

Nina Kohli-Laven, Pascale Bourret, Peter Keating and Alberto Cambrosio, "Cancer Clinical Trials in the Era of Genomic Signatures: Biomedical Innovation, Clinical Utility, and Regulatory-Scientific Hybrids." *Social Studies of Science*, 41 (2011), 487-513.

**Cancer clinical trials in the era of genomic signatures: Biomedical innovation, clinical utility, and regulatory-scientific hybrids**

**Nina Kohli-Laven**

Department of Social Studies of Medicine, McGill University

**Pascale Bourret**

Faculté des Sciences Économiques, Aix-Marseille Université  
UMR 912- SE4S, INSERM-IRD-Université de la Méditerranée

**Peter Keating**

Department of History, University of Quebec (Montreal)

**Alberto Cambrosio**

Department of Social Studies of Medicine, McGill University

**Abstract**

The paper examines two large-scale, North-American and European clinical trials designed to validate two commercially available genomic tumor *signatures* that predict a patient's risk of breast cancer recurrence and response to chemotherapy. The paper builds on empirical evidence from the two trials to explore the emergence of diverse regulatory-scientific hybrids, i.e. configurations of genomic practice and bioclinical work that depend on linkages between technical, commercial, patient, clinical, and legal interests and institutions. The development of the genomic signatures for each trial — *Oncotype DX* and *MammaPrint* — has followed quite different routes. *Oncotype* began as a commercial platform: the company that produced it did not discover a signature but rather constructed it by asking users at every step what clinical question they wanted the signature to answer and what data would be credible in that regard. The test has been designed to minimally disrupt existing clinical workflows. *MammaPrint*, on the other hand, began as a breast cancer signature: the researchers who discovered it, at the Netherlands Cancer Institute (NKI), established a company to commercialize it as a test after the fact. *MammaPrint* requires a change in pathologists' routines. Thus, while these two trials signify a new departure for clinical cancer trials on a number of levels — they both incorporate new models of interaction between biotech companies and public research, and they both aim to establish the clinical relevance of genomic markers — they also embody different socio-technical scripts: one attempts to accommodate established routines, while the other openly challenges prevailing evidential hierarchies and existing biomedical configurations.

**Keywords:** Clinical trials; Breast cancer; Genomic technologies; Clinical utility; Biomedical innovation; Socio-technical scripts; Diagnosis

**Contact Author:** Nina Kohli-Laven, Department of Social Studies of Medicine, McGill University, 3647 Peel, Room 207, Montreal, Quebec, H3A 1X1. Email: [nina.laven@mcgill.ca](mailto:nina.laven@mcgill.ca)

Social scientists have often interpreted the introduction of genomics into medicine as the first act of a doomsday scenario wherein technology evolves into ever more ethically suspect forms, reducing people and politics to biology. This paper uses evidence from two major clinical trials in the United States and Europe to suggest that genomics in medicine engenders something else: regulatory-scientific hybrids. These hybrids are configurations of persons and tools that marry genomic techniques to market, legal and regulatory strategies in ways that reframe conceptions of risk, diagnosis, prognosis, therapy, discovery, utility and validity. Rather than dystopic futures populated with molecular citizens lashed to the juggernaut of unbridled technology, we see a redefinition of biology, genomics, and medicine in relation to one another and to treatments and patients. The realignment of socio-technical networks crafts new technologies through the artful management of regulatory norms and organizational partnerships. Ethical questions may still be relevant here, but these practices beg a further, and in some regards more interesting set of questions: what specific forms envelop the reconfigured regulatory-technological science of genomic medicine? What consequences do these forms have for contemporary biomedicine? The two clinical trials we discuss exemplify

novel forms of clinical research designed to extend genomics into the clinic, with the stated goal of transforming major components of clinical practice.

The clinical trials that are the focus of this paper use genomic tools (tumor *signatures*) for the selection of treatment for breast cancer. The first trial, TAILORx — for **T**rial **A**ssigning **I**ndividualized **O**ptions for Treatment (**R**x) — conducted in the US and Canada, initially planned to enroll 10,000 patients (of which 4,500 were to be in the randomized arm) in 900 participating centers, a number subsequently increased due to non-compliance issues (interview with Dr. Joseph Sparano, Bronx, NY, 12 February and 21 May 2010, henceforth Sparano interview). The trial uses Oncotype DX, a diagnostic test based on a 21-gene breast cancer signature developed by the US company Genomic Health. The second trial, MINDACT — for **M**icroarray **I**n **N**ode-negative and 1 to 3 positive lymph node **D**isease may **A**void **C**hemo**T**herapy<sup>i</sup> — pursued in 11 European countries, projected the enrollment of 6,000 patients. Touted as the ‘first trial truly based on biology’ (Fieldnotes, 8ème Biennale de cancérologie, Monaco, January 2008, henceforth Monaco fieldnotes; see also Meunier, 2008), MINDACT uses MammaPrint, a 70-gene breast cancer signature commercialized by Agendia, a Dutch spin-off established by researchers from the Netherlands Cancer Institute. TAILORx compares outcomes between groups of patients who have been tested using Oncotype and who, depending upon the outcome of the test, may or may not receive chemotherapy, whereas MINDACT compares outcomes between groups of patients whose prognosis has been established using MammaPrint *versus* those who have been assessed using traditional clinico-pathological techniques.

The rationales for the trials and for the commercial development of their associated tests are largely the same. In the broadest sense, both trials exemplify ongoing efforts within biomedicine to ‘translate’ genomic biomarkers from bench to bedside. The term *translational research* first emerged in the early 1990s at the US National Cancer Institute (NCI) and came into common use after the turn of the century, spreading from cancer to other biomedical domains (Butler, 2008; see also Brown, 2007). In 2003 the NIH released the *NIH Roadmap* (Zerhouni, 2003, 2005), an ambitious project to re-engineer the clinical research enterprise and to provide even more ample space for translational research. The NCI had already developed several specialized translational research programs such as the 2001 Translational Research Initiative and the 2000 Program for the Assessment of Clinical Cancer Tests (PACCT) (Jessup et al., 2005) that launched the TAILORx trial. The European Organization for Research and Treatment of Cancer (EORTC), the principal pan-European cancer trialists’ organization, underwent a reform in the late 1990s that led to a similar reframing of its activities around the notion of translational research. European clinical researchers have gone so far as to speak of a ‘paradigm shift in the culture of cancer trials’ so that ‘it’s no longer considered ethically acceptable to support studies that do not include a translational research component’ (Fricker, 2007). For the researchers and clinicians discussed in this paper, translational research stands as the only strategy likely to avoid ‘undertreatment, overtreatment or wrong treatment’ (Wilkinson, 2009). The emergence of genomic biomarkers has prompted researchers to portray current ‘one size fits all’ clinical treatment strategies as deficient and, at best, stopgap solutions.

## **The TAILORx and MINDACT trials**

The organization of clinical trials with a translational design is no easy task. Small, underpowered studies — that is, studies with insufficient sample sizes — that provide conflicting results will obviously not do. Moreover, European clinical trialists face a particularly daunting task because legal, ethical and financial problems have to be solved on a country-by-country basis. National laws, for instance, govern the circulation of tissue samples, allowing, in some cases, for the shipment of slides but not tumor blocks, or of RNA extracted from samples but not whole tumors. When participating institutions located in different countries must send frozen samples for genomic analysis to a central location, the numerous problems entailed require creative solutions. In the case of MINDACT, in order to comply with EU funding regulations that finance new organizational structures but not clinical trials themselves, clinicians had to set up a new ad hoc organization (Fricker, 2007: 9), and the resultant consortium's legal negotiations on issues such as intellectual property rights took more than two years (EJC, 2008).

In addition to reflecting similar shifts in the research environment in Europe and the US, the specific aims of the TAILORx and MINDACT trials also overlap insofar as both trials address the same problem in the clinical treatment of breast cancer. The standard of care for patients with early stage breast cancer in both Europe and North America is surgery followed (depending on the extent of cancer spread) by radiation therapy, chemotherapy and/or hormonal therapy. Treatment guidelines from the US National

Comprehensive Cancer Network and the American Society for Clinical Oncology have, until recently, recommended chemotherapy for most early stage patients (Harris et al., 2007). There is a broad consensus, however, based on several decades of follow-up research with chemotherapy breast cancer patients, that not all patients benefit from post-surgery chemotherapy. Consequently, according to this view, many patients receive the physically taxing and financially costly therapy needlessly. Researchers and regulators claim that TAILORx and MINDACT, as well as their respective tests, *Oncotype* and MammaPrint, will mitigate chemotherapy ‘overtreatment’ by providing clinicians with more information about which patients potentially will benefit.

Clinicians, in turn, see *Oncotype* and MammaPrint as tests that spare women aggressive treatment (interview with Dr. Daniel Hayes, Ann Arbor, MI, 7-8 January 2010, henceforth Hayes interview). The chief commercial scientist who developed *Oncotype* has expressly described the test as ‘saving women from unnecessary and painful procedures’ (interview with Dr. Maureen Cronin, Redwood City, CA, 6 November 2009, henceforth Cronin interview). Clinicians and researchers have framed these claims as part of a longer tradition of reducing the extent of therapeutic intervention in this disease. At an early meeting during *Oncotype*’s development, the chair of the U.S. National Surgical Adjuvant Breast and Bowel Project (NSABP), an NCI-supported clinical trials cooperative group, showed gruesome surgical photos of mastectomies as a prelude to discussion (Cronin interview). The images harkened back to landmark 1970s studies that had shown radical mastectomy, a physically and psychologically mutilating procedure, to be no more effective than less invasive forms of surgery. Scientists and marketers of the

Oncotype test at Genomic Health have described the test to insurers and regulators as a means to eliminate unnecessary treatment, and the need to pay for it (interview with Cheya Pope, Redwood City, CA, 18 November 2009). Indeed, Genomic Health has sponsored economic evaluations of its test to demonstrate cost-effectiveness in spite of its elevated price relative to other diagnostics (Hornberger et al., 2005; Lyman et al., 2007), and Agendia has done the same for MammaPrint (Retèl et al., 2010).

