OUTCOME PREDICTION IN CANCER

癌症预后预测

Edited by

AZZAM F.G. TAKTAK

and

ANTHONY C. FISHER

Department of Clinical Engineering
Royal Liverpool University Hospital
United Kingdom

科学出版社

北京

图字：01-2007-5359

This is an annotated version of

OUTCOME PREDICTION IN CANCER

Azzam F.G. Taktak and Anthony C. Fisher

ISBN-13: 978-0-444-52855-1

ISBN-10: 0-444-52855-5

图书在版编目(CIP)数据

癌症预后预测＝Outcome Prediction in Cancer：英文/(英)塔克塔克（Taktak，A. F. G)等编著.—北京：科学出版社，2007

ISBN 978-7-03-020347-2

Ⅰ. 癌… Ⅱ. 塔… Ⅲ. 癌-预防(卫生)-英文 Ⅳ. R730.1

中国版本图书馆 CIP 数据核字(2007)第 155612 号

责任编辑：田慎鹏 贾明月/责任印制：钱玉芬/封面设计：耕者设计工作室

科学出版社 出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

中国科学院印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2008 年 1 月第 一 版 开本：787×1092 1/16

2008 年 1 月第一次印刷 印张：31 插页：1

印数：1—1 500 字数：735 000

定价：98.00 元

（如有印装质量问题，我社负责调换〈科印〉）

导　读

据 WHO 报告，目前全球每年新发的癌症病人超过 1000 万，因癌症死亡的人数超过 700 万，占全部死亡人数的 12%，在发达国家占 21%，在发展中国家占 9%。如果这一趋势得不到改善，预计到 2020 年每年新发的病人将达 1500 万；如果各国不采取有力措施，到 2020 年全球每年死于癌症的人数将达到 1000 万，发展中国家癌症总人数将增加 73%，在发达国家增幅为 29%。卫生部公布的 2006 年城乡居民主要死亡原因统计显示，恶性肿瘤已成为中国人的首要死亡原因。所以，积极采取必要的癌症防治措施是各国优先考虑的课题。

癌症预防在过去半个世纪已经取得一定效果，尤其在发达国家，北美和北欧肺癌发病率和死亡率都已经开始下降。但在发展中国家，由于各种因素，癌症预防还刚刚开始引起重视。当前，减少大气和环境污染、改善不良生活习惯，特别是对吸烟的控制和饮食习惯的改变，已经受到普遍的关注；如何开展其他预防措施，特别是疫苗注射（包括 HBV、HPV 在适龄人群的注射），成为各国研究的重要课题。在治疗方面，我们已经开展了半个世纪的新药和新治疗方法的临床研究，深切体会到创新的重要意义。追溯几十年来临床治疗的进展，多数和综合治疗是分不开的，而内科治疗在其中的地位越来越重要，也是当前最活跃的研究领域之一。

经过专家多年来的论证，最近 WHO 将肿瘤确定为可控慢性病，并将很多工作重点前移，例如强调预防，特别是减少人们生活和工作中的致癌物，重视癌前病变的处理和早期发现等。另外的启示是，多年来我们致力于将所有肿瘤细胞完全消灭，以达到“根治”肿瘤的目的，但是很多病人的实际情况是就诊时肿瘤已经远远超越了可能切除的范围，承受这类治疗的体力不足。这时，我们就将病人列为不能手术甚至“不治”，导致无所作为的思想。今天，我们认识到多数慢性病（例如高血压、糖尿病等）虽然不能根治，但病人能长期正常工作，并保持良好生活质量。而且已经有一些肿瘤（如慢性白血病、低度恶性淋巴瘤、浆细胞肿瘤，甚至少数老年的乳腺癌、前列腺癌）病人可以长期带瘤生存。像其他慢性病一样，我们或许可以通过最大限度地提高机体的抗病能力，尽可能地调理减少疾病负荷，控制和减少肿瘤对机体的危害，长期保持病人的良好生活质量，使其与肿瘤“和平共处”。在靶向治疗问世以后，这种观点已经为更多的临床医师所接受。进入 21 世纪，循证医学、诊疗规范化和个体化已经成为临床学术界公认的趋势。新世纪的临床医学需要脱离几千年经验医学的模式，发展为循证医学（Evidence Based Medicine，EBM）。

我几十年来从医的体会是：在临床肿瘤学实践中，我们还远远没有做到完美地将多数病人治愈，还有很多影响预后的因素需要探索。所以，无论预测癌症发生还是发生以后的预后都具有重要的作用。而实现这一愿望需要对大量的试验和临床资料进行严谨科学的分析。但是，无论是个人经验的还是循证医学的结论都必须经受长期临床实践的检验。因此，必须谦虚对待我们的每一个试验结果。

《癌症预后预测》是一本探讨我们对癌症发生发展的知识，从而预测发生的危险程度，

以及发生后经过治疗决定预后的各种可能因素的专著。但本书的主要编者不是临床医生,而是临床工程学的理学博士(Doctor of Philosophy, PhD,亦称哲学博士)。在43位编者中,医学博士(MD)只有5位,理学博士有36位,还有2位是医学博士和理学博士双学位的获得者,他们来自英国、美国、加拿大、意大利、比利时、荷兰、葡萄牙、芬兰和希腊等国家。这种安排的目的是试图从不同的角度,通过计算,得出可信程度比较高的模型,预测癌症的发生和预后。正如P.J.G. Lisboa在前言中说的:"我们生活在一个充满矛盾的时代,一方面科学技术进步很快,但另一方面我们对个别病人预后的预测仍然很茫然。无论规范治疗或替代治疗,病人治疗后结局的未知数常常很多。"

本书系统回顾了近年来分子标记物、组织病理学和临床表现等综合治疗的进展,临床试验的数据,以及精确的非线性数学和统计学推算的规范化治疗后果等资料。在最大限度收集病人生物数据(包括肿瘤的诊断、分化程度、关于死亡和复发的完整资料等)的基础上,希望建立一种预测病人预后的模式。这无疑对临床上如何处理病人具有决定性的重要意义。

本书分为五部分。第一部分是关于临床问题的讨论,包括口腔癌手术治疗、眼球内黑色素瘤手术治疗和最近有关相对生存率的分析等影响预后的资料;第二部分是生物和遗传因素对预后的影响,作者介绍了导致肺癌发生的环境和遗传因、头颈部癌影响生存率的资料和分歧;第三部分是预后因素的数学模型,主要介绍了淋巴结阴性乳腺癌的分析;第四部分名为机械学习方法的应用,介绍了人工神经网络用于癌症诊断和预后的判断,机械学习对解决医学预后问题的贡献,MRI、SPET资料对脑肿瘤的分类,根据家谱对遗传性非息肉大肠癌自动分析,微阵列技术在脑肿瘤的临床应用;第五部分则主要介绍网络信息的传播,包括网络和新一代医学信息系统,网络环境和多中心研究的Geoconda系统和医学预测模型的发展与临床应用等。

