Radboud Repository

Radboud University Nijmegen

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/36300

Perspective

Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive

M. Casparie, A.T.M.G. Tiebosch, G. Burger, H. Blauwgeers, A. van de Pol, J.H.J.M. van Krieken and G.A. Meijer *

Abstract. Since 1991, a nationwide histopathology and cytopathology network and archive is in operation in The Netherlands under the name PALGA, encompassing all sixty-four pathology laboratories in The Netherlands. The overall system comprises decentralized systems at the participating laboratories, a central databank, and a dedicated communication and information exchange tool. Excerpts of all histopathology and cytopathology reports are generated automatically at the participating laboratories and transferred to the central databank. Both the decentralized systems and the central system perform checks on the quality and completeness of excerpts. Currently, about 42 million records on almost 10 million patients are stored in the central databank. Each excerpt contains patient identifiers, including demographic data and the so-called PALGA diagnosis. The latter is structured along five classification axes: topography, morphology, function, procedure, and diseases. All data transfer and communication occurs electronically with encryption of patient and laboratory identifiers. All excerpts are continuously available to all participating pathology laboratories, thus contributing to the quality of daily patient care. In addition, external parties may obtain permission to use data from the PALGA system, either on an ongoing basis or on the basis of a specific permission. Annually, 40 to 60 applications for permission to use PALGA data are submitted. Among external users are the Dutch cancer registry, population-based screening programs for cancer of the uterine cervix and breast cancer in The Netherlands, and individual investigators addressing a range of research questions. Many scientific papers and theses incorporating PALGA data have been published already. In conclusion, the PALGA system is a unique system that requires a minimal effort on the part of the participating laboratories, while providing them a powerful tool in their daily practices.

Keywords: Central pathology archive, histopathology, cytopathology, information exchange, PALGA, cancer registry, cancer screening, biobanking

1. Introduction

Pathology laboratories associated with hospitals and clinics generate a host of data that are useful for decision making at the level of the individual patient, but also in a broader context. In order to facilitate the optimal use of histopathology and cytopathology data, a unique, nationwide network and archive has been set up in the Netherlands under the name PALGA ('Pathologisch Anatomisch Landelijk Geautomatiseerd Archief'; Pathological Anatomy National Automated Archive; www.palga.nl). PALGA was founded by a number of pathology laboratories in 1971. The aims of PALGA were, and still are, to promote communication and information exchange between participating histopathology and cytopathology laboratories and to provide potentially useful data to others in the health care field in the interest of patient care and research. This paper provides an outline of the PALGA system and its main usage.

*Corresponding author: G.A. Meijer, VU University Medical Center, Room 3E46, PO Box 7057, 1007 MB Amsterdam, The Netherlands. E-mail: ga.meijer@vumc.nl.
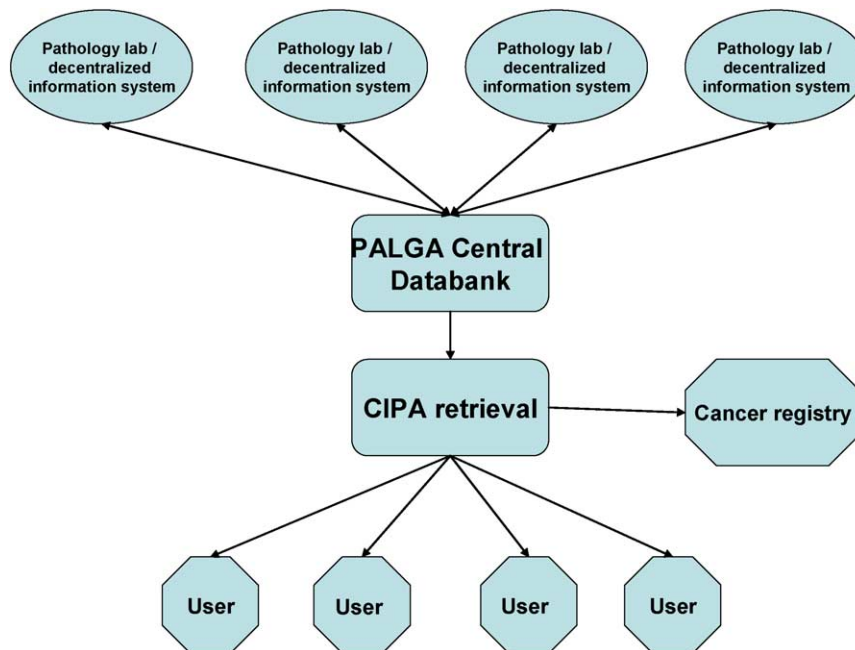
Fig. 1. Organizational scheme of the PALGA system with decentralized input by local laboratories, the central databank and information exchange technology providing authorized users access to the data. The CIPA retrieval box is the interface that allows queries based on diagnosis codes, demographic factors and conclusion texts of pathology reports.

