

Debating the Desirability of New Biomedical Technologies: Lessons from the Introduction of Breast Cancer Screening in the Netherlands

Marianne Boenink

Published online: 16 April 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract Health technology assessment (HTA) was developed in the 1970s and 1980s to facilitate decision making on the desirability of new biomedical technologies. Since then, many of the standard tools and methods of HTA have been criticized for their implicit normativity. At the same time research into the character of technology in practice has motivated philosophers, sociologists and anthropologists to criticize the traditional view of technology as a neutral instrument designed to perform a specific function. Such research suggests that the tools and methods of more traditional forms of HTA are often inspired by an ‘instrumentalist’ conception of technology that does not fit the way technology actually works. This paper explores this hypothesis for a specific case: the assessments and deliberations leading to the introduction of breast cancer screening in the Netherlands. After reconstructing this history of HTA ‘in the making’ the stepwise model of HTA that emerged during the process is discussed. This model was rooted indeed in an instrumentalist conception of technology. However, a more detailed reconstruction of several episodes from this history reveals how the actors already experienced the inadequacy of some of the instrumentalist presuppositions. The historical case thus shows how an instrumentalist conception of technology may result in implicit normative effects. The paper concludes that an instrumentalist view of technology is not a good starting point for HTA and briefly suggests how the fit between HTA methods and the actual character of technology in practice might be improved.

Keywords Health technology assessment · Methods · Values · Ethics · Neutrality · Evidence · Mammography · Breast cancer screening · Instrumental view of technology

M. Boenink (✉)

Department of Philosophy, Faculty of Behavioural Sciences, University of Twente,
P.O. Box 217, 7500 AE Enschede, The Netherlands
e-mail: m.boenink@utwente.nl

Introduction

The rapid growth of technological possibilities in health care is accompanied by a growing need to determine an emerging technology's desirability before it is introduced in the health care system. In the 1970s and 1980s Health Technology Assessment (HTA) was developed as a way to facilitate collective decision making on the desirability of new biomedical technologies. The birth of HTA is often dated in 1975, when the American Congress asked its Office for Technology Assessment (OTA) which justification was necessary to introduce a new medical technology [4, p. 2; 6, p. 276]. The report, published in 1976, defined technology assessment quite broadly as: "a comprehensive form of policy research that examines the short- and long-term social consequences (e.g. societal, economic, ethical, legal) of the application or use of technology" [33, p. 45].

From the start HTA has had a complicated relationship with ethics and politics—or with normativity in general [6, 36, 38, 44]. Some advocates approach HTA as the first step in a two-tiered process (e.g. [40]). In this conception HTA researchers collect evidence about the probable consequences of a technology ('assessment'), which is subsequently weighed by ethicists or political actors ('appraisal'). Others argue, in contrast, that HTA-reports should pass judgment on the value of a technology itself (e.g. [58]).

Although not everyone agrees that HTA should include normative conclusions as to what is or is not desirable, many authors have criticized the idea that factual assessment and normative appraisal can be clearly separated. They point out that HTA methods to collect evidence on a technology's performance are inevitably imbued with values. Therefore, its results are not as neutral as often purported [21, 27, 30, 36, 38, 43, 44, 58]. For example, the selection of consequences to take into account, as well as the choice of specific outcome measures to assess them both have a normative impact. The interpretation of apparently technical concepts like 'efficacy', 'safety' and 'cost-effectiveness' rests on evaluative choices; the same goes for 'health,' 'disease', 'disability' or 'quality of life' [20, 21, 32, 38].

At about the same time philosophers, sociologists and anthropologists conceptually and empirically investigated the character of technology. They criticized the common idea that technology is a neutral instrument designed to perform a specific function. Some argued that technology is a complex phenomenon, consisting of material, social and organizational elements [1, 26]. Others pointed out that, because of this complexity, technology does much more than it is intended to do [5, 31]. Many of these effects are indirect: technology mediates our experience of reality and the options for action we perceive [22, 62]. Such analyses imply that the effects of a novel technology are much more difficult to foresee than the instrumentalist view of technology suggests.

This conceptual and empirical work on the character of technology in practice seems to suggest that tools and methods of traditional forms of HTA may be problematic partly because they are based on an instrumentalist conception of technology that fits badly with the workings of technology in practice. In this paper, I will explore this hypothesis by reconstructing and discussing an early endeavor to systematically assess a technology's desirability: the assessments and deliberations

preceding the introduction of breast cancer screening in the Netherlands in the 1970s and 1980s. Although these activities were not yet labelled as ‘HTA’, the case is interesting precisely because it shows the ideal of HTA ‘in the making’ and set the stage for subsequent HTA activities.¹ In this period a model of HTA emerged that was widely used, both in the Netherlands and abroad, for many years (as is visible, for example, in the special issue on the history of HTA in the *International Journal of Technology Assessment in Health Care* 25, Supplement 1, 2009).

First I will reconstruct the processes preceding the decision to introduce a nation wide program for breast cancer screening in the Netherlands. This reconstruction is based on a systematic survey of: (1) the relevant advisory reports and policy documents in the years 1968–2005; (2) the minutes of Dutch Parliament in the years 1970–1995; (3) the medical professional journals *Nederlands Tijdschrift voor Geneeskunde* (1960–2005) and *Medisch Contact* (1960–2005); and (4) opinion articles on breast cancer screening in the Dutch national newspapers (1960–2005). Subsequently I will reconstruct the stepwise model of HTA that emerged during the process and analyze how this model was based on an instrumentalist conception of technology. I will proceed then with a more detailed discussion of three episodes. These episodes illustrate how the actors during the process were already confronted with or became aware of unanticipated effects, and how this in some cases led to adjustments in the HTA procedures. The historical analysis enables me, in the end, to argue how and why an instrumentalist conception of technology is not a good starting point for HTA. I will conclude with some suggestions for improving the fit between HTA methods and the actual character of technology in practice.

Debating the Desirability of Breast Cancer Screening: HTA in the Making

The history of breast cancer screening starts in the 1950’s and 1960’s, when in France and the USA a radiological device for diagnosing breast diseases was developed [19]. It was noticed soon that this device, called mammography, detected lesions that were not found by palpation (the current diagnostic standard). Moreover, these lesions were often smaller than those detected by palpation. Mammography thus seemed to offer promising possibilities for earlier diagnosis of breast cancer, which in turn might improve curability and survival rates.

