Medical Hypotheses 82 (2014) 412–420

Contents lists available at ScienceDirect

Medical Hypotheses

journal homepage: www.elsevier.com/locate/mehy

The efficacy of surgical treatment of cancer – 20 years later

Donald J. Benjamin *

Convenor/Research Officer, Cancer Information & Support Society, St Leonards, NSW, Australia

ARTICLE INFO

Article history:
Received 12 November 2013
Accepted 3 January 2014

ABSTRACT

Introduction: Cancer treatment will be effective only if it is be based on a valid paradigm of what cancer is and therefore capable of affecting the course of the disease. A review in 1993 found no evidence that surgery affected the course of the disease and an alternative paradigm was proposed. A review of mammography screening trials in 1996 found no benefits from breast cancer screening. This was predicted by this alternative paradigm. This review updates the evidence twenty years later.
Aim: To identify evidence that the primary treatment of cancer, surgery, has been shown to affect the course of the disease.
If there is no such evidence, then to identify the correct paradigm of what cancer is from other cancer treatments that have been shown to be effective.
Method: Because surgery has never been shown in a randomised controlled trial to affect the course of cancer seven other indirect methods were used to evaluate its efficacy.
Results: None of the seven indirect methods used showed that surgery clearly affects the course of the disease for any type of cancer. The lack of benefits from cancer screening now includes not only from breast cancer but also from bowel, lung, prostate and ovarian cancer screening. This confirms that cancer surgery is based on an invalid paradigm of what cancer is. Survival figures following treatments based on an alternative paradigm that assumes cancer is a systemic disease were found to be superior to those following surgery, reinforcing the conclusion that cancer is a systemic disease and that cancer surgery is unlikely to be of benefit in most cases.
Conclusion: No benefits can be expected to be achieved from using cancer surgery except in a few immediately life-threatening situations. Surgery appears to be based on an invalid paradigm of what cancer is. Cancer appears to be a systemic disease and therefore standard treatments need to be reassessed in this light.

Introduction

The current paradigm is that cancer is a local disease that sometimes later spreads, so removing the tumour means removing the disease. Based on this paradigm the accepted primary treatment for cancer once a solid tumour is diagnosed is surgery. Prior to the acceptance of this paradigm cancer was considered a systemic disease. Over 130 years have passed since surgery became accepted as the primary treatment for cancer. Despite the increasingly aggressive surgery used in the early years, culminating in the hemicorporectomy, little progress was made in saving lives. Over the intervening years surgery has become gradually less aggressive. For example Halstead's radical mastectomy applied the same aggressive principles with the aim of "getting it all" with breast cancer. This has since been essentially replaced by lumpectomy without any loss of survival. This change was achieved by using randomised controlled clinical trials (RCTs) to confirm that for each step in between, the simple mastectomy and the quadrantectomy, the treatment was just as effective in terms of survival as its earlier more radical version. Comparing lumpectomy with no treatment has not been done or even suggested.

Because surgery has been the accepted treatment for cancer for over a century there is little questioning of its efficacy. Most people would have no reason to question the claim of its efficacy because they accept the assumption that cancer starts as a local disease that later spreads; and surgery is clearly capable of removing the tumour. So "clinicians know that almost every type of cancer can be treated more successfully if discovered in a localized stage" [1]. This unproven belief is also generally accepted by the community, most of whom know of someone who has had a tumour removed and is now alive or lived many years before they died.

* Address: 1/38 View St, Chatswood, NSW 2067, Australia. Tel.: +61 416121140; fax: +61 299062189.
E-mail address: don.benjamin@bigpond.com

http://dx.doi.org/10.1016/j.mehy.2014.01.004
0306-9877

This anecdotal evidence does not necessarily reflect the reality of the situation.

It is only over the past forty years since the passing of the Cancer Act in 1971 in the US and the spending of hundreds of billions of dollars that many medical scientists have started asking why there has been so little progress in reducing cancer mortality in the intervening years [2].

To prove that surgery is an effective treatment for cancer would require running an RCT with two matched groups with a particular type of cancer at the same stage of development: the treated group would be given surgery to remove the tumour and the control group would be left untreated. This has never been done.

Because surgery became an 'accepted' treatment for cancer before the concept of the RCT was developed to demonstrate the efficacy of a treatment, it is now considered unethical and politically impossible to do that now. The argument is that it would be unethical to withhold an 'effective treatment' from a group of people with cancer just to confirm that it is effective. The illogical part of this argument is that surgery is assumed to be effective without such a proper evaluation ever having been done. Those who put forward this argument do so assuming that the tumour is the disease. Removal of the tumour, ipso facto, must remove the disease, so there is no need for further proof.

If, as suggested by some medical researchers, cancer is a systemic disease and the tumour is only a symptom or element of that disease, surgery would have little or no effect on the course of the disease. However, as is mentioned below, there can be no doubt that the tumour is often a problem that needs to be dealt with directly to save or to improve the quality of life of many patients with cancer.

Because a proper RCT is unlikely to be carried out to prove the efficacy of surgery, it is necessary to identify and use other indirect methods of evaluating the efficacy of surgery for cancer. This was done in 1993 [3]. No clear evidence of survival benefit was found. This review updates the evidence twenty years later with the advantage of additional evidence from several RCTs evaluating surgical intervention for prostate cancer and longer term follow-ups from RCTs evaluating screening for breast and other forms of cancer.

Aim

To identify new evidence showing that the primary treatment of cancer, surgery, has been shown to affect the course of the disease.

If there is no such evidence, then to identify the correct paradigm of what cancer is from treatments that have been shown effective in randomised controlled clinical trials.

Method

Because surgery has never been shown in a randomised controlled trial to affect the course of cancer, seven other indirect methods were used to evaluate its efficacy. These include

(a) the graphical method – plotting survival curves on log-linear graph paper to identify sub-populations with cancer whose mortality had been reduced following intervention;
(b) the dose–response method – comparing survival after different amounts of excision (e.g. radical mastectomy, simple mastectomy, quadrantectomy, lumpectomy) – efficacy requires an optimal degree of excision;
(c) analysing the results of RCTs evaluating the effects of prostatectomy compared to 'watchful waiting';
(d) changes in survival over time – e.g. improved percentage 5-year survival;
(e) comparison of incidence and mortality trend lines over time – efficacy requires the lines to diverge in a particular way;
(f) epidemiological analysis – identifying effects of interventions on mortality trends.
(g) comparing the results of early treatment with delayed treatment by analysing the results of RCTs evaluating screening for breast, prostate, bowel, lung and ovarian cancers.

