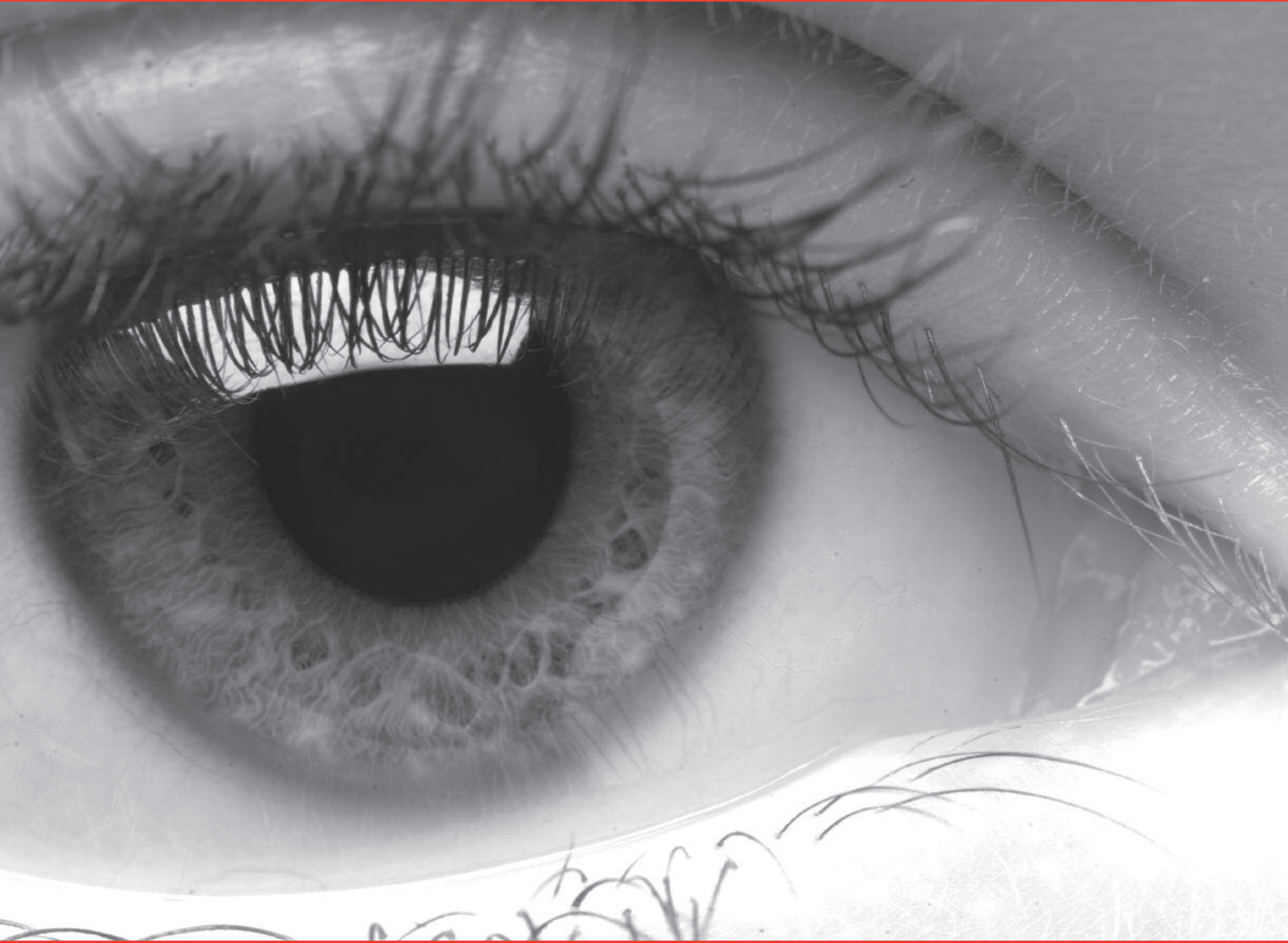


**A GENETIC SIGNATURE OF RELAPSE
IN STAGE II COLORECTAL CANCER...**



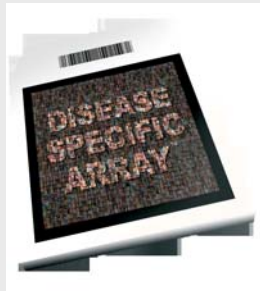
A Genetic Signature of Relapse in Stage II Colorectal Cancer Derived From Formalin Fixed Paraffin Embedded (FFPE) Tissue Using a Unique Disease Specific Colorectal Array.