When an American study showed that mammographic screening might reduce breast cancer mortality by 30% concerned scientists in the Netherlands set up several pilot projects to corroborate the effectiveness of mammographic screening in a Dutch setting. In 1975 projects started in Nijmegen, Utrecht and Leiden. The Leiden project was meant as a pilot project offering a novel health care service and lasted only 1 year. The projects in Nijmegen and Utrecht had scientific goals and were designed as case control studies.

In 1977 the Dutch Minister of Health asked the Health Council to report on the *desirability* and *acceptability* of a national breast cancer screening program. The

¹ Banta and Oortwijn [3], writing the history of HTA in the Netherlands, identify the first Dutch HTA reports as of 1988 and 1989, but they focus on the reports and activities explicitly labeled as HTA.

committee established by the Health Council consisted of a number of medical professionals and a sociologist; a psychologist was added later on. The committee soon concluded that the available evidence on the effectiveness of mammographic screening was limited and decided to wait for the results of the pilot projects in Utrecht and Nijmegen before formulating its advice. Since the Dutch projects were case control studies and thus liable to bias, results from randomised clinical trials (RCT's) in other countries were taken into account as well. The Health Council's committee issued two interim reports and a final report [15–17].

When the second interim report cautiously judged the available evidence for effectiveness of screening as promising, the Minister of Health ordered an independent study of the *cost effectiveness* of such a screening program—actually the first such a study to be ordered by Dutch government. This study, written by the Institute for health Technology Assessment of Erasmus University, was published in 1990 [9]. The study not only calculated the costs per saved life year; it also used QALY's (Quality Adjusted Life Years), a relatively new tool to take into account effects of screening on the screened individual's wellbeing.

To sum up, although the label 'HTA' was not used yet, mammographic breast cancer screening was the first health care technology in the Netherlands to be systematically assessed for its *efficacy, safety, cost effectiveness* and *impact on quality of life*, with the explicit aim to inform and facilitate political decision making. When the experts judged that the performance of mammographic screening on these items was acceptable, the Minister proposed to gradually introduce breast cancer screening for women aged 50–69 on a national scale, as the Health Council had advised. The Dutch Health Insurance Council was asked for advice on the financial and organizational aspects of such a screening program. It proposed to make screening free for participants and to organize it in a centralized way, closely resembling the set up of the pilot projects [63]. These suggestions were accepted by the Minister.

The desirability of the screening program was not only discussed in scientific reports, however. It was also the subject of public and political debate. Both during the preparatory phase and immediately after the Health Council published its final report, a small group of opponents voiced criticism. Most of these critics had a background in social science or social medicine. They wrote pieces in newspapers and medical journals arguing that a screening program would seriously contribute to medicalization [8, 41, 45–50, 61]. In response, others (often investigators of the pilot projects or members of the Health Council's committee) stressed the advantages of breast cancer screening and the evidence for the positive effects of such a program [10, 12, 37, 54].

The first extensive parliamentary debate took place in 1991, when the Parliamentary Commission on Health Care discussed the proposed national breast cancer screening program with the Minister of Health. Although no one completely opposed the screening program, several members of parliament took up criticism voiced earlier in newspapers and journals. Some suggested that if women were really to decide autonomously about their participation, the invitational letter should explicitly mention the uncertainties and potentially negative effects of screening. Others questioned the justification of the proposed age limits. And some argued that

the program might lead to overdiagnosis, overtreatment, and thus to unnecessary medicalization [51].

In his response, the Minister first pointed out that many of the potential negative effects of screening had been taken into account in the cost effectiveness study, in particular by the use of QALY's [52]. In a second debate between the Minister and the Parliamentary Commission on Health Care, in 1993, the Minister recognized, however, that there might be a tension between voluntary participation and the need for high participation rates. He promised to guarantee (1) respect for the freedom of choice and personal life of the women involved (by giving even handed information); (2) equal access (again by giving good information and by making participation free); and (3) a high quality of care (by installing an extensive monitoring system) [53, pp. 13–14, 18].

Thus, between 1989 and 1995 breast cancer screening was gradually introduced in all regions of the Netherlands. Since 1995, all Dutch women between 50 and 69 have been invited for breast cancer screening every 2 years. Effectiveness, efficiency and quality of care of the screening program continued to be monitored by an independent scientific committee. Participation rates (and thus apparent acceptance by the target group) slowly increased from about 75% in 1990 to 82% in 2007 [25].

Since then, public and political debate on breast cancer screening rekindled on two occasions. The first debate took place when practice with regard to the upper age limit was changed. Although the upper age limit was set at 69, initially participants reaching the age of 70 could continue screening if they liked to. In 1992 the Health Insurance Council announced that from now on the age limit would be strictly kept. This decision was justified mainly with budgetary reasons, but doubts about effectiveness and worries about the impact of screening on older women's lives were also mentioned [56, p. 139]. The decision initiated a heated debate, in which representatives of associations for the elderly claimed a *right* to be screened and accused the Council of age discrimination and paternalism. Medical professionals supported their claims (e.g. [55]). In the end, the upper age limit was increased to 75 years in 1998 when new evidence suggested that screening this age group could be effective after all [57].

A second surge of debate occurred when the Cochrane Collaboration in 2001 published a critical metareview of the available evidence for the effectiveness of breast cancer screening. The authors concluded that this effectiveness was doubtful and that screening might even produce harm, in terms of overdiagnosis and overtreatment [34, 35]. This led to a renewed public exchange of arguments pro and contra screening, in which the strengths and weaknesses of the scientific methods to establish effectiveness were also subjected to debate. Some critics argued once more that the information given to women when invited for screening was too biased to enable autonomous decision making [59]. The debate subsided after a Health Council committee, hastily asked by the Minister of Health to advise whether the Dutch program should be adjusted or even abolished in view of the Cochrane review, concluded that there was no need to do so. Effectiveness might be somewhat lower than previously expected, but it was sufficient to justify a national screening programme [18].