2. Organization and operations of PALGA

Between 1971 and 1991, an increasing number of laboratories joined PALGA such that virtually complete national coverage was achieved [1]. Currently, the globally unique PALGA databank contains about 42 million excerpts on nearly 10 million patients. It is growing at a rate of over two million excerpts per year by the online addition of new excerpts on a daily basis.

The PALGA network and registry comprise three components: decentralized information systems at all 64 participating laboratories, a central databank, and a dedicated communication and information exchange infrastructure (Fig. 1). The complete system is supervised and managed by the PALGA Foundation, supported by a small administrative office, based in Utrecht, and several external service providers. Apart from its executive committee, the organisation includes several special committees dealing with privacy issues, scientific exploitation of the databank, and other issues.

2.1. Decentralized information system

From the inception of PALGA, it has been the policy to tap the relevant information at the site of origin (i.e. the participating laboratories) on an ongoing

basis and avoid the need for additional efforts by the participating laboratories. This aim has been achieved by using decentralized information systems running on the local computer system of the pathologist's laboratory. This decentralized PALGA system automatically collects data from the local computer system and takes care of communication with the central PALGA databank. These decentralized systems also perform checks on data and reports entered by the local pathologist to ensure the quality and completeness of data going into the national databank.

When the local pathologist writes his/her pathology report, he/she adds to the standard description of macro/microscopic characteristics, ancillary techniques and conclusion, a series of standardized codes describing the nature of material analysed, the method of obtaining the material and the diagnosis. Subsequently, the decentralized PALGA system automatically generates a standardized, thesaurus-coded excerpt of the report containing a limited number of patient data (encrypted), a report identifier, (part of) the conclusions, and the so-called PALGA diagnosis, a coded diagnosis line based upon standard pathology terminology (Table 1). Excerpts are transferred to the central databank for retrieval on a daily basis. Patient identifiers include date of birth, gender and the first

Table 1
The standardized PALGA diagnosis

Element	Description
Thesaurus	PALGA thesaurus[a]
Classification axes	T: Topography (localization)
	M: Morphology (nature of tissue change)
	E: Etiology
	F: Function (functional abnormality)
	P: Procedure
	D: Diseases
Diagnosis line (format)	Topography*technique*diagnosis
	(e.g., uterus*extirpation*myoma';
	mamma*biopsy*adenocarcinoma)

[a] Derived from SNOMED (Systematized Nomenclature of Medicine), College of American Pathologists, version 1982.

eight letters of the patient's family name. These identifiers as well as those identifying the submitting laboratory are transferred in an encrypted form, thus securing the privacy of the patient and the laboratory. Encrypting takes place locally at the laboratory when its decentralized system contacts the central databank.

2.2. Central databank

Excerpts transferred to the central databank are checked automatically before they are submitted by the decentralized system as well as upon receipt by the central databank. These checks include the diagnostic terminology used and the structure and completeness of the diagnosis line. In addition, the corresponding classification code is added to each diagnostic term appearing in the excerpt. Each submitted excerpt is registered in the central databank, even if it contains errors. In case of errors, the submitting laboratory will receive an error notification requesting the re-submission of a corrected version of the excerpt. When a corrected version is submitted, the original excerpt in the central databank is overwritten by the corrected version.

Through the procedures described above, which require only a minimal effort on the part of the participating laboratories, PALGA's Central Information System (CIPA) is filled with a continuously increasing amount of information. This information is derived from all histopathological and cytopathological evaluations conducted in The Netherlands, and supplemented by evaluations conducted through the national screening program for cancer of the uterine cervix, and all post mortem investigations.