In addition to these it is also possible to compare survival or mortality figures achieved using surgery with treatments based on an alternative paradigm of what cancer is. If better results are achieved using surgery, this would support the validity of the current paradigm. If the opposite were observed and better results were achieved by treating cancer as a systemic disease, this would question the validity of the current paradigm on which surgery is based.

Results

The following is the result of analysing the benefits of surgery using these seven indirect methods.

Plotting survival curves to identify sub-populations with cancer whose mortality had improved following surgical intervention – the graphical method

As pointed out in the earlier review [3] in 1825 Gompertz established that, as a person ages from birth to death, the age specific mortality rate doubles about every 8½ years [4]. The same process applies for all living organisms but the time for doubling of mortality varies. A typical curve for humans is one with a formula such as $y = e^{0.08x}$ where x is the chronological age in years. When plotted on semi-logarithmic graph paper (i.e., age along a linear x-axis and mortality along a logarithmic y-axis) this gives a straight line sloping up to the right, a so-called 'constant' mortality rate line.

Jones [5] analysed statistics on mortality rates for different diseases from studies throughout the world and found that for chronic diseases, including both cancer and heart disease, people with these diseases lie on a constant slope mortality curve. Their mortality continues to double every 8½ years. Each disease has a different age-specific mortality rate line, but all the lines had the same slope. The line for cancer is higher than for healthy people of the same age. Alternatively it can be considered to be displaced to the left by about 15 years. In other words people with cancer behave as if the cancer had aged them by 15 years.

The graphical method is sensitive to the effects of any therapy that can affect mortality because groups of patients whose mortality is reduced by intervention fall onto a lower, healthier mortality rate line. Jones was unable to find any intervention for cancer that produced mortality or survival benefits.

A more common use of this graphical method involves plotting relative or percentage survival on a logarithmic scale against years from diagnosis. This produces a straight line sloping down from left to right. However in this case higher mortality rates produce a steeper slope.

Fox used this technique to measure efficacy of treatment for breast cancer [6]. He used data on the survival of untreated breast cancer patients at Middlesex Hospital at the turn of the 19th century. This group had received no surgery, radiotherapy or chemotherapy. Their death rate was 25% per year with few survivors after 7–8 years as would be shown by a steep downward straight line mortality curve.

Fox then took the survival data of patients treated by surgery, radiotherapy and chemotherapy during the period 1950–1973 and plotted it in the same way.

He found that the total treated group could be broken down into two sub-populations with a completely different prognosis. One group (40% of the total) had a poor prognosis and, like the untreated Middlesex Hospital group, few had survived after 7–8 years. The remaining 60% had a survival characteristic only slightly lower than that of women of similar age without evidence of cancer.

The usual conclusion drawn from this data is that breast cancer can be cured if detected early. However a second possible conclusion is that there are two groups of patients with breast cancer: 40% with a suppressed immune system for whom the cancer grows and spreads rapidly (the Middlesex Hospital group would have consisted mainly of advanced cases and therefore in this category) and 60% whose immune defences are still capable of keeping cancer under control. If the second conclusion is valid it is possible that treatment had no effect in either of the groups.

Another analysis helps clarify this ambiguity. A sub-group of 704 women first seen with breast cancer in the Cambridge area from 1947 to 1950 and with long-term survival was at first thought to have been identified as 'cured' after 26 years using the correct definition of cure, *viz a group of people treated for a particular type of cancer has the same mortality rate as a comparable group of healthy people in the community*.

However further follow up by Brinkley and Haybittle after 35 years [7] showed that the ratio of observed to expected deaths reached a minimum of about 2. 'Cure' would require the curve to reach 1, equivalent to a comparable group of healthy women. So surgery had had no 'curative' effect on these women with long-term survival.

So the graphical method of analysis of plotting survival curves used by Gompertz, Jones, Fox, Brinkley and Haybittle provides alternative interpretations of the data used to claim that surgery affects the course of the disease, suggesting no benefits from surgical intervention.

Comparing survival after different amounts of excision (e.g. radical mastectomy, simple mastectomy, quadrantectomy, lumpectomy) – the dose–response method

Another useful source of evidence for evaluating the efficacy of cancer surgery is the group of randomised controlled trials that compared survival after different degrees of surgery: radical mastectomy, simple mastectomy, quadrantectomy, lumpectomy, or excisional biopsy. This is analogous to the dose response approach used to determine the optimum dose of a relatively toxic medicine. A typical dose response is represented by a bell-shaped distribution curve. Too low a dose has no beneficial effect. The positive effect increases with dosage and reaches a maximum (optimal dose) after which the excessive toxicity causes the response to fall back down again to reach zero.

According to the current paradigm it is important to detect cancer early before cancer cells can spread. So the more healthy tissue around the tumour that is removed the less likely it is for malignant cells to be left behind to spread. Excessively traumatic surgery could adversely affect the body. So there should be an optimum degree of excision to maximise survival. If this paradigm is invalid and cancer is a systemic disease with the tumour only a late-state symptom, it would make no difference how extensive the surgery was unless the excessive intervention caused harm. Up to this point the same survival would be expected as without any surgery.

As mentioned in the Introduction, several such comparative studies have been carried out to measure the importance of radical surgery. All of these studies showed there was no difference in survival between women who underwent different degrees of surgery ranging from radical mastectomy to simple mastectomy to quadrantectomy to lumpectomy and finally excisional biopsy [8–13]. If the amount of excision were important one of these RCTs would have found an optimal degree of excision.

These results show the relevance of the comment made in 1963 by Shimkin that 'when many forms of treatment appear to yield the same results or lack thereof, suspicion should arise that none is really effective and a no-treatment group in subsequent comparisons may be acceptable' [14]. However that has never been done.

So the dose–response method provides further evidence questioning the claim that surgery affects the course of the disease.

Despite what Shimkin said, no such randomised controlled trials have yet been held to compare surgery with no treatment. As mentioned above such a trial is considered unethical today because it would involve withholding a treatment widely accepted as effective to demonstrate its efficacy. So the medical profession has locked itself into using an unproven method as the primary treatment for cancer.

However there are two RCTs that have been held that come close to a comparison of surgery with no surgery. They involved comparing survival or mortality after a radical prostatectomy with so-called "Watchful Waiting". The latter group still involved some treatment but the surgical treatment was delayed until more definite signs of progression were observed. This indirect method of comparing the effect of normal treatment with delayed treatment is considered under (c) below but could also have been considered similar to the effects of screening discussed under (g) below which compares early treatment with normal treatment.