Emerging HTA and Its Instrumentalist Presuppositions

In effect, then, mammographic breast cancer screening was the first health care technology in the Netherlands to be subjected to a more or less systematic assessment, with the aim to inform and facilitate political decision making. In the process a model of HTA emerged that now seems quite familiar and self-evident because since then it has been used widely, both in the Netherlands and in other countries. What is interesting about the historical case discussed here is that it shows the gradual ‘birth’ of this widespread model in the Dutch context.² In reconstruction, the process consisted of four stages, each considering a specific aspect of mammographic screening:

1. Health gains
2. Cost-effectiveness
3. Organizational conditions
4. Ethical considerations (like justice, autonomy, screening rights, medicalization)

The first three stages occurred in a chronological chain: the next stage was initiated only when the results of the preceding one appeared to be sufficiently positive. They were also linked to a clear division of roles and responsibilities. The Ministry of Health delegated the first three stages to scientific experts: the Health Council, independent economic scientists and the Health Insurance Council respectively. In contrast, the ethical considerations were (implicitly) left to public and political debate. They occasionally popped up during the process, with a clear increase in the final stages of decision making.

This procedure seems to have been inspired by an implicit instrumentalist view of technology in several respects. Such a view perceives technology as a *material device* that is *designed to perform a specific function*. It is an instrument (‘just a means’) to realize a specific goal; the instrument itself is value neutral. This view has been widespread in Western thinking [14, 62], so it is not surprising that it influenced ideas and tools for HTA. What is the task of (H)TA from an instrumentalist perspective? If technology is a device designed to fulfil a specific function, the first question is of course whether the proposed device is indeed capable to fulfil this function. Assessments should ask, then, to what extent the device realizes the intended effect. The goal (the intended effect) itself is supposed to be clear and is taken for granted for the time being, as is the device’s actual design.

Seen this way, it makes sense to separate the factual assessment of a device’s performance from the normative evaluation of the intended goal. Such a separation is clearly implied by the split between stages 1–3 and stage 4 in the mammographic screening case. Whereas establishing the facts was considered a task for scientific experts, evaluating goals was assigned to the political domain. The experts involved in the mammographic screening research endorsed this view. In its final report, the

² This does not mean, of course, that the Dutch were the only ones, or even the first to ‘invent’ this model. Nor does it mean that other countries imported the Dutch model. In all likelihood, the concept of technology assessment in health care was picked up by several countries at the same time, each importing insights from elsewhere and giving it a national flavor at the same time.

Health Council deliberately refrained from statements about the desirability of a national screening program, because this was considered a *political* issue [17, cover letter p. 2]. In a similar vein, the authors of the cost-effectiveness study stated that considerations other than cost-effectiveness could “of course” play a large role in the political deliberations [9, p. 7].

The instrumentalist view is also clearly visible in the set up of the first assessment stage. Mammography was approached as a material device designed to diagnose women with asymptomatic breast cancer, in order to increase the life expectancy of these women. To decide about the desirability of this device, the uncertainty with regard to its performance had to be reduced. Did it diagnose sufficient asymptomatic cancers, and did this early diagnosis lead to increased survival? The pilot projects and the report of the Health Council thus initially were set up to produce factual evidence on the actual magnitude of the *intended* effect only.³ The only additional effect that was taken into account was the potential causation of new tumours resulting from the radiation necessary to make the mammograms—a side effect, that is, that would directly detract from the intended effect. In the second and third stage of the assessment the financial and organizational efforts needed to realize the intended effect were charted, to enable an informed comparison with potential alternatives.

The HTA procedure developed in this historical episode thus was driven by instrumentalist presuppositions. This is visible in particular in (1) the separation of assessment of facts by experts from the deliberation on desirability by political actors and (2) the exclusive focus on the intended effect (i.e. mortality reduction). However, a closer look at the historical episode shows that the actors involved also encountered problems resulting from limitations and shortcomings of the instrumentalist view. Some of these problems were diagnosed already by the actors themselves, resulting in modifications of the assessment procedure or other attempts to compensate for the procedure’s weaknesses. Other problems can be identified with the benefit of hindsight, suggesting a need for further modification. Thus, while trying to invent a systematic procedure to assess mammographic screening, the actors involved learned important lessons about the lack of fit between an instrumentalist view of technology and the way technology works in practice. In the next part of this paper, I point out how difficulties arose with respect to three presuppositions inspired by an instrumentalist view of technology:

- (1) Technology is a means to realize intended effects;
- (2) Technology is a material device;
- (3) Morality may influence technology, but technology does not influence morality.

³ Here, an interesting difference between TA and HTA can be observed. Whereas HTA usually focuses at intended outcomes, TA from the start has tried to include unintended side effects, like environmental damage. Part of the explanation may be that most HTA is driven by financial motives and thus has a more narrow scope (the need to select the most cost-efficient technologies). In addition, the side effects of medical technology that are visible and feared most are the ones that directly counter the intended goal by causing infirmity or death. Such side effects can be assessed using the outcome measure that is also used to measure the intended effect.

Lesson 1: Technology Does More Than It is Intended To Do

As described above, the first publications on mammography present it as a tool to *improve the prognosis* of women with asymptomatic breast cancer. The actors involved in the assessment process are aware, however, that measuring improved prognosis is rather complicated. Counting cases of early detection does not suffice; the screening program is successful only if subsequent treatment is effective and results in prolonged healthy lives. Therefore the outcome measures proposed at first are 10 years—survival and mortality rates. However, because an increase of 10 years survival rates without a concurrent mortality reduction is considered undesirable, reduction of breast cancer mortality quickly becomes the central measure of effectiveness. The pilot projects in Utrecht and Nijmegen are explicitly set up to investigate to what extent mammographic screening reduces breast cancer mortality. The Health Council in its first report (1981) affirms this choice: this is the intended effect of mammographic screening and it should be the central focus of the assessment procedure.

The actors gradually become aware, however, that mammographic screening does more than it was intended to do and that these additional effects may also be relevant when assessing the technology. The Health Council's 1984 report, for example, mentions 'quality of life' as an additional outcome measure for the effectiveness of mammographic screening:

Therapeutic effectiveness in the above is translated as reduced breast cancer mortality. However, this is not the only measure of effectiveness. One could also think of an improved quality of life. It is important to point at the so called breast saving surgery techniques that have gained ground in the treatment of early stage breast cancer lately. [16, p. 25]

This development of breast saving surgery techniques was actually spurred by the increasing number of 'pre-clinical' tumors becoming available as a result of screening [13, p. 95]. The novel surgery techniques were thus co-produced by mammographic screening, resulting in an unforeseen (and up to then unintended) effect. Decreased mutilation and increased wellbeing of patients now emerged as a potential additional goal of breast cancer screening, next to mortality reduction.

