proficient CRCs, the discrepancy arising due to lack of tissue or tumour in several TMA punches. Immunoreactivity ranged from 0% to 100% (fig 1).

Selection of cut-off scores

The ROC curves for each clinicopathological feature (fig 2) clearly illustrate the point on the curve closest to (0.0, 1.0) which maximises both sensitivity and specificity for the outcome. The cut-off score for T stage, N stage, tumour grade and vascular invasion was 100% and that for survival 90%.

Reproducibility of selected cut-off scores

Figure 3 shows the distribution of cut-off scores obtained from 100 resamples of the data. The most frequently selected cut-off score was 100% for T stage, N stage, tumour grade, and vascular invasion, whereas that of survival was determined to be 90%. Table 1 summarises the AUCs (95% CI).

DISCUSSION

A common problem faced by researchers and pathologists involved with IHC is the determination of the extent of tumour positivity for a given marker which is clinically and biologically

Table 1 Area under the receiver operating characteristic curve (AUC) for each clinicopathological feature

Feature	AUC (95% CI)
T stage	0.54 (0.49 to 0.58)
N stage	0.56 (0.52 to 0.6)
Tumour grade	0.58 (0.52 to 0.65)
Vascular invasion	0.54 (0.50 to 0.58)
Survival	0.69 (0.65 to 0.73)

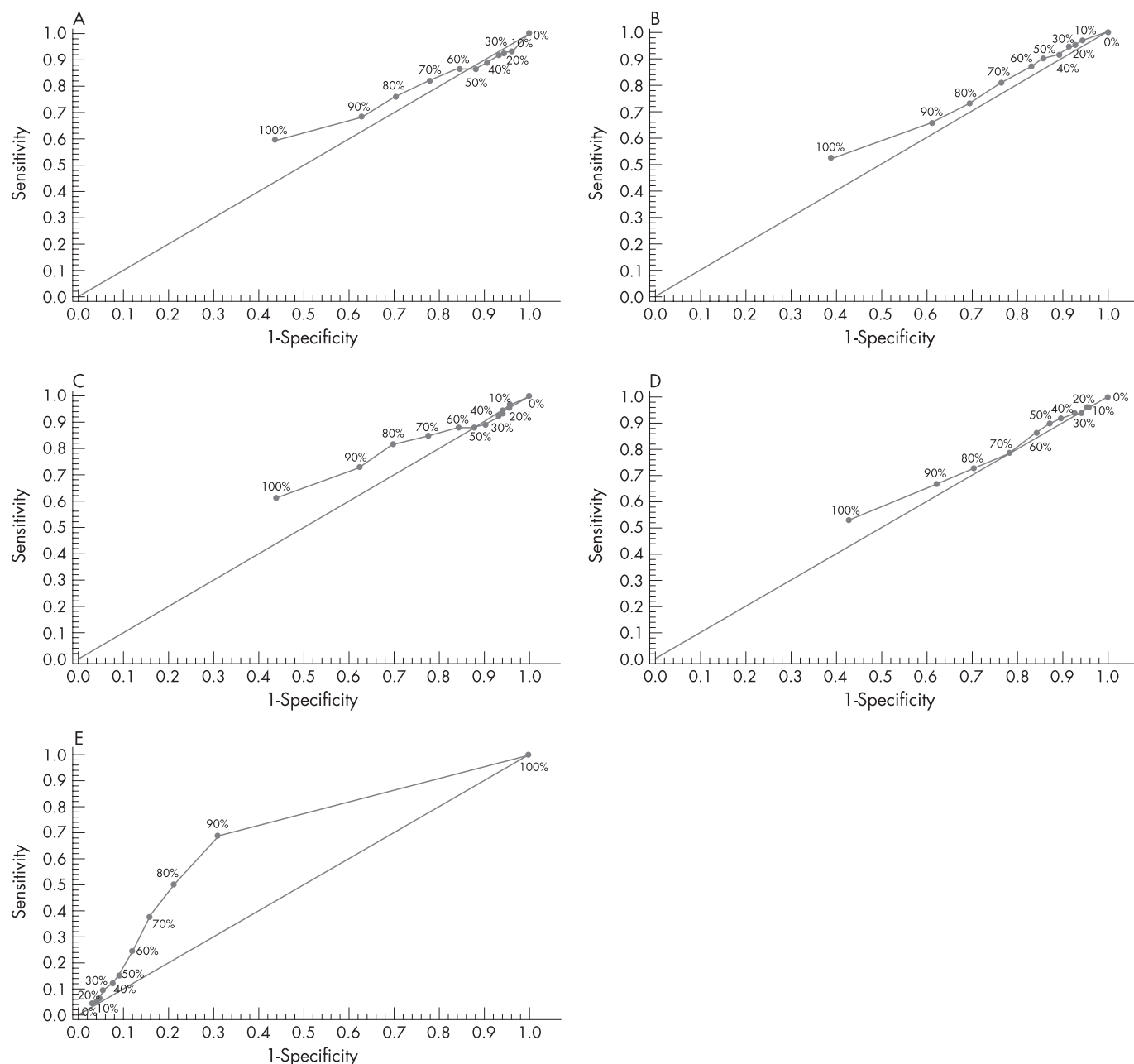


Figure 2 Receiver operating characteristic curves for receptor for hyaluronic acid mediated motility (RHAMM) and T stage (A), N stage (B), tumour grade (C), vascular invasion (D) and survival (E).