不言而喻,如何预测癌症病人的预后是当前具有重要意义的课题之一。除了传统的病理分类、临床分期和发展趋向以外,近年来分子生物学的发展提供了个体化治疗和决定预后的重要数据。本书总结了临床表现、组织病理学、恶性程度、治疗方法、肿瘤标记物和各种转移复发的统计学因素在决定预后中的重要地位和影响,并提出一些工作模型,但显然是不成熟和需要进一步完善的。作者强调,各种决定因素对病人预后的影响从统计学角度来看不是线性的,所以必须进一步深入研究才能更好地揭示癌症各个阶段决定预后的关键性因素。

本书内容新颖,视角独特,但并不是一本非常成熟的读物。虽然文中涉及的内容尚存在争议,但具有重要的参考意义,尤其是对我国开展创新的规范化和个体化治疗会有很多启发。

孙燕　院士
中国医学科学院　肿瘤医院
2007年6月

前　言

这是一个矛盾的时代，从遗传研究到惊人的计算机力量，科技高速发展，却依然无法满足社会对于个体化病人护理的期待。在癌症的临床处理中，对于标准疗法或替代疗法的可能预后进行预测尤其重要，而上述矛盾也表现得尤为明显。

本书综述了一系列学科分支的最新发展，其中包括综合的决策支持系统，它的组成包括组织病理学和临床标志的分子标志物等，后者与完善的非线性数学和统计学方法相结合，来精确地预测标准疗法的预后。这些进展为尽可能多地使用个体生物谱进行个体化推断开辟了道路，个体化推断的目的是研究人们所关注结局的协变依赖——典型例子为恶性诊断、癌症发展阶段，以及死亡与复发的事件发生时间(time-to-event)统计。

普通非线性算法的最新改进使复杂决策边界和生存曲线的明晰模型成为可能，而无需求助于有关参数线性或风险比例的限制假设。不过，对于复杂的非线性模型，在明确了用于预测推断的备选生物标志物的生物学意义后，还是应该采用一种现实的方法。原因是遗传指示物可用数量庞大，存在关联错误的风险，而这样的关联在不可见数据的分析中是站不住脚的。本书的部分章节推荐了强大的非线性处理过程，通过真实病例的研究对其进行说明。监察之下的事件发生时间数据的一些分析，对于模拟癌症预后来说至关重要，其基准为已确证的统计学方法。这显示了有力的新式分析框架的强大性、适应性及预测的精准性。获得可靠预测所必需解决的疑难问题也正在研究之中，其中包括模型选择和效用规则化。

实际的决策支持系统也需要有下层基础，用以支持标准化数据的获取，使其可来自远程中心并可远程访问。这些问题和高级分析方法的使用同等重要，因为基于数据的模型的可靠性取决于其基础数据的完备性、完整性及一致性。本书的一个重要信息是，来自数据分析的净增值证明，对数据获取和追踪的标准化操作流程的投入是有道理的。这使得多中心模型和评价研究在大尺度上展开，超越了可能只是单纯依赖已有患者历史记录的范畴。

总之，本书对实际结局预测的一些关键点的最新研究结果做了综述，这对于癌症的处理是极其重要的。它使得技术和科学的一些分支(如细胞遗传学、保健信息学和机器学习等)之间形成了协作的效果，并超越这些层面，延伸到了临床领域。

数据获取、实验室测量、临床操作以及数据记录的严格标准化实践，将会放大已经从完善的数学方法中提炼出的价值。它们会在“死数据”(dead data)、代表性的临床审查追踪以及预期数据之间架起桥梁，从而使人们获得癌症现象学的坚实知识。它们也可以作为实际临床工具，以患者病史为基础得到预测性推断的证据。我希望这本书可以在某种程度上开启这项重大的变革。

Paulo J.G. Lisboa, BSc, PhD, CEng, FIEE, FIMA
School of Computing and Mathematical Sciences
Liverpool John Moores University
United Kingdom

(贾明月　译)

引　言

数十年来，癌症的预后预测一直是临床医生、医务人员和患者所关心的主题。对于患者来说，生存是最重要的预后，因为生存可以帮助他们计划生活并为其家庭成员做出考虑。然而，还有其他同样重要的预后，比如功能丧失、畸形和生活质量。

癌症预后预测的传统方法包括卡-迈（Kaplan-Meier）非参数模型和 Cox 回归半参数模型。人们对人工神经网络（ANNs）在预后预测中的使用也表现出了浓厚的兴趣，因为 ANNs 可以提供大量的理论优势。当数据的模式不是那么明显时，ANNs 在使模型适应于数据方面可以提供更大（但并非无限）的弹性（Ripley，1996）。在建模中使用 ANNs 的主要益处有：第一，允许自变量和因变量之间任意的非线性关联；第二，允许因变量之间所有可能的相互作用；第三，ANNs 不需要明确的分布假设。

许多临床医生已经意识到 ANNs 作为一种辅助分析工具的应用潜力。对于“黑盒子”问题，许多研究者都进行了研究，并提供了 ANNs 的一个统计学框架（Biganzoli *et al.*，1998；Ripley and Ripley，2001；Lisboa *et al.*，2003；Bishop，2004）。

考察新技术是否成功应用于临床的一个实用标准，是文献中临床试验的数量。尽管有大量出版物描述了 ANNs 在医学中的应用，但这一领域的临床试验数量依然很少（Gant *et al.*，2001；Lisboa，2002；Lisboa and Taktak，2006）。临床医生不愿意把这些强大的工具融入日常实践的原因是多方面的。从前有许多研究者曾利用这些技术的“黑盒子”性质进行了实验。该特性的优点的确很有吸引力，但也可能成为这些强大工具的祸因。对其基础学科——数学的理解欠缺常常会导致不恰当的使用，而最终使研究者得出错误的结论。比如一个常见的错误就是，在测试中使用的样本量过少，结果数目有限（Bottaci *et al.*，1997；Das *et al.*，2003）。这种情况下，单独引用测试的准确性作为实践的测量标准就不是很有帮助，因为即使网络完全没有检测到结果，得到的数字也会很高（Ripley and Ripley，2001；Kaiserman *et al.*，2005）。