Thus, the assessors' initial focus on the intended effect only started to look too narrow. The concept of quality of life was proposed to broaden the scope of the assessment, to allow different types of presumably positive effects of screening to become visible. Although such effects were not explicitly intended at the start of the projects, the Health Council appears to have been willing to include this novel measure in the assessment. However, the investigators of the two Dutch pilot projects did not include 'quality of life' as an additional measure of effectiveness, presumably because they did not want to change the design of their long term studies.

It is only in the cost-effectiveness study (1990), therefore, that quality of life first appears as a measure of effectiveness. The authors of this report motivate the introduction of the concept in a way that is very different from the Health Council's earlier suggestion:

... next to mortality reduction, inevitably undesirable effects will occur. An undesirable effect is, for example, anxiety in women who are referred for further diagnostics and who in the end appear not to have breast cancer (the false positives (...)). Another undesirable effect is that a number of women now know they have breast cancer whereas for some of them the prognosis won't improve. (...) Weighing the amount of desired and undesired effects on the one hand and the expected costs and reduced expenses on the other is necessary. [9, p. 6]

Here 'quality of life' is not meant to measure additional positive effects of screening; on the contrary, it is a measure for any impact (both desirable and undesirable) screening might have on the lives of individual participants.

Thus, the actors involved in the assessment of breast cancer screening gradually realized that mammographic screening may have significant impacts besides the intended increase in life expectancy, and that these effects, moreover, need not always be positive. Voices from the public debate (on breast cancer screening, but also on other medical technologies) may have played a role here. The introduction of 'quality of life' was an attempt to accommodate a broader array of effects, in a way that apparently enabled a sensible comparison of these effects. This history corroborates the argument made by Armstrong and Caldwell, who claim that the emergence of 'quality of life' in the 1970s and 1980s among others expressed a general discomfort regarding the human costs of technological interventions [2, p. 36]. In the Dutch case, the concept helped to weigh the various effects of mammographic screening. The desire to include patients' perspectives (often mentioned as a justification to include quality of life) apparently hardly played a role.

It is clear from the cost-effectiveness report how the researchers struggled with the operationalization of 'quality of life' [9, pp. 113–119]. What should or should not be included? How to operationalize the more subjective types of impacts? And how to compare different types of impacts? Interestingly, the questionnaires to assess quality of life impacts were not filled out by women participating in the pilot projects, because the questions were considered too complex for these participants. Instead, 'stand ins' like economists and medical professionals involved in cancer screening and care acted as respondents. The first were considered to be more used to this type of questionnaire, and the latter were supposed to have a reliable view of how breast cancer may affect women's lives. Since then, struggles have continued, leading to a proliferation of instruments measuring quality of life, most of which do involve patients (or healthy subjects) as respondents. Quality of life, in addition to mortality reduction, has become a widely accepted dimension for assessing the impact of novel medical technologies. However, controversy over the tools to measure it persists.

Notwithstanding these difficulties, the history of assessing breast cancer screening shows why HTA should not focus on a technology's *intended* effect only. Technology does more than it is intended to do, both directly and indirectly. As for the direct effects, screening may increase one's life expectancy and decrease wellbeing at the same time. Focusing exclusively on the intended (hoped for) effects

inevitably biases the assessment and leads to (avoidable) implicit normativity of the results. Including a broad array of potential effects is important for even handed decision making on the desirability of a novel technology. Unintended effects are, of course, difficult to foresee. This is even truer for indirect effects, like the improved surgery techniques resulting from the emergence of early diagnostics. Good HTA should nonetheless try to anticipate and include such effects, both at the start and on the way, as the Dutch assessors tried to do.

Lesson 2: Technology is More Than a Material Device

As described above, the reduction of breast cancer mortality by mammographic screening not only requires a well functioning mammographic device, but also an effective therapeutic follow up. Breast cancer screening has, then, clear social and organizational conditions from the start. However, at first both the researchers and the political actors involved approach mammographic screening as a material device that processes input (asymptomatic women) to deliver a specific output (mortality reduction). The pilot projects are presented as a means to find out how effective this new diagnostic technology will be when applied to asymptomatic women. Practical conditions like setting age limits and frequency of screening do get some attention, but only in relation to the question how to increase the screening's effectiveness. It is interesting to see how the actors in the Dutch screening assessment seem to have become gradually aware of the importance of social and organizational conditions; an awareness that in the end led to a strict regulation of the national screening program and even to a law on population screening.

When the results of the pilot projects seem sufficiently promising to warrant serious consideration of a nation-wide screening program, the researchers start pointing out that the organizational setting of screening is crucial. The positive results can be repeated only, they state, when the national program is conducted in a way similar to the pilot projects. The Health Council agrees. In its second report it states:

The committee emphasizes that a potential screening program will succeed or fail depending on the way the introduction and design are organized. (...) Guarantees are needed for the quality of the work of participating centers (diagnostics, treatment and counseling of the women involved). In this respect it is important that strict protocols are kept. A newly instituted center of reference might serve quality control as well. [16, p. 28]

Both the scientific and the political actors thus gradually became aware that mammography is more than a material device and that breast cancer screening requires more than the distribution of mammography devices. As a result, they concluded that a national screening program should be as similar to the pilot projects as possible. The awareness that organizational conditions are crucial in realizing a certain effect motivated the Health Council's advice to centralize the organization of screening and make it a uniform practice. Moreover, since the pilot projects were designed as a separate, autonomous service, situated between GP's

and hospitals, the national program also became a separate service next to primary and secondary health care. Thus, because the sociotechnical character of mammographic screening came into the picture relatively late, the organizational choices of the scientific pilot project became indicative for the organization of the national program. As one critic stated, when the committee developed a refined constellation of protocols and procedures for national screening practice, it was actually “trying to approach the conditions of scientific research” [8, pp. 475–476]. This blocked serious consideration of alternative options. Even the question whether or not all actors agreed with the proposed division of roles and responsibilities was hardly discussed.