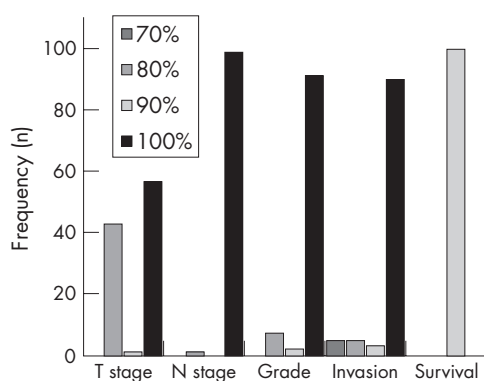


Figure 3 Distribution of cut-off scores obtained from 100 bootstrap replications of receptor for hyaluronic acid mediated motility (RHAMM).

relevant. This is often assessed using a predetermined cut-off score which, particularly for novel tumour markers, is often set arbitrarily and varies between different reports.¹⁻¹¹

In this study we propose a method for determining cut-off scores which should improve the clinical utility of IHC findings. ROC curve analysis is an established method¹⁸ in other areas of medical research, but has not previously been used in the context of IHC to select scores for positive protein expression. To demonstrate its application, we chose the protein RHAMM which we previously identified as a potential marker of tumour progression and prognosis in CRC.²⁷ However, its biological function has not been fully elucidated and so no criteria currently exist for determination of a biologically relevant IHC cut-off point.

The results of this study clearly show that the selected cut-off scores from ROC curve analysis are reproducible for each clinicopathological feature studied. The cut-off score leading to

Take-home messages

- Receiver operating characteristic (ROC) curve analysis is an established method in clinical oncology to evaluate sensitivity and specificity of diagnostic tests.
- The evaluation of immunoreactivity using percentage of positive tumour cells is a reproducible scoring method with a strong inter-observer agreement.
- ROC analysis can be used as an alternative method in the selection and validation of cut-off scores for immunohistochemical tumour positivity.
- The cut-off scores are selected such that the trade-off between sensitivity and specificity is the smallest. When investigating different outcomes such as response to treatment, it may be beneficial to choose a cut-off leading to higher sensitivity over specificity.
- ROC analysis shows that for the same tumour marker the cut-off for positivity can vary with different clinicopathological endpoints.

the best discrimination of tumours with and without the outcome was 100% (100% vs <100% staining) for T stage, N stage, tumour grade and vascular invasion and 90% ($\geq 90\%$ vs $<90\%$ staining) for survival.

The cut-off scores were selected such that the trade-off between sensitivity and specificity was the smallest, therefore leading to the greatest overall number of correctly classified tumours with and without the clinicopathological feature. However, it may be more beneficial when investigating different outcomes, such as response to treatment, to choose a cut-off leading to higher sensitivity rather than specificity. This would allow for the selection of the greatest number of potentially responsive candidates for treatment.

It should be emphasised that categorising protein expression around the selected cut-off score does not imply significant statistical associations with the outcome. However, significant associations may be more biologically meaningful and more likely to occur when appropriate cut-off scores are used to assess positivity.

The use of ROC curve analysis is based on the premise that the evaluation of immunoreactivity using the percentage of positive tumour cells is a reproducible scoring method. We have previously found strong inter-observer agreement using this scoring method in several tumour markers in rectal cancer.³² The intra-class correlation coefficient (ICC) is an accepted method for determining agreement for semi-continuous IHC scores.³³ We have investigated the reproducibility of this scoring method on the same TMA for proteins APAF-1 and EGFR and have found the scores to be highly consistent and reproducible among pathologists (ICC = 0.75 and 0.86 respectively) (unpublished data).

It should be mentioned that time-dependent ROC curves for analysing survival time have been established³⁴ and software recently developed to analyse these outcomes (survivalROC package in R software, The R Development Core Team, V.2.4.0, 2006). Using this method we determined that the AUC for RHAMM was 0.613 using the Kaplan–Meier estimator and 0.608 with the nearest neighbour estimator. Both these results are similar to the AUC we obtained in this study. Time-dependent ROC curves are advantageous as they take into account the number of months until censoring or death from CRC. Though the classic ROC curves illustrated in this study

categorise censored observations or death at the 5-year mark, they are considerably simpler to use.