Analyzing the results of RCTs evaluating the effects of radical prostatectomy compared to 'watchful waiting'

These two RCTs comparing radical prostatectomy with Watchful Waiting have been evaluated by the Cochrane Group [15].

Because of the difficulties caused by different types of bias in most RCTs and other confounding factors, the only reliable measure of efficacy is to compare the results from the whole treated group with the whole untreated (control) group to see if treatment reduced mortality [16]. In other words to see if all-cause mortality is reduced by treatment.

If disease specific mortality, i.e. prostate cancer (PC) deaths, is compared between the groups it is not easy to confirm if any observed reduction in prostate cancer deaths is real and saved lives, or rather was the result of a harmful treatment causing a man who would have died from prostate cancer to die instead from other causes. The latter would produce an apparent reduction in prostate cancer deaths unless it was correctly classified as a death from prostate cancer [16].

The Cochrane Review found that in the US study (the Veterans Administration Co-operative Urological Research Group trial, referred to as "the VACURG trial") the reduction in all-cause mortality after 23 years follow-up was not statistically significant. In the Scandinavian Trial (referred to as the SPCG-4 trial) the overall reduction in all-cause mortality was also not statistically significant. It observed that in men under 65 the reduction was statistically significant [17]. However, like many other RCTs there were confounding factors that could explain not only the reduction in prostate cancer deaths but also the small reduction in all-cause mortality.

As with the trials evaluating breast and prostate cancer screening (see (g) below) the protocols used guaranteed that treatments would be used differently in the two arms of the trial rendering the results of limited value [18]. Unfortunately confounding factors are built into many RCTs. In the Scandinavian trial (considered by the Cochrane Group to be of higher quality) the protocols included:

For men assigned to the radical-prostatectomy group, the surgical procedure started with a lymphadenectomy of the obturator

fossa; if no nodal metastases were found in frozen sections, the radical prostatectomy was performed. Radical excision of the tumour was given priority over nerve-sparing surgery. Men who were assigned to the watchful-waiting group did not receive any immediate treatment.

If signs of local recurrence (a palpable nodule or histologically confirmed recurrence) developed in a patient in the radical-prostatectomy group, hormonal therapy was initiated.

Men in the watchful-waiting group who had signs of obstructive voiding disorders were treated with transurethral resection.

Metastases detected by bone scan were managed with hormonal therapy.

In 2003, after the introduction of antiandrogens, clinicians were allowed to initiate hormonal treatment if there were signs of tumour progression, including elevations in PSA level, and if they believed that hormonal treatment would be beneficial to the patient.

These protocols guaranteed that several potentially harmful treatments, apart from the radical prostatectomy, would be used differently in the two trial arms, thus confounding the results.

Before 2003 there was no difference in all-cause mortality between the two groups. By 2010 a difference was observed. The following are some of the differences between the two trial arms:

- Because local recurrence would be delayed in the Radical Prostatectomy (RP) group because of the absence of tissue for cells to grow in, these men would receive less hormone treatment than those in the Watchful Waiting (WW) group.
- More men in the WW group would have received Transurethral Resection of the Prostate (TURP).
- Because there would probably be more tumour progression and increasing PSA levels in the WW group, men in this group would again receive more Androgen Deprivation Treatment (ADT), i.e. hormone treatment.
- There was more chemotherapy used in the WW Group.
- There was more radiotherapy used in the WW Group.

Table 1 summarises these differences. Because there was more radical surgery in the treated group there would have been more complications from the RP surgery in this group. Any deaths resulting from complications would not necessarily have been added to the prostate cancer deaths as they should have been. So this could have increased the deaths from other causes in the treated group, giving a false reduction in deaths from prostate cancer.

There were at least 96 more additional potentially harmful treatments in the WW Group.

TURP, radiotherapy and chemotherapy are known to cause harm including extra deaths. The death rate from a lumbar laminectomy used for metastatic prostate cancer is only about 1% and is unlikely to be observed among only 10 cases. The median survival is about 6 months so in any case it might be difficult to accurately specify the cause of death.

The hormone treatment (induced testosterone deficiency/Androgen Deprivation Treatment (ADT)) has been shown to lead to various side effects, including endocrine, hematological, musculoskeletal, and psychological changes, with negative impacts on Quality of Life.

Saylor et al. [19] described the detrimental effects of ADT on several metabolic end points and to bone health. ADT had been prospectively shown to cause decreased lean muscle mass, increased fat mass, weight gain, increased cholesterol and triglycerides, insulin resistance, and loss of bone mineral density. In population-based analyses it had been associated with an increased incidence of diabetes, clinical fractures, and cardiovascular disease.

Garnick et al. [20] presented results of three phase III randomised trials of ADT at the 2005 meeting of the American Society of Clinical Oncology with respect to the impact on the QT interval. The QT interval on a standard 12-lead electrocardiogram measures ventricular repolarization (see Fig. 1).

A prolonged QT interval, particularly above 450–500 ms, is associated with a higher incidence of arrhythmias and sudden death. Garnick's analysis suggested that use of a luteinising hormone-releasing hormone (LHRH) agonist, combined androgen blockade (with 1 of 2 LHRH agonists and a NSAA), or an LHRH antagonist were all associated with 9–21 ms increases in the corrected QT interval among a total of 476 patients. So hormone (ADT) treatment could have resulted in deaths.

Efstathiou et al. have confirmed the deaths from hormone treatment for early prostate cancer [21].

Actual results in the radical prostatectomy trial above were as follows:

The difference in deaths from all causes between the two groups was 35, i.e. 166 in the RP group and 201 in the WW group. The difference in deaths from Prostate Cancer was 26, i.e. 55 in the RP group and 81 in the WW group.

This means that deaths from other causes were 111 in the RP group and 120 in the WW group, a difference of 9 (or about 8%) (see Fig. 2a). Because the two groups were almost exactly matched in numbers (347 vs 348) it is valid to compare the deaths in the two groups.

Table 1 shows that there were about 96 extra potentially harmful treatments, TURP, hormone (ADT), chemotherapy and radiotherapy in the WW group, so any deaths as a result would have biased the results of the trial in favour of RP treatment in two ways:

- Those deaths resulting from harmful treatments should have been included in the deaths from prostate cancer in the WW group. If they were, they could have contributed to the 26 extra PC deaths in the WW shown in Fig. 2a and produced an apparent reduction in deaths from PC in the RP group.
- If they were not, they could have contributed to the 9 extra deaths from other causes in the WW Group shown in Fig. 2a, again favouring the RP group.