2.3. Communication and information exchange

A dedicated communication and information exchange infrastructure constitutes the 'wiring' of the PALGA system. Communication and information exchange between authorized parties within this system takes place via telephone lines, not web-based, contributing to the security of data exchange. Parties involved in communication and data exchange are the participating pathology laboratories and certain screening organizations in the Netherlands, such as the organizations conducting the national screening programs for cancer of the uterine cervix and breast cancer.

Other individuals or groups do not get direct access to the PALGA system. However, if an application for a search question, submitted by an external party is approved, the relevant data will be made available to that party through an authorized intermediate having access to the PALGA system. Patients' identity will not be released to external parties, however.

3. Applications of the PALGA system

With millions of excerpt in its central databank, including excerpts of all pathology reports generated in The Netherlands as of 1991, the PALGA system is a powerful tool that could potentially serve various purposes, including daily patient care, disease registries, population-based studies, and epidemiological and scientific research.

3.1. Patient care

One of the core functions of PALGA is to support patient care on a daily basis. When a local pathologist is consulted on a specific patient and enters a new evaluation, the decentralized computer system automatically checks the central PALGA databank for information possibly available on that patient. The pathologist will then see all previous pathology evaluations the patient has undergone, their outcome and the identity of the laboratories that conducted these evaluations to the extent available in the databank. Thus, PALGA may contribute to optimal patient care by providing the patient's complete pathology history to the consulting pathologist. Conversely, the availability of an easily accessible (to participating pathologists) central archive provides a strong motivation for pathologists to continue contributing data and to maintain the quality of data in the PALGA system.

3.2. Cancer registry and cancer screening

The PALGA databank is a very useful source of information for disease registries and population-based health studies in The Netherlands. For instance, PALGA is an important supplier of data to the Dutch cancer registry [40]. The cancer registry receives overviews of all new histologically or cytologically confirmed malignant tumours on a weekly basis. In addition, PALGA plays a central role in the population-based screening program for cancer of the uterine cervix [4,12,42]. Results of all PAP smears are recorded in the PALGA system in a standardized manner and invitations for repeat evaluations are generated automatically. Similarly, the Dutch breast cancer screening program uses the PALGA system to collect follow-up data from regular healthcare sources on women testing positive in the screening program. PALGA will also have an important role in the planned screening program for colorectal cancer in the Netherlands [9].

3.3. Research questions form third parties

Numerous scientific papers and many dissertations have been, and continue to be, published for which the PALGA databank has provided data. In principle, anyone with a relevant question may request permission for the use of data in the PALGA databank. Annually, forty to sixty applications for permission are being received, mainly from investigators in The Netherlands. Questions may be related to epidemiology, but may also address pathological diagnosis and follow-up. Applications must be made in writing and be motivated by the applicant. They are reviewed by PALGA's Scientific Advisory Board for validity of the research question(s) to be addressed and by PALGA's Privacy Protection College. If both committees issue a positive advice, permission may be granted by the Board of PALGA. Depending on the background of the applicant, more or less detailed data will be made available. For certain questions, approval by a medical-ethical review board may be required. A user fee is charged for services to others than the pathologists contributing excepts to PALGA. Examples of recent research questions addressed by means of data from the PALGA databank can be found in the references [2,3,5,7,8,10, 11,13–27,30,32–39,41,43–45].

4. Structured reporting

Traditionally, pathology reports are relatively well structured and contain information on the clinical question, macroscopic characteristics and microscopic evaluation of the specimen and the concluding diagnosis. Yet, pathologists have been relatively free in the way they present relevant information under each of these headings, but this is gradually changing. Multiple initiatives exist to determine the minimally required datasets to be reported for different types of pathology specimens [28,31]. Increasingly, such standards are developed through multidisciplinary initiatives. Obviously, for evidence-based patient care standardized pathology reports are of major importance. Also for research applications, standardized pathology reporting holds great promises. In addition to standards on which items should be reported, standardization also concerns data formats in which these items should be reported, and the databases in which these data should be stored.