In the TAILORx trial, all patients recruited provide a tumor tissue sample for testing with Oncotype. The test measures the level of expression of 21 genes in the sample and, through application of a proprietary algorithm, provides a result in the form of a score between 0 and 100, called a *Recurrence Score* (RS). That score expresses the likelihood of breast cancer recurrence within 10 years of initial diagnosis. For the clinical trial, patients are separated into three categories based on the score: a low RS (10 and below), a middle RS (11-25), and a high RS (26 and above). The trial cut-points differ from those used by the company in its commercial package (less than 18, 18-31, 31 and above): this change in cut-points, as we will see below, was subject to much discussion during trial preparation due to concerns about undertreatment. The middle group is the primary focus of the TAILORx trial for, unlike the clearly high-risk and low-risk patients, these cases are difficult to characterize, standardize, and treat effectively. In the trial, patients in the middle group are randomized to receive chemotherapy and hormonal therapy together or hormonal therapy alone. By determining ‘who with intermediate risk will respond to chemotherapy’ (Ross et al., 2008), investigators hope to ascertain if the test will be helpful in treatment planning for this group. More generally, the trial uses Oncotype to

test whether and how genomics can prescribe optimum care. In that sense, it is not only a trial of *Oncotype*, but of the applicability of current standards of care in genomically re-defined patient groups.

In the MINDACT trial, women recruited into the trial are assigned to high- and low-risk categories using *both* standard clinical-pathological features *and* the results of the 70-gene MammaPrint test. Clinical-pathological features include patient information such as age, co-morbidities, tumor size, number of positive nodes, etc. An open-access computer program, Adjuvant!Online — developed in the US and widely used by breast cancer clinicians — uses this information to estimate the outcome in terms of relapse and survival with or without chemotherapy. By confronting the predictions of MammaPrint and Adjuvant!Online, the trial directly compares these two prognostic tools: women whose Adjuvant!Online and MammaPrint results are discordant (when clinical/pathological features indicate high risk of recurrence when MammaPrint indicates low risk, or vice-versa) are then randomized for chemotherapy. Commenting on the MINDACT design, co-PI Martine Piccart noted that the primary goal of the trial was ‘to prove that if we use this new tool we are going to treat [10%-15%] fewer women with chemotherapy’ (Tuma, 2005: 333).<sup>ii</sup> The stated aims of the MINDACT trial thus differ in an important way from those of TAILORx: the latter makes no formal comparison between the genomic test and clinical-pathological prognostic tools, although such a comparison was briefly contemplated at an early design stage (see the next section). Indeed, TAILORx PIs take great pains to emphasize that the *Oncotype* test does not threaten traditional indicators but instead adds to them: this is sometimes explained as



necessary to procuring ‘buy-in’ from pathologists (Goldstein and Sparano, 2009; Sparano, interview; Cronin interview; interview with Dr. Tracey Lively, Bethesda, MD, 16 December 2009, henceforth Lively interview). The more ‘confrontational attitude’ of the MINDACT leaders toward traditional clinico-pathological tools has resulted in a very different trial design and led pathologists at one meeting in Monaco to ask with tongue-in-cheek whether they should begin lining up at the employment office (Monaco fieldnotes).

By examining TAILORx and MINDACT *together*, we can identify elements of a new departure for clinical cancer trials, and, more broadly, genomic medicine. We will see that the trial designs emerged from protracted deliberations between numerous private, public and joint bodies over legal concerns, concepts of care, appropriate statistical reasoning, and calculations about clinical behavior. From a *comparative perspective*, we can use the trials to generate insights into two related but distinct configurations of bio-clinical work. Rather than lumping these two trials together and glossing over them as simply further examples of the geneticization or molecularization of patients, we shall examine how each trial was designed and why their questions were so different.

### **From Genomic Health’s 21-gene signatures to TAILORx**

To fully understand TAILORx, we must begin with the two coterminous events that precipitated its establishment: the emergence of the biotech company Genomic Health

that developed the *Oncotype* test, and the establishment of the NCI's PACCT group and its decision to use *Oncotype* in its first clinical trial.

Three veterans of the biotechnology firm Genentech founded Genomic Health in 2001 with the aim of developing a product that would allay 'critical inefficiencies in oncology healthcare delivery'. Starting from this economic vantage point, the company then decided to focus on breast cancer, reasoning that 'likelihood of recurrence information for breast cancer would provide the most savings to the healthcare system ... and thus provide the most opportunity for value creation' (Interview with Patrick Terry, Arlington, VA, 16 December 2009, henceforth Terry interview). To develop the prototype of their 21-gene panel, Genomic Health proceeded in four steps that, in their own account, formed a highly structured process: first, they developed a method (RT-PCR) to 'quantify gene expression with the use of sections of fixed, paraffin-embedded tumor tissue'; second, they selected 250 'candidate genes' from a heterogeneous set of sources (in particular, published literature and genomic databases); third, in order 'to test the relation between expression of the 250 candidate genes and the recurrence of breast cancer' they analyzed data obtained from previous clinical studies of breast cancer, and, finally, they used these results 'to select a panel of 16 cancer-related genes and 5 reference genes and designed an algorithm, based on the levels of expression of these genes, to compute a recurrence score for each tumor sample' (Paik et al., 2004: 2817).

Given the technical nature of this process, we will review its key elements in a more user-friendly fashion. Let us begin with the tumor samples. The company's Clinical

Development Group, an internal committee dedicated to finding banked specimens suitable for use in commercial development, had identified stored samples from three sources, including a clinical trial that had compared patients submitted to a post-operative tamoxifen-only regimen (hormone therapy) with patients who had received both hormone and chemotherapy (NSABP trial B-20). In other words, samples were derived, in part, from studies whose design involved the kind of comparison that would be used in the TAILORx set-up. During his previous tenure with Genentech, one of the founders of Genomic Health had worked closely with the director of the NSABP, the NCI-supported clinical trials group, on a trial testing the much vaunted targeted-therapy drug Herceptin. This facilitated access to the NSABP samples while shaping the company's 'therapeutic approach to diagnostics' (Cronin interview).

Genomic Health developers used the samples from 447 patients as the 'discovery' or 'training' set to select the 21 genes eventually included in the *Oncotype* test (Ross et al., 2008: 484). According to the team's lead scientist, 'the 21 we chose were always at the top of the list' (Cronin interview), an unsurprising statement given the prominence of those genes in the published literature. Company researchers then applied an algorithm to the results of the tests and developed the aforementioned RS score. The company believes that the score is one of the strengths of the *Oncotype* test: as a single number on a continuous 0-100 scale and not a category (i.e. yes/no, good/poor), it is supposed to provide clinicians with information that is 'actionable' (Cronin interview) in the dual sense that it provides clinicians with 'useful' information — a basis on which to act —

while preserving clinical decision-making as a clinician's prerogative, since by not providing a categorical answer it does not entail a specific intervention.

Making the results of the *Oncotype* test 'actionable' is a recurrent theme amongst the scientists who developed the test and the clinical researchers who designed the TAILORx trial (Sparano and Solin, 2010; Jessup et al., 2005). An important decision made early on by Genomic Health in this regard concerned the format of patient specimens used for the test. As noted above, the first step in the process concerned the production of a platform, an RT-PCR assay that worked with tissue stabilized with a formalin fixative and embedded in paraffin (FFPE). This is the most common form of pathological tissue preservation, and pathologists in both Europe and the US have developed a number of ancillary techniques and markers that work in paraffin. Genomic Health scientists decided early on to develop a test that could process such samples rather than fresh-frozen samples since using the latter would have required 'going upstream against a system and infrastructure that has been in place for a couple hundred years.' Moreover, '[a]sking pathologists, and surgeons, to give up their traditional markers that mean a lot to them ... was something we decided was less preferable' (Cronin interview).

Genomic Health thus designed the platform for the test based on calculations about how to integrate it into the routine workflow of clinicians and pathology laboratories. Their move was calculated *not* to produce a disruptive innovation. The choice of FFPE shaped the downstream regulatory, scientific, and marketing tasks involved in developing *Oncotype* in far-reaching ways. To begin with, FFPE samples have long been presumed

to contain DNA and proteins so damaged that they are not usable for gene expression profiling. The team at Genomic Health thus devoted vast resources for countering this presumption. ‘We had to show [pathologists] that paraffin could match fresh tissue results, that there was a high concordance.’ They experimented, published and presented at international and national breast cancer meetings, arguing for the feasibility of RNA extraction, amplification, and analysis using FFPE samples (Cronin et al., 2004, 2009; Baker et al., 2005). In 2003, Maureen Cronin chaired a National Institute of Standards & Technology-sponsored workshop involving representatives from government, industry and academia to define standards for RNA-based microarray and quantitative RT-PCR technologies (Cronin et al., 2004).

Having developed the platform and a test for the platform, in 2003 the company’s scientific team once again partnered with the NSABP, this time to *validate* the *Oncotype* prototype. They used banked specimens from a 1982-1988 NSABP study, B-14, and were able to obtain samples and data from 668 breast cancer patients whose cancer fit the profile of *Oncotype*’s intended patient population. Genomic Health researchers generated an RS for each sample and classified patients based on the RS score. Following this assignment, the group tested a series of hypotheses against the categories: Was the proportion of patients free of 10-year post-surgery recurrence significantly greater in the low-risk than the high-risk RS group? Was there a statistically significant relationship between RS and 10-year recurrence, a relationship that ‘went beyond’ the indications provided by existing ‘standard measures’ such as patient age and tumor size? They

concluded that the answer to both questions was affirmative, thus demonstrating the test's *prognostic* value (Paik et al., 2004).

A second validation study using single-center samples from the MD Anderson Cancer Center failed (Esteva et al., 2005), but a subsequent study of 790 patient samples from 14 Northern California Kaiser Permanente Hospitals produced positive results (Habel et al., 2006). Then, once again using patient samples from the NSABP B-20 trial (but, this time, from both the experimental and control arms rather than the control arm alone) Genomic Health researchers sought to demonstrate *Oncotype's* power to *predict* benefit from chemotherapy. This move raised some eyebrows since discovery and validation sample sets are supposed to be independent and, in this case, patients from the B-20 control arm had been used in both studies (Ross et al., 2008: 484).

