# "Risk-adjusted Cancer-Screening"

A paper drafted by the sub-working group on "Risikoadaptierte Krebsfrüherkennung" (Risk-adjusted Cancer-Screening) of Working Group 1 "Weiterentwicklung der Krebsfrüherkennung" (Further Development of Early Cancer Detection) under the **German National Cancer Plan**

Speaker of the Sub-Working Group: Prof. Rita Schmutzler

Members: Dr. Dominik Dietz, Prof. Karl-Heinz Jöckel, Dr. Rüdiger Greinert, Prof. Wolfgang Hoffmann, Prof. J. Riemann, Prof. Bernt-Peter Robra, Prof. Peter Albers, Dr. Stefanie Thomas

**4th October 2011**

## Contents

# Risk-adjusted screening

## 1. General problems and objectives

Cancer screening (CS) aims to identify certain types of cancer or their precancerous stages in individuals who consider themselves healthy, i.e. who do not recognise any symptoms of disease as such. By definition, these can be at a presymptomatic stage or may include a disease stage that, while already symptomatic, has not been identified as such (Wilson and Jungner, 1968). The rationale of advancing the time of diagnosis is to improve patient outcomes and to stop the disease from progressing including the prevention of possible complications. Ultimately, the aim is to reduce the mortality and morbidity associated with the type of cancer screened for and to improve quality of life. Cancer screening programmes implemented so far under section 25 of Social Code Book V (§ 25 SGB V) have been designed to test a target population which is mainly made up of people with an average risk to develop the disease Therefore, the screening test being used is appropriate for the majority of individuals in this population.

The recent advances in the sequencing of the human genome and the discovery of genes associated with an increased risk of cancer, paved the way for identifying risk groups. Moreover, it is likely that additional risk indicators will be identified in the future and that the option of genetic screening will become increasingly important. By adopting the Act on Genetic Testing (*Gendiagnostikgesetz - GenDG*) on 24 April 2009, the German legislator sought to effectively address the issues around genetic testing. In view of the developments in human genome research, citizens are to be empowered to exercise their right to informational self-determination, that is the right of the individual to decide what information about him-/herself should be communicated to others and under what circumstances. The Act aims to preclude the risk of genetic discrimination that might result from the diagnosis of certain genetic traits etc. while, at the same time, securing the potential of genetic testing for the individual. The Act includes, *inter alia*, specific provisions on informed consent and genetic counselling and also specifies requirements for good genetic testing practice.

This paper specifically refers to the procedures which the Act defines as 'genetic screening' - insofar as routine genetic screening is to be routinely offered as part of a risk-adjusted screening programme. Genetic screening is defined as 'genetic testing for medical purposes that is systematically offered to the entire population or specific

segments of the entire population without the person affected necessarily having reason to believe that he or she has the genetic traits tested for" (section 3 (9), 'Definitions'). In line with the requirements for statutory programmes for the early detection of diseases laid down in section 25 of the Social Code, the Genetic Testing Act stipulates that "genetic screening may only be performed if its purpose is to identify whether or not the individuals concerned have genetic traits associated with a disease or health disorder that is avoidable or treatable or that can be prevented" (section 16(1) of the Genetic Testing Act. Moreover, the Act clearly says that "genetic screening may only be undertaken after the Committee on Genetic Testing has assessed the screening programme involved in writing (section 16(2) Genetic Testing Act). "Their opinion is not binding but is a recommendation" (Special Part B. of the Act, p. 45).
An exception exists if diagnostic genetic testing is done in individuals who are already ill in order to "diagnose a symptomatic disease or health disorder" (section 3(7) Genetic Testing Act). This might be the case where a breast cancer patient's personal or family history strongly suggests a hereditary predisposition, so that genetic testing would be indicated. In the language of the Genetic Testing Act, however, these would be "diagnostic genetic tests" but not "screening tests" even if, under certain conditions, this test could actually be part of a risk-adjusted screening programme.

For individuals known to be at an elevated risk for certain tumour diseases, general cancer screening programmes are not appropriate or start too late in life. Depending on the circumstances, these individuals might benefit from risk-adjusted or targeted CS in order to reduce the morbidity and mortality risk associated with the disease. This requires high-risk individuals to be identified which can be done by screening for certain risk indicators[1].

Beyond the generally recognised principles for the implementation of CS programmes (e.g. Wilson and Jungner 1968), additional requirements have to be fulfilled to do justice to the complex concept of risk-adjusted CS (Andermann *et al*. 2008). Specifically, according to the Genetic Testing Act, any risk-adjusted CS programmes that involve genetic testing must be part of an appropriate and target group specific counselling concept. This also means that, in line with non-directive counselling, the "right not to know" must be respected and any discrimination, as set out in the Genetic Testing Act, due to the identification and documentation of certain genetic characteristics must be

[1] For the distinction risk indicator vs. risk factor see 4.2

prevented. Apart from valid diagnostic risk indicators there has to be a typical clinical presentation (e.g. histopathological subtype, age of onset, course of disease). Furthermore, before targeted CS tests can be justified, these tests require a sufficient evidence base for their effectiveness in this particular risk group.

Part I of this Paper relates risk-adjusted cancer screening to the overall concept of population-based screening. However, it does not consider certain constitutional characteristics (such as number of pregnancies) or behavioural and environmental risks. These should be addressed primarily by primary prevention activities. On the other hand, the methodological framework and the clinical-epidemiological discussion do not exclusively apply to genetic risk indicators. The approach taken is also valid with regard to constitutional, behavioural and environmental risk indicators.

Furthermore, the paper systematically presents the specific aspects of and requirements for risk-adjusted screening and the potential risk and benefits inherent in such an approach. Thus, it goes on to establish general criteria for assessing the potential risks and benefits associated with risk-adjusted screening (Part II).