多数临床试验研究都把 ANNs 的性能和其他方法作了比较，如临床指示剂（Stephan *et al.*，2003）和统计学分析（Remzi *et al.*，2003）。子宫癌中有许多例子使用了广为人知的 PAPNET 体系，这种体系是 ANN 系统中得到 FDA 允许可以临床使用的为数不多的体系之一。它利用 ANNs 从阴道涂片中提取出异常的细胞表观，然后从组织学的角度对其进行描述（Boon and Kok，2001）。另一种更常见的方法是在显微镜下对涂片进行重筛。Mango 和 Valente（1998）已证明，相比于传统的显微镜重筛，PAPNET 体系会检测出更多的阳性结果。这一点也被细胞学家 Sherman 等人（1997）所证实，他们在最初的筛选未得出结论的 200 个样本中观察了 PAPNET 的结果，并将其与传统显微镜、DNA 分析以及组织活检进行了比较。结果证明这些病例中，PAPNET 有可能减少不必要的活检，但同时需要付出假阳性增加的代价。Parekattil 等人（2003）在膀胱癌的一个临床试验中证明，ANN 模型在识别需要膀胱镜检查的患者时更为精确，从而可以节省开支。

这本书从不同角度提供了生存分析的观点。旨在将不同学科的专家集中到一起，这

些专家从完全不同的角度来研究这一问题，但目标是一致的。本书分为以下五个主要部分：

临床难题

本书的第一部分包括着重介绍提供预后的传统方法的章节。这些方法中包括广泛使用的 TNM 分期体系，这种体系以癌症在原发位点的侵袭程度（T）、淋巴结侵袭程度（N）和转移程度（M）为基础。该体系提供反映癌症发展阶段的一个数字，而癌症的发展阶段将影响到预后和疗法的选择。许多文献研究了不同类型癌症中这种体系的真实价值。第一章中，Woolgar 对口腔癌中传统和当代病理特征的预后评价作了全面的综述，并建议使用实用的技巧来帮助病理学者做出评估。Damato 和 Taktak 在第二章中强调了传统方法的一些局限性，包括基线变量的不恰当分类、竞争性结局、对于结局的过少或过多报道引起的偏差结果，以及对结局数据的猜测性阐释等。在癌症特异的存活事件中，若将年龄作为输入变量之一，竞争性风险就变得很明显。因为年纪大的患者退出研究的比例比年轻患者高，这就为模型引入了偏差。第三章中，Hakulinen 和 Dyba 阐述了如何处理竞争性风险的问题。

生物和遗传因素

第四章中，Cassidy 和 Field 对研究肺癌的过程中不同的风险因素以及它们之间的相互作用作了概述。这章内容着眼于发展一种个体的分子遗传和流行病学风险评估模型，用以鉴别高风险的个体，这些个体可能随后被征集参加适当的介入项目。该部分的下一章中，Jones 提出了一种用于阐明癌症细胞无规则性质的细胞通路模型，并解释了为什么一个从基因到表型组织严密（每种都使用严格的通路）的体系（比如生物化学模型），远远不如一个允许流动性的相互连接的体系（比如神经网络）。该章中还提及了一项关于 1000 位喉癌患者的初步研究。

预后模型的数学背景

这一部分对预后模型特别是 ANNs 的数学背景作了详细描述。第六章中，Biganzoli 和 Boracchi 阐述了生存模型中，说明性变量非线性相互作用的数学背景。在以超高通量数据将这些模型应用于遗传和预后数据时，这一方面就显得极为重要。Eleuteri 等人在第七章中观察到了如下事实：传统方法的使用中可能作了太严格的假设。他们还描述了 ANN 模型中如何从数学角度进行特征选择。第八章中，Arsene 和 Lisboa 深入分析了神经网络在统计学方法和参数技术背景中的作用，并把这种模型应用在淋巴结阴性乳腺癌患者中。

机器学习方法的应用

这一部分包括了使用不同类型机器学习法则的许多应用。第九章中，Marchevsky 对应用 ANN 模型的实践方面作了一个实用的综述，并讨论了确证这些模型精确性的一些困难。Baronti 等人在第十章中描述了机器学习方法在头颈部鳞癌中的应用，以及遗传因

素如何改变个体风险，比如烟草致癌物的代谢和 DNA 修补机制中涉及的酶的多态性。Devos 等人在第十一章中讨论了磁共振波谱成像（magnetic resonance spectroscopic imaging，MRSI）的应用，以及大脑肿瘤的自动化表征中其与传统磁共振成像（MRI）的联合使用。这章还讨论了一种用于临床的医疗决策支持系统的重要性，该系统融合了来自一些 MR 和非 MR 技术的数据。第十二章中，Kokuer 等人提出了研究遗传性非息肉性结直肠癌的多种统计学和人工智能模型，目的在于更有规律地筛查那些高风险的人群。该部分的最后一章中，Kounelakis 等人综述了脑癌分析的一些基于基因组学的方法，着重介绍了 DNA 微阵列技术。

信息传播

有一个经常被忽略但相当重要的方面是传播信息并共享知识。为了在临床医生、医务人员和患者之间达到有效沟通，这一点是至关重要的。最完善体系的成败有时取决于信息转化为临床实践的方式，比如建立用户友好的界面工具。互联网便利传播提供了理想媒介，使得临床医生可以很方便的获取信息。第十四章中，Fonseca 等人综述了最新的智能医疗信息系统，与其发展相关的主要问题，以及目前采用的解决方案。最后两章列举了目前所使用的系统的例子。Setzkorn 等人在第十五章中描述了基于网络的标准化和信息共享系统的发展，信息的标准化与共享在多中心的协作中是很重要的，而 Kattan 等人则在第十六章中介绍了一种叫做列线图的简单而有效的沟通工具。列线图是多元模型的一种图形描述，已被使用很长时间，然而考虑到其优势，它其实并没有得到如期待般广泛的应用。

Azzam F.G. Taktak
Department of Clinical Engineering
Royal Liverpool University Hospital, UK

（贾明月　译）

参 考 文 献

Biganzoli, E., P. Boracchi, L. Mariani and E. Marubini, 1998, Feed forward neural networks for the analysis of censored survival data: a partial logistic regression approach. Stat. Med., 17(10), 1169-1186.

Bishop, C.M., 2004, Error functions. In: Neural networks for pattern recognition. Oxford University Press, Oxford, pp. 230-236.

Boon, M.E. and L.P. Kok, 2001, Using artificial neural networks to screen cervical smears: how new technology enhances health care. In: G.V. Dybowski R. (Ed.), Clinical Applications of Artificial Neural Networks Cambridge University Press, Cambridge pp. 81-89.

Bottaci, L., P.J. Drew, J.E. Hartley, M.B. Hadfield, R. Farouk, P.W. Lee, I.M. Macintyre, G.S. Duthie, and J.R. Monson, 1997, Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions. Lancet, 350(9076), 469-472.

Das, A., T. Ben Menachem, G.S. Cooper, A. Chak, M.V. Sivak, J.A. Gonet Jr. and R.C. Wong, 2003, Prediction of outcome in acute lower—gastrointestinal haemorrhage based on an artificial neural network: internal and external validation of a predictive model. Lancet, 362(9392), 1261-1266.