The growing awareness of the social conditions for realizing the aimed for mortality reduction also influenced the approach of and the role ascribed to female participants. In its final report the Health Council explicitly mentions a high participation rate as a condition for attaining the aimed for effect: “The committee emphasizes the importance of a maximal participation in screening for breast cancer. If the participation rate is too low, the aimed for mortality reduction on the population level may also become too low” [17, pp. 39–40, see also 50–54]. Recommendations were formulated to make screening as accessible as possible: the invitational system should function well, screening should be affordable, the location of screening centers should be accessible, counseling should be available, and education should be as persuasive as possible: “Careful education of the female population is necessary. Research into the effectiveness of education with regard to the participation of the women involved in screening is necessary” [17, p. 72]. Here, the need for a high participation rate to achieve the aimed for effects actually led to a neglect of a potentially conflicting value: the autonomy of participants.

The implicit tension between effectiveness and women’s autonomy came to the fore once more in 1991, when the Minister of Health in a letter to the Health Insurance Council in 1991 simply states that invitations, leaflets and campaigning material should contribute to high participation rates as well as enable considered choices by the women in the target group [51, p. 11]. In the parliamentary debates on screening, however, MP’s stressed that autonomy was the most important value. This position in the end was confirmed by the Minister in a memo from 1993, stating that information on screening should include negative aspects of screening, even if this would lead to a decrease of participation rates [53, pp. 13–14].

This episode shows how the implicit view that technology is a material device led the assessors of mammographic screening to neglect the social and organizational elements of this technological practice for a long time. In the end they did stress the importance of social and organizational conditions, but because the outcomes of the assessment by then were quite promising, this did not result in discussion on the pros and cons of the specific design of the screening practice in the pilot studies. On the contrary, the practice investigated in the pilot projects now was put forward as the most desirable. Moreover, these conditions were approached only in terms of effectiveness; issues of feasibility and desirability did not come to the fore.

These episodes show once more how the view that technology assessment is about the intended effects of a material device may lead to implicit normative effects of the assessment procedure, in this case with regard to the social and

organizational conditions. It tends to limit or even close off the space to discuss the pros and cons of alternative designs of the practice at hand. Moreover, it may lead to roles for both professionals and lay users that are not acceptable to them or that severely diminish the realization of alternative values (like autonomy). Conceiving novel technology in terms of a sociotechnical practice might help to avoid this implicit normativity by including such social and organizational aspects in the assessment procedures right from the start.

Lesson 3: Technology Does Affect Morality

The third and last presupposition of the assessment procedure for mammographic screening I want to discuss here is that the development of mammographic screening itself would not affect morality. As pointed out above, the instrumental view of technology implies that humans can first set goals and then decide whatever technology serves their goals best. In the case of mammographic screening the starting point was that health is an important value to many people. The assessment, then, need not focus at the goal aimed for, but only at the question to what extent screening would realize this goal. Such reasoning presupposes, however, that the meaning of ‘health’ is clear, widely accepted and stable. The history of the HTA process and the resulting screening practice show, however, that both the meaning of and the weight assigned to ‘health’, changed *as a result of* the technology development. Since this is a long term process, it is easier to discern with the benefit of hindsight. The actors involved in the assessment touched on these issues only incidentally.

First, the meaning of the value ‘health’ seems to have shifted. Interestingly, some actors involved in the process actually anticipated this shift. One of the first Dutch professional publications on early detection of cancer pointed out that ‘health’ was usually conceived of as *not* having to notice your body. The authors explicitly asked:

Is it desirable indeed to ask people’s attention for their own body as a source of disease time and again? Or would a healthy life rather consist in living without having to worry about one’s breasts, intestinal tract or stomach?
[24, p. 67]

The issue was both conceptual (a shift might occur in the way ‘health’ is generally interpreted) and psychological (this conceptual shift might have impact on people’s lived experience). Typically, the authors of this report declared they were not competent to address such questions. They did not indicate who would be either.

Not surprisingly, then, the issue was hardly taken into consideration in the subsequent assessment and deliberation processes. As discussed earlier the scientists first defined health exclusively in terms of survival and later also in terms of quality of life. Considerations on the impact of screening on views and experiences of health incidentally reappeared in newspaper contributions arguing against screening. They were also briefly mentioned in the parliamentary debates. However, since empirical evidence for such ‘soft impacts’ [7] was absent (and hard to investigate

anyway), they were easily dismissed in view of the positive evidence for the effectiveness of screening.

In retrospect it seems nonetheless justified to conclude that the introduction of breast cancer screening, albeit in conjunction with many other technologies for early detection of disease, did contribute to a new understanding of what it means to be healthy or diseased. The long standing view that ‘health’ means absence of complaints has been exchanged for the view that disease can be a-symptomatic. As a result, most people now probably accept the view that they can have a disease without noticing anything. This is actually a generalization of the shift that mammography brought about earlier in the medical domain. The starting point of the debates on screening was, after all, the observation that this technology enabled doctors to identify lesions in women who had up to then seemed perfectly healthy, both to themselves and to their doctors.

This novel view of disease and health has brought along new norms as well. On the one hand, the new technological possibilities have led to new *obligations*. Now that it is possible to diagnose a-symptomatic disease, people are expected to subject themselves to these technologies. Even though Dutch women are not legally required to participate in the mammographic screening program, those who do not participate often have to justify their choice. They may even be accused of irresponsible behavior, especially if they are diagnosed with breast cancer later on. Participation in screening, that is, has become a social norm.

On the other hand, novel *rights* have been claimed. At the introduction of a nation wide program for breast cancer screening the age limits of the target group were set as 50 and 69 years. Evidence for the effectiveness of screening older women was thought to be lacking (another example of how the design of the pilot projects became normative for the national program). As described above, in 1992 the Health Insurance Council (responsible for the funding of the national program) decided to keep a strict hand at the upper age limit, partly motivated by overspending. The Council was confronted with an indignant public response from elderly women as well as medical professionals. They argued that women had a *right* to be screened, and that denying this right to a subgroup of women on the basis of their age amounted to age discrimination [55]. Apparently women above 70 who had been screened for some years felt that they were denied something valuable.

Of course one may question the legitimacy of this claim. However, just the fact that it was actually claimed shows that a substantial group of people had come to perceive screening as a ‘good’ that women are entitled to and that should be distributed justly. Thus, the screening providers and politicians were confronted with moral considerations they had not anticipated when assessing and deliberating the desirability of mammographic screening. As a result they were at a loss: how to deal with this novel argument?