## 2. Risk-adjusted versus general cancer screening

In Germany, the statutory screening programmes implemented in line with section 25 of Social Code Book V, the "Early Cancer Detection Guideline" (*Krebsfrüherkennungs-Richtlinie*) and the "Health Check-up Guidelines" (*Gesundheitsuntersuchungs-Richtlinien*) drafted by the Joint Federal Committee are based on a "risk population" that is currently defined only in terms of age and sex. All individuals covered by the statutory health insurance who are at the relevant age and have not been diagnosed as having the target condition, are eligible or encouraged to take up the gender-specific screening tests. There is no regular risk stratification beyond that, e.g. in the form of history-taking or laboratory tests. This type of screening will hereafter be referred to as "general cancer screening".

Unlike 'general cancer screening', 'risk-adjusted cancer screening' relies on further factors in addition to age and gender in order to identify healthy individuals whose risk for the presence or development of a specific tumour disease is clearly above the normal risk solely associated with their age or gender. These are mostly individuals with a hereditary risk for certain cancers. It is true that only a few high-risk genes for some tumours have been identified so far. Recent scientific developments, however,

open up possibilities for detecting a host of additional risk genes through genome-wide high-throughput analyses, although their contribution to raising cancer risks may vary and their benefit for risk-adjusted screening remains to be confirmed on a case-by-case basis. Moreover, evidence suggests that several such genes interact with one another and with environmental factors to produce a cumulative overall risk. Once these links are unravelled, individual risk calculations may become possible. It would open up many new aspects for CS that range all the way from the chance of targeted prevention in high-risk groups to the danger of social discrimination against individuals who are at an elevated risk.

The aims of risk-adjusted CS are, as with general CS, reducing both mortality from the target condition and overall mortality in the risk group, lowering morbidity rates, enhancing quality of life and optimising the efficiency of the health care system. Prior to the introduction of risk-adjusted CS, there must be demonstrable proof of its capacity to attain these aims. The common approach to obtaining proof is prospective randomised studies. If evidence of benefit can be provided, accompanying prospective process and/or outcome-oriented evaluations of risk-adjusted CS may be required depending on the structure of the CS involved.

Generally, it must be borne in mind that any benefit of a screening test in the general population cannot be generalised to a group whose risk profile is different. That is a greater benefit of a "stepped up" screening activity in a risk group must not automatically be taken for granted. The very specific problems in evaluating risk-adjusted CS will be discussed below.

### 3. Multi-step concept of risk-adjusted CS

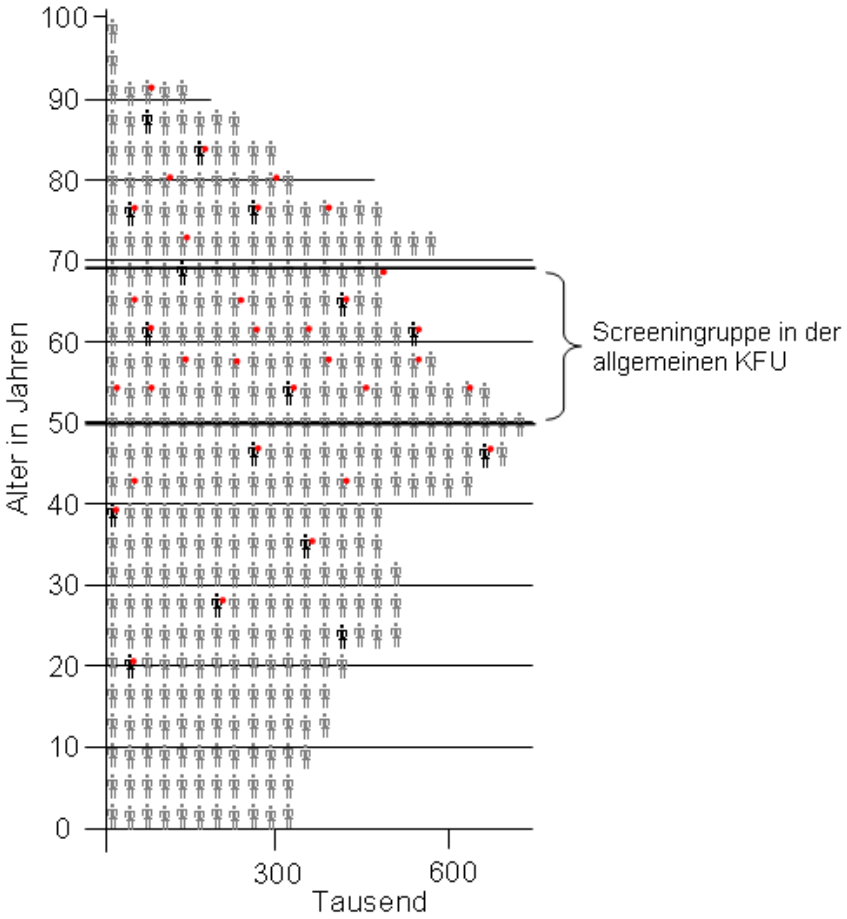
Risk-adjusted screening involves, first of all, the analysis of risk indicators to identify a relatively small number of individuals at risk who - without intervention - would account for a relatively large number of all disease cases occurring in the entire population (Fig. 1). This 'filtering process' that precedes the actual screening test can take place at various points throughout the individuals' lifetime and differ in terms of both intensity and duration. At its most simple, risk stratification is done, for instance, merely by documenting details of the individual's (personal and family) history, e.g. in form of a validated questionnaire, to assess familial colorectal cancer risk. In more complex cases, this is followed up by additional measures to quantify the risk involved, maybe

down to specific genetic testing. So far, this systematic approach to risk-adjusted screening is not part of the statutory health insurance system. An exception are certain joint projects (such as familial breast-ovarian cancer or familial colorectal cancer) and the respective contractual agreements between purchasers and providers of the service.

The fact that risk-adjusted CS is a two-step process and/or includes an additional filtering process, opens up new opportunities, but also new risks compared to conventional screening programmes. In addition to the risks and benefits inherent in early disease detection - that tends to be more intensive once an individual has been included in a 'high-risk group' - the upstream filtering process can imply additional risks, but also opportunities. This is largely due to the fact that individuals are no longer 'automatically' assigned to a 'risk group' merely by reaching a gender-specific age, thus joining their entire age group among the population, but specifically by means of certain indicators. This highlights the 'exceptionality' or 'deviation from the norm' for the individuals who find themselves at high risk compared to the average person.