In conclusion, ROC curve analysis can be used as an alternative method in the selection and validation of cut-off scores for determining the most clinically relevant threshold for immunohistochemical tumour positivity. We recommend that this method be used not only for novel tumour markers, but also to re-evaluate protein expression in established biomarkers that often yield contradictory results.

ACKNOWLEDGEMENTS

We thank Privatdozent Dr Hanspeter Spichtin, Institute of Clinical Pathology Basel, Switzerland and Professor Dr Robert Maurer, Institute of Pathology, Stadtspital Triemli, Zurich, Switzerland for providing the cases, as well as Dr Nilima Nigam, Dr Sanjo Zlobec and Kristi Baker for their input with editing this manuscript.

Authors' affiliations

Inti Zlobec, Jeremy R Jass, Alessandro Lugli, Department of Pathology, McGill University, Montreal, Québec, Canada
Russell Steele, Department of Mathematics and Statistics, McGill University, Montreal, Québec, Canada
Luigi Terracciano, Institute of Pathology, University Hospital of Basel, Basel, Switzerland

Funding: This study was supported by the Faculty of Medicine, McGill University, by a grant from the Swiss National Foundation (grant no PBB5B-110417) and the Novartis Foundation, formerly Ciba-Geigy-Jubilee-Foundation.

Competing interests: None declared.

REFERENCES

- 1 **Galizia G**, Lieto E, Ferraraccio F, *et al.* Determination of molecular marker expression can predict clinical outcome in colon carcinomas. *Clin Cancer Res* 2004;**10**:3490–9.
- 2 **Garrity MM**, Burgart LJ, Mahoney MR, *et al.* Prognostic value of proliferation, apoptosis, defective DNA mismatch repair, and p53 overexpression in patients with resected Dukes' B2 or C colon cancer: a North Central Cancer Treatment Group Study. *J Clin Oncol* 2004;**22**:1572–82.
- 3 **Kang SM**, Maeda K, Onoda N, *et al.* Combined analysis of p53 and vascular endothelial growth factor expression in colorectal carcinoma for determination of tumor vascularity and liver metastasis. *Int J Cancer* 1997;**74**:502–7.
- 4 **Khorana AA**, Ryan CK, Cox C, *et al.* Vascular endothelial growth factor, CD68, and epidermal growth factor receptor expression and survival in patients with stage II and stage III colon carcinoma: a role for the host response in prognosis. *Cancer* 2003;**97**:960–8.
- 5 **Nehls O**, Klump B, Holzmann K, *et al.* Influence of p53 status on prognosis in preoperatively irradiated rectal carcinoma. *Cancer* 1999;**85**:2541–8.
- 6 **Resnick MB**, Routhier J, Konkin T, *et al.* Epidermal growth factor receptor, c-MET, beta-catenin, and p53 expression as prognostic indicators in stage II colon cancer: a tissue microarray study. *Clin Cancer Res* 2004;**10**:3069–75.
- 7 **Rosati G**, Chiacchio R, Reggiardo G, *et al.* Thymidylate synthase expression, p53, bcl-2, Ki-67 and p27 in colorectal cancer: relationships with tumor recurrence and survival. *Tumour Biol* 2004;**25**:258–63.
- 8 **Saw RP**, Morgan M, Koorey D, *et al.* p53, deleted in colorectal cancer gene, and thymidylate synthase as predictors of histopathologic response and survival in low, locally advanced rectal cancer treated with preoperative adjuvant therapy. *Dis Colon Rectum* 2003;**46**:192–202.
- 9 **Schwandner O**, Schiadeck TH, Bruch HP, *et al.* p53 and Bcl-2 as significant predictors of recurrence and survival in rectal cancer. *Eur J Cancer* 2000;**36**:348–56.
- 10 **Suzuki T**, Sadahiro S, Fukasawa M, *et al.* Predictive factors of tumor shrinkage and histological regression in patients who received preoperative radiotherapy for rectal cancer. *Jpn J Clin Oncol* 2004;**34**:740–6.
- 11 **Takahashi Y**, Bucana CD, Cleary KR, *et al.* p53, vessel count, and vascular endothelial growth factor expression in human colon cancer. *Int J Cancer* 1998;**79**:34–8.
- 12 **Atkins D**, Reiffen KA, Tegtmeier CL, *et al.* Immunohistochemical detection of EGFR in paraffin-embedded tumor tissues: variation in staining intensity due to choice of fixative and storage time of tissue sections. *J Histochem Cytochem* 2004;**52**:893–901.
- 13 **Goldstein NS**, Armin M. Epidermal growth factor receptor immunohistochemical reactivity in patients with American Joint Committee on Cancer Stage IV colon adenocarcinoma: implications for a standardized scoring system. *Cancer* 2001;**92**:1331–46.
- 14 **Italiano A**, Vandenbos FB, Otto J, *et al.* Comparison of the epidermal growth factor receptor gene and protein in primary non-small-cell-lung cancer and metastatic sites: implications for treatment with EGFR-inhibitors. *Ann Oncol* 2006;**17**:981–5.