At about the same time that Genomic Health was founded in 2000, several scientists at the NCI launched PACCT, one of whose aims was to introduce standards and procedures — from statistical design of trials to levels and types of evidence — for the development of molecular diagnostic tests. The program's strategy group, a committee of oncologists, patient advocates, FDA representatives, and NCI scientists from various divisions, soon began discussing a debut pilot project. According to PACCT's leader, Sheila Taube, the group had asked: 'What are the most pressing clinical questions that might be informed by an assay?' Based on their understanding that 'chemo only helped in twenty to thirty percent of [breast cancer] cases', they decided that early stage breast cancer should be a priority. Like Genomic Health, PACCT considered adaptation of the test to existing

clinical routines to be paramount in the design of a new genomic technology to increase the likelihood of its use. The strategy pursued ultimately conformed to that outlined in PACCT's guide to the development and testing procedures for molecular diagnostics in 2005: 'When a promising marker is identified ... the intended clinical use of the marker has to be defined because evaluation of both the assay characteristics ... and the clinical utility have to be assessed in the context for intended use' (interview with Sheila Taube, Bethesda, MD, 18 December 2009, henceforth Taube interview).

At this stage, we must briefly discuss some terminology that, in the case of genomic tests, has provoked conflicting claims, namely the (regulatory) distinction between analytical validity, clinical validity and clinical utility (Teutsch et al., 2009; for a sociological discussion see Hedgecoe, 2008). Analytical validity refers to the test *per se* — Does the test work the way it is supposed to work? Does it measure accurately and consistently whatever it is supposed to measure? — and is maintained via practices such as quality control, calibration, etc. The next two items are more difficult to define. Clinical validity refers to the extent to which the entity measured by the test is consistently related to a clinical variable both in terms of specificity and sensitivity; in other words, it 'refers to the accuracy with which a test predicts a given ... clinical disorder or outcome' (SACGHS, 2008: 32). Clinical utility, in contrast, 'refers to ... [a test's] ability to inform clinical decision-making [sic], prevent adverse health outcomes (e.g., morbidity, mortality), and predict outcomes considered important to patients and their families' (SACGHS, 2008: 32-33). Variables such as effectiveness and cost are also commonly included, explicitly or implicitly, in the definition of clinical utility (Smart, 2006), as are

practical considerations concerning whether the test ‘fits’ existing routines. In the present case, clinical, commercial and regulatory actors on both sides of the Atlantic made contradictory claims about the clinical validity and/or utility of the two competing signatures, *Oncotype* and *MammaPrint*, and of the evidence necessary to establish that validity and/or utility. In a promotional brochure, Agendia went so far as to claim that ‘clinical validity equals clinical utility’.

From the perspective of PACCT, the key to ensuring ‘clinical utility’ was the use of multiple retrospective *and* prospective validation studies. This type of validation prescribed for diagnostic development is significant because it requires a specific infrastructure and set of institutional relationships. Retrospective studies — studies using collections of biological samples that are already on file, whether in general purpose bio-banks or from previous studies — are less costly but are often deemed less rigorous because, by their very nature, they lack an inquiry-specific data collection design. Clinicians consider retrospective studies using prospective data and samples (i.e., those collected by a clinical trial already completed and designed with a different purpose) superior to studies using ‘convenience samples’, such as those contained in a general-purpose bio-bank. Clinicians, however, also mistrust retrospective studies insofar as patient selection bias cannot be excluded (Cardoso, 2005). Prospective studies like TAILORx and MINDACT thus remain the gold standard for testing the value of a genomic tool such as *Oncotype* or *MammaPrint*. In prospective studies, the protocol is designed prior to data collection: collection infrastructures, training, and groups can be aligned and calibrated to answer the specific inquiry at hand. At the US FDA, while data



collected from retrospective studies have sufficed in the past for *diagnostics* validation and approval, developers of drug *therapies* are almost always required to submit data from monitored prospective clinical trials.

Throughout the 1980s and 1990s, the NCI's therapy and diagnosis programs had been administrated separately, with cancer therapy receiving the lion's share of funding, but work on improving the research profile for diagnostics began as early as 1988, with the establishment of a Cancer Diagnosis, Decision, and Implementation Committee. Concurrently with the founding of the PACCT program in 2000, cancer diagnosis research was elevated within NCI to the same status as therapy research. By then, the diagnosis team had established several major bio-banks. PACCT subsequently pursued the establishment of general-purpose banks of searchable, comprehensively tagged and indexed tissue specimens to service validation programs in the field of genomic technologies (Sparano interview; Taube interview). The production and testing of assays like *Oncotype* is thus embedded in the regulatory networks that give rise to standards (e.g., for evaluating validity) and infrastructures (e.g., for enabling such standardized evaluations). These networks combine both formal and informal work routines. One of the co-PIs in the TAILORx trial described the informal nature of some of this work: '[because of the lack of emphasis and funding for them] we often try to get marker studies done on the side — I happen to have some samples, I happen to have an assay, I have lunch with somebody who does [too], we throw them together ... show clinical validity ... but people don't understand it' (Hayes interview).

The intersection of Genomic Health and PACCT, then, was not simply the junction of a genomic product with a diagnostic research initiative. It was the hybridization of a genomic test that had been attuned to address a range of clinical, laboratory, and interactional issues (clinical hierarchies, clinical decision-making, etc.) and an initiative that sought to reorganize biological collections and repositories, procedural norms, and ultimately, institutional alignments in order to make a new kind of research possible. When PACCT founder Sheila Taube recalled her first meeting with Genomic Health at a 2003 cooperative oncology group meeting, she said that she immediately recognized that: ‘[they] had done everything right’. Another company had presented a competing test at the meeting but it ‘was too focused on technology and developing a platform. Steve Shak [Genomic Health’s chief scientist] was a *clinician*’ (Taube interview). By ‘everything right’, Taube meant several things. First, Genomic Health had chosen a feasible format (FFPE). Second, they had developed the gene panel for the assay according to defined clinical endpoints. Finally, they had tested and refined the assay and scoring system in a retrospective study of specimens from clinical trials.

PACCT formally reviewed Genomic Health’s data and canvassed cooperative oncology groups about using *Oncotype* in a pilot trial. *Oncotype* was not accepted immediately or wholesale — two related questions about the test lingered. Were the links between the RS score and observed recurrence stable? Which patients in a trial using the genomic test should have chemotherapy withheld on the basis of the *Oncotype* results? In addition, confusion surrounded the clinical validity of the test: even though the B-14 samples had shown that *Oncotype* was a significant predictor of recurrence (a *prognostic* marker),

clinicians were not convinced that *Oncotype* predicted reaction to chemotherapy (a *predictive* marker). The B-14 study had therefore not addressed one of the main practical questions that clinical oncologists in the cooperative groups said they faced when treating patients: ‘The physicians looked at the data ... showing that RS could predict low versus high [risk of recurrence] and they said, ‘not good enough’. [They] wanted to know, ‘will they benefit if I give them chemo or not?’ (Lively interview). Genomic Health responded by initiating and completing the aforementioned retrospective-prospective study with B-20 samples (Paik et al, 2006; see also Sparano and Paik, 2008: 724).

On the specific issue of using *Oncotype* to withhold chemotherapy from women for whom it would otherwise be a standard of care, PACCT and the cooperative groups engaged in prolonged debate, as some physicians expressed hesitation about such a bold step. There were further debates about where to set the cut-off points for the intermediate range of women who would be randomly selected to not receive chemo. ‘Clinicians [said], how do we know that if we don’t treat someone we aren’t depriving them?’ (Taube interview). The Breast Cancer Intergroup, which had taken over the project from PACCT, ultimately decided to lower the RS for the women randomly selected for chemotherapy (from 18-31 to 11-25), meaning that women with scores higher than the new threshold would be assured chemotherapy. The *Oncotype* test accepted and integrated into TAILORx was thus an object that had been refashioned in order to make it congruous with physician conceptions of both *clinical utility* and *adequate care*.

The design of both the TAILORx and MINDACT trials led to protracted discussions of

topics ranging from the implementation of randomization, to which groups should be randomized and compared. In the end, TAILORx adopted a relatively simple protocol that compared outcomes in the middle RS-score patients randomly assigned to chemotherapy. An initial design proposal had embodied an explicit comparison of *Oncotype* with traditional markers. That protocol had two arms, one tracking outcomes in women whose treatment was assigned based on the RS score, and a second following women who had been assigned to chemotherapy according to clinico-pathological markers. The committee overseeing the trial dispensed with this earlier design since the protocol entailed withholding the test from half of the participants. Had they done so, ‘[b]ecause the test was commercially available ... people would just test outside the context of the study’ (Sparano interview).

### **From the 70-gene signature to MammaPrint and Agendia**

*Oncotype* was conceived and developed in-house by a commercial enterprise. In contrast, MammaPrint, also known as the Amsterdam 70-gene signature, originated in a research project carried out at the Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital (NKI-AVL), a publicly funded institution with a 180-bed hospital, a large outpatient clinic, a cancer research center (the NKI) and over 500 scientists and research personnel (Van Luenen and van Harten, 2006). Since 1983, the NKI-AVL has maintained a bio-bank of frozen (in contrast to FFPE) tumor specimens. By the late 1990s, the bio-bank held approximately 20,000 such samples, a vital resource for molecular genetics projects

that rely on RNA extraction from tumor cells. The fact that the specimens also bore high-quality, long-term clinical follow-up information on patients who had, moreover, routinely signed consent forms for research use, raised their bio-clinical value significantly, turning the NKI bio-bank into a rather unique resource for cancer genomic studies.