A distinction, though, must be made between the concept of risk-adjusted screening and the medical management of patients whose disease is associated with a clearly increased risk of certain (secondary) conditions (e.g. higher prevalence of malignant skin tumours among immunosuppressed patients). Here, dealing with this risk is important and has to be taken into account by the physician in charge at all stages of "routine" care.

*Figure 1:* Distribution of females at high and normal breast cancer risk and incident cases in Germany 2011




Legend: As far as the high risk population is concerned, not surprisingly, here the disease occurs more frequently compared to the average risk population. Also breast cancer in the high risk population develops more often before and after the target age range of the routine mammography screening programme (50 to 70 years).

(Assumptions: The prevalence of high risk is 5% across all age groups; the proportion of incident breast cancers among the general population are as follows: 25% under the age of 50 years, 50% between 50 and 70 years, 25% above the age of 70 years)

Symbols: Individuals at normal risk, individuals at high risk and thus members of the target group of risk-adjusted CS, ● incident cases

## 4. Approach and methods for the selection of suitable risk indicators

*Methodology*

To be able to identify individuals with an increased risk of disease there have to be clearly defined risk indicators. These have to be of sufficient validity and quantifiable with reference to the penetrance of the disease, the course of disease and with reference to the specific requirements for diagnosis and therapy.

A multi-step process is commonly used to characterise risk indicators:

1) Identification of one or several risk indicators.
2) Definition and validation of each individual risk indicator with respect to the occurrence of the target condition and/or its course of disease.
3) Analysis of the significance of each risk indicator in the context of all known risk indicators and additional determinants.
4) Prospective validation of the predictive value of the first "filtering test" (e.g. family tree) for the presence of one or more risk indicators; if appropriate, prospective validation of the predictive value of a second filtering test (one or several risk indicators, e.g. a risk gene) with reference to the manifestation of the target diseases.
5) (Piloting) the implementation and prospective validation of risk-adjusted CS in the target group.
6) Providing proof of a significantly higher benefit of risk-adjusted CS – on the basis on one or several risk indicators - compared to general CS.

*Risk indicators and risk factors*

A risk indicator must be distinguished from a risk factor. While a risk indicator only informs about the association with a disease, a risk factor also involves a causal link. Risk indicators can, but need not, be concurrent risk factors for the manifestation of the given disease. Generally, the identification of a risk indicator only allows the mere probability of the future manifestation of a disease to be assessed through population-based observations. The task of a risk indicator is to identify the risk group among the population. A risk factor is a causal factor. Risk factors can be grouped into *environmental, behavioural and genetic* (constitutional) factors. Of these, behavioural and environmental risk factors can be modified primarily through primary prevention. By contrast, genetic factors are currently not or hardly modifiable and can be associated with a clearly elevated penetrance of the given disease. If the penetrance of the disease

is high, so is the prevalence of the disease in the risk group characterised by the genetic factor. These properties uniquely qualify genetic risk factors as risk indicators in the context of risk-adjusted screening.

In this context, the Genetic Testing Act differentiates between predictive-deterministic and predictive-probabilistic tests. Predictive-deterministic gene mutations are those that will almost certainly cause a disease to become clinically apparent in the carrier's lifetime. A case in point would be Huntington's disease. By contrast, predictive-probabilistic tests identify genetic mutations that carry a far smaller likelihood of a disease manifestation (penetrance). Such testing can at best assess the probability of a later manifestation of a disease (Genetic Testing Act b) Special Part p. 17).

*Suitable risk indicators for risk-adjusted cancer screening*

Risk indicators to be used in the context of risk-adjusted screening should fulfil the following requirements:

1) Reliable identification of individuals with an elevated - to be defined - likelihood of testing positive for the risk indicator, e.g. by family and family tree analysis (filtering process).
2) Strong association of the risk factor with the clinical manifestation of the disease, e.g. there is a strong positive and negative predictive value (PPV and NPV) for the disease[2].
3) The risk indicator is reliable and allows valid testing.
4) The stress on the subject from being tested is negligible
5) Testing is cost-effective compared to the cost of cancer screening.

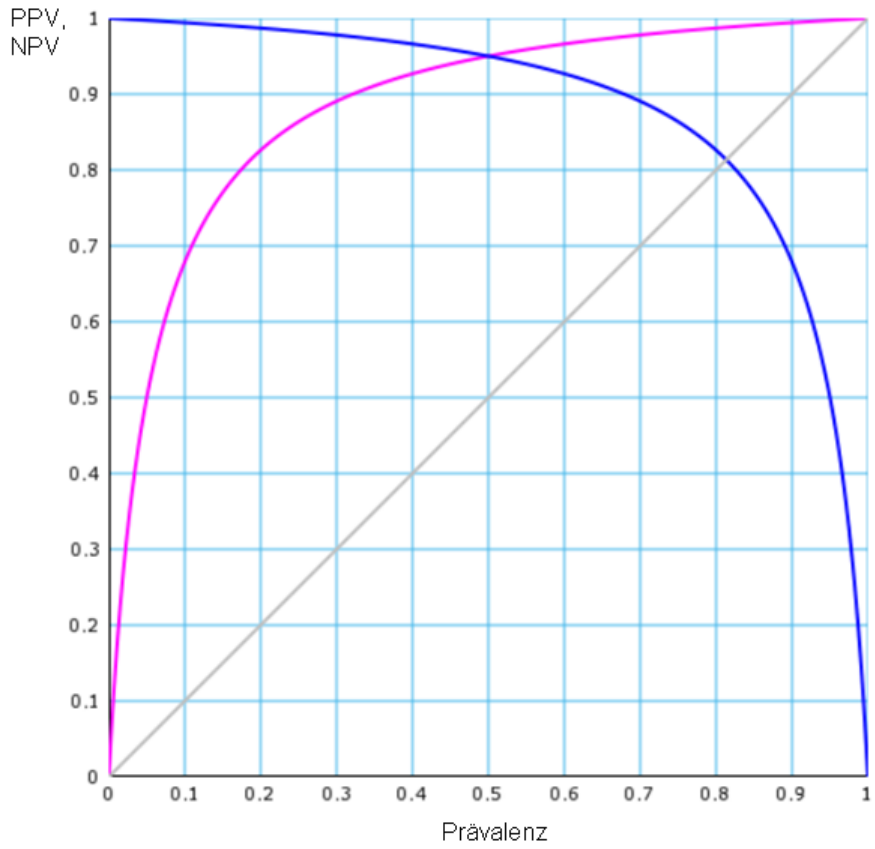
*The screening strategy: Prevalence as a key variable*