In practical terms there is likely to have been a mixture of these two factors.

Table 1
Differential use of additional potentially harmful treatments.

Variable	$n = 347$	$n = 348$	
	Radical prostatectomy group	Watchful waiting group	Possible increase in deaths?
Radical prostatectomy	Yes	No	Yes
TURP	No	Yes	Yes
Hormone (ADT)	132	205	Yes
Radiation	33	47	Yes
Laminectomy	3	10	
Chemotherapy	6	8	Yes
Total	174	270	

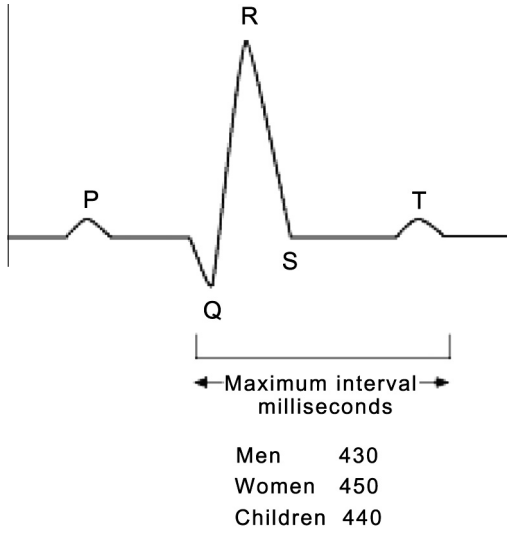

Fig. 1. An electrocardiogram showing the QT interval. This represents the duration between the onset of depolarisation and the completion of repolarisation of the myocardium. The QTc is the QT corrected for heart rate. Androgen Deprivation Treatment (ADT) can increase the QTc interval by 9–21 ms [20]. A prolonged QTc interval increases the risk of arrhythmias and sudden death [20,21]. Figure from: *http://www.australianprescriber.com/magazine/25/3/63/5*

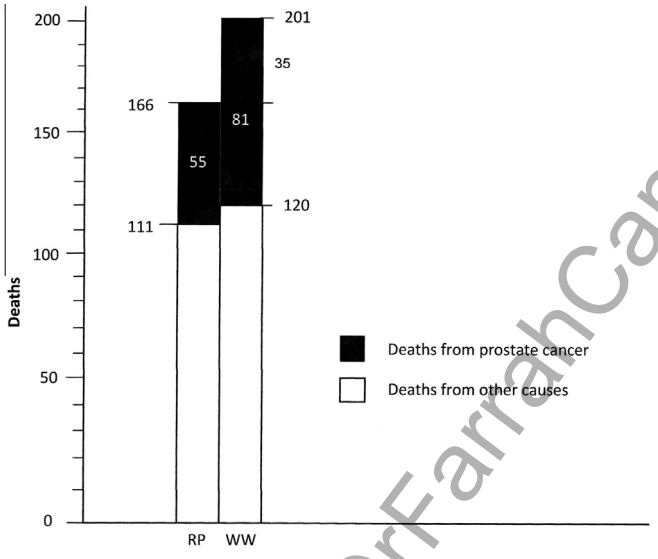

Fig. 2a. Outcomes from SPGC-4 trial.

Black et al. [16] describe what the outcomes of a properly run randomised controlled trial should look like. This is shown in Fig. 2b.

As can be seen from Fig. 2a the non-cancer mortality or deaths from other causes in the screen and control groups in the SPGC-4 trial differed by about 8% (120–111)/111. There should be no significant difference in a properly run RCT that has the same treatments in screen and control groups other than the treatment being evaluated. This suggests that the effect of the above medical interventions in one arm of the trial had played a part as confounding factors in biasing the results.

Analysis of sub-groups, those 65 years or older and those younger than 65 suggested that the reduction in all cause mortality of those under 65 reached significance. However the Cochrane review suggested that these findings should be interpreted very cautiously as this sub-group analysis was not pre-specified in the trial protocol, the trial was not powered to investigate age groups separately

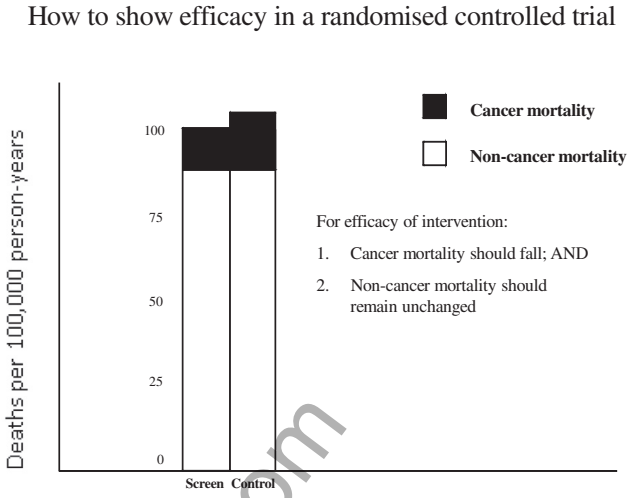

Fig. 2b. Expected outcomes from a properly run RCT without confounding factors.

and there was no *a priori* biological hypothesis underpinning the sub-group analysis [15]. The reported benefits of RP over WW in the SPCG-4 trial must be carefully weighed against the side effects of surgical and other invasive treatments.

So again this indirect method of using results of RCTs provides evidence suggesting that surgery does not affect the course of the disease.

Changes in survival over time – e.g. improved percentage 5-year survival

A fourth and less reliable method of evaluating the efficacy of surgery is to see how percentage 5-year survival figures have varied over time for the different types of cancer. This could identify the effects of improvements in surgical treatment. For example there was an increase in percentage 5-year survival for all cancer types between 1960 and 1975.

According to Enstrom and Austin [22] these figures are unreliable as a measure of progress in cancer control for several reasons including:

- increased 5-year survival can result from death happening later – real progress; or from making an earlier diagnosis – there is no progress here. Death still occurs at the same time but the existence of cancer has been known for a longer time leading to an 'apparent' increase in survival.
- Earlier detection affects the stage of a cancer when detected. For example stage III in 1990 is not the same thing as Stage III in 1970, and has a better prognosis irrespective of the effects of therapy.
- The 'Will Rogers phenomenon' shows that early detection results in an improved survival for all stages of a cancer, even if there is no overall increase in survival. This occurs because healthier cases move into later stages boosting the apparent survival of the later stage. This is similar to the previous point.
- Comparison between different years results in comparison of unmatched groups.