Gant, V., S. Rodway and J. Wyatt, 2001, Artificial neural networks: practical considerations for clinical applications.In: G.V. Dybowski R. (Ed.), Clinical Applications of Artificial Neural Networks. Cambridge University Press, Cambridge pp. 329-356.

Kaiserman, I., M. Rosner and J. Pe'er, 2005, Forecasting the prognosis of choroidal melanoma with an artificial neural network. Ophthalmology, 112(9), 1608.

Lisboa, P.J., 2002, A review of evidence of health benefit from artificial neural networks in medical intervention. Neural Netw., 15(1), 11-39.

Lisbosa, P.J., H. Wong, P. Harris and R. Swindell, 2003, A Bayesian neural network approach for modelling censored data with an application to prognosis after surgery for breast cancer. Artif. Intell. Med., 28(1), 1-25.

Lisboa, P.J. and A.F.G. Taktak, 2006, The use of artificial neural networks in decision support in cancer: A Systematic Review. Neural Netw., 19, 408-415.

Mango, L.J. and P.T. Valente, 1998, Neural—network—assisted analysis and microscopic rescreening in presumed negative cervical cytologic smears. A comparison. Acta Cytol., 42(1), 227-232.

Parekattil, S.J., H.A. Fisher and B.A. Kogan, 2003, Neural network using combined urine nuclear matrix protein-22, monocyte chemoattractant protein—1 and urinary intercellular adhesion molecule-1 to detect bladder cancer. J. Urol., 169(3), 917—920.

Remzi, M., T. Anagnostou, V. Ravery, A. Zlotta, C. Stephan, M. Marberger and B. Djavan 2003, An artificial neural network to predict the outcome of repeat prostate biopsies. Urology, 62(3), 456-460.

Ripley, B.D., 1996, Pattern Recognition and Neural Networks. Cambridge University Press, Cambridge.

Ripley, B.D. and R.M. Ripley, 2001, Neural networks as statistical methods in survival analysis. In: G.V. Dybowski R. (Ed.), Clinical Applications of Artificial Neural Networks. Cambridge University Press, Cambridge, pp. 237-255.

Sherman, M.E., M.H. Schiffman, L.J. Mango, D. Kelly, D. Acosta, Z. Cason, P. Elgert, S. Zaleski, D.R Scott, R.J Kurman, M. Stoler and A.T. Lorincz, 1997, Evaluation of PAPNET testing as an ancillary tool to clarify the status of the "atypical" cervical smear. Mod. Pathol., 10(6), 564-571.

Stephan, C., B. Vogel, H. Cammann, M. Lein, V. Klevecka, P. Sinha, G. Kristiansen, D. Schnorr, K. Jung and S.A. Loening, 2003, An artificial neural network as a tool in risk evaluation of prostate cancer. Indication for biopsy with the PSA range of 2-20 mg/l. Urologe A, 42(9), 1221-1229.

目　录

（贾明月　译）

To my wife
Diane

for all the help and support in preparing this book

A. Taktak

Foreword

We live in an era of contrasts, when rapidly developing science and technology from genetic research to availability of astounding computing power, still fall short of addressing societal expectations of individualized patient care. This is especially relevant in the clinical management of cancer, where prediction of the likely outcome from standard or alternative treatments is criticial.

This book reviews recent advances across a range of disciplines, including integrated decision support systems comprising molecular markers, from histopathology to clinical signs, which combine with sophisticated non-linear mathematical and statistical methods to accurately predict outcome from standard treatment. These developments make way for personalized inferences using as much as possible of the individual's bioprofile, in order to explore the covariate dependence of outcomes of interest—typically, diagnosis of malignancy, tumour grading, or time-to-event statistics for mortality and recurrence.

Current improvements in generic non-linear algorithms enable explicit modelling of complex decision boundaries and survival curves, without resorting to limiting assumptions regarding parameter linearity or hazard proportionality. Nevertheless, it is advisable to adopt a realistic approach to complex non-linear modelling, with a clear understanding of the biological significance of the candidate biomarkers for predictive inference, since the availability of vast numbers of genetic indicators, for example, runs the risk of identifying spurious correlations that would not stand up in the analysis of unseen data. Recommendations for robust non-linear processing, illustrated by real-world case studies, have been made in several chapters of this book. Several analyses of censored time-to-event data, so crucial to modelling cancer outcomes, are benchmarked against proven statistical methods. This demonstrates the robustness, flexibility and predictive accuracy achieved by powerful new analytical frameworks. Difficult issues essential for obtaining reliable predictions are addressed, including model selection and efficient regularization.

Practical decision support systems also require an infrastructure supporting standardized data acquisition from remote centres and for remote access. These issues are equally as critical as the use of advanced analysis methods, since reliance on data-based models depends on the completeness, integrity and consistency of the underlying data. An important message of the book is that the added value from data analysis now justifies an investment on standardized protocols for data acquisition and monitoring. This enables multicentre modelling and evaluation studies to take place on a large scale, beyond what is possible solely on the basis of existing historical patient records.

In conclusion, this book is an up-to-date review of the state-of-the-art in several key elements for practical outcome prediction, which is of special importance for the management of cancer. It makes the case for collaborative efforts between technical and scientific disciplines, such as cytogenetics, healthcare informatics and machine learning, and, beyond them, into the clinical arena.

Strictly standardized practices in data acquisition, laboratory measurements, clinical protocols and data recording will magnify the value already abstracted through the use of sophisticated numerical methods. They will bridge the gap between "dead data", representing clinical audit trails, and prospective data, from which solid insights are gained into the phenomenology of cancer. They also serve as practical clinical instruments for predictive inference evidence on the basis of previous patient histories. My hope is that this book will, in some way, inspire the beginning of this momentous transformation.