[2] The positive predictive value (PPV) is the proportion of persons correctly tested positive (= manifest disease within a defined period of time) among all screening participants tested positive (= risk indicator positive) The negative predictive value (NPV) is the proportion of persons correctly tested negative among all screenees tested negative. In the case of a genetic risk factor, the PPV equals the penetrance of the genotype (see also the "*Richtlinien*" of the Committee on Genetic Testing (*Gendiagnostikkommission*).

A key parameter for the implementation of a CS programme is the prevalence of the condition to be screened for ( which is by definition preclinical and has not caused any symptoms yet).

The prevalence essentially determines the relationship between false-positive and true-positive screening results and, implicitly, the risk-benefit ratio and the cost-benefit ratio of each screening strategy as well. This is because the PPV rapidly increases as prevalence rises when using a test with a fixed test validity (sensitivity, specificity). The graph in Figure 2 shows the ratio between prevalence and PPV and, inversely, between prevalence and NPV at a given sensitivity and specificity of the screening test.
More specifically, the graph demonstrates that in the area of low or very low prevalence rates of the condition to be screened for the PPV is particularly sensitive to any changes. In other words the prevalence of the condition of interest has got a strong influence on issues around costs, benefits and efficiency of screening activities. Therefore, the rationale of restricting screening to a population at risk with high prevalence rates is to improve the cost benefit ratio.

*Figure 2: Positive and negative predictive value in relation to disease prevalence at a given validity (here: test sensitivity 95% and specificity 95%)*




Before establishing a risk-adjusted screening programme cut-off levels of the filtering test and the screening test proper have to be defined clearly. The above text illustrates that defining cut-off levels of prevalence rates in different populations with reference to possible cost-benefit ratios is one of the biggest challenges in establishing a risk-adjusted cancer screening programme. These analyses have to be scientifically sound, evidence-based and transparent. Also they require the involvement of the relevant stakeholders and decision makers to achieve the necessary consensus.

*The definition of risk groups*

The most important method for defining a risk group is stratification of the population into (not less than) two groups that (compared to the general average) have a lower risk and a higher risk (risk group). For age-specific target conditions, this is done along the age-specific prevalence gradient (with a lower and an upper age limit). High-prevalence individuals can also be identified by looking for additional risk indicators. Once the clinical presentation (phenotype) of these risk indicators is known, more targeted,

comprehensive or closely spaced CS tests may be appropriate to lower the mortality risk associated with the disease. Therefore, this kind of approach requires risk indicators to be defined and the predictive value of each risk indicator to be quantified with reference to the risk of disease.

Risk groups can be identified by performing a test that precedes the actual screening test: by asking about a positive family history and/or determining a genetic marker that is elevated in populations with an elevated disease risk (prevalence marker). Technically, stratification by prevalence amounts to a diversification of possible care pathways. The various groups differ in terms of their respective baseline prevalence and thus (if the intervention has a given effectiveness) in terms of the risk-benefit ratio that is possible in each case. Different risk-benefit quotients can lead to a situation where clinical care is differentiated according to specific target groups by, for instance, requiring a higher diagnostic sensitivity for high-prevalence groups (i.e. settling for lower specificity), while reducing the aggressiveness of follow-up diagnostic procedures in low-prevalence groups. In extreme cases, only one of these (two) groups is defined as eligible.

If prevalence rates and test properties are altered, this can change the detectable preclinical phase and, as a result, other characteristics of the target condition as well, such as treatability depending on the aggressiveness of changes and disease progression in the specifically defined CS target groups.

Another option would be stratification not merely by *prevalence markers*, but additionally by "*progression markers"*. These are indicators of fast (or slow) growing changes, such as cytological or molecular markers of the degree of cellular differentiation, certain metabolic profiles of suspect cells, presence of chronic inflammation etc. Stratification by progression markers differs from stratification by prevalence markers. Specifically, estimates on the intervention effect (fast growing tumours tend to be less amenable to treatment than slow growing tumours) and follow-up intervals have to be adjusted accordingly.

This paper focuses on populations at risk due to constitutional or genetic factors. However, it can also be applied to other groups at an increased risk, such as individuals in whom the sensitivity/specificity of the test is affected (e.g. radiopaque breast tissue in mammography, patients with known comorbidities such as haemorrhoids in FOBT, respectively) or where the responsiveness to therapy is lower.

Where robust evidence from randomised clinical studies is not sufficiently differentiated to assign valid effect estimates to all plausible variants of a population-based screening programme, simulation models should be used to supplement it. This requires first of all sound information on the various epidemiological data, test properties and, finally, on the effectiveness of preventive and therapeutic interventions in the various target groups. On this basis, both disease progression at an individual level and the purpose of a target group-specific screening programme at population level can be extrapolated as precisely as possible using all available information. In the final analysis, the effectiveness and efficiency of a screening programme can be assessed and followed up with appropriate evaluation tools after its implementation (where appropriate using data from cancer registers).

## 5. Societal, ethical and legal aspects of risk-adjusted CS

The statutory health insurance does not yet provide for routine risk-adjusted cancer screening. The issues that are implicated in connection with such an approach also include societal and socio-legal aspects. These can directly concern the individual directly, but can also have implications for their relatives, society or members of sickness funds/insurance companies in general.

While conventional CS only uses age and gender to classify an individual as being at risk, so that each member of the statutory health insurance system beyond a certain age is considered a risk person, risk-adjusted screening draws on additional factors. Consequently, the resulting classification as a risk person puts a spot light on being "deviant" from the reference age or gender cohort.