J von Frese^{1*}, JM Black¹, E Kay², A Tanney¹, R Cummins², S Moore¹, M Doherty¹, V Proutski¹, C Fulton¹, U McDermott³, R Wilson³, H Mulcahy⁴, J O'Sullivan⁴, K Sheehan⁴, D O'Donoghue⁴, K Mulligan¹, D Harkin¹, PG Johnston^{1,3}.

¹ Almac Diagnostics, 19 Seagoe Industrial Estate, CRAIGAVON, N IRELAND; ² Department of Pathology, RCSI ERC, Beaumont Hospital, DUBLIN 9, IRELAND; ³ Centre for Cancer Research and Cell Biology (CCRCB), Belfast City Hospital, BELFAST, N IRELAND; ⁴ Centre for Colorectal Disease, Education & Research Centre, St Vincents University Hospital, Elm Park, DUBLIN 4, IRELAND.

* Juergen.vonFrese@almacgroup.com

Introduction



There is an increasing interest in using archival formalin-fixed paraffin-embedded material (FFPE) for post-genomic diagnostic studies or drug research. Based on the Affymetrix GeneChip[®] technology we have developed a novel microarray platform for this setting. Our range of **Disease Specific Array[™] (DSA[™])** research tools provide a unique, comprehensive coverage of the disease and tissue specific transcriptome based on our in-house high-throughput sequencing and bioinformatics efforts. The 3' bias of the chosen probe sequences allows for a significantly improved performance of **DSA[™]** research tools also for degraded RNA from FFPE samples.

We present a technical assessment of our **Colorectal Cancer DSA[™]** and its first successful application in a clinical application as prognostic tool for stage II colorectal cancer.

Colorectal Cancer DSA™ Research Tool Development

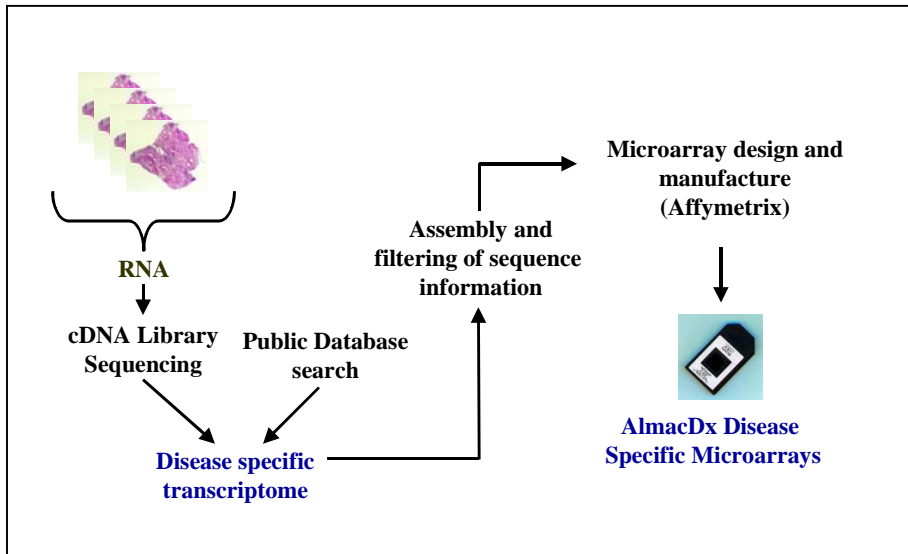


Figure 1. Schematic overview of the generation of the **Colorectal Cancer DSA™** research tool

- The **Colorectal Cancer DSA™** research tool content was generated by a combination of high throughput in-house sequencing, public database mining and experimental investigation.
- The design was optimized to enable profiling of RNA extracted from fresh and paraffin-embedded tissue samples.
- The **Colorectal Cancer DSA™** research tool is designed and manufactured based on Affymetrix GeneChip® technology, i.e. using the same proven standard industry platform, measurement equipment and laboratory protocols.