- 15 Kim JS, Kim JM, Li S, et al. Epidermal growth factor receptor as a predictor of tumor downstaging in locally advanced rectal cancer patients treated with preoperative chemoradiotherapy. *Int J Radiat Oncol Biol Phys* 2006;**66**:195–200.
- 16 Spano JP, Lagorce C, Atlan D, et al. Impact of EGFR expression on colorectal cancer patient prognosis and survival. *Ann Oncol* 2005;**16**:102–8.
- 17 Anwar S, Frayling IM, Scott NA, et al. Systematic review of genetic influences on the prognosis of colorectal cancer. *Br J Surg* 2004;**91**:1275–91.
- 18 Hanley J. Receiver operating characteristic (ROC) methodology: the state of the art. *Critical Rev Diagn Imagin* 1989;**29**:307–37.
- 19 Punglia RS, D'Amico AV, Catalona WJ, et al. Impact of age, benign prostatic hyperplasia, and cancer on prostate-specific antigen level. *Cancer* 2006;**106**:1507–13.
- 20 Chen Y, Hyrien O, Williams J, et al. Interleukin (IL)-1A and IL-6: applications to the predictive diagnostic testing of radiation pneumonitis. *Int J Radiat Oncol Biol Phys* 2005;**62**:260–6.
- 21 Reid JF, Lusa L, De Cecco L, et al. Limits of predictive models using microarray data for breast cancer clinical treatment outcome. *J Natl Cancer Inst* 2005;**97**:927–30.
- 22 Al-Homoud S, Purkayastha S, Aziz O, et al. Evaluating operative risk in colorectal cancer surgery: ASA and POSSUM-based predictive models. *Surg Oncol* 2004;**13**(2–3):83–92.
- 23 Lind PA, Wennberg B, Gagliardi G, et al. ROC curves and evaluation of radiation-induced pulmonary toxicity in breast cancer. *Int J Radiat Oncol Biol Phys* 2006;**64**:765–70.
- 24 Linke SP, Bremer TM, Herold CD, et al. A multimarker model to predict outcome in tamoxifen-treated breast cancer patients. *Clin Cancer Res* 2006;**12**:1175–83.
- 25 Carpelan-Holmstrom M, Louhimo J, Stenman UH, et al. CEA, CA 242, CA 19-9, CA 72-4 and hCGbeta in the diagnosis of recurrent colorectal cancer. *Tumour Biol* 2004;**25**:228–34.
- 26 Bipat S, Glas AS, Slors FJ, et al. Rectal cancer: local staging and assessment of lymph node involvement with endoluminal US, CT, and MR imaging—a meta-analysis. *Radiology* 2004;**232**:773–83.
- 27 Lugli A, Zlobec I, Gunthert U, et al. Overexpression of the receptor for hyaluronic acid mediated motility is an independent adverse prognostic factor in colorectal cancer. *Mod Pathol* 2006;**19**:1302–9.
- 28 Sauter G, Simon R, Hillan K. Tissue microarrays in drug discovery. *Nat Rev Drug Discov* 2003;**2**:962–72.
- 29 Lugli A, Zlobec I, Minoo P, et al. Role of the mitogen-activated protein kinase and phosphoinositide 3-kinase/AKT pathways downstream molecules, phosphorylated extracellular signal-regulated kinase, and phosphorylated AKT in colorectal cancer—a tissue microarray-based approach. *Hum Pathol* 2006;**37**:1022–31.
- 30 Hampel H, Stephens JA, Pukkala E, et al. Cancer risk in hereditary nonpolyposis colorectal cancer syndrome: later age of onset. *Gastroenterology* 2005;**129**:415–21.
- 31 Efron B. *An introduction to the bootstrap*. Boca Raton, FL: Chapman & Hall/CRC, 1994.
- 32 Zlobec I, Steele R, Michel RP, et al. Scoring of p53, VEGF, Bcl-2 and APAF-1 immunohistochemistry and interobserver reliability in colorectal cancer. *Mod Pathol* 2006;**19**:1236–42.
- 33 Kirkegaard T, Edwards J, Tovey S, et al. Observer variation in immunohistochemical analysis of protein expression, time for a change? *Histopathology* 2006;**48**:787–94.
- 34 Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 2000;**56**:337–44.

BNF for Children 2006, second annual edition

In a single resource:

- guidance on drug management of common childhood conditions
- hands-on information on prescribing, monitoring and administering medicines to children
- comprehensive guidance covering neonates to adolescents

For more information please go to bnfc.org