There is a host of highly complex issues to be addressed in this context. One important question is to what extent a healthy person after having been labelled as belonging to a risk group can be protected from any social and socio-legal disadvantages. The identification of risk indicators, such as deleterious changes (mutations) in risk genes, allow the likelihood of manifestation to be predicted long before the disease actually manifests itself. Moreover, genetic analyses might allow conclusions to be drawn about the genetic status of relatives who never consented to genetic testing. The Genetic Testing Act that was adopted by the German *Bundestag* on 24 April 2009 and supersedes the voluntary moratorium of insurance carriers and clearly stipulates that insurance companies may not use any information derived from genetic testing. The

only exeption are insurances with a policy cover of more than 300,000 euros or an annual annuity of more than 30,000 euros. With respect to labour law, too, the Genetic Testing Act clearly stipulates that the right of self-determination and the protection of the rights of the individual take precedence over the employer's interest in having productive staff that continue to be healthy in the future (prohibition of discrimination).

Currently, there are specific and intensified screening programmes only for well-defined high-risk groups on the basis of individual risk indicators (e.g. mutation in a high-risk gene for hereditary breast-ovarian cancer or hereditary colorectal cancer). It is likely that several other, especially genetic risk indicators will be identified that only confer a moderately elevated risk. In this context, it begs the question who, and on what basis, should define the cut-off level that classifies a healthy individual as belonging to a risk group. Furthermore, what will be the policy towards those people whose individual risk, while above average, does not exceed the specifically defined cut-off level.

Defining risk groups is likely to have implications on screening practice in the general population. For example, the pooling of individuals with a high risk might alter the predictive values of the screening test in the average risk group requiring for example changes of the screening intervals.

In the case of prostate cancer, the definition of a high-risk group could lead to the reduction of the currently rampant uncontrolled ("grey") screening using prostate-specific antigen (PSA) tests. Here, large retrospective studies suggest that a single test with an elevated baseline PSA value performed in men between 44 - 50 years of age is of relevance for the definition of a high-risk group. Men with an increased baseline PSA level account for the majority of prostate cancer deaths later in life. Currently, however, healthy men from the age of 40 are recommended to undergo PSA screening without any prospectively tested risk evaluation. This PSA testing is not part of the statutory health insurance system and is covered by out of pocket payments. As a consequence there is a high rate of false positive results which is associated with needless anxiety and significant costs for the statutory health insurance system due to follow-up treatment etc.. The prospective evaluation and potential introduction of a screening strategy could help prevent unnecessary screening tests and, thus, ultimately ease the financial strain on the statutory health insurance funds.

Nor is it currently known how best to address the problem of behavioural or lifestyle-related risk indicators or factors in the long term. Risk factors such as lack of physical

exercise, obesity and smoking tend to raise the individual (baseline) risk of certain diseases. As a result, individuals with certain risk indicators, can, theoretically, carry a similar risk for a given disease as individuals in some risk groups that are defined by family history/genetic traits. It is also conceivable that, although the elevated risk revealed by testing is still below the cut-off level, this risk is raised further by the individual's behavioural risk factors. Even if behavioural risks are not (initially) considered in the context of risk-adjusted screening, this problem will have to be addressed in the future. Risk-adjusted CS touches on a number of additional ethical and legal issues that are not discussed here in depth but are relevant nonetheless (e.g. non-inclusion in the risk group due to inconclusive family history, right to early detection examinations while rejecting genetic testing, right to early detection examinations while invoking the right not to know).

## 6. Validated diagnostic procedures for RACS

Diagnostic tests are usually described by stating both sensitivity and specificity. Generally speaking, sensitivity is the proportion of individuals (correctly) tested positive of all people with the disease, while specificity is the proportion of individuals with (correctly) unremarkable test results of all people without the disease. When evaluating diagnostic measures for early disease detection, the test results of 'individuals correctly identified as ill' should be specified. Attention should be drawn to the fact that the detected early stage can either become clinically manifest or stay inconsequential (cf Table 1: six-cell table for cancer screening, based on Raffle & Gray 2007). Only in the first case would a relevant disease be screening-preventable, in the second case, there would be no benefit.

| | | **'Truth'** | | |
|---|---|---|---|---|
| | | *pathological finding of early stage cancer that would become clinically manifest* | *Pathological finding of early stage cancer that would remain in a latent state* | *No pathological finding* |
| **Out-come of CS test** | *positive* | true positives (TP) | 'true' positives ('TP')* | false positives (FP) |
| | *negative* | false negatives (FN) | 'false' negatives ('FN') | true negatives (TN) |

Table 1: Six-cell table for cancer screening tests
* 'overdiagnoses' (individuals testing positive whose early stage cancer would remain in a latent state)

*6.1. – Incorrect generalisation of test quality estimates – the validation of CS tests in risk groups is necessary*

In connection with risk-adjusted screening, it is plausible that screen-detectable tumours in the risk population might be particular tumour subtypes. Depending on the diagnostic procedure used for CS, this might influence its test quality, i.e. sensitivity and specificity in the specific risk group would differ from what would 'usually' be assumed with this test. For instance, a tumour might, in principle, be harder (or easier) to diagnose in a risk person than in the 'the average' individual, simply due to morphological reasons. Specifically, hereditary breast cancers frequently mimic benign criteria in imaging and are therefore more unlikely to be detected early than sporadic tumours. Moreover, hereditary subtypes can have different growth patterns and/or respond differently to therapy. For instance, BRCA1-associated breast cancers have a markedly higher growth rate than sporadic breast cancers which can influence the interval cancer rate. Moreover, they seem to respond differently to standard therapy than sporadic breast cancers. Consequently, the phenomenon that sensitivity and specificity can differ in various groups, may, if study results from other populations are uncritically transferred, lead to errors in estimating test quality in the risk population[3]. Therefore, the CS test must be validated in the (high-risk) population in which it is intended to be used. The magnitude and direction of transfer errors cannot be predicted in general terms, but depend on the type of target condition and the way the test works. This problem is particularly complex in the context of screening tests, since they must consider not only 'ill' and 'healthy' individuals, but theoretically also those who are diagnosed with early stage disease that would never become clinically manifest ('Overdiagnoses', Table 2). Here, specific knowledge of the special prognosis for the tumour disease in the risk population is necessary. Specific studies on the test quality of CS tests could suggest, for instance, that the tests used in risk populations should differ from those in used in general CS tests.