Array Content

The **Colorectal Cancer DSA™** research tool contains 61,528 probesets covering 52,306 colorectal expressed transcripts:

- 21,968 (42%) are present in the human RefSeq database
- 26,676 (51%) of transcripts are not present in the human RefSeq database
- 7% of the content represents expressed antisense transcripts to annotated genes
- 20,000 transcripts are not detected by current generic microarray platforms

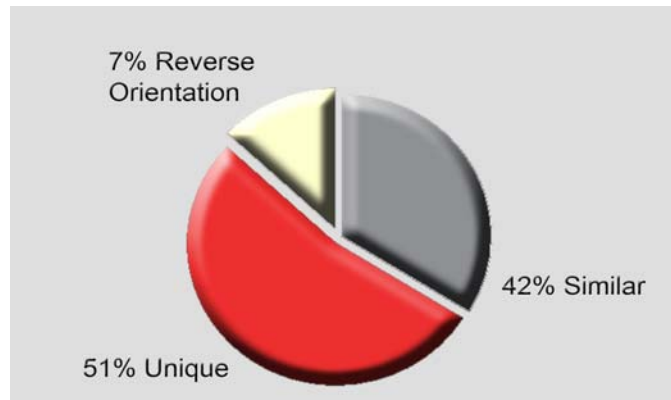


Figure 2. Chart showing the 52,306 **Colorectal Cancer DSA**[™] research tool transcripts, compared to the human RefSeq database by BLAST analysis. RefSeq provides a comprehensive and non-redundant collection of known transcripts.

Sequence content analysis confirms that the **Colorectal Cancer DSA**[™] research tool offers the most comprehensive platform available for the study of colorectal cancer.

Technical Assessment Study

The performance of the array was assessed by profiling a series of technical replicates on both the **Colorectal Cancer DSA**[™] research tool and the Affymetrix U133 Plus 2.0 Human Genome array (Figure 3).

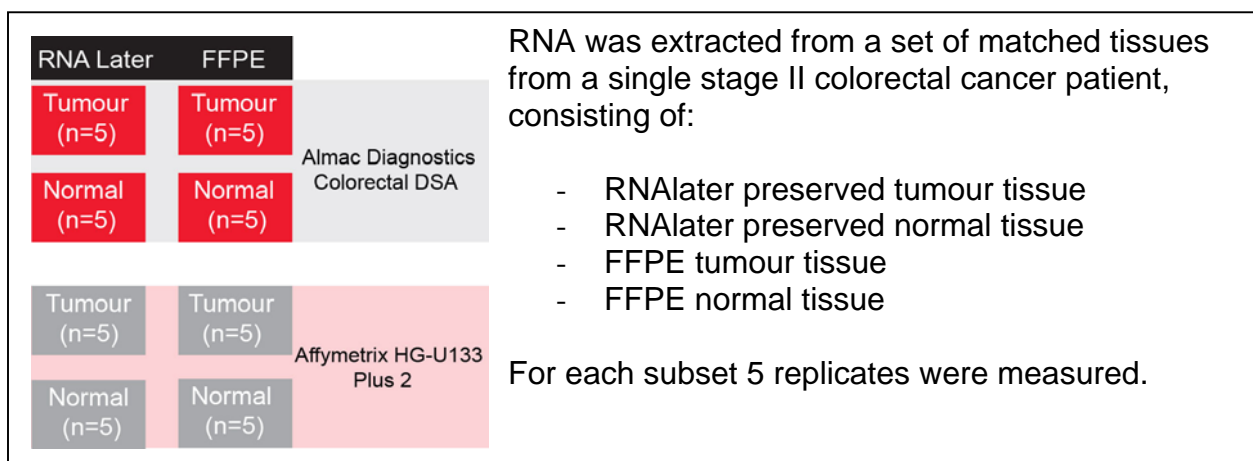


Figure 3. Experimental design of the technical assessment study.