In 1993, as part of a drive to restore its scientific visibility,<sup>iii</sup> the NKI established a position in ‘molecular pathology’. The chosen candidate, molecular biologist Laura van’t Veer, was given a mandate to conduct what would soon be called translational research, allying, for instance, studies of hereditary breast cancer susceptibility genes with the establishment of a familial-cancer clinic (Interview with Laura J. van’t Veer, Amsterdam, 8 December 2006, henceforth van’t Veer interview; see also Beishon, 2006). Following a Dutch Ph.D., van’t Veer had received training in molecular oncology as a post-doc at Harvard Medical School where she met the future NKI head of the molecular carcinogenesis group, René Bernards, and the future co-founder of Rosetta Inpharmatics, Stephen Friend, in collaboration with whom the 70-gene signature was eventually developed (interview with René Bernards, Amsterdam, 23 June 2010, henceforth Bernards interview). Around 2000, Bernards and van’t Veer turned to microarrays, a novel technology that had become available in the second half of the 1990s and that allows for the simultaneous analysis of the activity of thousands of genes, thus resulting in so-called *gene expression profiles*, also known as *genomic signatures*. In the present case, the signature was the particular pattern of gene activation that characterizes a patient’s tumor cells.

The NKI project favored the use of innovative genomic tools rather than seeking accommodation with existing pathological laboratory practices. Both the NKI and the Genomic Health teams planned, of course, to develop not only a clinically *valid* test, but also a clinically *useful* one, but they differed in the extent to which they expected that existing practices could and should make room for the new tool. NKI sought replacement rather than addition. The NKI-AVL bio-bank and the related clinical information that it held acted as a condition of possibility for this project for several reasons. First, as already mentioned, microarray analysis requires the use of fresh-frozen tissue. Second, the NKI team decided to focus on the biological underpinnings of breast cancer prognosis and for this they needed specimens of untreated tumors. Their aim, in other words, was to spotlight the natural history, as opposed to the treated history, of breast tumors: Were some tumors intrinsically (i.e., prior to any treatment) more likely to produce metastases to other organs. If so, could one differentiate between high-risk and low-risk patients? Since Dutch oncologists did not offer systemic therapy to node-negative, premenopausal patients prior to 1995, the NKI bio-bank contained the required specimens. Last but not least, the clinical information attached to the samples covered an extended time frame, thus allowing for correlations between the gene expression profile and patient outcomes such as relapse-free survival or disease recurrence and death.

Using these selection criteria, the NKI team procured and analyzed 78 tumors with whole-genome microarrays developed by a Seattle company, Rosetta Inpharmatics. The NKI team processed the tumor specimens in Amsterdam, while the microarray and

bioinformatics work was carried out in Seattle, with frequent meetings between the two teams to make sense of the resulting data (Bernards interview). Through supervised data analysis (i.e., researchers inspected the microarray data by taking into account the clinical information linked to the samples), the team selected 231 genes that appeared to be linked to prognosis out of the approximately 25,000 human genes present on the microarray. This number was subsequently reduced to the 70 genes that provided a reproducible pattern, hence the name *70-gene signature*. The project had started in 2000 and the first results were published in a 2002 *Nature* article (Van't Veer et al., 2002) that quickly became a citation classic: less than four years later it had garnered 1400 citations.

The Amsterdam signature was designed to distinguish between low-risk, 'good signature' patients and high-risk, 'poor signature' patients. In other words, contrary to *Oncotype's* 21-gene signature, whose results are expressed in terms of a continuous variable, the 70-gene signature was construed as a dichotomy. How was such a distinction produced? Inspection of the distinctive microarray pattern of the signature — frequently reproduced in articles and commercial brochures since its original publication in the 2002 *Nature* article — shows the conventional nature of the threshold between the good and poor signature. Some of the patients who fall into the good signature category developed metastases, meaning that the boundary separating the good from the poor signature was drawn so as to include a small number of false negatives. The rationale for doing so emerged from discussions between the research team and clinicians, who insisted that the main goal of the test should be to avoid overtreatment. To accomplish this end, the low-risk group had to be defined inclusively. At the same time, the test developers felt that

clinicians expected a clear answer (good/poor signature) from the test, hence the dichotomy (Bernards interview; interview with Dr. Emiel Rutgers, Amsterdam, 23 June 2010, henceforth Rutgers interview). This position, once again, contrasts with Genomic Health's decision to report their data analysis as a continuous variable that leaves room for clinical judgment.

The NKI team had developed their signature using retrospective data from 78 patients. The signature thus needed validation on a larger, independent patient population. Given that the test had to be performed with frozen tissue, the Amsterdam group could not adopt the Genomic Health solution of using FFPE samples obtained from the tissue banks of previous clinical trials. The primary validation of the Amsterdam signature was thus carried out via another retrospective study that used samples from 295 patients held in the same NKI bio-bank. This choice would come to haunt debates about the relative merits of the 21-gene and the 70-gene signatures, both because retrospective studies carry less weight than prospective studies and because of the overlap between the two sample sets. A team of TRANSBIG researchers (the institution that eventually launched MINDACT) later published an independent, but once again retrospective, validation study in 2006 using 307 patient samples from five European institutions.

As we saw, one of the obstacles faced by Genomic Health was to convince practitioners that RNA could be reliably extracted from FFPE samples. For the Amsterdam team, the problem was less RNA extraction than the microarray analysis itself. Compared to RT-PCR, microarray analysis was a relatively novel, non-standardized technology and as



such, it raised a number of logistic and statistical challenges (Rogers and Cambrosio, 2007; Keating and Cambrosio, 2010). As a result, in addition to the validation studies of the signature *per se*, researchers conducted a number of other studies to show that sample collection for the test (as distinct from the centrally performed test itself) was feasible and reproducible in a routine context, for instance in ‘community-based settings’ (Bueno-de-Mesquita et al., 2007).

Following validation, the NKI team faced several choices. The first concerned commercial development. Until the mid-1990s, the NKI had not been particularly concerned with intellectual property and technology transfer. Since then, however, its leaders had become aware of the importance of these issues and set up procedures to deal with them and educate Institute members as to their relevance (Van Luenen and van Harten, 2006: 213-15). Although they initially favored licensing the technology, the NKI team found no viable taker. So, in 2003 van’t Veer and Bernards, in consultation with the NKI board of directors, established a spin-off company using private venture capital and European Union funding, and convinced the director of oncology at a leading diagnostic company, Amersham, to join them as CEO (Bernards interview). They called the company Agendia, for Amsterdam **Genetic Diagnostics**. Note that the (developing) test led to the company, and not vice versa as in the case of Genomic Health and Oncotype.

Secondly, while the Agendia team had a signature, they did not have a test. In other words, it was not immediately obvious how to convert the 70-gene signature into what eventually became MammaPrint, a ‘high-throughput diagnostic test’ (Glas et al., 2006).

The original signature had been developed using microarrays containing 25,000 oligonucleotides, a highly impractical platform for routine use. The company therefore developed a customized microarray containing a reduced set of probes, whose production was entrusted to Agilent, to whom Rosetta had, in the meantime, sold its technology. As with *Oncotype*, MammaPrint is performed ‘in-house’ at Agendia’s lab: in addition to protecting proprietary procedures, this obviates inter-laboratory differences. The MammaPrint kit contains a ‘puncher’ (a small instrument designed to procure a fresh tumor sample), a test tube with a proprietary solution called RNARetain to preserve the sample’s RNA for several days (thus avoiding the need to immediately freeze the sample and making sampling doable in community settings) and a special envelope to send the preserved sample to Agendia via courier.

Armed with the test, Agendia now sought to clarify its regulatory status, in particular in the US market, and to confirm its clinical validity and usefulness in prospective clinical trials. Let’s begin with the regulatory issue. US law gives the FDA the power to regulate drugs and devices, with tests falling under the less rigorous medical devices statute. As explained in more detail elsewhere (Bourret et al., 2010), the FDA has traditionally exercised ‘enforcement discretion’ by leaving the actual performance of ‘in-house’ tests to be regulated by a different mechanism defined by the Clinical Laboratory Improvement Amendments (CLIA), a set of federal regulatory standards that fall under the authority of the Centers for Medicare and Medicaid Services and that are intended to ensure the reliability and accuracy of clinical laboratory testing. Confronted, however, with the development of tests such as *Oncotype* and MammaPrint, FDA regulators

suggested that these tests may constitute an entirely new regulatory category.

Contrary to more traditional tests, *Oncotype* and *MammaPrint* provide a clinical judgment as opposed to a laboratory result that has to be interpreted by a clinician. Moreover, the algorithm used to reach this judgment is proprietary and thus inaccessible to the clinician. In a move that remains controversial at the time of this writing, in 2006 and 2007 the FDA published two versions of a ‘Draft Guidance’, signaling the Agency’s inclination to step in and take direct responsibility for the novel test category. While, to date, the FDA has not issued a legally binding guidance, in 2007, Agendia decided to submit its *MammaPrint* test to the agency and successfully obtained FDA ‘clearance’ after only 30 days. An ‘FDA cleared’ button promptly (and prominently) appeared on all commercial *MammaPrint* material. Given the non-binding nature of the FDA draft guidance, Genomic Health chose not to pursue this regulatory route, preferring instead to gain ‘official’ recognition via inclusion in the clinical practice guidelines of professional oncology organizations. As one of the founders of Genomic Health explained, the company viewed pursuit of FDA clearance as much more costly and time-consuming than simply lobbying professional organizations of clinicians — many of whom the founders already knew through their previous work at Genentech (Terry interview). The American Society of Clinical Oncology (ASCO) included *Oncotype* in its 2007 guidelines and the U.S. National Comprehensive Cancer Network (NCCN) followed suit in its 2008 guidelines, moves that the Agendia scientific director and TRANSBIG clinicians involved with *MammaPrint*’s clinical testing criticized, claiming the organizations had deployed inconsistent criteria for inclusion that had favored the US

product (Bernards, 2008; Cardoso et al., 2008).<sup>iv</sup>

TRANSBIG is a translational spin-off of the Breast International Group (BIG). Founded in 1996 by leading European clinicians, the BIG network is an attempt to establish a European-led equivalent of the NCI-sponsored Breast Cancer Intergroup of North America, an international network of cooperative groups conducting breast cancer clinical research. While TRANSBIG's official objective was to promote international collaboration in translational research, in practice its creation in 2004 was an *ad hoc* response to the perceived need for a first large-scale European trial involving a genomic signature for breast cancer and, in particular, a way of securing European funding for such a trial, since EU regulations restricted funding to new organizations. In keeping with the regulatory challenges posed by this bio-clinical enterprise, TRANSBIG also sought to establish a stable framework for dealing with issues such as intellectual property and different national regulations governing biological samples.