Interestingly, the issue was approached in a way resembling the original assessment process: judgment on desirability was postponed pending the production of evidence on the effectiveness of screening older women. Novel research was started; in the meantime the age limits were strictly kept. In 1998 a Dutch case control study claimed to have shown that reduction of mortality for 70–75 years olds was possible and that screening could significantly improve the quality of life

of this group. Some actors argued that this evidence was not sufficiently conclusive, since it was not based on RCT's and thus could be biased. However, the political decision makers thought it would be unacceptable to let older women wait any longer for the outcomes of such (usually very long term) studies [57, p. 1098]. In this respect, the claim that a moral right was at stake mitigated the standard of scientific rigor used in the political decision making.

In sum: although the initial assessment procedure was predicated on the assumption that mammographic screening would not affect the value of health, both common views of health and disease, as well as related obligations and rights were affected by the development and introduction of mammographic screening. To be sure, mammography was not the only source of these changes. The concurrent emergence of many other technologies for early diagnosis definitely played a role as well. Overall, this historical episode illustrates how technology and morality mutually interact. Morality may steer technology development, but technology also influences morality.

If so, HTA had better acknowledge this mutual interaction. Partly for practical reasons: if the actors involved reflect on the potential impact of technology on morality, they will be confronted with unanticipated changes in morality less often. Another reason is that such changes may affect judgment on the desirability of a technology. Since it is usually very difficult to withdraw or prohibit a technology already in use (as the controversy on age limits shows), it makes sense to try to anticipate an emerging technology's long term impacts. This presupposes, however, that HTA discards the assumption that morality is stable and independent from technology.

Tailoring HTA Methodology to Technology: Some Suggestions

To sum up: the actors involved in the assessment of mammographic screening in the Netherlands not only developed a general procedure for HTA. In the process of assessing mammography, they were also confronted with several limitations and weaknesses of this procedure. With the benefit of hindsight we can trace even more. These limitations and weaknesses are related to instrumentalist presuppositions implicit in the assessment procedure. My reconstruction of the case displays how at various moments a misfit emerged between the instrumentalist presuppositions of the procedures and the way technology actually works. These misfits produced implicit normative effects because important considerations were not or only marginally taken into account. This significantly reduces the moral and political legitimacy of the judgment on the desirability of mammographic screening—even though mammographic screening since its introduction has been a widely accepted technological practice in the Netherlands.

The historical case is interesting because it reminds us that HTA procedures and tools should fit with the way technology works. This may sound self-evident, but it is easier said than done. Too many assessments today still resemble the procedure emerging in this early example. How to tailor HTA methodology to technology, then? In this section I will summarise the lessons for HTA that can be drawn from

the case of mammographic screening, while briefly suggesting which tools or measures might help to avoid the weaknesses from the past.

First, technology does more than it is intended to do: it has a *broad array of effects*. Many of them are hard to measure in a quantitative way. Moreover, complex interactions may occur. As a result, many effects are difficult to foresee. It would be naïve to suggest that all effects can be anticipated and included in HTA of emerging technologies. However, one might at least attempt to be as inclusive as possible. A first step is to engage users, or more generally stakeholders, early in the assessment process [23, 39, 60]. Including actors with different backgrounds in the assessment process will help to chart and imagine the wide variety of effects a technology might have. This helps to reduce unpleasant surprises later on. It also contributes to the democratic legitimacy of the deliberation process.

To avoid foreclosing or steering of such an inventory, it may be advisable to start with explorative, qualitative research [28, 29]. This would allow stakeholders to respond as much as possible from their own perspective and in their own words. Stakeholders could also be asked to participate in the selection of effects to be measured and to comment on proposed outcome measures, to ensure these do justice to the rich meaning of anticipated effects. Scientific research all too often results in knowledge that only partially covers stakeholders' considerations. Quality of life measurement is a case in point.

Furthermore, particular attention needs to be paid to long term and so called 'soft' impacts.⁴ Technology does not only affect our health, environment or safety. It also influences our (to list just a few) health care practices, roles, responsibilities and identities. Because such effects usually emerge only gradually and do not lend themselves for quantitative research, they are easily neglected [7]. Even stakeholders may not think of them. To remedy this, historical and comparative analysis could be used as a source of imagination. Earlier examples, ongoing trends and patterns can help to make us more aware of the soft impacts. In addition, they can be used to enhance our imagination of what might happen in the future (for an example see [7]). Historical and sociological insights can be incorporated, for example, in scenarios of possible futures. These may in turn serve to feed stakeholders with rich material for their deliberations.

The second lesson from the case discussed above, partly related to first one, is that technology should not be approached as a device, but as a *sociotechnical practice*. This means that the social and organizational conditions for making the material device work have to be included right from the start of the assessment [39]. The aim here is not to assess how effective they are with regard to the intended outcome, but to explore additional impacts, for example on the division of roles and responsibilities of the actors involved. To avoid normative effects resulting from the experimental design, it might help to experiment with alternative sociotechnical designs at the same time (for a similar suggestion see [11]). This would enable comparisons and thus result in informed debate on the acceptability and desirability of proposed practices.

⁴ The term 'soft impact' has been coined by Tsjalling Swierstra, see <http://www.onderzoekinformatie.nl/nl/oi/nod/onderzoek/OND1337080/>.

Finally, the third lesson is that morality is not a stable phenomenon; it is evolving and technology is an important driver of this evolution. When assessing and debating the desirability of an emerging technology one should, then, *take into account the mutual interaction of morality and technology* and try to anticipate such effects. A tool to support this is, again, historical analysis, which could then be used to construct plausible scenarios that serve as input for public and political debate [42].

As I indicated above, all this will not suffice to foresee and predict a new technology's impact. Technology is a complex phenomenon and so is society, and the confrontation of the two is bound to produce events and effects no one would have been able to imagine in advance. If this is true, it makes sense to construct assessment procedures as flexible as possible. It should be possible to incorporate unexpected effects or adjust outcome measures later on.

Overall, these lessons and the suggested tools aim to broaden the number and type of considerations that can be put forward in deliberation and debate on the desirability of novel biomedical technologies. The underlying thought is that this both improves the quality of the assessment (since fewer aspects will be neglected) and enhances the democratic legitimacy of the process (since all views can be seriously considered). I have hardly paid attention to the division of roles and responsibilities in the deliberation and decision making processes. Much more should be said, for example, about the relation between scientific experts, clinicians, lay people and politicians, but this will have to wait for another occasion.