Sensitivity

The number of probesets assessed as “present” by the MAS 5 algorithm can be used as a measure for the detection sensitivity of the arrays. The high detection rates in both RNAlater and FFPE and high degree of data retention clearly demonstrate the power of the Almac Diagnostics **Colorectal Cancer DSA™** when used in FFPE-related studies (Figure 4).

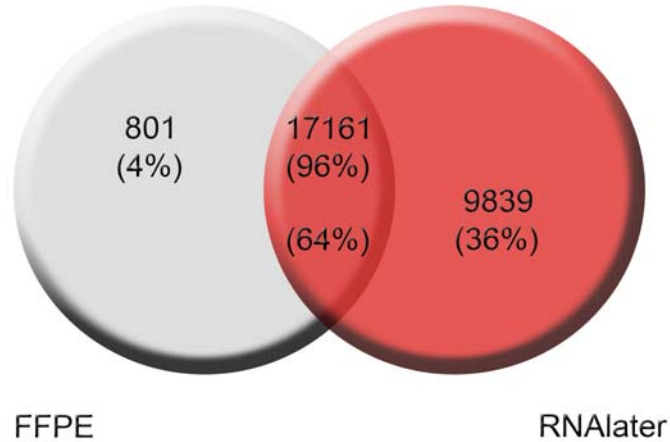


Figure 4. Sensitivity comparison for the **Colorectal Cancer DSA™** research tool between RNAlater® and FFPE. 64 % of the transcripts detected in RNAlater can also be detected in FFPE, whereas only 4 % of the transcripts detected in FFPE cannot be verified in RNAlater.

Precision

Tissue storage method	Array	Joint replicate variation (R ²)	Coefficient of Variation
RNAlater	Colorectal Cancer DSA™ research tool	94.8 %	4.5 %
	U133 Plus 2	95.0 %	4.8 %
FFPE	Colorectal Cancer DSA™ research tool	87.6 %	9.2 %
	U133 Plus 2	86.1 %	10.0 %

Table 1. Precision of the **Colorectal Cancer DSA™** research tool.

- The technical assessment shows a good detection sensitivity and precision of the **Colorectal Cancer DSA™** research tool even for partially degraded RNA from archival FFPE material.

Generation of a Prognostic Signature for Stage II Colorectal FFPE Tissues

Colorectal cancer is currently the second leading cause of cancer death with an estimated 530,000 people dying from this disease each year [1]. Approximately 30% of all colorectal cancer patients are diagnosed with stage II disease [2]. The 5-year survival for patients with stage II colorectal cancer is approximately 75-80%, demonstrating that the majority of patients are cured by surgery alone [3]. However, approximately 20-25% of these patients will develop recurrent disease within their lifetime [4].

Therefore, there is a need to accurately identify those patients who are at high risk of relapse and who would potentially benefit from adjuvant chemotherapy.

Patient Samples

78 stage II colorectal cancer FFPE specimens, collected at a range of centers between 1990 and 2002, from patients who did not receive pre- or post chemotherapy treatment, were obtained after ethical approval.

From the 78 patient specimens:

- 53 were classified as low risk of relapse (or 'good' prognosis), on the basis of five years disease free survival post-resection of the primary tumour
- 25 were classified as high risk of relapse (or 'poor' prognosis), on the basis of relapse (cancer recurrence) within five years of surgical resection of the primary tumour.

Materials and Methods

Total RNA was extracted from a single 10 µm FFPE section from each of the 78 tissue specimens, using an in-house protocol. Biotinylated ARNA targets were generated using the Affymetrix Two Cycle Target labeling kit then hybridized to **Colorectal Cancer DSA™** research tools and scanned using the GeneChip® 7G Scanner.

All arrays were normalized with dChip using Invariant Set Normalization and the Expression Based Model Index was fitted using a PM-only model. Recursive Feature Elimination Support Vector Machines (RFE-SVM) feature selection [5] was performed, which removed the 10% most uninformative genes in every iteration. The optimal number of transcripts, as well as the final performance of the signature, was determined using 10-fold cross-validation including the gene selection within the cross-validation procedure [6].