Once established, TRANSBIG faced the issue of which genomic product to choose. While the Amsterdam signature announced in the 2002 *Nature* publication was an obvious candidate, it had not yet been independently validated. Thus, the first step undertaken by the consortium was to carry out the aforementioned validation of the signature using frozen material from British, Swedish and French patients (Buyse et al., 2006). The study's conclusion, that the Amsterdam signature outperformed commonly used clinical-pathological criteria to assess the recurrence risk of node-negative breast cancer patients, was considered by TRANSBIG as the 'green light to move forward with

the trial' (Straehle et al., 2009: S23).<sup>v</sup> The team now had to develop a suitable trial design and protocol. While cancer clinical trials have recourse to relatively routine designs, no standard design was available for the comparison of a biomarker with clinical tools to assess the risk of cancer recurrence and thus reach a treatment decision. An initially simple, somewhat classical design was prepared in 2004. In 2005, however, the MINDACT steering committee announced a radical transformation of the trial's initial design. A second, incremental design change subsequently added randomized comparisons of different chemotherapy and hormone therapy regimens to the protocol. These latest additions drew pharmaceutical firms and the related intellectual property issues into the trial, which as a result had its start date moved from 2005 to February of 2007. Let us have a closer look at these design changes.

The initial, 2004 design was not so different from a classical clinical trial: breast cancer patients (early stage, node-negative) were to be randomly assigned either to an arm in which the decision to give or withhold chemotherapy would be based on routine clinical-pathological criteria (low/high risk), or to an arm in which the decision would be based on the MammaPrint verdict (good/poor signature). By comparing the clinical outcomes between the two arms, the design aimed to show that decreasing the number of patients in the signature arm who received chemotherapy did not increase the rate of distant relapse: in other words selection using the signature was not inferior in terms of efficacy. This design, aptly termed a 'non-inferiority design', soon attracted criticism. The main objection was that it did not take into account that the clinical-pathological and the signature methods reached concordant conclusions for many patients, and that 'any

benefit of either approach would be greatly diluted in such an overall comparison’ (Bogaerts et al., 2006: 545).

Simulations carried out in the wake of this criticism confirmed the reality and scope of the problem. After considering and rejecting two other options, the emerging consensus among TRANSBIG researchers was to focus the trial on a core group of ‘discordant’ patients, the group of patients for whom the signature and traditional clinical-pathological criteria produced opposite results. The trial would be designed so as to identify and compare all the patients who would be treated differently depending on which of the two prognostic assessments (clinical or signature) was followed. To make this comparison, all the patients would have to be tested with MammaPrint, but only discordant patients would be randomized to chemotherapy vs. no chemotherapy. According to TRANSBIG’s head, Martine Piccart, the new design quickly garnered widespread support: ‘Everybody thinks now we have a much stronger design and we will be able to exactly answer all of the questions that we were interested in’ (Tuma, 2005: 332). While we cannot discuss here the details of this choice and the debates or even the outspoken criticism (e.g. Koscielny, 2008) that it raised, it is worth mentioning that it was a *pragmatic* choice. Rather than being the optimal solution from a statistical point of view, it was a ‘satisficing’ (Simon, 1956) one, given the material and institutional constraints the trialists faced, and what the clinicians perceived as the primary question to be answered by the trial (interview with Jan Bogaerts, Brussels, 30 June 2010)

The new protocol was now much more expensive given that genomic testing had to be performed on all patients. The trial leaders, moreover, had also decided to use a full-genome version of the Agendia microarray, instead of the commercial 70-gene MammaPrint pack. Thus, in contrast to TAILORx, MINDACT became a discovery platform that, in addition to testing the 70-gene signature, was designed to provide data on the involvement of additional sets of genes in patients' response to therapies (telephone interview with Martine Piccart, 7 October 2010). Finally, for statistical reasons, the number of patients had to be increased to 6000 patients. Although the EU had awarded €7 million in funding, the cost of the trial was now estimated at €40 million. So, TRANSBIG added a drug-testing component — a 'drug question' (Wilkinson, 2009: 390) — in order to obtain funding from pharmaceutical companies. As TRANSBIG's Scientific Director confessed, 'we've been forced to bring in commercial partners, which has brought complexity to the trial design and forces us to share the legal rights to any data generated' (Fricker, 2007: 9). She and fellow clinicians have since come to see this development as a bonus, insofar as the trial will produce a wealth of genomic information concerning patients' response to specific drugs (Cardoso interview; Rutgers interview).

As already noted, MammaPrint requires fresh-frozen tissue. The commercial kit distributed by Agendia contains the aforementioned RNARetain solution, but this was insufficient for MINDACT. In addition to performing a set of ancillary genomic and proteomic studies, a very important component of the trial is the constitution of a frozen tissue bio-bank as a resource for future studies. This required changes in surgeons' and pathologists' routines and created the challenge of circulating bio-specimens around

Europe. While all RNA extraction and microarray analysis is performed in the central laboratory of Agendia in Amsterdam, thus reducing operator and technical variability in the course of gene expression analysis, serum samples are sent to Wales for proteomics analysis, and additional tissue, including FFPE specimens, to the European Oncology Institute in Milan for central pathological review. Last but not least, remaining tissue and RNA goes to Brussels where it is stored in the TRANSBIG bio-bank. As a result, and in contrast to the more traditional logistics required by TAILORx, ‘the logistics of MINDACT have been one of the most challenging and expensive part of the trial’ (Bogaerts et al., 2006: 549). The TRANSBIG team thus decided to perform a pilot study in six European hospitals to test the collecting, freezing and shipping of specimens, and to define in painstaking detail – right down to the glue to be used so that labels do not fall off from sample containers when shipped with dry ice (Cardoso interview) – the standard operating procedures prescribed by the MINDACT protocol (Mook et al., 2007: 154). When presenting the results of the pilot phase of the trial in 2010, the TRANSBIG team announced that the principal finding, to date, was that such a trial could actually be carried out in real-world conditions (Rutgers et al., 2010).

### **The stakes of the trials**

As noted above, MINDACT and TAILORx share a number of commonalities: they both participate in the redefinition of cancer clinical research as *translational* research. Moreover, both trials seek to establish the clinical relevance of genomic signatures, rather



than to test a particular drug or drug combination, as in more traditional trials. In this sense, both trials are ‘proof of concept’ trials, adapting the clinical trial methodology developed in the Rx domain to the Dx (diagnostic) domain. In addition, the signatures at the core of the two trials are proprietary tests, which are performed in the central laboratories of the commercial enterprises that produce and market them. This explains the MINDACT trialists’ claim (Cardoso et al., 2007: 249-50) that the novelty of these studies lies also in the fact that they experiment with new models of interaction between pharmaceutical companies, academic and public clinical institutions, cancer agencies, regulatory bodies, professional segments within the field of oncology, and patient groups. Finally, the technologies tested in the two trials have the potential to rearrange clinical decision-making by shifting relations between biologists, pathologists and clinicians, insofar as they use prognostic/predictive markers (more below on this emerging distinction) that no longer rely on the trained eye of the pathologist and can dictate (or at least direct) therapeutic decisions. And yet, as readers will by now have surmised, the two trials and the signatures they test differ significantly in their approach to therapeutic decisions, as they embody two different ‘scripts’ of the relation between genomic innovation and existing clinical activities, and two different understandings of what should count as *clinical utility*. In this final section, we will concentrate on this last issue.

Madeleine Akrich (1992) has argued that technical objects, from their initial design phase, embody an understanding — which, borrowing a term from the film industry, she calls a *script* — of their future context of use. Scripts, from this point of view, may embody diverging representations (de-*scriptions*) of the context within which the new

tools are expected to be deployed and different perceptions of the extent to which a given organizational and epistemic configuration can (and/or ought to) be transformed by the new tool (*pre-scriptions*). According to Akrich, objects co-produce the context of their use, to which they cannot therefore be reduced, as well as the norms that regulate such a use (see also Callon, 1986; Jasanoff 2004). In addition, these understandings, representations and perceptions may evolve. *Oncotype* and MammaPrint are excellent illustrations of script-laden objects.

The promoters of *Oncotype* have envisioned a medical system consisting of more or less rigidly interconnected sectors — clinicians, pathologists, clinical laboratory specialists and regulators — into which genomic tools must be introduced by adapting them as much as possible to existing routines. In particular, as we saw, the assay was designed to work with FFPE tumor samples because this is how a large majority of pathologists preserve patient specimens: a considerable amount of work went into developing an RT-PCR technology that allows *Oncotype* to use such samples. Moreover, the individual biomarkers included in *Oncotype*, while measured using RT-PCR rather than more traditional techniques, include many of the routine biomarkers used by pathologists and clinicians to establish a patient's prognosis. Candidate genes for developing the test were selected in part from the published literature and, as argued during a Genomic Health session at a recent breast cancer conference, clinicians using *Oncotype* need not to worry about replacing trusted indicators with a novel tool, because those indicators are included in the test; indeed, the report sent by Genomic Health provides, in addition to the RS

score, individual values for traditional biomarkers (Fieldnotes, EBCC-7, Barcelona, 24-27 March 2010, henceforth, Barcelona fieldnotes).

The TAILORx team has, similarly, emphasized the *complementary* nature of the information provided by *Oncotype* (Goldstein et al., 2008). At the aforementioned conference session, Genomic Health organizers distributed a glossy brochure with case presentations of ‘real patients’ to participants who were then asked to vote, case by case, on which treatment they would suggest based on the standard clinico-pathological criteria, and then to vote again, after disclosure of the patients’ RS scores. While the manifest function of this exercise was to demonstrate the test’s clinical utility, its latent function was to show how the test could easily ‘fit into’ existing clinical routines (Barcelona fieldnotes). The very fact that RS has an intermediate zone allows Genomic Health to claim that the test provides clinicians with information that they can integrate into existing clinical information in order to exercise their judgment, rather than receiving a yes or no that leaves them no other choice than to follow or ignore such a verdict.