According to Greenberg [23], another factor that can distort the survival figures is that

- Earlier figures with lower survival applied when more aggressive and therefore more harmful treatments were reducing survival. An improved survival might simply reflect less harm done.

In 1987 the US General Accounting Office confirmed that claims of increased survival have been overstated [24].

More recent evidence for the unreliability of an increasing percentage 5-year survival as a measure of cancer progress comes from a study headed by Welch et al. [25]. In this study they calculated the change in 5-year survival from 1950 to 1995 for the 20 most common solid tumour types and correlated these changes in survival with changes in incidence and mortality for each tumour type. They found that from 1950 to 1995 there was an increase in 5-year survival for each of the 20 tumour types. The absolute increase ranged from 3% for pancreatic cancer to 50% for prostate cancer. During the same period mortality rates declined for 12 types of cancer and increased for the remaining 8 types. There was little correlation between changes in 5-year survival for a specific tumour and the change in tumour-related mortality. On the other hand the change in 5-year survival was positively correlated with the change in the tumour incidence rate. They concluded that 5-year survival is a valid measure for comparing therapies in a randomised trial, but in other cases changes in 5-year survival bear little relationship to changes in mortality. Instead they appear related to changing patterns of diagnosis.

So again this indirect method of comparing 5-year survival rates provides evidence questioning the claim that surgery affects the course of the disease.

Comparison of incidence and mortality trend lines over time – efficacy requires the lines to diverge in a particular way

A fifth and fairly reliable measure of progress in cancer treatment comes from comparing the change of incidence of a particular type of cancer with the change in mortality over time. According to Enstrom and Austin [22] progress in cancer control requires that the mortality rate decline more rapidly or rise more slowly than the incidence.

There have been large changes in incidence and mortality over time for many cancers but in most cases the two graphs have changed by the same amount. In the few cases where incidence has risen faster than mortality, such as prostate and breast cancers, this can be explained by the fact that for prostate cancer there was an increasing rate of autopsies and operative procedures for other causes leading to an increasing recorded incidence of occult tumours most of which would not have caused serious symptoms during the man's lifetime. This effect accelerated with the widespread introduction of the PSA test. For breast cancer the increased incidence resulted from the widespread introduction of breast cancer screening.

As Cairns pointed out [26] the incidence discovered by autopsies on 70-year old men who had died of other causes was up to 100 times as great as the incidence based on diagnosis of such patients presenting with symptoms. For this reason incidence figures are not a reliable measure for comparing with mortality trends for assessing progress in prostate cancer control.

Similarly with breast cancer, the rising use of mammographic screening has identified many ductal carcinomas in situ, most of which would not have produced symptoms nor have been life-threatening [27]. Other research has shown that even a significant proportion of invasive breast tumours identified by screening are not life-threatening and disappear if left untreated [28]. This means that up to 40% of tumours identified by breast cancer screening are not life-threatening and provide an artificial increase in incidence [28,29] – and also a false increase in survival.

Fox, mentioned above [6], plotted the change in incidence of breast cancer from the 1930s to the mid 1970s and compared this curve with that of breast cancer mortality over the same period. He found that both lines remained essentially flat and parallel until the mid 1960s when the introduction of mammographic screening resulted in an apparent increase in the incidence of breast cancer. This continued to rise as breast cancer screening became more widespread. Many more early breast cancers were detected and treated but the mortality rate curve remained unaffected. Autier et al. [30] and Jørgensen et al. [31] observed the same phenomenon over the last 30 years in Europe.

So again this indirect method of comparison of incidence and mortality curves provides evidence that surgery does not affect the course of the disease.

Epidemiological analysis – identifying effects of interventions on mortality trends

A sixth method of evaluating medical intervention is to observe mortality figures over time before and after the introduction of a new therapy or procedure. If the therapy or procedure is effective in reducing mortality the curve should show a significant drop in the mortality rate after its introduction.

One procedure widely claimed to have reduced cancer mortality is the Pap smear (named after its inventor George Papanicolaou). It is claimed that after its widespread introduction in the early to mid 1950s the mortality rate for cancer of the uterus, which includes cervical cancer, fell and has been falling ever since as a result of the test's continuous use.

Fig. 3 suggests that this claim is questionable. The death rate figures come from the American Cancer Society publication: Cancer Fact and Figures, 1984. The arrow has been added to show the time of introduction of the Pap test. It is clear from the slope of the graph that the mortality had already peaked in the 1930s and had already started its sharp decline before 1950.

The mortality from cervical cancer in the United States decreased by 70 percent over 37 years between 1947 and 1984, a rate of about 2% per annum after the introduction of the Pap test. The fall from 31 per 100,000 in 1936 to 24 in 1948 was a fall of 22.6% over 12 years, again about 2% per annum *before* the Pap test was introduced. So there was no change in slope of the graph after the widespread introduction of the Pap test.

In contrast with this there is some ambiguous evidence of the benefit of early surgery for cervical cancer. In the mid-1960s, Finland, Sweden and Iceland implemented PAP screening programs

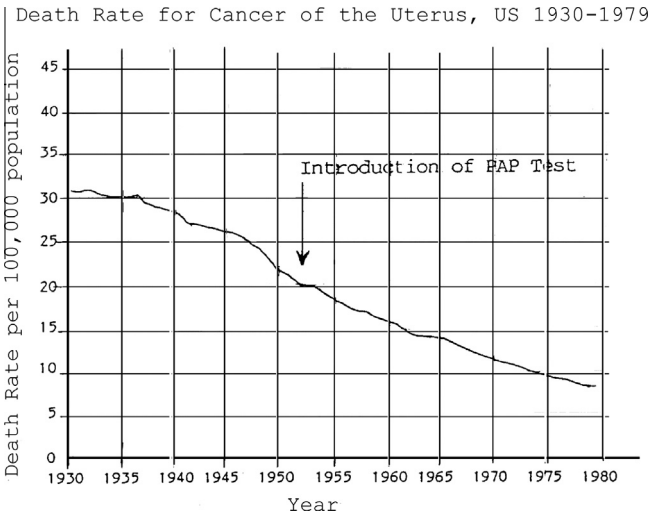

Fig. 3. The declining mortality of uterine cancer in the United States. This death rate includes death from cervical cancer. The falling rate was not affected by the introduction of the Pap test.

in which more than 80 percent of women participated, whereas Norway performed screening in only one county, which comprised 5 percent of the population. All four countries noted a similar incidence of cervical cancer in 1960, which is not surprising given the homogeneity of their populations. During the ensuing 20 years, the incidence of cervical cancer did not change in Norway but decreased by approximately 50 percent in the other three countries, where mortality from cervical cancer also decreased [32]. Likewise, with the implementation of Pap-smear screening programs in British Columbia, the incidence of cervical cancer *decreased* by 85 percent between 1955 and 1988 [33].

