Predictive testing of early CIN behaviour by molecular biomarkers

Jan P.A. Baak a,b,c,∗, Arnold-Jan Kruse a, Emiel Janssen a and Bianca van Diermen a

a Department of Pathology, Stavanger University Hospital, Stavanger, Norway
b The Gade Institute, University of Bergen, Norway
c Free University, Amsterdam, The Netherlands

1. Introduction

Cervical Intraepithelial Neoplasias (CIN) is a very frequent disease caused by Human Papilloma Virus (HPV). In early CIN, grade (as 1 or 2) indicates the progression-risk to CIN-3 and determines treatment but the Positive Predictive Value (PPV) is low (10% and 20% for CIN-1 and -2 respectively, 16% on average). This results in an enormous number of overtreatments. Moreover, grade reproducibility is of particular concern. Certain molecular biomarkers such as Ki-67 have a higher PPV (30% – an improvement of 14%), which in Europe alone could improve treatment for many thousands of women per year with considerable work-load and cost reduction for the health care system. The quantitative Ki-67 prognostic model has been validated in independent retrospective and prospective studies from different laboratories.

Elsewhere in this issue of Cellular Oncology, several important questions pertaining to the validity, usefulness and daily use of quantitative molecular characteristics in predicting the behaviour of early CIN lesions [6–8] have been raised [12] which we answer here.

2. The progression risks of early CINs is so low that prognostication is clinically not important

This view reflects the widespread opinion amongst pathologists, however a closer analysis shows that this is not correct. Accepted Positive Predictive Values (PPVs) of prediction of progression to CIN-3 and invasive cancer by CIN grade [10] are 10% for CIN-1 and 20% for CIN-2. With the Ki-67 model alone, the PPV is 30% (it should be kept in mind that a 1% improvement of the PPV already regards 1200 women per year!). Table 1 shows the significance of this improved PPV for the European Union (based on estimations of the incidence figures in the catchment areas of our laboratories in The Netherlands and Norway). For CIN-2 alone, each year over 8000 patients would receive better treatment by Ki-67 quantitative image analysis than by subjective grade.

We do agree that a PPV higher than 30% is desirable. A multivariate analysis of many different molecular markers showed that combining Ki-67 with retinoblastoma protein quantisation in the deep layer of the epithelium increases the PPV to 48% [9]. General use of these combined biomarkers, to CIN-2 alone, would mean improved treatment for over 23000 patients per year in the EU. Additional analysis of cytokeratins-14 and -13 can achieve even higher PPVs [9] (see Fig. 1).

3. Reproducibility of the quantitative pathology assessments

Emphasis of reproducibility is important as few Pathology Laboratories do have formal protocols for daily pathology determinations. Unfortunately, however many pathology laboratory tests are introduced without any formal Quality Control and Assurance protocols. The tests are described in a rather intuitive
Table 1
The estimated annual number of patients with CIN, and overtreatments in the European Union (EU)

|                              | EU                     |
|------------------------------|------------------------|
| Nr of inhabitants           | 375 million            |
| Nr of CIN-1, 2 and 3 in EU   | 330,000                |
| Nr of CIN-2 in EU            | 82,500                 |
| Annually over-treated CIN-2s in EU by |                |
| CIN grade                    | 66,000                 |
| Quantitative Ki67           | 57,750                 |
| Quantitative Ki67&Rb        | 47,025                 |
| Number of improved treatments due to |            |
| Quantitative Ki67/year       | 8,250                  |
| Quantitative Ki67 + Rb/year  | 23,100                 |