Paulo J.G. Lisboa, BSc, PhD, CEng, FIEE, FIMA
School of Computing and
Mathematical Sciences
Liverpool John Moores University
United Kingdom

Contents

Contributors

Corneliu T.C. Arsene, PhD, School of Computing and Mathematical Sciences, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK

Flavio Baronti, MSc, Dipartimento di Informatica, University of Pisa, Largo Bruno Pontecorvo, 3, 56127 Pisa, Italy

Pedro Barroso, BSc, UNINOVA, Campus da FCT/UNL, Monte de Caparica, 2829-516 Caparica, Portugal

Elia Biganzoli, PhD, Unità di Statistica Medica e Biometria, Istituto Nazionale per lo Studio e la Cura dei Tumori, Via Venezian 1, 20133 Milano, Italy

Patrizia Boracchi, PhD, Istituto di Statistica Medica e Biometria, Università degli Studi di Milano, Milano, Italy

Lutgarde M.C. Buydens, MSc, PhD, Laboratory for Analytical Chemistry, Radboud University Nijmegen, PO Box 9010, 6500 GL Nijmegen, The Netherlands

Adrian Cassidy, BSc, MSc, Roy Castle Lung Cancer Research Programme, The University of Liverpool Cancer Research Centre, The University of Liverpool, 200 London Road, Liverpool L3 9TA, UK

Bertil E. Damato, MD, PhD, FRCOphth., Ocular Oncology Service, Royal Liverpool University Hospital, Liverpool L7 8XP, UK

Andy Devos, MSc, PhD, SCD-SISTA, Department of Electrical Engineering, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, 3001 Heverlee (Leuven), Belgium

Tadeusz A. Dyba, Sc.D., Finnish Cancer Registry, Liisankatu 21 B, FIN-00170, Helsinki, Finland

Antonio Eleuteri, PhD, Dipartimento di Scienze Fisiche, Università degli Studi di Napoli "Federico II", Italy

John K. Field, MA, PhD, BDS, FRCPath, Roy Castle Lung Cancer Research Programme, The University of Liverpool Cancer Research Centre, The University of Liverpool, 200 London Road, Liverpool L3 9TA, UK

José Manuel Fonseca, PhD, UNINOVA, Campus da FCT/UNL, Monte de Caparica, 2829-516 Caparica, Portugal

Mithat Gönen, PhD, Department of Biostatistics and Epidemiology, Memorial Sloan-Kettering Cancer Center, New York, NY, USA

Marinette van der Graaf, MSc, PhD, Department of Radiology, Radboud University Nijmegen Medical Centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands

Roger Green, PhD, Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland, Canada

Timo R. Hakulinen, Sc.D., Finnish Cancer Registry, Liisankatu 21 B, FIN-00170, Helsinki, Finland

Arend Heerschap, MSc, PhD, Department of Radiology, Radboud University Nijmegen Medical Centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands

Peter Jančovič, PhD, Electronic, Electrical & Computer Engineering, University of Birmingham, Birmingham, UK

Andrew S. Jones, MD, FRCS, Section of Head and Neck Surgery, Faculty of Medicine, University of Liverpool, Clinical Sciences Centre, University Hospital Aintree, Longmoor Lane, Liverpool L9 7AL, UK

Michael W. Kattan, PhD, Department of Quantitative Health Sciences, The Cleveland Clinic Foundation, 9500 Euclid Avenue / Wb-4, Cleveland, Ohio 44195, USA

Münevver Köküer, PhD, Biomedical Computing Research Group (BIOCORE), School of Mathematical and Information Sciences, Coventry University, Priory Street, Coventry CV1 5FB, UK

Xenofon Kotsiakis, MD, Department of Neurosurgery, General Hospital of Chania, Chania, Crete, Greece

Michail G. Kounelakis, M.A.Sc, PhD, Department of Electronic and Computer Engineering, Technical University of Crete, Kounoupidiana, Chania, Crete, Greece – PC 73100

Michele de Laurentiis, MD, Dipartimento di Endocrinologia ed Oncologia Molecolare e Clinica, Università degli Studi di Napoli "Federico II", Italy

Paulo J.G. Lisboa, BSc, PhD, CEng, FIEE, FIMA, School of Computing and Mathematical Sciences, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK

Alberto Mario Marchevsky, MD, Department of Pathology, Cedars-Sinai Medical Center, 8700 Beverly Blvd, Los Angeles CA 90048, USA

Alessio Micheli, PhD, Dipartimento di Informatica, University of Pisa, Largo Bruno Pontecorvo, 3, 56127 Pisa, Italy

Leopoldo Milano, PhD, Dipartimento di Scienze Fisiche, Università degli Studi di Napoli "Federico II" Napoli, Italy

André Damas Mora, MSc, UNINOVA, Campus da FCT/UNL, Monte de Caparica, 2829-516 Caparica, Portugal

Raouf N.G. Naguib, PhD, DIC, CEng, CSci, Biomedical Computing Research Group (BIOCORE), School of Mathematical and Information Sciences, Coventry University, Priory Street, Coventry CV1 5FB, UK

Alessandro Passaro, MSc, Dipartimento di Informatica, University of Pisa, Largo Bruno Pontecorvo, 3, 56127 Pisa, Italy

Peter T. Scardino, MD, Department of Urology, Memorial Sloan-Kettering Cancer Center, New York, NY, USA

Christian Setzkorn, MSc, PhD, Department of Clinical Engineering, Royal Liverpool University Hospital, 1st Floor, Duncan Building, Daulby Street, Liverpool L7 8XP, UK

Arjan W. Simonetti, MSc, PhD, SCD-SISTA, Department of Electrical Engineering, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, 3001 Heverlee (Leuven), Belgium

Antonina Starita, PhD, Dipartimento di Informatica, University of Pisa, Largo Bruno Pontecorvo, 3, 56127 Pisa, Italy

Roberto Tagliaferri, PhD, DMI, Università degli Studi di Salerno, Napoli, Italy

Azzam F.G. Taktak, BEng, PhD, CEng, CSci, Department of Clinical Engineering, Royal Liverpool University Hospital, 1st Floor, Duncan Building, Daulby Street, Liverpool L7 8XP, UK

Sabine van Huffel, MEng, PhD, SCD-SISTA, Department of Electrical Engineering, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, 3001 Heverlee (Leuven), Belgium

Julia A. Woolgar, PhD, FRCPath, FDS RCS Eng., Oral Pathology, Liverpool University Dental Hospital, Pembroke Place, Liverpool L3 5PS, UK

H. Banfield Younghusband, PhD, Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland, Canada

Michalis Zervakis, PhD, Department of Electronic and Computer Engineering, Technical University of Crete, Kounoupidiana, Chania, Crete, Greece – PC 73100

Introduction

Outcome prediction in cancer has been the subject of interest to clinicians, healthcare workers and patients for many decades. Survival is the most important outcome to patients since it helps them plan their lives and provide care to their family members. However, there are other outcomes of equal importance such as loss of functionality, disfigurement and quality of life.

Traditional methods of outcome prediction in cancer include the Kaplan–Meier non-parametric model and the Cox regression semi-parametric model. There has also been considerable interest in the use of artificial neural networks (ANNs) in outcome prediction due to the number of theoretical advantages they offer. ANNs can provide much wider (but not infinite) flexibility in fitting models to data where patterns are not so obvious (Ripley, 1996). The main advantages of using ANNs in modelling are: first, they allow arbitrary non-linear relationships between independent and dependent variables, second, they allow all possible interactions between dependent variables and third, ANNs do not require explicit distributional assumption.