The problem with this data is that it runs counter to all other screening programs where it is found that screening results in a significant *increase* in the incidence of the cancer. One explanation for this is that unlike other screening Pap testing prevents cervical cancer by detecting mainly pre-malignant lesions. The screening might increase the incidence of pre-malignant lesions and as a result of surgical intervention reduce the incidence of malignant cervical cancer. The reduction in mortality observed in most developed countries since the 1940s would therefore appear to be made up of two components: a continuing reduction from the 1930s due to factors unknown, possibly lifestyle factors, together with a reduction in incidence that led to a reduction in mortality.

Despite the lack of randomised trials, these data provide some evidence for a possible effect of Pap-smear screening on the *prevention* of cervical cancer and thereby on mortality. However cervical cancer is a special case in that it is claimed by some that a significant proportion of it is caused by a human papillomavirus (HPV) infection and hence preventable. Pap smear screening allows for the early detection of pre-cancerous lesions. Surgery then removes the area of the cervix where the pre-malignant lesion appeared, so that it takes a lot longer for new lesions to form and then possibly become malignant. This could explain the decrease in incidence which would result in decreased mortality. This suggests that surgery is effective when it is used on pre-malignant lesions but not on real cancer.

Where Pap smear screening has not been subjected to any RCTs, there is one area where the effect of screening has been measured in RCTs and its impact on long-term mortality has also been measured. This is the introduction of mammographic breast cancer screening where a lack of effect can be shown. This suggests the rather novel idea that surgery is effective when it is used on pre-malignant lesions but not on real cancer.

A similar lack of effect can be shown in relation to the introduction of mammographic breast cancer screening.

In recent studies researchers from the Nordic Cochrane group analysed trends in mortality from breast cancer to measure the impact of the introduction of mammographic screening in those areas of the country where it was introduced [31]. They compare the annual percentage change in breast cancer mortality in areas of Denmark where screening was used with the change in areas where it was not used – during 10 years before screening was introduced and for 10 years after screening was in practice (starting five years after introduction of screening). The hypothesis is that, if earlier treatment made possible by screening had an effect, the mortality would fall more in areas where widespread screening was introduced (Copenhagen and Funen county) than in areas where it was not.

They found that in women who were claimed to benefit from screening (ages 55–74 years) there was a decline in mortality of 1% per year in the screening areas during the 10 year period when screening could have had an effect (1997–2006). In women of the same age in the non-screening areas, there was a decline in mortality of 2% per year in the same 10 year period [31].

So again this indirect method does not provide consistent evidence that surgery affects the course of the disease once it reaches the malignant stage.

Comparing the results of early treatment with delayed treatment – by analysing the results of RCTs evaluating screening for breast, bowel, lung, prostate and ovarian cancers

RCTs have been carried out to evaluate the benefits of early detection in terms of increased survival or reduced mortality with five different types of cancer - breast, bowel, lung, prostate and ovarian cancers. The following are the results from these studies:

Screening for breast, bowel and lung cancer

Black et al. [15] analysed the results of trials evaluating these three types of cancer screening. As mentioned in (c) above they pointed out that proof of reduction in deaths as a result of earlier treatment made possible by screening healthy people requires at least two requirements to be fulfilled in the trial as shown in Fig. 2b above:

(i) The deaths from the particular cancer must fall in the group offered screening compared to those in the control group not offered screening; and
(ii) The deaths from other causes must not rise.

In other words, deaths from all causes must also fall. This is to ensure that any harmful post-screening treatments did not cause deaths attributed to other causes that would give the appearance of a reduction in the deaths from the particular cancer. The authors were not able to identify any trials where these two requirements were satisfied that did not have other methodological flaws.

For example there was an increase in deaths from other causes in most breast cancer screening trials, presumably as a result of the harmful effect of radiotherapy that was used differently in the two arms of the trial [18]. This produced an apparent reduction in up to 25% of deaths from breast cancer. The mechanism was identified as being mainly due to damage by radiotherapy to the left anterior descending artery to the heart in women whose left breasts were irradiated [34].

More recent analyses of breast cancer screening by the Nordic Cochrane Group have confirmed this lack of benefit from mammography screening. The authors found that although trials with suboptimal randomisation showed a significant 25% reduction in breast cancer mortality they also found that breast cancer mortality was an unreliable outcome that was biased in favour of screening, mainly because of differential misclassification of cause of death. The trials with adequate randomisation did not find an effect of screening on total cancer mortality, including breast cancer, after 10 years [35].

Screening for prostate cancer

Five RCTs with a total of 341,342 participants have evaluated the benefits of prostate cancer screening. All involved prostate-specific antigen (PSA) testing, with or without digital rectal examination (DRE). In all trials participants were randomly assigned to be screened or not with the PSA test.

The methodological quality of three of the studies was assessed by the Cochrane Group as posing a high risk of bias [36]. The other two, the European Randomized Study of Screening for Prostate Cancer (ERSPC) and the US Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial were assessed as posing a low risk of bias, but provided contradicting results. In these two studies the two groups were followed for more than a decade.

In the American (PLCO) study run by the US National Cancer Institute the deaths from all causes in the group offered screening

was higher than that in the control group although the difference was not statistically significant. Boyle and Brawley [37] commented that this was a concerning feature of this and two earlier RCTs. It "raises the possibility that it could be due to a small increased risk of non-prostate cancer death from prostate cancer treatment. It has been reported that hormonal therapies in men with locoregional prostate cancer with gonadotropin-releasing hormone analogs increase the risk of diabetes, cardiovascular disease, and stroke".

The European Study (ERSPC) involved seven European countries with 72,890 men aged 50–74 in the group offered screening and 89,353 in the control group. Unlike the American study the details of deaths from all causes were not published and had to be obtained from the trial leaders. Bias from the effects of harm from hormone treatment used differently in the two arms of the trial have since been confirmed by Haines and Gabor [38].