The PALGA databank and its management structure provide a great opportunity in this respect. For a specific subset of samples, i.e. cervical cytology specimens, standardized reporting using a defined dataset, and databanking using fields with coded, searchable variables, has been operational for many years already. At present, initiatives are underway to implement this type of standardization for other types of samples, including surgical oncology specimens of the large bowel, breast, lung and other organs. To facilitate the use of these pro formas in combination with PALGA, a dedicated software module has been developed that will appear on top of the client software when the pathologist reports a specimen of this type. This software module will provide standardised entry forms. This will yield a standardised report for routine patient care as well as storage of fully searchable data in the central databank. On a nationwide basis this will allow specific queries like "all breast cancers smaller than one centimetre with more than three positive lymph nodes in a given year". The current use of excerpts and diagnosis codes is much more prone to errors in data entry and coding and does not allow such detailed queries.

5. Biobanking

Pathology archives contain millions of tissue samples, either frozen or formaldehyde-fixed and embed-

ded in paraffin blocks. A subset of these samples has specifically been collected for research purposes with informed consent of the patients involved. For the larger part, this is material left over from completed diagnostic procedures. Such material may later on prove clinically relevant for the patient and/or his/her relatives. In addition, this material provides a tremendously powerful and important resource for scientific research as it allows correlations to be made between phenotypic features and biological characteristics of disease. The latter can be obtained by analysing tissue samples with state-of-the-art molecular technologies. In the Netherlands, the adequate secondary use of left-over tissue samples for scientific research is regulated by of a self-regulatory code of conduct, developed by the Federation of Dutch Medical Scientific Societies [6].

The PALGA databank has a dual role in this type of research. Firstly, PALGA serves as the index of the tissue archive, either within the local institution or nationwide. Secondly, the pathology reports in the databank provide a major source of information on disease phenotype in all cases. Obviously, the implementation of structured reports in the PALGA databank will benefit both functions. The great potential of the PALGA system in this respect is underscored by a recent report of the Netherlands Academy of Sciences. Recommendations are made to involve PALGA in projects for improving the research infrastructure for translational biomedical research in The Netherlands [29].

6. Concluding remarks

The PALGA system is a unique, highly automated pathology archiving and communication tool that provides considerable support to pathologists in their daily practices. The technology and procedures of PALGA were designed from the beginning to minimize the amount of effort required by the participating laboratories. This approach has resulted in the joining of the PALGA network by all 64 histopathology and cytopathology laboratories in The Netherlands and a continuously expanding automated archive of excerpts of pathology reports with currently about 42 million excepts on nearly 10 million patients in the central databank.

The PALGA system not only supports pathologists in daily practice by giving them easy access to current and historical data on their patients, it also supports the Dutch cancer registration, national screening organi-

zations, and investigators wishing to address research questions in various domains, including epidemiology, pathology, and molecular biology of disease.

Acknowledgements

References

[1] F. Bijlsma and E.P. van der Esch, [PALGA, of labor pains of the Pathological Anatomical District Automatized Archives], *Ned. Tijdschr. Geneeskd.* **116** (1972), 902–903.

[2] A.A. Botterweck, P.A. van den Brandt and R.A. Goldbohm, A prospective cohort study on vegetable and fruit consumption and stomach cancer risk in The Netherlands, *Am. J. Epidemiol.* **148** (1998), 842–853.

[3] S. Bulk, F.J. van Kemenade, L. Rozendaal et al., The Dutch CISOE-A framework for cytology reporting increases efficacy of screening upon standardisation since 1996, *J. Clin. Pathol.* **57** (2004), 388–393.

[4] S. Bulk, F.J. van Kemenade, L. Rozendaal and C.J. Meijer, The Dutch CISOE-A framework for cytology reporting increases efficacy of screening upon standardisation since 1996, *J. Clin. Pathol.* **57** (2004), 388–393.

[5] A.M. Cleton-Jansen, M.C. Timmerman, M.J. van de Vijver et al., A distinct phenotype characterizes tumors from a putative genetic trait involving chondrosarcoma and breast cancer occurring in the same patient, *Lab. Invest.* **84** (2004), 191–202.

[6] Code for adequate secondary use of data (2004), Code for adequate secondary use of tissue (2001), http://www.federa.org/ [accessed 31 December 2006].

[7] C.G. Csizmadia, M.L. Mearin, B.M. von Blomberg et al., An iceberg of childhood coeliac disease in the Netherlands, *Lancet* **353** (1999), 813–814.