Many clinicians have realised the potential of ANNs as an aid tool in the analysis. The main concern over the "black box" issue has been addressed by a number of researchers who have provided a statistical framework for ANNs (Biganzoli et al., 1998; Ripley and Ripley, 2001; Lisboa et al., 2003; Bishop, 2004).

A useful measure for the success of new technologies in integrating into clinical practice is the number of clinical trials in the literature. Despite a large number of publications describing the use of ANNs in medicine, the number of clinical trials in this area remains small (Gant et al., 2001; Lisboa, 2002; Lisboa and Taktak, 2006). The reluctance of clinicians to readily embrace these powerful tools in everyday practice can be attributed to many factors. In the past, a number of researchers have experimented with these techniques taking advantage of their "black-box" nature. Whilst the benefit of such feature does have its appeal, it can also be the curse on these powerful tools. Lack of understanding of the underpinning mathematical science often leads to inappropriate use of the technique which ultimately leads to wrong conclusions. A common mistake, for example, is the use of far too few samples with a limited number of events in the test set (Bottaci et al., 1997; Das et al., 2003). In such cases, quoting the accuracy alone of the test set as a measure of performance is not very useful since this figure would be high even if the networks did not detect the event at all (Ripley and Ripley, 2001; Kaiserman et al., 2005).

The majority of clinical trial studies compared the performance of ANNs with other methods such as clinical indicators (Stephan et al., 2003) and statistical analysis (Remzi et al., 2003). In cervical cancer, there are many examples on the use of the widely known PAPNET system, one of very few ANN systems to gain FDA approval for clinical use. The system uses ANNs to extract abnormal cell appearance from vaginal smear slides and describe them in histological terms (Boon and Kok, 2001). The alternative more

conventional way is to re-screen the slides under the microscope. Mango and Valente (1998) have shown that the PAPNET system has uncovered a higher proportion of false negatives than conventional microscopic re-screening as confirmed by cytologists. Sherman et al. (1997) looked at the results of PAPNET in 200 specific cases where initial screening was inconclusive and compared them with conventional microscopy, DNA analysis and biopsy. The study showed that for these cases, PAPNET would have reduced unnecessary biopsies but at the expense of increasing false positives. Parekattil et al. (2003) showed in a clinical trial on bladder cancer that their ANN model was more accurate in identifying patients who required cystoscopy thereby providing possible savings.

This book provides an insight into survival analysis from different perspectives. It is aimed at bringing together specialists from different disciplines who deal with the problem from an entirely different angle but share a common goal. The book is organised into the following five main sections:

The clinical problem

The first section of this book contains chapters highlighting the traditional methods for providing prognosis. Such methods involve the widely used TNM staging system based on the extent of tumour involvement at the primary site (T), lymph node involvement (N) and metastasis (M). The system provides a number which reflects the stage of the tumour which influences the prognosis and choice of treatment. A number of studies in the literature have looked into the true value of this system for different types of cancer. In Chapter 1, Woolgar provides an overall review of the prognostic value of traditional and contemporary pathological features in oral cancer and suggests practical tips to aid reporting pathologists in producing their assessment. In Chapter 2, Damato and Taktak highlight some of the limitations of traditional methods including inappropriate categorisation of baseline variables, competing outcome, bias resulting in under- or over-reporting of outcomes and speculative interpretation of outcome data. Competing risks becomes obvious when using age, for example, as one of the input variables in tumour-specific survival. As older patients withdraw from the study at a higher rate than younger ones, this introduces a bias in the model. In Chapter 3, Hakulinen and Dyba explain how to deal with the issue of competing risks.

Biological and genetic factors

In Chapter 4, Cassidy and Field outline various risk factors and the interactions between them in studying lung cancer. The chapter looks at developing an individual molecular genetic and epidemiological risk assessment model to identify high-risk individuals who may subsequently be recruited into an appropriate intervention programme. In the next chapter in this section, Jones proposes a model for cellular pathways illustrating the chaotic nature of cancerous cells and explains how a “top-down” system from gene to phenotype (such as biochemistry models), each employing rigid pathways is far too inferior against a system which allows for fluid interconnections such as a neural network. A pilot study in this chapter involving 1000 patients with laryngeal carcinoma is also described.

Mathematical background of prognostic models

In this section, the mathematical background of prognostic models and ANNs in particular are described in detail. In Chapter 6, Biganzoli and Boracchi explain the mathematical background of non-linear interactions of the explanatory variables in survival models. This aspect is of great importance in applying these models to genetic and proteomic data with very high throughput of data. In Chapter 7, Eleuteri et al. observe the fact that the use of conventional models may involve making too strict assumptions and they describe how feature selections can be carried out mathematically in ANN models. In Chapter 8, Arsene and Lisboa provide an in-depth analysis into the role of neural networks within the context of statistical methods and parametric techniques and apply the model developed in node-negative breast cancer patients.

Application of machine learning methods

A number of applications using various types of machine learning algorithms are included in this section. In Chapter 9, Marchevsky provides a useful overview on the practical aspects of applying ANN models and discusses some of the difficulties in validating the accuracies of these models. In Chapter 10, Baronti et al. describe the application of machine learning methods in head and neck squamous cell carcinoma and how the individual risk is modified by genetic factors, such as polymorphisms of enzymes involved in the metabolism of tobacco carcinogens and in the DNA repair mechanisms. Devos et al. discuss the use of magnetic resonance spectroscopic imaging (MRSI) and the combination with conventional magnetic resonance imaging (MRI) for the automated characterisation of brain tumours in Chapter 11. The importance of a medical decision support system for clinical purposes fusing data from several MR and non-MR techniques is also discussed. In Chapter 12, Kokuer et al. propose various statistical and artificial intelligence models in studying hereditary non-polyposis colorectal cancer with the view of screening those at higher risk more regularly. In the final chapter of this section, Kounelakis et al. review several genomic-based methods for brain cancer analysis with the emphasis on DNA microarrays technology.

Dissemination of information

A very important aspect which is often overlooked is disseminating the information and sharing the knowledge. This is crucial in order to achieve effective communication between clinicians, healthcare workers and patients. The make-or-break of the most sophisticated systems is sometimes dependent on the manner in which information is translated into clinical practice such as building user-friendly interface tools. The internet provides the ideal medium for ease of dissemination making information readily available to clinicians, literally at their fingertips. In Chapter 14, Fonseca et al. review the state-of-the-art intelligent medical information systems, the main problems associated with their development and the currently adopted solutions. The final two chapters provide examples of current systems. Setzkorn et al. describe the development of a web-based system for standardising and sharing information which is essential in multi-centre collaboration in Chapter 15, whilst Kattan et al. focus on simple and effective communication tool called

the nomogram; in Chapter 16. A nomogram is a graphical depiction of a multivariable model which has been used for a long time, but (sadly) not as widely as one might expect given its advantages.

