# Individualised cost-effectiveness analysis in risk-based screening: practical and ethical?

05/12/2015

Christopher Sampson[1], Marilyn James[1], David Whynes[2]

1. *School of Medicine, University of Nottingham*

2. *School of Economics, University of Nottingham*

## Abstract

Personalised medicine presents the opportunity to improve outcomes by accounting for patient heterogeneity and providing individualised care. However, this raises practical and ethical dilemmas and is a challenge for health economists. We argue that there is one form of personalised medicine - risk-based screening - that can be practically and ethically implemented. A relationship exists between an individual's risk of disease onset and the cost-effectiveness of screening them for that disease. The expected cost-effectiveness of screening is likely to have a positive relationship with an individual's level of risk. Assuming that risk increases with time, intervals for recurrent screening should be shorter for people with a higher level of risk. We demonstrate that, by including individual risk in net benefit calculations, it is possible to estimate individualised cost-effectiveness results. These can be used to optimise policy by selecting the shortest screening recall period at which net benefit is positive, according to an individual's level of risk. By using a risk calculation engine and automatic generation of invitations to screening, such a programme is implementable. We consider different approaches to risk-based screening and their ethical implications, with particular reference to non-discrimination and allocation according to need. We argue that non-discrimination rules should not apply to individual risk in the case of screening because screening itself does not confer health benefit. Risk-based screening can operate within existing standards of distributive justice and the prevailing philosophy of health economics. We discuss implications for personalised medicine and individualised cost-effectiveness analysis more broadly.

Health Economists' Study Group meeting, Manchester, January 2016

# Introduction

Screening programmes can improve people's longevity and quality of life, and can be cost-saving for health services in the long run. However, it is unlikely that all of these goals could be achieved in a publicly funded programme without rationing and setting eligibility criteria. As such, screening programmes tend to discriminate based on individuals' characteristics - such as age and sex - and implement limited use criteria accordingly. The challenge of capturing heterogeneity of treatment effects is a long standing issue [1]. In economic evaluations, the issue is even more pronounced [2]. Although accounting for patient heterogeneity is a means of improving outcomes [3], it is rarely considered in economic evaluations [4].

Screening is one field in which efforts are being made to acknowledge heterogeneity and to differentiate care accordingly. It is now possible to use algorithms to estimate an individual's risk of disease onset, and the use of such tools has become known as 'predictive medicine'. Recently there have been calls for individualised screening and assertions that, in the future, a greater emphasis must be placed on risk-based screening [5–15]. In risk-based screening, individuals are only invited to attend screening if their risk of disease onset is deemed sufficiently high. Such an approach offers potential for improved outcomes and lower costs compared with standardised programmes, but presents new challenges.

One challenge for risk-based screening is to appropriately define limited use criteria; the grounds on which people should or should not be invited to screening. Additionally, it might also be necessary to differentiate the frequency of screening based on risk. Both individual risk and time are continuous and therefore infinitely divisible. As such, an optimal screening programme could lie anywhere between the extremes of no screening for anybody to constant screening for the entire asymptomatic population. In this study we present a means of setting optimal criteria for risk-based screening using cost-effectiveness analysis (CEA). We discuss some of the practical and ethical challenges presented by such an approach and argue that risk-based screening can improve the effectiveness, efficiency and fairness of screening programmes.

# Motivating example

The need for a rational and optimal approach to risk-based screening became clear through our work with an ongoing study of risk-based screening for diabetic retinopathy; the ISDR study [16]. The purpose of the ISDR study is to develop and evaluate a programme of personalised risk-based screening for diabetic retinopathy in Liverpool, UK. The study includes a large randomised controlled trial [17] and model-based economic evaluation [18].

Diabetic retinopathy is the leading cause of blindness among the working age population in England [19], and has been shown to negatively impact individuals' quality of life [20–22]. In the UK, the NHS Diabetic Eye Screening Programme (NHSDESP) has determined that all people with diabetes over the age of 12 should be invited for eye screening annually. There is good evidence for the cost-effectiveness of this screening programme for detecting disease and preventing loss of vision [23]. However, the implementation of an annual screening interval was based on expert opinion rather than empirical evidence. Recently, a large number of studies have identified that individuals with a low risk of developing sight-threatening retinopathy are being screened with unnecessary frequency [24–36].