MammaPrint embodies a different script, one that envisions a rather more disruptive role for genomic tools vis-à-vis established routines. The test uses microarrays and thus necessitates recourse to fresh-frozen tissue. Although the addition of an RNA-preserving solution for fresh tissue represents an attempt to facilitate test use by delaying the required freezing procedure, it remains a departure from standard routines. The message is that to fully exploit the potentialities of genomics, surgeons and pathologists must adapt to the requirements of the new tool, rather than *vice versa* (Bernards, 2010). The

development of the test, as we saw, followed a markedly different route when compared to *Oncotype*: rather than selecting genes with known prognostic value, the NKI team began by first examining all the genes included on a whole-genome chip and letting the bioinformatics tools and statistical algorithms decide which genes would be of prognostic relevance for breast cancer patients. Clinical researchers were thus unfamiliar with several of the genes of the resulting 70-gene signature. Rather than a score derived from the weighted expression of individual genes, it is the signature as a whole that provides the result or, to be more exact, the correlation between a patient's tumor profile and the MammaPrint template profile. The test thus provides a qualitative, dichotomous answer (good/poor signature) to the question of whether a patient is likely to develop distant metastases. In turn, such a test design presupposes that clinicians expect clear answers from genomics, rather than discretionary information, and that this does not prevent them from exercising their clinical judgment (Rutgers interview).

Last but not least, the MINDACT protocol stages a head-to-head confrontation between traditional clinico-pathological indicators and MammaPrint: the goal of the trial, in other words, is to show that the latter is superior to the former. In order to avoid any misunderstanding, we should immediately add that we are not suggesting that the dividing lines runs between genomic researchers and clinicians. At the aforementioned European breast cancer conference, for instance, one session was devoted to a debate between two groups of clinicians: those who maintained that a classical and somewhat obvious indicator, 'nodal status' (i.e., the degree of spread of tumor cells to lymph nodes), should continue to be used as a major prognostic factor and those who argued that

genomic parameters should replace it, with a sizeable portion of the largely clinical audience favoring the latter (Barcelona fieldnotes). At yet another conference a leading US pathologist exhorted his colleagues to convert from FFPE to microarrays, arguing that they could easily do so thanks to technical developments such as RNA-preserving solutions (Monaco fieldnotes).

The different regulatory routes taken by Agendia and Genomic Health (FDA clearance vs. clinical practice guidelines) show that genomic signatures, at this stage, are still located in a grey zone that lies between the objectifying procedures used for drug approval and the professional domain of clinical expertise. In turn, the two different scripts embodied in *Oncotype* and *MammaPrint* flow, in part, from the existence of differences in the procedures characterizing the activities of *clinical* networks in the US and Europe. One should not over-generalize, since different configurations coexist within these two *technological zones* (Barry, 2001); the distinction is thus a matter of degree and distribution rather than categorical. Still, the US situation is characterized by the presence of strong central institutions such as the NIH and, more specifically, the NCI that directly or indirectly control, financially and administratively, a large portion of clinical research activity (including oncology cooperative groups). Within this system, the balance between clinical and basic research activities is presently in flux, as evidenced by both the hype and the concrete initiatives surrounding that fuzzy entity, translational research. On the clinical side lie a set of *actuarial* procedures, in the sense that key clinical institutions — ranging from the aforementioned public agencies to powerful organizations, such as ASCO and NCCN — act as obligatory passage points. They do so

by defining a sequence of steps that, while sometimes perceived as obstacles, also provide opportunities and a platform for promoters of new tests who follow the rules (thus the PACCT leader's comment on Genomic Health's development procedure: 'They did everything right!'). These institutions set out the norms, guidelines, and pathways that allow entrepreneurs to informally calculate risk and to undertake an actuarial view of the road ahead. This is what allowed Genomic Health to construct a signature by asking users at every step of the actuarial network what clinical questions they wanted the signature to answer and what data would be credible in that regard.

At the pan-European level, due to factors that range from different national legislations to the lack of a central cancer agency such as the NCI, the situation is presently much more fluid, leaving researchers no choice but to pursue ad hoc arrangements and other forms of tinkering that, on the one hand, create multiple obstacles to the development of bio-clinical projects but, on the other hand, provide a degree of flexibility and open-endedness. When the NKI team discovered a signature — only later turned into a commercial product — it went looking for clinical validation by using the researchers' own understanding of what should count as proper clinical validation and as clinical utility. We are not claiming that the institutional and organizational contexts account for differences in the development of *Oncotype* and *MammaPrint*. Rather, we argue that test developers built into their product and development procedures (more obviously so in the case of Genomic Health) provisions that corresponded to their own representation of the existing network of actors and practices and of the opportunities it provided. *MammaPrint* has in the meantime been included in a more recent North-American breast

cancer clinical trial called I-SPY 2 that lies closer to the basic side of the translational continuum (Freiherr, 2010), showing the heterogeneity of existing US bio-clinical networks.

The differences just mentioned are most evident in the debates surrounding the *clinical utility* of each test. While regulators and authors of methodology articles offer seemingly clear-cut definitions of notions such as clinical utility and clinical validity, there is a good degree of slippage in actual use, prompting a former FDA regulator to claim that '[c]linical utility is a lot like beauty, it's in the eye of the beholder' (Ray, 2010). Similar considerations apply to the distinction between *prognosis* and *prediction*, two notions that, in the present case, were closely connected to discussions of clinical utility. This should come as no surprise insofar as concepts are always defined, enacted, performed and otherwise used *in relation* to a specific set of practices (Mol, 2002). In particular, genomic tools, given their innovative potential, confound existing distinctions such as prediction and prognosis as the recent proliferation of papers and conference presentations devoted to these notions clearly attests. Much work is presently devoted to the attempt to re-specify their meaning in relation to the new tools. At the same time, building in part on the resulting ambiguities, promoters of *Oncotype* and *MammaPrint* have used their own interpretation of these in the scientific, clinical and commercial confrontation between the two tests.

Recall that clinical validity is generally defined as the degree to which a test identifies a patient population in terms of sensitivity and specificity, and that clinical utility is

supposed to refer to the ability of a test to benefit patients. While, in theory, one can distinguish between, on the one hand, identifying patients whose tumors express, say, a good signature and thus will not develop distant metastases (clinical validity), and, on the other hand, demonstrating that a good signature directs clinicians to avoid chemotherapy that is unlikely to benefit patients (clinical utility), the transitive link between the two becomes obvious when one says: patients whose tumors express a good signature will not develop distant metastases and can thus avoid unnecessary chemotherapy. This is why Agendia claims that clinical validity = clinical utility.

That utility, however, is not immediately evident to all. A co-PI of TAILORx and former Genomic Health consultant noted in an interview, ‘I don't know how to apply [MammaPrint] in my clinic’ (Hayes interview). While he believed that the test effectively segregated women into risk groups (clinical validity), he also felt that it did not provide clear information on how to then choose treatments for these particular patients (clinical utility). Quoting approvingly his ASCO colleagues who decided not to include the MammaPrint test in their guidelines, he said: ‘they felt MammaPrint had analytic validity, they felt MammaPrint had clinical validity, but they didn’t feel it had clinical utility, because there had never been a true clinical question asked’ (Hayes interview). Both he and many of his TAILORx collaborators consider that a key feature of clinical utility is the ability of the assay to provide information that can guide therapeutic decision-making. They simultaneously maintain that this kind of information is more powerful when, in addition to prognosis, it includes prediction; that is, it forecasts how a patient will react to treatment. Traditionally, therapeutic decisions have been made solely



on the basis of prognostic information: a ‘high risk’ patient would be treated more aggressively. By providing independent information about how a patient will react to a given treatment, the genomic test introduces a disjunction between prognosis and prediction and thus redefines clinical utility. On the other hand, the TRANSBIG team that validated the 70-gene signature clearly saw clinical utility in the *prognostic* nature of that test, which extended to therapy. The title of their paper is revealing in this respect: ‘Validation and *clinical utility* of a 70-gene *prognostic* signature’ (Buyse et al., 2006; our emphasis). The authors defined clinical utility as referring, in this specific case, to evidence that the test could ‘reliably identify patients in whom adjuvant chemotherapy is definitely not indicated despite having poor clinical risk factors and patients who need chemotherapy despite having good clinical risk factors’ (Buyse et al., 2006: 1184). This definition corresponds to the design of the MINDACT trial that stages a head-to-head confrontation between clinical and genomic prognostic factors, while maintaining the traditional link between prognosis and therapeutic decisions.

Like clinical utility and validity, the distinction between prognosis and prediction is in principle clear: prognosis refers to the future behavior of a tumor (‘Is the disease likely to recur?’) independently of any treatment, while prediction refers to a tumor’s reaction to a specific treatment. In practice, the distinction is often clouded.<sup>vi</sup> Recall that the NKI team, in order to develop their 70-gene signature, selected patient samples from untreated women in order to explore the biology of tumors and produce a *pure* prognostic marker. The modifier ‘pure’, as used here, refers to the aforementioned transition from clinical validity to clinical utility. In contrast, the Genomic Health team used samples from the

NSABP B-20 trial whose patients had all been treated with tamoxifen, leading Agendia representatives to argue that the prognostic validity of *Oncotype* was limited to the tamoxifen-treated patient population. Be that as it may, B-20 also had a chemotherapy arm that allowed Genomic Health to argue that the test had been validated for *predictive* use: its ability to provide both prognostic and predictive information is what, in their view, made it uniquely appropriate for use in TAILORx (Sparano and Solin, 2010). The debate, as can be easily seen, moves back and forth between these diverging definitions of related concepts, whereby notions such as prediction and prognosis are reshaped as part of the ongoing debate. But this process is not confined to the semantic sphere; it is intimately connected to the practical deployment of specific materials and procedures.