Azzam F.G. Taktak
Department of Clinical Engineering
Royal Liverpool University Hospital, UK

References

Biganzoli, E., P. Boracchi, L. Mariani and E. Marubini, 1998, Feed forward neural networks for the analysis of censored survival data: a partial logistic regression approach. Stat. Med., 17(10), 1169–1186.

Bishop, C.M., 2004, Error functions. In: Neural networks for pattern recognition. Oxford University Press, Oxford, pp. 230–236.

Boon, M.E. and L.P. Kok, 2001, Using artificial neural networks to screen cervical smears: how new technology enhances health care. In: G.V. Dybowski R. (Ed.), Clinical Applications of Artificial Neural Networks. Cambridge University Press, Cambridge pp. 81–89.

Bottaci, L., P.J. Drew, J.E. Hartley, M.B. Hadfield, R. Farouk, P.W. Lee, I.M. Macintyre, G.S. Duthie, and J.R. Monson, 1997, Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions. Lancet, 350(9076), 469–472.

Das, A., T. Ben Menachem, G.S. Cooper, A. Chak, M.V. Sivak, J.A. Gonet Jr. and R.C. Wong, 2003, Prediction of outcome in acute lower-gastrointestinal haemorrhage based on an artificial neural network: internal and external validation of a predictive model. Lancet, 362(9392), 1261–1266.

Gant, V., S. Rodway and J. Wyatt, 2001, Artificial neural networks: practical considerations for clinical applications. In: G.V. Dybowski R. (Ed.), Clinical Applications of Artificial Neural Networks. Cambridge University Press, Cambridge pp. 329–356.

Kaiserman, I., M. Rosner and J. Pe'er, 2005, Forecasting the prognosis of choroidal melanoma with an artificial neural network. Ophthalmology, 112(9), 1608.

Lisboa, P.J., 2002, A review of evidence of health benefit from artificial neural networks in medical intervention. Neural Netw., 15(1), 11–39.

Lisbosa, P.J., H. Wong, P. Harris and R. Swindell, 2003, A Bayesian neural network approach for modelling censored data with an application to prognosis after surgery for breast cancer. Artif. Intell. Med., 28(1), 1–25.

Lisboa, P.J. and A.F.G. Taktak, 2006, The use of artificial neural networks in decision support in cancer: A Systematic Review. Neural Netw., 19, 408–415.

Mango, L.J. and P.T. Valente, 1998, Neural-network-assisted analysis and microscopic rescreening in presumed negative cervical cytologic smears. A comparison. Acta Cytol., 42(1), 227–232.

Parekattil, S.J., H.A. Fisher and B.A. Kogan, 2003, Neural network using combined urine nuclear matrix protein-22, monocyte chemoattractant protein-1 and urinary intercellular adhesion molecule-1 to detect bladder cancer. J. Urol., 169(3), 917–920.

Remzi, M., T. Anagnostou, V. Ravery, A. Zlotta, C. Stephan, M. Marberger and B. Djavan 2003, An artificial neural network to predict the outcome of repeat prostate biopsies. Urology, 62(3), 456–460.

Ripley, B.D., 1996, Pattern Recognition and Neural Networks. Cambridge University Press, Cambridge.

Ripley, B.D. and R.M. Ripley, 2001, Neural networks as statistical methods in survival analysis. In: G.V. Dybowski R. (Ed.), Clinical Applications of Artificial Neural Networks. Cambridge University Press, Cambridge, pp. 237–255.

Sherman, M.E., M.H. Schiffman, L.J. Mango, D. Kelly, D. Acosta, Z. Cason, P. Elgert, S. Zaleski, D.R Scott, R.J Kurman, M. Stoler and A.T. Lorincz, 1997, Evaluation of PAPNET testing as an ancillary tool to clarify the status of the "atypical" cervical smear. Mod. Pathol., 10(6), 564–571.

Stephan, C., B. Vogel, H. Cammann, M. Lein, V. Klevecka, P. Sinha, G. Kristiansen, D. Schnorr, K. Jung and S.A. Loening, 2003, An artificial neural network as a tool in risk evaluation of prostate cancer. Indication for biopsy with the PSA range of 2–20 mg/l. Urologe A, 42(9), 1221–1229.

Section 1

The Clinical Problem

Chapter 1

The Predictive Value of Detailed Histological Staging of Surgical Resection Specimens in Oral Cancer

Julia A. Woolgar

Clinical Dental Sciences, University of Liverpool, Liverpool, UK
Email: jaw@liverpool.ac.uk

Abstract

It is well known that the outcome of squamous cell carcinoma of the oral cavity and oropharynx is related to the stage (i.e. extent) of the tumour, and detailed histopathological assessment of the surgical resection specimen provides information that is central to determining the post-operative treatment needs and prognosis for an individual patient. This chapter reviews in detail the prognostic value of traditional and contemporary pathological features of the primary tumour and the cervical lymph node metastases; and outlines general patient factors such as age, gender and co-morbidity, and considers their relative importance. Practical tips to aid the reporting pathologist in producing a standardized pathological staging assessment are included. The value of the current pathological TNM staging classification is considered and possible amendments and alternatives are explored. The chapter ends with "the way ahead"? – a brief review of molecular and biological markers in oral and oropharyngeal squamous cell carcinoma.

Contents

Outcome Prediction in Cancer
Edited by A.F.G. Taktak and A.C. Fisher

1. INTRODUCTION

Oral cancer – squamous cell carcinoma arising from the surface epithelium lining the mouth and oropharynx (OSCC) – is an important and serious disease. It ranks among the ten most common cancers in the world, accounting for 3–5% of all malignancies (Silverman, 2001). In Europe, the incidence has risen sharply in recent years, particularly in females and younger age groups, apparently due to changing patterns of exposure to tobacco and alcohol, the main aetiological factors which act on a genetically susceptible individual (Bettendorf et al., 2004). Survival has remained at a disappointingly stable level despite significant development in multimodality treatment (Silverman, 2001; Bettendorf et al., 2004), and in the United Kingdom, the death: registration ratio is 0.6 (1400 deaths and 2500 new cases per year) (Hindle et al., 1996). In addition to the high mortality, the disease causes great morbidity, with patients having to cope with both the aesthetic and functional changes resulting from the disease and its treatment.

The extent of the disease at presentation has a major influence on outcome and survival. The disease begins within the surface epithelium and invades the surrounding tissues. In addition to the local spread, metastatic deposits develop in the regional (cervical) lymph nodes in the neck in at least 50% of cases. Blood-borne systemic metastases, mainly to the lungs, liver and bone, are common in the later stages of the disease but death usually occurs as a result of uncontrolled locoregional disease and malignant cachexia (Woolgar et al., 1999; Funk et al., 2002). Outcome is usually measured by actuarial (life tables) survival analysis but consideration of only disease-specific deaths probably underestimates the true impact of the disease due to the frequency of deaths due to cardiovascular and respiratory diseases in the post-operative period, and also deaths indirectly related to the disease, such as suicide (Woolgar et al., 1999).