[8] E. de Bree, F.A. Zoetmulder, R.B. Keus et al., Incidence and treatment of recurrent plantar fibromatosis by surgery and postoperative radiotherapy, *Am. J. Surg.* **187** (2004), 33–38.

[9] M. de Visser, M. van Ballegooijen, S.M. Bloemers et al., Report on the Dutch consensus development meeting for implementation and further development of population screening for colorectal cancer based on FOBT, *Cell. Oncol.* **27** (2005), 17–29.

[10] W.L. Diemont, O.J. ten Thije and C.H. ten Napel, Unsuspected herpes esophagitis: properties of eight non-HIV-related cases, *Neth. J. Med.* **37** (1990), A54–A55.

[11] M.J. Dirx, P.A. van den Brandt, R.A. Goldbohm et al., Energy restriction in childhood and adolescence and risk of prostate cancer: results from the Netherlands Cohort Study, *Am. J. Epidemiol.* **154** (2001), 530–537.

[12] M. Geertsen, A.G. Bais, H. Beerman and T.J. Helmerhorst, Follow up after an abnormal pap smear: time interval acceptable, nature of follow up leaves room for improvement, *Ned. Tijdschr. Geneeskd.* **147** (2003) 2430–2434 [Article in Dutch].

[13] E.K. George, M.L. Mearin, H.C. Franken et al., Twenty years of childhood coeliac disease in The Netherlands: a rapidly increasing incidence?, *Gut* **40** (1997), 61–66.

[14] R.W. Giard and J. Hermans, Use and accuracy of fine-needle aspiration cytology in histologically proven thyroid carcinoma: an audit using a national nathology database, *Cancer* **90** (2000), 330–334.

[15] W.G. Goettsch, S.D. Bos, N. Breekveldt-Postma et al., Incidence of gastrointestinal stromal tumours is underestimated: results of a nation-wide study, *Eur. J. Cancer* **41** (2005), 2868–2872.

[16] A. Hanselaar, M. van Loosbroek, O. Schuurbiers et al., Clear cell adenocarcinoma of the vagina and cervix. An update of the central Netherlands registry showing twin age incidence peaks, *Cancer* **79** (1997), 2229–2236.

[17] E.I. Hauben, J. Arends, J.P. Vandenbroucke et al., Multiple primary malignancies in osteosarcoma patients. Incidence and predictive value of osteosarcoma subtype for cancer syndromes related with osteosarcoma, *Eur. J. Hum. Genet.* **11** (2003), 611–618.

[18] E.I. Hauben, G. Jundt, A.M. Cleton-Jansen et al., Desmoplastic fibroma of bone: an immunohistochemical study including β-catenin expression and mutational analysis for β-catenin, *Human Pathology* **36** (2005), 1025–1030.

[19] M.P.W.A. Houben, J.W.W. Coebergh, R.M.C. Herings et al., The association between antihypertensive drugs and glioma, *Br. J. Cancer* (2006), 752–756.

[20] F. Loeve, M. van Ballegooijen, R. Boer et al., Colorectal cancer risk in adenoma patients: a nation-wide study, *Int. J. Cancer* **111** (2004), 147–151.

[21] E.M. Mathus-Vliegen, H. van Halteren and G.N. Tytgat, Malignant lymphoma in coeliac disease: various manifestations with distinct symptomatology and prognosis?, *J. Intern. Med.* **236** (1994), 43–49.

[22] A.E. Odink, C.J. van Asperen, J.P. Vandenbroucke et al., An association between cartilaginous tumours and breast cancer in the national pathology registration in The Netherlands points towards a possible genetic trait, *J. Pathol.* **193** (2001), 190–192.

[23] S.J. Otto, I. van der Cruijsen, M.K. Liem et al., Effective PSA contamination in the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer, *Int. J. Cancer* **105** (2003), 394–399.

[24] P.N. Post, E.J. Kuipers and C.J. Meijer, Declining incidence of peptic ulcer but not of its complications: a nation-wide study in The Netherlands, *Aliment. Pharmacol. Ther.* **1** (2006), 1587–1593.

[25] F.H. Rampen, J.I. Casparie-van Velsen, B.E. van Huystee et al., False-negative findings in skin cancer and melanoma screening, *J. Am. Acad. Dermatol.* **33** (1995), 59–63.