As we saw, the Amsterdam team used convenience samples for validation purposes, whereas Genomic Health undertook retrospective studies of samples collected as part of prospective trials. These two strategies were obviously due to the different designs of the two tests, which in one case gave access to FPPE samples and in the other did not. But this proposition also works in the other direction: *Oncotype* was designed with a specific kind of validation in mind. Genomic Health obviously felt that the PACCT program and its stringent requirements provided the best opportunity for the clinical and commercial success of their test. There seems to be widespread agreement that only a large, randomized prospective clinical trial can establish the predictive value of a biomarker (Simon et al., 2009), but since such a trial is costly and complex to organize, PACCT leaders wanted to maximize the chosen biomarker's chance of success. Cooperative group clinicians, moreover, wanted to have a reasonable assurance that, in accordance

with the severe ethical strictures that regulate US clinical trials, they were not exposing patients to risks. Data provided, even if retrospectively, from samples that had been collected for previous trials seemed to offer the best evidence and played a key role in convincing clinicians that TAILORx could be done without exposing patients to undue risks. According to the trial's PI, 'the B-20 study really sealed the birth of the TAILORx trial' (Sparano interview). Finally, PACCT's definition of clinical utility included a pragmatic dimension: to what extent was a novel test likely to be widely adopted? Genomic Health obviously emphasized how seamlessly *Oncotype* fit into existing clinical arrangements, prompting the former PACCT leader to describe Genomic Health as a company with a clinically oriented 'attitude' (Taube interview).

In contrast to the incremental nature of the procedures characterizing the PACCT/TAILORx project — the *actuarial*, consensually established, professionally authorized evidentiary procedures embodied in documents such as the guidelines drafted by PACCT — the MINDACT and the NKI teams emphasize the radically innovative potential of their endeavor and its research-laden nature, chastising Genomic Health for having 'picked a set of genes based on the literature' instead of harnessing the discovery potential of genomic technologies (Van't Veer interview). And the MINDACT trialists did not hesitate to select a design that stages a head-to-head confrontation with existing clinico-pathological practices. While some US practitioners no doubt consider that MINDACT was launched without sufficient and substantive clinical validation of the test, the TRANSBIG team appears less willing to sacrifice innovation on the altar of existing clinical-regulatory conventions. This is consistent with the general attitude of EORTC

leaders who, faced with the rise of Contract Research Organizations that carry out routine trials for pharmaceutical companies (Mirowski and Van Horn, 2005), have been left to redefine the scope and activities of the pan-European organization as an Academic Research Organization that should privilege ‘pivotal studies’ addressing ‘fundamental issues’ and whose ‘outcome will influence oncological health care’ (Van Oosterom and Meunier, 1995).

In instancing these differences — present in different combinations on both sides of the Atlantic, rather than neatly opposing North America and Europe — our point is less to engage in a comparative study than to emphasize how biomedicine remains a problematic domain. Indeed, biomedicine’s components — ranging from basic biological research to clinical treatment via an increasingly complex set of translational activities, to which one should add hybrid configurations such as regulatory science — are in need of constant re-alignment. And this is even more so with the growing role played by biotechnology startups in biomedical research (Cambrosio et al. 2009). As noted by one of the PACCT members, in the search for an appropriate test for the PACCT pilot trial during 2005-2008, ‘it could only have been a private company’ that sourced the selected test (Lively interview).

The development of *Oncotype* was never strictly a matter of ‘getting the science right’: the test’s shape hinged on a range of factors, including the usability of *Oncotype* within current laboratory infrastructures, practices, and attitudes. In developing *Oncotype* this way, Genomic Health (and PACCT’s guidelines) have in a sense been turning

conventional modes of biotechnology ‘discovery’ on their heads. Bench work is not shrouded in the mystique of unknown possibilities. It proceeds along the lean path of inevitability — the market and the clinical endpoint are what determine the cast of ‘discovery’ rather than the other way around. NKI researchers, in contrast, have adopted a model that embodies a more open-ended and impromptu approach, while also integrating the role of private companies (MammaPrint is produced and marketed by Agendia, which is presently engaged in a fierce competition with Genomic Health). In other words, while both MINDACT and TAILORx deploy new models of interaction between biotechnology/pharmaceutical companies and public agencies and, accordingly, between technology development and regulatory norms, the specific modes of interaction that have unfolded within each trial and many of the differences between the trials and the genomic technologies they incorporated, flowed from this broader distinction.

### **Acknowledgments:**

Research for this paper was made possible by grants from the Canadian Institutes of Health Research (MOP-93553), the Fonds québécois de la recherche sur la société et la culture (SE-124896), the Social Sciences and Humanities Research Council of Canada (410-2008-1833), the French Institut National du Cancer (0610/3D1418/SHS08), and the Conseil Régional PACA. We would like to thank all the scientists and clinicians who kindly accepted to be interviewed and to comment on early drafts of the test. A first version of this paper was presented at the 2010 Gordon Cain Conference *Personalized medicine here and now: Empirical studies of post-genomic medicine*, sponsored by the Chemical Heritage Foundation in Philadelphia.

---

### **Notes**

<sup>i</sup> The original meaning of the MINDACT acronym (‘Microarray In Node-negative Disease may Avoid ChemoTherapy’) was amended to include node-positive patients.

<sup>ii</sup> The article mentions different percentages (20%-28%) but, according to Dr. Piccart (telephone interview, 7 October 2010) the journalist misquoted her.

---

<sup>iii</sup> For a short history of the NKI see:

<http://www.nki.nl/Research/About+the+Netherlands+Cancer+Institute/The+history+of+The+Netherlands+Cancer+Institute/History+NKI.htm>

<sup>iv</sup> US clinicians denied the charge citing past decisions that had favored European products (Harris et al., 2008).

<sup>v</sup> In the meantime, two other prognostic signatures for breast cancer had emerged in Europe. Validation studies of the newcomers concluded that the three signatures performed equally as well, but TRANSBIG decided to stick with its initial choice given the more advanced development of the Amsterdam signature.

<sup>vi</sup> To further complicate matters, some clinicians use ‘prognosis’ in relation to the decision whether to treat or not, and ‘prediction’ for the choice of a *specific* treatment, whereas other clinicians speak of non-specific vs. drug-specific prediction (Monaco fieldnotes).

## References

- Akrich M (1992) The de-scription of technical objects. In: Bijker WE and Law J (eds) *Shaping Technology/Building Society: Studies in Sociotechnical Change*. Cambridge, MA: MIT Press, 205-224.
- Baker J, Salter J, Liu ML, et al. (2005) Measurement of gene expression using tissue microarray cores of paraffin embedded breast cancer tissue. Paper presented at the 28<sup>th</sup> Annual San Antonio Breast Cancer Symposium (Abstract 310). San Antonio, TX, 8-11 December.
- Barry A (2001) *Political Machines: Governing a Technological Society*. London: Athlone Press.
- Beishon M (2006) Laura van't Veer: The person behind personalized treatments. *Cancer World* 12 (May-June): 6-13.
- Bernards R (2008) Reaction to American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *Journal of Clinical Oncology* 26: 2057-2058.
- Bernards R (2010) It's diagnostic, stupid. *Cell* 141: 13-17.
- Bogaerts J, Cardoso F, Buyse M, et al. (2006) Gene signature evaluation as a prognostic tool: Challenges in the design of the MINDACT trial. *Nature Clinical Practice Oncology* 3: 540-551.
- Bourret P, Cambrosio A, and Keating P (2010) Regulating diagnosis in post-genomic medicine: Re-aligning clinical judgment? Manuscript under review.
- Brown H (2007) Translational cancer research in the USA. *Molecular Oncology* 1: 11-13.

- Bueno-de-Mesquita JM, van Harten WH, Retèl VP, et al. (2007) Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: A prospective community-based feasibility study (RASTER). *Lancet Oncology* 8: 1079-1087.
- Butler D (2008) Crossing the valley of death. *Nature* 453 (12 June): 840-842.
- Buyse M, Loi S, van't Veer LJ, et al. (2006) Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *Journal of the National Cancer Institute* 98: 1183–1192.
- Callon M (1986) The sociology of an actor-network: The case of the electric vehicle. In: Callon M, Law J, and Rip A (eds) *Mapping the Dynamics of Science and Technology*. London: Macmillan, 19-34.
- Cambrosio A, Keating P, Bourret P, Mustar P, and Rogers S (2009) Genomic platforms and hybrid formations. In: Atkinson P, Glasner P, and Lock M (eds) *Handbook of Genetics and Society: Mapping the New Genomic Era*. London: Routledge, 502-520.
- Cardoso F (2005) Show me the genes – I will tell you who/how to treat! *Breast Cancer Research* 7: 77-79.
- Cardoso F, Piccart-Gebhart M, Van't Veer L, and Rutgers E (2007) The MINDACT trial: The first prospective clinical validation of a genomic tool. *Molecular Oncology* 1: 246-251.
- Cardoso F, Saghatchian M, Thompson A, and Rutgers E (2008) Inconsistent criteria used in American Society of Clinical Oncology 2007 update of recommendations for the



use of tumor markers in breast cancer. *Journal of Clinical Oncology* 26: 2058-2059.

Cronin M, Ghosh K, Sistare F, Quackenbush J, Vilker V, and O'Connell C (2004) Universal RNA reference materials for gene expression. *Clinical Chemistry* 50: 1464-1471.

Cronin MT, Dutta D, Pho M, Nguyen A, Jeong J, and Liu ML (2009) Tumor marker discovery by expression profiling RNA from formalin fixed paraffin embedded tissues. *Methods in Molecular Biology* 520: 177-193.

EJC (2008) Agreements reached with cancer centres. *European Journal of Cancer* 44: 330.

Esteva FJ, Sahin AA, Cristofanilli M, et al. (2005) Prognostic role of a multigene reverse transcriptase-PCR assay in patients with node-negative breast cancer not receiving adjuvant systemic therapy. *Clinical Cancer Research* 11: 3315 –3319.

Freiherr G (2010) Trial aims to shave years off cancer drug development. *Oncology News International* 19(5) (13 May 2010). Available at: <http://www.cancernetwork.com/display/article/10165/1568563>.

Fricker J (2007) Translational cancer research in Europe. *Molecular Oncology* 1:8-10.

Glas AM, Floore A, Delahaye LJMJ, et al. (2006) Converting a breast cancer microarray signature into a high-throughput diagnostic test. *BMC Genomics* 7: 278.