[3] "Spectrum-related sources of variation", cf. Whiting et al.., Annals of Internal Medicine 2004

| | | **Risk indicator present ('individuals at risk')** | | | **Risk indicator absent ('individuals without elevated risk')** | | |
|---|---|---|---|---|---|---|---|
| | | **'truth' for at-risk group regarding a pathological finding** | | | **'truth' for individuals without elevated risk regarding a pathological finding** | | |
| **Result of cancer screening test** | *positive* | true positives (TP) | 'true' positives ('TP')* | false positives (FP) | true positives (TP) | 'true' positives ('TP')* | false positives (FP) |
| | *negative* | false negatives (FN) | 'false' negatives ('FN') | true negatives (TN) | false negatives (FN) | 'false' negatives ('FN') | true negatives (TN) |

Table 2: Two six-cell tables showing CS of individuals at risk and individuals without elevated risk. The frequencies of the six cells vary depending on the ratio of individuals with or without risk indicator in a diagnostic study (TP, 'RP', FN, 'FN', TN).
* 'Overdiagnoses, (individuals testing positive for early stage cancer that would remain in the latent state)

Apart from issues around test quality considerations, it must be borne in mind that, in risk populations, histopathological parameters and clinical behaviour of tumour subtypes tend to differ considerably. These circumstances must be taken into consideration when adjusting adequate screening intervals for the regular implementation of CS to limit the risk of so-called interval tumours. Deliberations about appropriate screening intervals are all the more important since, within the framework of risk-adjusted CS, individuals younger than usual might be enrolled that will have to be repeatedly rescreened over the coming decades. It must be factored in that an individual's lifetime disease risk might change in ways other than normal.
Consequently, the selection of appropriate tests for risk-adjusted CS urgently requires specific, high-quality diagnostic studies in the target risk population.

*6.2. Treatability*

Moreover, the tumour disease's treatability in the risk group can differ from usual patient outcomes in the non-risk population, adding to the foregoing considerations regarding the specific properties of the target condition in the risk population that are liable to influence the test quality of CS. Hence, the benefit of risk-adjusted CS cannot be judged

merely on the basis of good detection rates (good sensitivity and specificity). Instead, it is necessary to verify patient outcomes in the risk population or the comprehensive strategy of risk-adjusted CS in terms of efficacy (i.e. reduction of mortality from the target condition) in randomised controlled studies prior to the introduction of such a strategy.

*Use of special tests in a small population with elevated prevalence*

While difficulties may arise for risk-adjusted CS due to unfavourable prognostic properties of the tumour disease and potentially inferior diagnostic properties of CS tests, the major advantage of the strategy is its focus on relatively small populations with an elevated prevalence.

Due to the elevated prevalence in a risk population, the positive predictive value is usually higher than in general cancer screening tests. In practice, this means that the use of a test with a fixed sensitivity and specificity is less likely to confront individuals with a false positive test result than would it be the case with general CS. As a consequence, other diagnostic tests may be used in the risk population than in general CS. In the context of comprehensive population-based CS, particular care must be taken to ensure that the test has a sufficiently high specificity to minimise the proportion of false positives among all positive test results (1-PPV). Too low a specificity would, in widely used tests and at low prevalence rates, lead to intolerable rates of false positive findings. This optimisation of specificity is usually associated with a loss of sensitivity. Enhanced PPV, combined with the less resource-intensive CS in a relatively small population, opens up possibilities for the use of less specific, but more sensitive tests. To achieve this, several diagnostic procedures can be used in parallel (**additive approach**). To further investigate positive results from a particularly sensitive risk adjusted test strategy, specific confirmatory tests might be warranted that, because of too large cohort sizes, would be too resource-intensive for general CS, and too stressful for those affected.

To sum up, it can be concluded that focusing on a special spectrum of individuals at risk can influence test quality and the target condition's amenability to treatment that together determine the potential benefit of risk-adjusted screening.

Possible aspects of populations at risk relate to:

1.) the specific test quality of the CS test (compared to the test quality in the normal population),

2.) the changed prognostic characteristics of the target condition, incl. growth rate (adjustment of special screening intervals),

3.) the potential for optimisation of the test strategy through improved framework conditions (prevalence, small size of the target population) by combining sensitive procedures with specific confirmatory tests,

4.) the differential effectiveness of available therapies in various patient groups.

Due to the factors set out here, implementation of specific studies is indispensable for the comprehensive evaluation of benefit in risk-adjusted CS, since the results from studies in unselected populations are not transferable to the risk population of interest.

## 7. Advantages and disadvantages of RACS compared to general CS

Like every CS, risk-adjusted CS has potential advantages and disadvantages. However, the screening risks involved in assignment to a high-risk group, are not fundamentally different from those involved in general CS. These include, in addition to the risks related to the test procedure itself, having to live longer with the disease without changing its prognosis, overdiagnosis and overtreatment of questionable findings, increased costs and false reassurance for participants with false-negative results. However, it must be borne in mind that predictive values rise due to the higher prevalence of the target condition in the high-risk group. Moreover, the more tightly spaced screening intervals and potential use of additional screening tests can cause differences compared to general CS. In addition tumours in the high-risk population may have a different tumour biology - tumours developing in this population might, for instance, have a higher growth rate and a different metastasizing potential.

Additional benefits, but also risks, arise from the **filtering process**, which is meant to assign individuals with a clearly higher risk for a certain type of cancer to the high-risk population. Here, it must be borne in mind that even apparently straightforward questions about disease in the individual's personal or family history may yield false positive or negative answers leading to a wrong classification. For instance, a systematic review on the use of family histories concludes that the specificity (i.e. correct information about the absence of cancer cases in the family) is relatively high

across all reviewed cancers (91-99%, according to type of cancer) while sensitivity rates are far more variable (e.g. breast cancer 58-90%, bowel cancer 57-90%)[4].