The Cochrane Review meta-analysis of the five included studies showed no statistically significant difference in prostate cancer-specific mortality or overall mortality between men randomised to the screening and control groups [36].

The ERSPC was the only study of the five included in the review that reported a significant reduction in prostate cancer-specific mortality (16%) which is now being questioned [38]. The PLCO study found no significant benefit.

In both trials there was also a problem of a significant number of men in the control groups seeking PSA testing thus reducing the power of the trials to identify significant effects.

Screening for ovarian cancer

An RCT was carried out among more than 70,000 women in the general US population to see if early detection of ovarian cancer could reduce mortality. All of the women were post-menopausal and aged 55–74. The screening was done as part of the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial mentioned above. The 35,000 in the screened group were offered a blood test that measured a cancer marker CA-125 and at the same time offered screening using transvaginal ultrasound to detect enlarged or abnormal ovaries. Another 35,000 women not offered screening made up the control group.

This screening did not result in any reduction in deaths from all causes. In fact there was a non-significant 18% increase in deaths from ovarian cancer attributed to harm from post-screening surgical intervention [39].

So again this indirect method does not provide evidence that surgery affects the course of the disease by early detection using screening for breast, bowel, lung, prostate or ovarian cancer.

Anecdotal evidence

There is much anecdotal evidence showing that surgery can produce short- or medium-term increased survival when used to remove tumours that are immediately life-threatening by obstructing vital organs such as the bowel or pressing on the brain. However this process could also apply to benign growths threatening vital organs and is therefore not necessarily evidence that surgery affects the course of the disease. Unfortunately this temporary benefit has been assumed to also apply to the more than 95% of tumours that are not immediately life-threatening.

The alternative paradigm

As there appears to be no clear evidence from any of these seven indirect methods of evaluating the effects of surgery on the course of cancer, this suggests that the current paradigm is invalid. It is therefore necessary to identify a valid paradigm of what cancer is from treatments that are more effective. Only then can a correct diagnosis be of use and unnecessary treatment be avoided.

There are several different alternative paradigms that have been developed to explain the lack of benefits observed from current surgical and other interventions for cancer.

One developed by Demicheli et al. [40] in relation to breast cancer suggests a systemic relationship between primary tumours and distant metastases such that surgical removal can interrupt tumour dormancy (a state of 'peaceful co-existence' with the host) and accelerate metastatic growth. It suggests that the tumour–host relationship needs to be considered before surgery is contemplated.

A second developed by Devitt [41] in relation to breast cancer suggested a different explanation for the lack of validity of the 'seed and soil' hypothesis underlying the concept of metastasis. He suggested that 'metastases' are not in fact the result of spread from a primary tumour but rather a later expression of the same systemic cancer process in tissue that is less susceptible to tumour growth than that where the primary is first seen. In this concept breast cancer is simply defined as the type of cancer for which breast tissue is the most susceptible site for growth. This would explain why some metastases appear similar to primary breast tumours.

A third alternative paradigm as described by Issels [42] sees the tumour as a late-stage symptom or element of a systemic disease resulting from a gradual breakdown of several body systems including metabolic, endocrine, digestive and immune system triggered of by several factors, including some prior to birth. Issels saw 'psychic' factors as a relatively minor contributory factor to this breakdown. In a more recent analysis Benjamin has explored the influence of these psychological factors and suggested including chronic emotional stress, particularly at the subconscious level as a major contributory factor to this breakdown, with nutritional factors as possible additional minor factors [43]. The mechanism for this systemic degeneration could include the influence of chronic stress on the shortening of telomeres, possibly also involving the immune system [44,45].

Several RCTs by Eysenck and Grossarth-Maticek have evaluated treatments based on this paradigm. Those with positive results show benefits far exceeding those achieved using treatments based on the current paradigm. For psychotherapy the suggested increased 5-year survivals range from 32% to 64% [46]. Similar but lower increases were observed with immunotherapy [47].

Apart from RCTs there is also good evidence that other forms of immunotherapy, such as the Issels Wholebody therapy [42] had produced better survival results in the 1960s than anywhere else in the world. It is claimed to have produced 16.7% 5-year survivals and 15% 15-year survivals among late stage cancer patients, 5–8 times higher than could be achieved at the time using surgery or other therapies based on the current paradigm [48].

Discussion

It would therefore appear that the figure of about 60% five-year survival often quoted for cancer intervention in general, and claimed to have increased in recent years as a result of improved treatment, represents what survival would have been without intervention, i.e. the natural survival without treatment.

Diagnosis of cancer under a situation where the diagnosis is to be followed by surgery, particularly a diagnosis following cancer screening, can only result in over-treatment and lead to harm [49].

This can only be avoided if a therapy is used based on a valid paradigm of what cancer is.

Conclusion

There have been no randomised controlled clinical trials (RCTs) showing the survival benefits of surgery in the treatment of cancer.

None of the seven indirect methods of evaluating the efficacy of surgery showed any clear survival benefits although one involving the removal of pre-malignant tissue suggested that surgery might be of benefit in the *prevention* of cancer.

Evidence from other RCTs evaluating treatments for cancer based on an alternative paradigm shows significant benefits. Non-randomised evidence supports such benefits.

It would therefore appear that the current paradigm of what cancer is, a local disease that later spreads, is invalid.

Until better evidence is produced to support the use of surgery, apart from for tumours that are immediately life-threatening, any treatments used should be restricted to those shown in randomised controlled trials to have a significant benefit in terms of increased survival or reduced mortality.

Conflict of interest statement

None.