Conclusion

The tools, methods and procedures for HTA are the evolving products of historical learning processes. The first attempts to perform HTA in the Netherlands started with the endeavour to provide decision makers with scientific evidence on the effectiveness of mammographic screening. The resulting procedure was to a large extent informed by an instrumentalist view of technology. In the process, the actors performing and using the results of the assessments were confronted with the weaknesses and limitations of the instrumentalist view. Technology in practice acted rather differently from what was anticipated. The case shows how the actors involved partially responded to these surprises by adjusting their assessment methods. They learned by doing, inventing and developing HTA along the way.

Today, HTA can pride itself on quite some history in many countries. As a result those performing HTA need not start from scratch. However, the history of HTA is not only a source of standard models, tools and methods. It is also a source of experiences in the workings of technology in practice. HTA could take advantage of these latter experiences more often, to improve the fit between methodology and technology. Improving this fit will ultimately contribute to the social and moral legitimacy of HTA results.

Acknowledgments The historical research for this paper was carried out as part of the project 'Developing scenarios of moral controversies concerning new (biomedical) technologies', funded by the

Dutch Organization for Scientific Research (NWO) in its program Ethics, Research and Governance, for which my sincere thanks. My colleagues in this project: Tsjalling Swierstra, Dirk Stemerding, Margo Trappenburg and Hester van de Bovenkamp commented on an early version of this paper. Annemarie Mol provided me with feedback that was crucial for arriving at my current take on the historical material. Many colleagues at the Department of Philosophy of the University of Twente, in particular the participants of the research group Ethics and Politics of Emerging Technologies, commented on subsequent versions of the paper. Federica Lucivero and Tsjalling Swierstra kept encouraging me to finish it. Thanks to all of them for their support and feedback.

Open Access This article is distributed under the terms of the Creative Commons Attribution Non-commercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Akrich, M. (1992). The de-scription of technical objects. In W. E. Bijker & J. Law (Eds.), *Shaping technology/building society* (pp. 205–224). Cambridge, MA: MIT Press.
2. Armstrong, D., & Caldwell, D. (2004). Origins of the concept of quality of life in health care: A rhetorical solution to a political problem. *Social Theory & Health*, 2, 361–371.
3. Banta, D. H., & Oortwijn, W. (2001). Health technology assessment and screening in the Netherlands. *International Journal of Technology Assessment in Health Care*, 17(3), 369–379.
4. Banta, D., Kristensen, F. B., & Jonsson, E. (2009). A history of health technology assessment at the European level. *International Journal of Technology Assessment in Health Care*, 25, 68–73.
5. Berg, M., & Mol, A. (Eds.). (2001). *Ingebouwde normen. Medische technieken doorgelicht*. Utrecht: Van der Wees.
6. Blume, S. S. (2009). Assessing health technologies in a changing world. *International Journal of Technology Assessment in Health Care*, 25, 276–280.
7. Boenink, M., Swierstra, T., & Stemerding, D. (2010). Anticipating the interaction between technology and morality: A techno-ethical scenario study of experimenting with humans in bionanotechnology. *Studies in Ethics, Law and Technology*, 4(2), 1–38.
8. Coebergh, J. W., Crommelin, M. A., & Kluck, H. M. (1990). Bevolkingsonderzoek naar borstkanker—Een pleidooi voor behoedzaamheid. *Medisch Contact*, 45(15), 475–477.
9. de Koning, H. J. (1990). *De kosten en effecten van bevolkingsonderzoek naar borstkanker*. Rotterdam: Instituut Maatschappelijke Gezondheidszorg, Erasmus Universiteit Rotterdam.
10. de Koning, H. J., Boer, R., van der Maas, P. J., van Ineveld, B. M., Collette, H. J. A., & Hendriks, J. H. C. L. (1990). Effectiviteit van bevolkingsonderzoek naar borstkanker; sterftereductie in binnen- en buitenland. *Nederlands Tijdschrift voor Geneeskunde*, 134(46), 2240–2245.
11. de Vries, G. H., & Horstman, K. (Eds.). (2008). *Genetics from laboratory to society*. Houndmills Basingstoke: Palgrave/Macmillan.
12. de Waard, F. (1978). Argumenten voor bevolkingsonderzoek op borstkanker. *Tijdschrift voor Sociale Geneeskunde*, 56, 9–11.
13. Dietvorst, A. M. (2003). *Borstkankerscreening in historisch perspectief (1950–2002). Op het grensvlak van wetenschappelijk en publiek domein*. Medical History. Amsterdam: Vrije Universiteit.
14. Feenberg, A. (1991). *Critical theory of technology*. New York: Oxford University Press.
15. Gezondheidsraad. (1981). Interim-advies inzake de vroegtijdige opsporing van borstkanker. 's Gravenhage.
16. Gezondheidsraad. (1984). Tweede interimadvies inzake de vroege opsporing van borstkanker. 's Gravenhage.
17. Gezondheidsraad. (1987). Vroege opsporing borstkanker. 's Gravenhage: Gezondheidsraad.
18. Gezondheidsraad. (2002). *Het nut van bevolkingsonderzoek naar borstkanker*. Den Haag: Gezondheidsraad.
19. Hendee, W. R. (1995). History and status of X-ray mammography. *Health Physics*, 69(5), 636–648.
20. Hoedemaekers, R., & Oortwijn, W. (2003). Problematic notions in dutch health care package decisions. *Health Care Analysis*, 11(4), 287–294.