Goldstein LJ, Gray R, Badve S, et al. (2008) Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. *Journal of Clinical Oncology* 26: 4063-4071.

- Goldstein LJ and Sparano JA (2009) In reply. *Journal of Clinical Oncology* 27: 1338-1339.
- Habel LA, Shak S, Jacobs MK, et al. (2006) A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. *Breast Cancer Research* 8: 25.
- Harris L, Fritsche H, Mennel R, et al. (2007) American Society of Clinical Oncology 2007 update recommendations for the use of tumor markers in breast cancer. *Journal of Clinical Oncology* 25: 5287-5312.
- Harris LN, Hayes DF, Bast RC (2008) In reply. *Journal of Clinical Oncology* 26: 2060-2061.
- Hedgecoe A (2008) From resistance to usefulness: Sociology and the clinical use of genetic tests. *BioSocieties* 3: 183–194.
- Hornberger J, Cosler LE, and Lyman GH (2005) Economic analysis of targeting chemotherapy using a 21-gene RT-PCR assay in lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer. *American Journal of Managed Care* 11: 313–324.
- Jasanoff S (2004) Ordering knowledge, ordering society. In Jasanoff S (ed) *States of Knowledge: The Co-Production of Science and Social Order*. London: Routledge, 13-45.
- Jessup JM, Lively TG, and Taube SE (2005) Program for the assessment of cancer clinical tests (PACCT): Implementing promising assays into clinical practice. *Expert Review of Molecular Diagnostics* 5: 271-273.

- Keating P and Cambrosio A (2010) Too many numbers: Microarrays in clinical cancer research. Manuscript under review.
- Koscielny S (2008) Critical review of microarray-based prognostic tests and trials in breast cancer. *Current Opinion in Obstetrics and Gynecology* 20: 47–50.
- Lyman GH, Cosler LE, Kuderer NM, and Hornberger J (2007) Impact of a 21-gene RT-PCR assay on treatment decisions in early-stage breast cancer: An economic analysis based on prognostic and predictive validation studies. *Cancer* 109: 1011–1018.
- Meunier F (2008) The challenges and opportunities of cancer clinical research at pan-European level. Paper presented at the meeting Science & Industry Lunch: New Medicine Research Collaborations (22 April). Available at: [http://crosstalks.vub.ac.be/past\\_events/2006\\_futmed/fmeunier\\_220408.pdf](http://crosstalks.vub.ac.be/past_events/2006_futmed/fmeunier_220408.pdf)
- Mirowski P and Van Horn R (2005) The contract research organization and the commercialization of scientific research. *Social Studies of Science* 35: 503-548.
- Mol A (2002) *The Body Multiple: Ontology in Medical Practice*. Durham, NC: Duke University Press.
- Mook S, Van't Veer LJ, Rutgers EJT, Piccart-Gebhart MJ, and Cardoso F (2007) Individualization of therapy using Mammaprint: From development to the MINDACT trial. *Cancer Genomics & Proteomics* 4: 147-155.
- Paik S, Shak S, Tang G, et al. (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *New England Journal of Medicine* 351 (30 December): 2817-2826.
- Paik S, Tang G, Shak S, et al. (2006) Gene expression and benefit of chemotherapy in

women with node-negative estrogen receptor positive breast cancer. *Journal of Clinical Oncology* 24: 3726-3734.

Ray T (2010) FDA keeping all ideas in play in soliciting public comments on LDT regulation. *Pharmacogenomics Reporter* (21 July). Available at: <http://www.genomeweb.com/dxpgx/fda-keeping-all-ideas-play-soliciting-public-comments-ldt-regulation>

Retèl VP, Joore MA, Knauer M, Linn SC, Hauptmann M, and van Harten WH (2010) Cost-effectiveness of the 70-gene signature versus St. Gallen guidelines and Adjuvant Online for early breast cancer. *European Journal of Cancer* 46: 1382-1391.

Rogers S and Cambrosio A (2007) Making a new technology work: The standardization and regulation of microarrays. *Yale Journal of Biology and Medicine* 80: 165-178.

Ross JS, Hatzis C, Symmans WF, Pusztai L, and Hortobágyi GN (2008) Commercialized multigene predictors of clinical outcome for breast cancer. *The Oncologist* 13: 477-493.

Rutgers E, Piccart M, Delaloge S, et al. (2010) The EORTC 10041/BIG 03-04 MINDACT trial is feasible: First results of the pilot phase (Abstract 444). *European Journal of Cancer Supplements* 8(3): 188.

SACGHS (2008) *Realizing the Potential of Pharmacogenomics: Opportunities and Challenges: Report of the Secretary's Advisory Committee on Genetics, Health, and Society*. Bethesda, MD: Department of Health and Human Services.

Simon HA (1956) Rational choice and the structure of the environment. *Psychological Review* 63(2): 129-138.

- Simon R, Paik S, and Hayes DF (2009) Use of archived specimens in the evaluation of prognostic and predictive biomarkers. *Journal of the National Cancer Institute* 101: 1446-1452.
- Smart A (2006) A multi-dimensional model of clinical utility. *International Journal for Quality in Health Care* 18: 377-382.
- Sparano JA and Paik S (2008) Development of the 21-gene assay and its application in clinical practice and clinical trials. *Journal of Clinical Oncology* 26: 721-728.
- Sparano JA and Solin LJ (2010) Defining the clinical utility of gene expression assays in breast cancer: The intersection of science and art in clinical decision-making. *Journal of Clinical Oncology* 28: 1625-1627.
- Straehle C, Cardoso F, de Azambuja E, et al. (2009) Better translation from bench to bedside: Breakthroughs in the individualized treatment of cancer. *Critical Care Medicine* 37(1 Suppl): S22-S29.
- Teutsch SM, Bradley LA, Palomaki GE, et al. (2009) The evaluation of genomic applications in practice and prevention (EGAPP) initiative: Methods of the EGAPP working group. *Genetics in Medicine* 11: 3-14.
- Tuma RS (2005) Trial and error: Prognostic gene signature study design altered. *Journal of the National Cancer Institute* 97: 331-333.
- Van Luenen HGAM and Van Harten WH (2006) Quality management and stimulation of technology transfer in a research institute. *Creativity and Innovation Management* 15: 207-217.

Van Oosterom AT and Meunier F (1995) Reflections from Prof. Van Oosterom and Françoise Meunier about the core mission of the EORTC (20 October). Brussels: *EORTC Archives*.

Van't Veer LJ, Dai H, van de Vijver MJ, et al. (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415 (31 January): 530-536.

Wilkinson E (2009) Tailoring cancer therapy. Validating basic science with the 'supertrial'. *Molecular Oncology* 3: 389-391.

Zerhouni EA (2003) The NIH roadmap. *Science* 302 (3 October): 63-64, 72.

Zerhouni EA (2005) US biomedical research: Basic, translational, and clinical sciences. *JAMA* 294: 1352-1358.

### Biographical notes

**Nina Kohli-Laven** is currently a Postdoctoral Fellow in the Department of Social Studies of Medicine at McGill University and at the Institute for Health Policy Studies, at the University of California, San Francisco. Her research explores the production of emergent health technologies, combining anthropology, history, and science and technology studies. She has published on risk and narrative, looking at genetic diagnoses of orphan diseases. Her present work investigates the politics of breast cancer screening in the purview of new genomic technologies, looking at how regulatory and bioclinical infrastructures connect to styles of reasoning about risk and therapy.

**Address:** Department of Social Studies of Medicine, McGill University, 3647 Peel Street, Montreal, QC, Canada H3A 1X1; email: [nina.laven@mcgill.ca](mailto:nina.laven@mcgill.ca)

**Pascale Bourret** is Maître de Conférences (Associate Professor) at Aix-Marseille Université (Université de la Méditerranée) where she teaches sociology. She is also researcher at the INSERM-IRD-Université de la Méditerranée UMR SE4S unit (Economic and Social Sciences, Health Systems and Societies). She has published articles on bio-clinical collectives in the domain of BRCA testing. Her present work centers on the transformation of biomedical practices in connection to the development of genomic tools, with a focus on clinical work, the production of clinical judgment and clinical decision-making.

**Address:** UMR SE4S, INSERM, Institut Paoli-Calmettes, 232 boulevard Sainte-Marguerite, 13273 Marseille Cedex 9, France; email: [pascale.bourret@univmed.fr](mailto:pascale.bourret@univmed.fr)

**Peter Keating** is Professor of History at the Université du Québec à Montréal. His area of expertise combines the history of medicine, and the sociology of science and technology. His publications include *Biomedical Platforms: Realigning the Normal and the Pathological in Late-twentieth-century Medicine* (MIT Press, 2003) and *Exquisite Specificity: The Monoclonal Antibody Revolution* (Oxford University Press, 1995), both co-authored with Alberto Cambrosio. A new book (also co-authored with Alberto Cambrosio) entitled *Cancer On Trial: The Rise of Oncology as a New Style of Practice* will be published by the University of Chicago Press in early 2012.

**Address:** Department of History, University of Quebec at Montreal, Case Postale 8888, succursale centre-ville, Montreal, QC, Canada H3C 3P8; email: keating.peter@uqam.ca

**Alberto Cambrosio** is a Professor in the Departments of Social Studies of Medicine and of Sociology at McGill University. His work centers on biomedical practices and innovation, in particular at the clinical–laboratory interface, with a focus on the application of genomic technologies to the diagnosis and the therapy of cancer and the development of cancer clinical trials as a new style of practice. His publications include *Biomedical Platforms: Realigning the Normal and the Pathological in Late-twentieth-century Medicine* (MIT Press, 2003) and *Exquisite Specificity: The Monoclonal Antibody Revolution* (Oxford University Press, 1995), both co-authored with Peter Keating. A new book (also co-authored with Peter Keating) entitled *Cancer On Trial: The Rise of Oncology as a New Style of Practice* will be published by the University of Chicago Press in early 2012.

**Address:** Department of Social Studies of Medicine, McGill University, 3647 Peel Street, Montreal, QC, Canada H3A 1X1; email: alberto.cambrosio@mcgill.ca