Potential benefits:

1) Focusing screening tests on populations at risk may reduce the need for general screening tests.
2) Including individuals in the risk population who - without intervention - will actually go on to develop the disease (by increasing the prevalence) will lead to an increase in the proportion of true positive test results.
3) Risk stratification in screening may increase the informed willingness to participate of those affected.
4) Assignment to a high-risk population may heighten awareness of the importance of primary prevention and a 'healthy' lifestyle.

Risks:

1) False assignment to a high-risk group ('false positive')
2) False assignment to the 'normal group' ('false negative')
3) Psychological stress as a result of being assigned to a high-risk group
4) 'Stigmatisation' through assignment to a high-risk group

Classifying a healthy person as belonging to a high-risk group can involve major stress and strain. In addition to the concomitant psychological (and sometimes physical stress), it might entail social and societal disadvantages in particular, e.g. when taking out high-benefit life insurances.

## 8. Summary

Since the western industrialised countries introduced various cancer screening tests and programmes in the 1960s and '70s, there have been many advances both in the methodological realm, e.g. improvement of diagnostic certainty due to new tests, the

[4] Evidence report/Technology Assessment No. 159, Collection and Use of Cancer Family History in

methodological consideration of lead-time-bias, overdiagnosis and overtherapy, and in the conceptional realm, e.g. identification of new risk indicators. Especially the identification of genetic risk indicators and the establishment of appropriate molecular genetic test procedures for the identification of risk individuals open up the possibility of risk-adjusted cancer screening. Focusing on risk groups goes along with the potential of establishing more efficient cancer screening programmes with fewer side effects. However, the evaluation of the entire process is a fundamental requirement: It has to start with the identification of the risk indicator and finish with the proof that the screening programme does reduce mortality in the target population.

More specifically, for the establishment of risk-adjusted CS the following requirements have to be fulfilled:

1. The morbidity and mortality associated with the disease develop at a *known and high probability* in the target population.

2. A *risk profile* is unequivocally *defined*, reliable and can be documented at a *reasonable* cost (usually two-step filtering test: 1. information from history-taking , 2. genetic test).

3. The target population can be unmistakably *identified* and classified through a *risk profile*. Also,an threshold level that has been consented by the relevant decision makers exists for the implementation of specific screening measures.

4. Risk classification does not lead to societal or socio-legal disadvantages.

5. Reliable *diagnostic procedures* are available that have been validated in the risk population (sensitivity, specificity, PPV, NPV, specific phenotype of the target condition in the risk population).

6. Earlier diagnosis due to the screening can improve the prognosis and, especially, lower mortality from the target condition among the population at risk (e.g. tumour biology).

7. The potential *disadvantages*, especially due to false positive results and overdiagnosis, are known among the population at risk and are acceptable with reference to a lower mortality.

8. Appropriate and *non-directive counselling* is provided that enables those affected to make an *informed decision* to participate in or refuse CS (counselling in two-steps).

## 9. Outlook and recommendations for implementation

Risk-adjusted CS carries advantages and opportunities, but also possible disadvantages and risks. Therefore, this paper offers a conceptional and methodological framework for the development and validation of population-based risk-adjusted cancer screening measures. Moreover, it aims to raise the awareness of possible limitations and the technical requirements for risk-adjusted cancer screening and to prevent the rash and uncritical implementation of insufficiently validated measures of risk-adjusted cancer screening.

This paper is addressed to the medical scientific community, particularly authors and editors of guidelines (in Germany: AWMF), scientific opinion leaders, institutions and bodies representing purchasers and providers, who are responsible for the implementation of cancer screening, (e.g. Joint Federal Committee) and developers of genetic tests or other tools that can be used for risk-adjusted CS (e.g. from industry and science). In order to reach the foregoing target groups, the publication of this paper is recommended, e.g. in the *Deutsches Ärzteblatt*, and internationally (e.g. BMJ or Lancet).

So far, the aspects set out here that must be considered when implementing a risk-adjusted screening have been largely theoretical. Therefore, the working group recommends that research activities in the field of risk-adjusted CS be intensified. In view of the enormous resources required for the development and evidence-based validation of risk-adjusted CS measures, available research and development resources should be pooled, e.g. in the framework of joint European projects. This also includes the joint development of European and/or international standards for the development and validation of risk-adjusted CS measures. This paper could provide an important German contribution towards such a European co-operation project. Therefore, consideration should be given to including this topic into the European Partnership for Action Against Cancer (e.g. Work Package 6 'screening and early detection') as Germany's contribution. For instance, a workshop could be held with European early cancer detection specialists.
In addition to the concrete application and validation of the criteria outlined here using

specific tumour entities, discussion about the related social political and ethical aspects should be advanced, as well. This includes, for instance, questions of who should be authorised to define a cut-off level and on what basis, how to deal with moderate risk, and a life-style-related increase in risk, what health implications are linked to being included in a risk group and what economic consequences ensue. This calls for discussion at population-wide level and a multidisciplinary approach.

**Annex, part of the recommendations for implementation**

## II. Synopsis on the assessment of risk-adjusted (cancer) screening measures

### Preamble

The concept of risk-adjusted screening focuses on risk factors that do not refer to or are not modifiable by personal behaviour or lifestyle. This set of health determinants should be tackled primarily through primary prevention measures.

However, it must be borne in mind that lifestyle-related and environmental risk factors can interact with hereditary risk factors and contribute to the overall risk. Conversely, lifestyle can also have a protective effect and offset hereditary risk factors, at least in part.

Risk-adjusted screening is based on the identification of risk groups. Setting cut-off levels is fraught with difficulties both when testing for the presence of risk factors and as part of the overall screening strategy. These cut-off levels have to be agreed upon by the relevant stake holders (e.g. the Joint Committee) - and they have to be transparent.