Prognostic Signature

- A 42-gene prognostic signature was derived from the gene expression data of the FFPE colorectal samples.
- 12 out of 42 transcripts (~29%) were unique to the **Colorectal Cancer DSA™** research tool.
- The Almac Diagnostics 42-gene prognostic signature displayed an overall classification accuracy of 63% with a sensitivity of 80% and a specificity of 55% (Table 2 and Figure 5).

Overall classification accuracy	Sensitivity	Specificity
63%	80%	55%
63% of all stage II colorectal cancer patients will be correctly classified as either low risk or high risk of relapse	80% of high risk patients will be correctly classified	55% of low risk patients will be correctly classified

Table 2. Overall performance of the 42-gene stage II colorectal cancer prognostic signature.

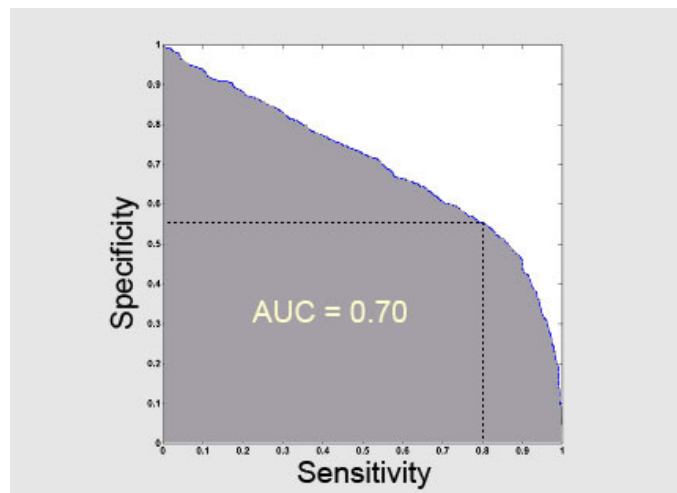


Figure 5. The prognostic signature gave an AUC (area under the curve) value of 0.70 from the ROC (Receiver Operating Characteristic) curve, indicating a high predictive power for stage II colorectal cancer patient outcome.

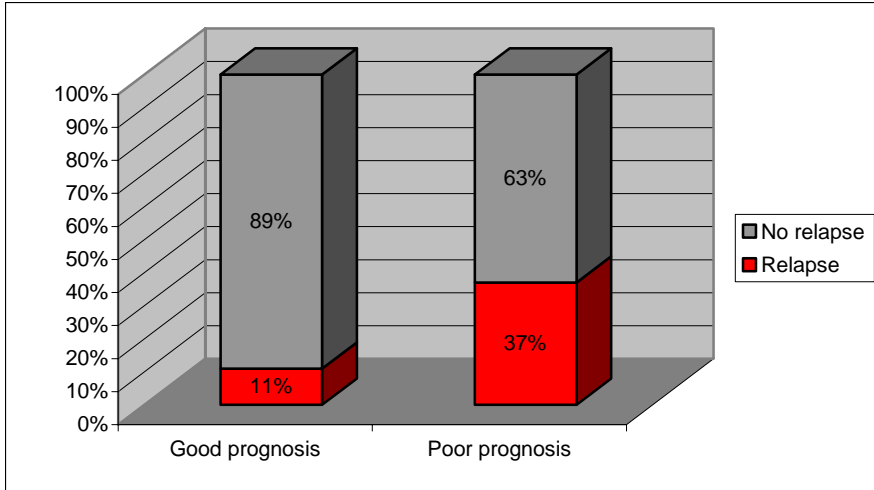


Figure 6. Graphical representation of the % chance of developing recurrent disease (relapse) within 5 years post-surgery. Of the patient groups identified as good or poor prognosis by the 42-gene signature, those identified as 'poor' prognosis have a 3.6 times higher chance of developing recurrent disease compared to those identified as 'good' prognosis.