References

[1] Cutler DM. Are we finally winning the war on cancer? J Econ Perspect 2008;22(4):3–26.
[2] Bailar JC, Gornik HL. Cancer undefeated. NEJM 1997;336(22):1569–74.
[3] Benjamin DJ. The efficacy of surgical treatment of cancer. Med Hypotheses 1993;40(2):129–38.
[4] Gompertz B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining value of life contingencies. Philos Trans R Soc 1825;115:513.
[5] Jones HB. Demographic considerations of the cancer problem. Trans NY Acad Sci 1956;18(4):298.
[6] Fox MS. On the diagnosis and treatment of breast cancer. JAMA 1979;241(5):489–94.
[7] Brinkley D, Haybittle JL. Long-term survival of women with breast cancer. Lancet 1984;1:1118.
[8] Fisher B et al. Ten year results of a randomised clinical trial comparing radical mastectomy and total mastectomy with or without radiation. N Engl J Med 1985;312(11):674.
[9] Fisher B et al. Five year results of a randomised clinical trial comparing total mastectomy and segmental mastectomy with or without radiation in the treatment of breast cancer. N Engl J Med 1985;312(11):665.
[10] Fisher B et al. Eight year results of a randomised clinical trial comparing total mastectomy and lumpectomy with or without irradiation in the treatment of breast cancer. N Engl J Med 1989;320(822):828.
[11] Fisher B et al. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. N Engl J Med 2002;347(16):1233–41.
[12] Veronesi U et al. Comparing radical mastectomy with quadrantectomy, axillary dissection and radiotherapy in patients with small cancers of the breast. N Engl J Med 1981;305(1):6.
[13] Sarrazin L et al. Conservative treatment versus mastectomy in breast tumors with macroscopic diameter of 20 millimeters or less. Cancer 1984;53(5):1209–13.
[14] Shimkin MB. The numerical method in therapeutic medicine. Public Health Rep 1964;79(1):1–12.
[15] Hegarty J., et al. Radical prostatectomy versus watchful waiting for prostate cancer (Review). Copyright©; The Cochrane Collaboration. John Wiley & Sons, Ltd; 2012.
[16] Black W, Haggerstrom D, Welch HG. All-cause mortality in randomized trials of cancer screening. J Natl Cancer Inst 2002;94(3):167–73.
[17] Bill-Axelson A, Holmberg L, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. N Engl J Med 2011;2011(364):1708–17.
[18] Benjamin DJ. The efficacy of surgical treatment of breast cancer. Med Hypotheses 1996;47(5):389–97.
[19] Saylor PJ, Smith MR. Adverse effects of androgen deprivation therapy: defining the problem and promoting health among men with prostate cancer. J Natl Compr Canc Netw 2010;8(2):211–23.
[20] Garnick MB, Pratt CM, Campion M, et al. The effect of hormonal therapy for prostate cancer on the electrocardiographic QT interval: phase 3 results following treatment with leuprolide and goserelin, alone or with bicalutamide, and the GnRH antagonist abarelix. J Clin Oncol 2004;22.
[21] Efstathiou JA, Bae K, Shipley WU, et al. Cardiovascular mortality after androgen deprivation therapy for locally advanced prostate cancer: RTOG 85-31. J Clin Oncol 2009;27:92–9.
[22] Enstrom JE, Austin DF. Interpreting cancer survival rates. Science 1977;195:847–51.
[23] Greenberg DS. Medicine and public affairs. "Progress" in cancer research–don't say it isn't so. N Engl J Med 1975;292(13):707–8.
[24] U.S. Congress General Accounting Office. Cancer Patients Survival; What progress has been made? PEMD-87-13, (3/31/87).
[25] Welch HG et al. Are increasing 5-year survival rates evidence of success against cancer? JAMA 2000;283(22):2975–8.
[26] Cairns J. The cancer problem. Sci Am 1975;233(5):64–78.
[27] Baum M. False premises, false promises and false positives–the case against mammographic screening for breast cancer. Int J Epidemiol 2004;33(1):66–7.
[28] Morrell S, Barratt A, Irwig L, Howard K, Biesheuvel C, Armstrong B. Estimates of overdiagnosis of invasive breast cancer associated with screening mammography. Cancer Causes Control 2010;21(2):275–82.
[29] Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med 2012;367:1998–2005.
[30] Autier P et al. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. BMJ 2011;343:d4411.
[31] Jørgensen KJ, Zahl P-H, Gøtzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. BMJ 2010;340:c1241.
[32] Johannesson G et al. The effect of mass screening in Iceland, 1965–74, on the incidence and mortality of cervical carcinoma. Int J Cancer 1978;21:418–25.
[33] Benedet JL et al. A comprehensive program for cervical cancer detection and management. Am J Obstet Gynecol 1992;166:1254–9.
[34] Cuzick J, Stewart H, Rutqvist L, et al. Cause-specific mortality in long-term survivors of breast cancer who participated in trials of radiotherapy. J Clin Oncol 1994;12(3):447–53.
[35] Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2013;6:CD001877.
[36] Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. Cochrane Database Syst Rev 2013;1:CD004720.
[37] Boyle P, Brawley OW. Prostate cancer: current evidence weighs against population screening. CA Cancer J Clin 2009;59:220–4.
[38] Haines IE, Gabor Miklos GL. Prostate-Specific Antigen Screening Trials and Prostate Cancer Deaths: The Androgen Deprivation Connection. J Natl Cancer Inst 2013;105(20):1534–9.
[39] Buys SS et al. Effect of screening on ovarian cancer mortality: the prostate, lung, colorectal and ovarian (PLCO) cancer screening randomized controlled trial. JAMA 2011;305(22):2295–303.
[40] Demicheli R, Retsky MW, Hrushesky WJ, Baum M, Gukas ID. The effects of surgery on tumor growth: a century of investigations. Ann Oncol 2008;19(11):1821–8.
[41] Devitt JE. Breast cancer: have we missed the forest because of the tree? Lancet 1994;344:734–5.
[42] Issels J. Cancer: a second opinion. London: Hodder and Stoughton; 1975.
[43] Benjamin D. The role of psychotherapy in the treatment of cancer. J Altern Med Res 2010;2(4):429–38. https://www.novapublishers.com/catalog/product_info.php?products_id=26210.
[44] Epel ES, Blackburn EH, et al. Accelerated telomere shortening in response to life stress. Proc Natl Acad Sci USA 2004;101(49):17312–5.
[45] Cawthon RM et al. Association between telomere length in blood and mortality in people aged 60 years or older. Lancet 2003;361(9355):393–5.
[46] Eysenck HJ, Grossarth-Maticek R. Creative novation behaviour therapy as a prophylactic treatment for cancer and coronary heart disease: part ii – effects of treatment. Behav Res Ther 1991;29(1):17–31.
[47] Grossarth-Maticek R, Kiene H, Baumgartner SM, Ziegler R. Use of Iscador, an extract of European mistletoe (*Viscum album*), in cancer treatment: prospective nonrandomized and randomized matched-pair studies nested within a cohort study. Altern Ther Health Med 2001;7(3):57–66 (68–72, 74–76).
[48] Issels J. Immunotherapy in progressive metastatic cancer – a fifteen-year follow-up. Clin Trials J 1970:357–65. with editorial by Phillips S. Dr Joseph Issels and the Ringberg Klinik. Clin Trials J 1970: 355–356.
[49] Welch GH. Overdiagnosed: making people sick in pursuit of health. Boston: Beacon Press; 2011.