### Structure

The following text analyses the concept of risk-adjusted screening with reference to the **four categories "disease", "test", "therapy" and "screening programme**" based on an approach that has first been proposed by Wilson & Jungner. A novelty is the introduction of a 'filtering test' that precedes the screening test proper and classifies individuals as a member of a risk group . There are four scenarios/options for implementing risk adjusted screening in routine practice:

1. 'General screening' for the target condition is already established on a routine basis. The introduction of risk adjusted screening would entail extending routine screening services to certain groups at high risk (e.g. commencing screening for familial colorectal and breast cancer screening at a younger age).

2. Groups at high risk of disease are being offered additional screening interventions to complement routine screening practice'; alternatively, the screening interval might be modified or both (e.g. multimodal screening in familial breast cancer, shortening the screening interval in familial colorectal cancer).

3. If 'general screening' has not been established on a routine basis, the possibility of introducing specific risk-adjusted screening for certain groups at risk "from scratch" has to be discussed and analysed.

4. Introducing screening for groups at high risk might render screening of the "general population" inefficient. As the main burden of disease concentrates within the risk groups, the altered cost benefit ratio due to stratification might justify the discontinuation of the screening activity in the general population (where incidence rates will be extremely low).

The following paragraph will address *specific criteria* for the evaluation of risk-adjusted screening that are, to a considerable degree, similar to those issues that apply to general screening.

**Morbidity in the population at risk:**

- **Risk indicators:** There is a clearly defined risk profile (risk indicators and the risk of disease they predict). Where several risk indicators are considered, their interactions are known and quantifiable.
- **Frequency:** The frequency (i.e. penetrance and age-specific incidence/prevalence rates) of the target condition as a function of risk indicators is known for the defined risk population. Moreover, the prevalence rates of risk indicators and risk of disease among the general population must be known.
  **Clinical progression of the disease:** The natural history of the disease among the risk population and the general population is known. Especially, any differences between these populations in terms of diagnosability, prognosis and the course of the disease have been investigated and are quantifiable (e.g. direction and magnitude of a spectrum bias with respect to the diagnostic validity of the screening/the test for the prodromal or early stages of the disease and/or in relation to the effectiveness of therapeutic measures).

**Test for elevated risk:**

Usually, a two-step filtering test is used for this purpose which first looks for familial/genetic traits based on the clinical history. As an additional second step, genetic testing can be appropriate. For both steps, inclusion criteria, i.e. cut-off levels, should be defined within the framework of a transparent process agreed upon by the relevant stakeholders.

**Clinical inclusion criteria**: The discriminatory strength of clinical and history-based criteria in respect of the target condition is known.

Where no risk genes are known and/or routinely detectable, at least the statistical disease risk with reference to clinical inclusion criteria has been evaluated in cohort studies in comparable populations.

Where a genetic test is available, there is a known likelihood of detecting a genetic mutation when the clinical and history-based inclusion criteria are met (=PPV and NPV of the clinical inclusion criteria with respect to the anticipated likelihood of the detection of a deleterious mutation).

- **Genetic testing for the presence of an elevated risk:** Testing for genetic risk factors that lead to inclusion in the risk population is objective, valid and reliable (quality criteria). Especially, sensitivity and specificity of the genetic test must be known. Moreover, the mutation spectrum should be known and the detected mutations must be classifiable with respect to their pathogenic relevance . In addition, disease penetrance in case of a positive test result must be known and high (=clinical validity).
- **Stress due to testing**: The test/s is/are acceptable for the target population (as little invasive and stressful as possible).
- The **cost involved** in testing is acceptable.
- **Cut-off level:** The cut-off level for inclusion in a risk group which entails being offered preventive measures has been agreed by the relevant stakeholders. It is clearly defined, transparent, and ethically acceptable.
- **Timing of the test:** There is an evidence-based age range in which the filtering test should be done.
- **Active involvement:** Active involvement of the eligible individual is required, depending on the test applied .

**Diagnostic procedures for risk-adjusted screening:**

- **Diagnostic procedures:** These must be objective, valid and reliable. Particularly, sensitivity, specificity, PPV and NPV of all tests must be validated and known, also in the risk population, and advance the time of diagnosis. At the same time, a possible spectrum bias must be excluded (clinical benefit).
- **Potential disadvantages:** The potential harm of diagnostic procedures (e.g. subjects experience stress related to testing, false-positive results, overdiagnosis) is known and quantifiable.
- **Prognosis:** Advancing the time of diagnosis lowers mortality from the target condition also in the risk population. If these data are not available yet, at least activities to quantify mortality reduction have been initiated, i.e. outcome-oriented documentation and quality control.

**Therapy:**

- There are established and effective therapies for the management of the target condition.
- Their effectiveness has also been demonstrated in the risk population (exclusion of spectrum bias regarding treatability). Alternatively, other effective therapeutic approaches are available.

In addition to the above mentioned criteria, there are *general* requirements that should be fulfilled in connection with the implementation of (risk-adjusted) screening programmes. These criteria will be outlined below:

**Requirements of a population-based risk-adjusted screening programme:**

- **Avoiding social and societal disadvantages**: Inclusion in the risk population does not lead to inacceptable societal or social disadvantages (as stipulated in the Genetic Testing Act).

- **Implications for the normal population:** The implications of a risk-adjusted screening programme for existing or potential screening activities in the 'normal risk group' will be investigated.
- **Benefit/Harm analysis:** An evidence-based analysis ensures that the potential benefit of risk-adjusted screening outweighs its potential harm.
- **Quality assurance/Evaluation:** Benefit and harm for various subpopulations participating in the risk-adjusted screening activities are being documented and evaluated on the basis of appropriate parameters.
- **Information:** Before taking the test(s), potential participants are informed in a balanced manner about all advantages and disadvantages that may result from testing and, if subsequent genetic diagnostic testing is to be done, must give their written informed consent after having been allowed an appropriate time for considering their decision (as required in the Genetic Testing Act).
- **Cost:** The cost of a risk-adjusted screening programme is known and reasonable compared to the overall resources spent on the specific target condition and health care expenditure overall.