OSCC can be treated by surgery, radiotherapy or chemotherapy, either alone or in combination, depending on the site and stage of the disease and general factors such as co-morbidity. Clinical staging of the extent of disease at both the primary site and in the neck is notoriously inaccurate, and the value of CT, MRI and SPECT imaging remains uncertain (Woolgar et al., 1995a; Woolgar, 1999a; Chong et al., 2004). The importance of pathological staging of resection specimens, both in selecting patients for adjuvant

therapies and in predicting survival, has been increasingly recognized in recent years. Although the TNM staging classification (UICC, 2002) is widely used throughout the world, it is too crude to offer an accurate prediction in an individual patient, since it considers only the surface diameter of the primary tumour (T); the number, laterality and size of positive lymph nodes (N); and the presence or absence of systemic metastases (M). In recent years, interest has been focused on the histological features of the deep invasive tumour front, and molecular and genetic markers (Bryne et al., 1992, 1995; Martinez-Gimenco et al., 1995; Po Wing Yuen et al., 2002; Sawair et al., 2003).

The main objective of this chapter is to discuss the predictive value of detailed histological assessment of routine surgical specimens from patients with OSCC and to highlight practical considerations including the development of minimum datasets. In addition, it will provide a brief overview of molecular and biological markers, and look at current predictive models and consider future possibilities.

2. PREDICTIVE FEATURES RELATED TO THE PRIMARY TUMOUR

2.1. Surface greatest dimension (tumour diameter)

Surface greatest dimension – "tumour diameter" – is the feature used to indicate tumour size in both the clinical (cTNM) and pathological (pTNM) arms of TNM staging classification system (UICC, 2002). The prognosis of oral cancer worsens as the size at presentation increases and several independent reports in the 1980s showed that large size at presentation is predictive of poor survival (Platz et al., 1983; Crissman et al., 1984; Maddox, 1984).

The diameter (and T stage) of the primary tumour affects both the choice and outcome of treatment. The size of the primary tumour is an important factor in determining the surgeon's ability to obtain tumour-free margins (Scholl et al., 1986; Sutton et al., 2003), and a higher rate of local recurrence is associated with tumours of increasing diameter and T stage (Scholl et al., 1986; Woolgar et al., 1999; Sutton et al., 2003). In patients treated by radiotherapy, tumour size is an important determinant of the dose necessary to effect a cure (Bentzen et al., 1991).

Tumour size is an important predictor of cervical metastasis, and this is a major factor in the correlation between diameter and outcome (Maddox, 1984; Woolgar et al., 1999). Hibbert et al. (1983) attempted to study the prognostic effect of diameter alone, and their results showed that in patients without cervical metastasis, diameter was not significantly related to the 5-year survival. This finding may reflect the poor correlation between tumour diameter and tumour thickness seen in patients without metastasis (Woolgar and Scott, 1995), since tumour thickness rather than diameter appears to be the more important size criterion in relation to both metastasis and survival (see below).

Diameter has the advantage that its clinical assessment is relatively simple compared to the clinical assessment of tumour thickness, and this explains its pivotal role in the TNM staging system. In the routine pathological staging assessment, no account is made for tissue shrinkage during fixation and processing – around 15% of the fresh tissue volume (Batsakis, 1999) – and the maximum diameter of invasive (not merely intraepithelial) carcinoma is measured to the nearest millimetre using an optical micrometer to supplement the

macroscopic inspection of the resection specimen (Woolgar and Scott, 1995; Helliwell and Woolgar, 1998, in press). In addition to tissue shrinkage, discrepancies between the clinical and pathological assessment of tumour diameter may occur due to the inability to distinguish between premalignant lesions and invasive carcinoma without microscopy, and the presence of a poorly cohesive invasive tumour front with extensive undermining of intact mucosa and satellite islands ahead of the main tumour mass, again features that are only detectable on microscopy.

2.2. Tumour thickness

Tumour thickness measurement as a prognostic indicator was first introduced by Breslow (1970) in relation to cutaneous malignant melanomas and the measurement proved to be more objective and reproducible than an assessment of the Clarke level of invasion (Clark et al., 1969) in which histological depth is expressed by reference to the anatomical deep structures reached by the advancing edge of the tumour. The technique was soon applied to squamous cell carcinomas of the skin (Friedman et al., 1985), lip and intra-oral mucosa (Frierson and Cooper, 1986; Mohit-Tabatabai et al., 1986; Shingaki et al., 1988), and the superiority of thickness over diameter was soon recognized. Several independent studies (Shingaki et al., 1988; Nathanson and Agren, 1989; Po Wing Yuen et al., 2002), have shown that tumour thickness is the only size criterion to have independent predictive value on multivariate analysis, particularly when the tumours are from a single intra-oral site or restricted to TNM T1 and T2 categories (diameter less than 40 mm), and it is now widely accepted that thickness is a more accurate predictor of sub-clinical nodal metastasis, local recurrence and survival than diameter (Po Wing Yuen et al., 2002). Nevertheless, the critical thickness differs widely in different reports and it is highly site dependent. For example, the critical thickness in relation to metastasis in floor-of-mouth tumours was only 1.5 mm in the study by Mohit-Tabatabai et al. (1986) compared to 6 mm for tumours of the buccal mucosa (Urist et al., 1987). In tumours of the oral tongue, the critical thickness is less for tumours of the ventral aspect than the lateral border, possibly due to differences in the depth, calibre and richness of the lymphatic vessels at the two sites (Woolgar and Scott, 1995). The reconstructed thickness – which compensates for both nodular and ulcerative growth by measuring to an imaginary reconstructed mucosal surface (Moore et al., 1986; Woolgar and Scott, 1995; Helliwell and Woolgar, 1998, in press) – is recommended as a more accurate and robust predictor than actual tumour thickness (Woolgar and Scott, 1995; Woolgar et al., 1999; Po Wing Yuen et al., 2002). In the study of Woolgar et al. (1999), the tumours were from diverse sites within the mouth and oropharynx, yet the mean thickness in patients dying of/with OSCC was twice that of survivors/patients dying free of OSCC.

Accurate pathological assessment of the thickness measurement (and other measurements such as diameter and excision margins) relies on thorough sampling of the surgical specimen by slicing the complete resection specimen into thin (3–4 mm) slices to ensure that any streaks and satellites, for example, due to vascular or neural invasion, are not overlooked. The micrometer measurement must include all tumour islands, including those well ahead of the main advancing tumour front. Immunohistochemical staining for pan-cytokeratins is useful for highlighting stray islands and individual tumour cells in difficult cases.