[26] J.M. Rhodes, Cholesterol crystal embolism: an important "new" diagnosis for the general physician, *Lancet* **347** (1996), 1641.

[27] R.M. Roumen, Anorectal melanoma in The Netherlands: a report of 63 patients, *Eur. J. Surg. Oncol.* **22** (1996), 598–601.

[28] Royal College of Pathologists, General introduction, Datasets for histopathology reporting on cancers, http://www.rcpath.org/index.asp [accessed 31 December 2006].

[29] Royal Netherlands Academy of Arts and Sciences, Foresight Committee on Multifactorial Diseases in the Genom cs Era (2006), http://www.knaw.nl/publicaties/pdf/20041002.pdf.

[30] L.J. Schouten, R.A. Goldbohm and P.A. van den Brandt, Anthropometry, physical activity, and endometrial cancer risk: results from the Netherlands Cohort Study, *J. Natl. Cancer Inst.* **96** (2004), 1635–1638.

[31] Templates for pathology reporting of common malignancies, http://www.pathology.nl/ [accessed 31 December 2006].

[32] T. van Dalen, H.J. Hoekstra, A.N. van Geel et al., Locoregional recurrence of retroperitoneal soft tissue sarcoma: second chance of cure for selected patients, *Eur. J. Surg. Oncol.* **27** (2001), 564–568.

[33] M.E. van den Akker-van Marle, M. van Ballegooijen, G.J. van ortmarssen et al., Cost-effectiveness of cervical cancer screening: comparison of screening policies, *J. Natl. Cancer Inst.* **94** (2002), 193–204.

[34] M.J. van der, J.L. Blaauwgeers, T.G. Sutedja et al., Granular cell tumors of the tracheobronchial tree, *J. Thorac. Cardiovasc. Surg.* **126** (2003), 740–743.

[35] B.A. van Dijk, L.J. Schouten, C.A. Hulsbergen-van de Kaa et al., Risk factors, genetic predisposition and mutations in the VHL gene in sporadic renal cell carcinoma: The Netherlands Cohort Study, *IARC Sci. Publ.* **156** (2002), 509–510.

[36] M.C. van Dijk, P.D. Rombout, W.J. Mooi et al., Allelic imbalance in the diagnosis of benign, atypical and malignant Spitz tumours, *J. Pathol.* **197** (2002), 170–178.

[37] J.J. van Lanschot, J.B. Hulscher, C.J. Buskens et al., Hospital volume and hospital mortality for esophagectomy, *Cancer* **91** (2001), 1574–1578.

[38] F.E. van Leeuwen, J. Benraadt, J.W. Coebergh et al., Risk of endometrial cancer after tamoxifen treatment of breast cancer, *Lancet* **343** (1994), 448–452.

[39] K.C. Veenhuizen, P.E. de Wit, W.J. Mooi et al., Quality assessment by expert opinion in melanoma pathology: experience of the pathology panel of the Dutch Melanoma Working Party, *J. Pathol.* **182** (1997), 266–272.

[40] Vereniging van Integrale Kankercentra (VIKC) / Association of Comprehensive Cancer Centres, Incidence of cancer in the Netherlands 1999/2000, Eleventh report of the Netherlands Cancer Registry, 2003.

[41] G.P. Vooijs, I.A. Casparie-van Velsen, F.A. Peters et al., National registry of cervical cytologic diagnoses in The Netherlands, *Acta Cytol.* **33** (1989), 825–830.

[42] G.P. Vooijs, I.A. Casparie-van Velsen, F.A. Peters and H.L. Beck, National registry of cervical cytologic diagnoses in The Netherlands, *Acta Cytol.* **33** (1989), 825–830.

[43] W. Willemsen, R. Kruitwagen, B. Bastiaans et al., Ovarian stimulation and granulosa-cell tumour, *Lancet* **341** (1993), 986–988.

[44] P.C. Witt Hamer, M.J. Verstegen, R.J. de Haan et al., High risk of acute deterioration in patients harboring symptomatic colloid cysts of the third ventricle, *J. Neurosurg.* **96** (2002), 1041–1045.

[45] L.H. Zaal, W.J. Mooi, H. Klip et al., Risk of malignant transformation of congenital melanocytic nevi: a retrospective nationwide study from The Netherlands, *Plast. Reconstr. Surg.* **116** (2005), 1902–1909.