21. Hofmann, B. M. (2008). Why ethics should be part of health technology assessment. *International Journal of Technology Assessment in Health Care*, 24(4), 423–429.
22. Ihde, D. (1990). *Technology and the life world: From garden to earth*. Bloomington: Indiana University Press.
23. Klüver, L. K., Nentwich, N., Peissl, W., Torgersen, H., Gloede, F., Hennen, L., et al. (2000). *European participatory technology assessment. Participatory methods in technology assessment and technology decision-making*. Brussels: European Commission.
24. Koningin Wilhelmina Fonds. (1968). *Vroege opsporing en behandeling van kanker*. Amsterdam: Koningin Wilhelmina Fonds/Nederlandse Organisatie voor de Kankerbestrijding.
25. Landelijk Evaluatie Team voor Bevolkingsonderzoek naar Borstkanker. (2007). *Tussenrapportage 2007*. Rotterdam: Erasmus Medisch Centrum.
26. Latour, B. (1996). *Aramis, or, the love of technology*. London: Harvard University Press.
27. Lehoux, P., & Blume, S. (2000). Technology assessment and the sociopolitics of health technologies. *Journal of Health Politics, Policy and Law*, 25(6), 1083–1120.
28. Leys, M. (2003). Health care policy: qualitative evidence and health technology assessment. *Health Policy*, 65(3), 217–226.
29. Leys, M. (2003). Health technology assessment: The contribution of qualitative research. *International Journal of Technology Assessment in Health Care*, 19(2), 317–329.
30. May, C. (2006). Mobilising modern facts: health technology assessment and the politics of evidence. *Sociology of Health & Illness*, 28(5), 513–532.
31. Mol, A. (2000). What diagnostic devices do: the case of blood sugar measurement. *Theoretical Medicine and Bioethics*, 21, 9–22.
32. Norheim, O. F. (2002). The role of evidence in health policy making: A normative perspective. *Health Care Analysis*, 10(3), 309–317.
33. Office of Technology Assessment. (1976). *Development of medical technology: Opportunities for assessment*. United States Congress.
34. Olsen, O., & Gotzsche, P. C. (2001). *Screening for breast cancer with mammography (Cochrane review)*. *The Cochrane library*. Oxford: Update Software.
35. Olsen, O., & Gotzsche, P. C. (2001). Cochrane review on screening for breast cancer with mammography. *The Lancet*, 358, 1340–1342.
36. Oortwijn, W., Reuzel, R., & Decker, M. (2004). Ethical inquiry and (health) technology assessment: The social shaping perspective. *Poiesis & Praxis*, 2, 91.
37. Peeters, P. H. M., Verbeek A. L. M. & Hendriks J. H. C. L. (1987). *Massaal onderzoek borstkanker zinvol. de Volkskrant*.
38. Reuzel, R., Wija, O., Michael, D., Christian, C., Pedro, G., John, G., et al. (2004). Ethics and HTA: Some lessons and challenges for the future. *Poiesis & Praxis*, 2, 247–256.
39. Rip, A., Misa, T. J., & Schot, J. (Eds.). (1995). *Managing technology in society: The approach of constructive technology assessment*. London: Pinter.
40. Stevens, A., & Milne, R. (2004). Health technology assessment in England and Wales. *International Journal of Technology Assessment in Health Care*, 20(1), 11–24.
41. Sturmans, F. (1978). Screening op borstkanker: Het standpunt van een andersdenkende. *Tijdschrift voor Sociale Geneeskunde*, 56, 12–17.
42. Swierstra, T., Boenink, M., & Stemerding, D. (2009). Exploring techno-moral change: The case of the obesity pill. In P. Sollie & M. Düwell (Eds.), *Evaluating new technologies. Methodological problems for the ethical assessment of technology developments* (pp. 119–138). Dordrecht: Springer.
43. Ten Have, H. A. M. J. (1995). Medical technology assessment and ethics. *Hastings Center Report*, 25(5), 13–19.
44. Ten Have, H. (2004). Ethical perspectives on health technology assessment. *International Journal of Technology Assessment in Health Care*, 20(1), 71–76.
45. Tijmstra, T. (1984). Screening en bevolkingsonderzoek. Over de (on)mogelijkheid van actieve vroege opsporing van afwijkingen. *Medisch Contact*, 39(18), 561–567.
46. Tijmstra, T. (1987a). Bevolkingsonderzoek borstkanker verwerpelijk. *de Volkskrant*.
47. Tijmstra, T. (1987). Bevolkingsonderzoek op borstkanker: Een groot aantal vragen. *Medisch Contact*, 42(39), 1233–1235.
48. Tijmstra, T. (1988). Screeningsdiagnostiek is niet alles. *NRC Handelsblad*.
49. Tijmstra, T. (1989). Nut screening op borstkanker onduidelijk. *Medisch Contact*, 44(10), 323–326.
50. Tijmstra, T. (1990). Borstkankerscreening blijkt bij nader inzien een anachronisme. *NRC Handelsblad*.

51. Tweede Kamer der Staten Generaal. (1990–1991a). 18833 nr. 6, Verslag van een mondeling overleg.
52. Tweede Kamer der Staten Generaal. (1990–1991b). 18833 nr. 7, Brief van de staatssecretaris van Welzijn, Volksgezondheid en Cultuur.
53. Tweede Kamer der Staten Generaal. (1992–1993). 18833 nr. 8, Brief van de staatssecretaris van Welzijn, Volksgezondheid en Cultuur.
54. van Bekkum, D. W. (1990). Vrouwen moeten zelf over borstscreening kunnen besluiten. *NRC Handelsblad*.
55. van Bekkum, D. W., & Paulides, J. J. (1993). Leeftijdscriminatie bij het bevolkingsonderzoek op borstkanker. *Medisch Contact*, *48*(5), 135–137.
56. van der Kooij, S. (1993). Borstkankerscreening ook boven 70 jaar? *Medisch Contact*, *48*(5), 137–139.
57. van der Maas, P. J. (2000). Bevolkingsonderzoek naar borstkanker: een tussenbalans. *Nederlands Tijdschrift voor Geneeskunde*, *144*(2), 1096–1099.
58. van der Wilt, G. J., Rob, R., & Banta, H. D. (2000). The ethics of assessing health technologies. *Theoretical Medicine and Bioethics*, *21*, 103–115.
59. van Maanen, H. (2002). Eenzijdige voorlichting voor vrouwen over het bevolkingsonderzoek naar borstkanker maakt geïnformeerde en weloverwogen keuze onwaarschijnlijk. *Nederlands Tijdschrift voor Geneeskunde*, *146*(22), 1026–1028.
60. van Merkerk, R. O., & Ruud Smits, E. H. M. (2008). Tailoring CTA for emerging technologies. *Technological Forecasting & Social Change*, *75*, 312–333.
61. van Zonneveld, R. J. (1978). De mythe van de preventie. *Tijdschrift voor Sociale Geneeskunde*, *56*(16), 495.
62. Verbeek, P.-P. (2005). *What things do: Philosophical reflections on technology, agency and design*. University Park, PA: Penn State University Press.
63. Ziekenfondsraad. (1992). Advies inzake de organisatie en de kwaliteitsbewaking van het bevolkingsonderzoek borstkanker. Amstelveen.