and vague manner without any logical formal standard operating procedure, with vague hints only as to the reproducibility ("the methods were well reproducible with double blind assessments") and without clear decision thresholds ("staining results were interpreted by independent observers as negative or positive", "the tumours were classified as low and high proliferating"). This gives a gloomy picture of pathology, which is even worse when considering that many laboratory methods in surgical- and cyto-pathology are only very rarely validated in formal independent phase 2–4 studies as required for GLP [2,4]. As a result, many pathology tests are poorly reproducible both by the same and different observers. It is difficult to understand why (inter)national societies and health care officials at the governmental level take these essential points so sloppily. CIN grade is a typical example of such a laboratory test [5,11]. "Good reproducibility" of a quantitative assessment, in the sense of a high correlation coefficient between two independent assessments, is important but not enough. Validation should include the whole procedure of tissue processing, measurement and interpretation. We have therefore analyzed the predictive value in independent retrospective and prospective studies in different laboratories, also when used routinely by many different pathologists, without ever changing the decision thresholds of Ki-67 [6,8]. These studies confirmed the prognostic value of the Ki-67 prognostic model in early CIN lesions. In this manner, the strictest possible GLP criteria have been fulfilled.

4. Spread of the methods

The evidence in favour of the use of quantitative methods in pathology in neoplasias from many different organ sites is massive and well validated. Typical examples are proliferation factors in breast cancer [1,3,13]. Why then, is the application of quantitative image analysis such as morphometry and DNA ploidy, amongst other techniques, not more wide spread? There are no scientific explanations anymore but a variety of arguments and excuses can be heard:

1. Administrations do not give the necessary resources.
2. Pathologists just do not believe that quantitative image analysis helps. This is difficult to believe, in view of the massive literature available.

3. Other methods are more popular and get higher priority. A typical example is Her2-neu. However, accurate quantitative assessment of the percentage of Her2-neu positive cells is important to get the best information regarding therapy effect. Molecular and quantitative methods must therefore be combined to get the best results.

4. Pathologists are not always aware of the value of quantitative image analysis, as they may not be up to date with the literature. (Interestingly, Dr. George Mutter, Department of Pathology, Brigham and Women’s Hospital, Harvard Medical School, Boston, has a richly illustrated educational public homepage www.endometrium.org, on endometrial hyperplasia, which gets hundreds to thousands hits per day. Are homepages better communication media than classical journals?)

5. Quantitative image analysis is too expensive. This is certainly not true; the cost-benefit analysis is excellent and well documented.

6. The techniques are too time-consuming. Not true, 3–15 minutes per case.

7. Quantitative image analysis techniques do not fit in the culture of a pathology lab; the same holds for PCR and other molecular methods, which are not easily introduced either. This is a psychologically understandable but wrong argument which is clear from the abundance of literature.

8. The equipment is not user friendly enough. As long as pathologists buy innovative equipment as infrequently as they do, the industry will be reluctant to further improve their machines.

Consequently, we did not quite understand the causes preventing the spread of quantitative pathology methods. To get better insight, one of us (J.B.) in 1995, did the following study. After more than 25 years at different academic departments mostly spent in developing and implementing quantitative pathological methods, the spin-off outside the ivory tower of the University environment was disappointing. Then it was decided to do a formal implementation study and investigate on the non-academic work floor itself where the obstacles were. An 80% position was accepted in a large classical pathology practice (20,000 surgical specimens, 210 autopsies, 23,000 cervical and 1500 clinical cytological specimens, 4 pathologists, 18 other positions) in a teaching hospital in Alkmaar, The Netherlands.

A detailed treatise of all the factors discovered in that implementation study that prevented innovation, has been published before [2]. There were many random factors, but in the end it all came down to the following essential factor:

Implementation of modern techniques requires constant, active, motivation by pathologists. Lack of primary innovative leadership of the pathologists
is the dominant factor that prevents pathology modernisation.

As long as there is no penalty for poor performance, and high turn-over of simple inexpensive poorly reproducible determinations is rewarded by insurance companies or the national health care system, there is little hope that the situation will change.

The cost-benefit analysis of implementing and routinely using modern pathology methods is very favourable for the health care system as a whole. The responsibility thus now lies with the surgical and cytopathologists!