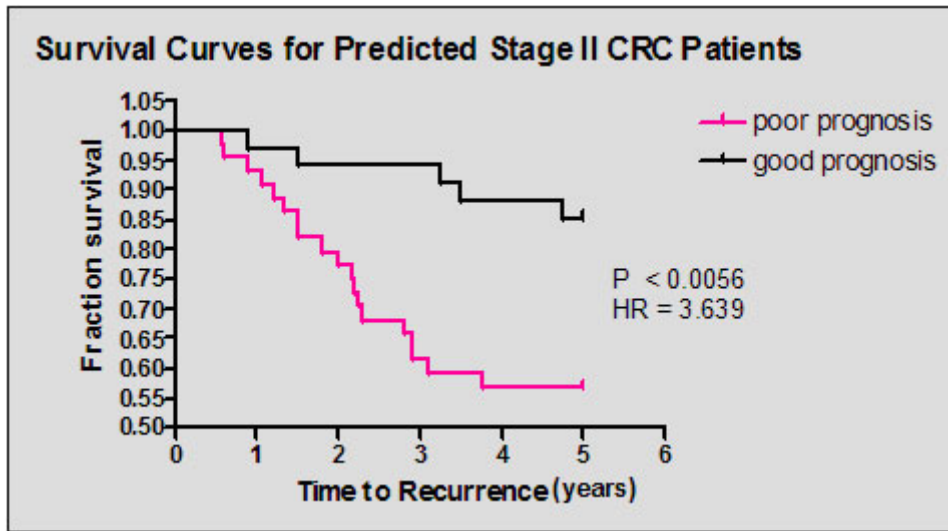


Figure 7. Survival curves for the predicted stage II colorectal cancer patients. Survival curves for the predicted stage II colorectal cancer patients demonstrate significant differences ($p < 0.0056$). The calculated Hazard Ratio (HR) demonstrates the higher risk of relapse among patients predicted to have poor prognosis compared to those predicted to have good prognosis, within 5 years post-surgery.

Conclusions

Almac Diagnostics has developed the first transcriptome based microarray for colorectal cancer (the **Colorectal Cancer DSA™** research tool)

- For a given tissue and disease setting, the **Colorectal Cancer DSA™** research tool provides the most comprehensive insight into the transcriptome.
- The tailored probe design provides a good detection sensitivity and precision for RNA derived from archival formalin-fixed paraffin-embedded tissue.
- Using the **Colorectal Cancer DSA™** research tool we have developed an initial prognostic signature from FFPE tissue that accurately predicts for relapse for stage II colorectal cancer.
- **Colorectal Cancer DSA™ research tool technology now permits the use of paraffin embedded tissue to derive prognostic and predictive signatures in colorectal cancer.**
- **The range of DSA™ research tools is under development and will enable comprehensive analyses of archival material also for other tumors, e.g. breast, lung or ovarian.**

References

- [1] Parkin, DM., Bray, F., Ferlay, J., Pisani, P. Global Cancer Statistics, 2002. *CA Cancer J Clin* (2005); **55**: 74.
- [2] Jemal, A., Tiwari, RC., Murray, T., Ghafour, A., Samuels, A., Ward, E., Feuer, EJ., Thun, MJ. Cancer Statistics, 2004. *CA Cancer J Clin* (2004); **54**: 8.
- [3] Nauta, R., Stablein, DM., Holyoke, ED. Survival of patients with stage B2 colon carcinoma. The Gastrointestinal Tumour Study Group Experience. *Arch Surg* (1989); **124**: 180.
- [4] Gill, S., Loprinzi, CL., Sargent, DJ. et al. Pooled analysis of fluorouracil-based adjuvant therapy for stage II and III colon cancer: who benefits and by how much? *J Clin Oncol* (2004); **22**: 1797.
- [5] Guyon, I., Weston, J., Barnhill, S., Vapnik, V. Gene selection for cancer classification using support vector machines. *Machine Learning* (2002); **46**: 389.
- [6] Ambrose, C., McLachlan, GJ. Selection bias in gene extraction on the basis of microarray gene-expression data. *Proc. Natl. Acad. Sci. USA* (2002); **99**: 6562.

Acknowledgements

We would like to acknowledge the following for their contribution:

Steven Walker
Robert Holt
Joan Hughes
Roisin Scullion
Mark Totten
Donna Mcllwaine

Stewart McWilliams
Maryalice McGurk
Gavin Oliver
Cecilia McCluskey
Peter Kerr