References

[1] D. Amadori, O. Nanni, M. Marangolo, P. Pacini, A. Ravaiolii, A. Rossi, A. Gambi, G. Catalano, D. Perroni, E. Scarpi, D.C. Giunchi, A. Tienghi, A. Becciolini and A. Volpi, Disease-free survival advantage of adjuvant cyclophosphamide, methotrexate, and fluorouracil in patients with node-negative, rapidly proliferating breast cancer: a randomized multicenter study, J. Clin. Oncol. 18 (2000), 3125–3134.

[2] J.P. Baak. The framework of pathology: good laboratory practice by quantitative and molecular methods, J. Pathol. 198(3) (2002), 277–283.

[3] J.P. Baak, P.J. van Diest, F.J. Voorhorst, E. van der Wall, L.V. Beex, J.B. Vermorken and E.A. Janssen, Prospective multicenter validation of the independent prognostic value of the mitotic activity index in lymph node-negative breast cancer patients younger than 55 years, J. Clin. Oncol. 23(25) (2005), 5993–6001.

[4] P.A. Hall and J.J. Going, Predicting the future: a critical appraisal of cancer prognosis studies, Histopathology 35(6) (1999), 489–494.

[5] A.J. Kruse, J.P. Baak, T. Helliesen, K.H. Kjellevold, M.G. Bol and E.A. Janssen, Evaluation of MIB-1-positive cell clusters as a diagnostic marker for cervical intraepithelial neoplasia, Am. J. Surg. Pathol. 26(11) (2002), 1501–1507.

[6] A.J. Kruse, J.P. Baak, E.A. Janssen, M.G. Bol, K.H. Kjellevold, B. Fianne, K. Lovslett and J. Bergh, Low- and high-risk CIN 1 and 2 lesions: prospective predictive value of grade, HPV, and Ki-67 immuno-quantitative variables, J. Pathol. 199(4) (2003), 462–470.

[7] A.J. Kruse, J.P. Baak, E.A. Janssen, K.H. Kjellevold, B. Fianne, K. Lovslett, J. Bergh and S. Robboy, ki67 predicts progression in early CIN: validation of a multivariate progression-risk model, Cell. Oncol. 26(1–2) (2004), 13–20.

[8] A.J. Kruse, E. Guðlaugsson, T. Helliesen, E.A. Janssen, B. van Diermen, S. Sandvik and J.P. Baak, Evaluation of prospective, routine application of Ki-67 immunoquantitation in early CIN for assessment of short-term progression risk, Anal. Quant. Cytol. Histol. 26(3) (2004), 134–140.

[9] A.J. Kruse, I. Skaland, E.A. Janssen, S. Buhr-Wildhagen, J. Klos, M.J. Arends and J.P. Baak, Quantitative molecular parameters to identify low-risk and high-risk early CIN lesions: role of markers of proliferative activity and differentiation and Rb availability, Int. J. Gynecol. Pathol. 23(2) (2004), 100–109.

[10] A.G. Ostor, Natural history of cervical intraepithelial neoplasia: a critical review, Int. J. Gynecol. Pathol. 12(2) (1993), 186–192.

[11] E.C. Pirog, R.N. Baergen, R.A. Soslow, D. Tam, A.E. DeMattia, Y.T. Chen and C. Isacson, Diagnostic accuracy of cervical low-grade squamous intraepithelial lesions is improved with MIB-1 immunostaining, Am. J. Surg. Pathol. 26(1) (2002), 70–75.

[12] R. Silvestrini, Letter to the editor: Ki67 predicts progression in early CIN: Validation of a multivariate progression-risk model, Cell. Oncol. (2005), 355.

[13] R. Silvestrini, M.G. Daidone, A. Luisi, M. Mastore, M. Leitner and B. Salvadori, Cell proliferation in 3,800 node-negative breast cancers: consistency over time of biological and clinical information provided by 3H-thymidine labelling index, Int. J. Cancer 74 (1997), 122–127.