A number of efforts have been made to 'optimise' screening intervals for diabetic retinopathy. Aspelund et al. developed a model based on a fixed number of diagnoses across screening intervals [37]. Their model suggested that optimal screening intervals range from 6 to 60 months, dependent on an individual's level of risk. Van der Heijden et al. recently sought to validate Aspelund et al.'s model using Dutch data [38]. The authors found that the model enabled a 23% reduction in screening frequency compared with biennial screening, and a 61% reduction compared with annual screening, and suggested that the use of such a model could help reduce the costs of diabetes care. Mehlsen et al. adopted a similar method, using a fixed risk margin of 0.5% chance of event [39]. They used multiple logistic regression to find the optimal screening interval for low-risk diabetic retinopathy patients, and found that screening intervals should be extended for most low-risk individuals. Using a discrete event simulation, Day et al. found that the risk of vision loss increased as the screening interval was extended from 1 to 5 years [40], though the authors assert that extending the interval from 1 to 2 years is safe. Similarly, Chalk et al. found that extending the screening interval for low-risk individuals was a cost-effective strategy [41]. Davies et al. suggested that simulation methods could be used to stratify individuals based on their risk level, and that screening levels could be adjusted accordingly, but the authors believed this to be impractical [42]. These studies did not define optimality in terms of cost-effectiveness. Scanlon et al. recently sought to 'optimise' the screening interval in terms of cost-effectiveness, but only within a restricted set of pre-specified options; all of which might be suboptimal [24].

The ISDR study is concerned with using information about an individual's risk to differentiate the frequency of recall for screening for diabetic retinopathy. A possible outcome in the near future is that individuals will be stratified into a number of groups based on their level of risk as estimated by a risk calculation engine. However, screening programmes could be individualised on a per-person basis and could be specifically tailored to the individual.

The risk estimation process makes it possible to define subgroups based on individual risk level and to apply limited use criteria as outlined by Coyle et al [43]. Thresholds might be defined in a number of

ways, such as acceptability, safety, affordability or cost-effectiveness. However, to define clinically meaningful subgroups based on individual risk is not a simple task. Given the interval properties of individual risk, we do not believe it possible. In this study we demonstrate a means of identifying optimality within an individualised risk-based screening programme that does not require arbitrary definition of subgroups. The application of clinically meaningless risk thresholds has ethical implications that will be discussed later in this article.

# Individualised cost-effectiveness analysis

Risk-based screening might be based on risk cohorts, such as low-, medium- and high-risk subgroups. In this case, limited use criteria can be determined using stratified cost-effectiveness analysis, as developed by Coyle et al. [43]. Stratified CEA estimates the net benefit of an intervention within subgroups as

$$NB_i = \lambda(E_{1i} - E_{0i}) - (C_{1i} - C_{0i}) \qquad \text{Equation 1}$$

where $NB_i$ is the net monetary benefit of the intervention within subgroup $i$, $E_{1i}$ and $E_{0i}$ are the outcomes for subgroup $i$ in the treatment group and the comparator group respectively, $C_{1i}$ and $C_{0i}$ are the costs for subgroup $i$ in the treatment group and the comparator group respectively, and $\lambda$ is the willingness to pay threshold per unit of outcome. Stratified CEA can be used to determine limited use criteria, such that

$$TNB = \sum_i NB_i \quad \forall_i \quad where\ NB_i > 0\,, \qquad \text{Equation 2}$$

where $TNB$ is the total net monetary benefit. However, there is not likely to be any reasonable basis on which to define risk groupings. Any threshold between low-, medium- or high-risk can only be arbitrary. Stratified CEA can only be used where it is possible to define clinically meaningful subgroups. Where thresholds are based on an estimated level of risk, this is not possible.

Individualised cost-effectiveness analysis involves the estimation of expected costs and outcomes (and their combination) at the individual level, rather than the aggregate for a population or subgroup. It can be implemented as a generalisation of stratified CEA. Individualised CEA involves the use of stratified CEA under two real-world conditions. Firstly, in individualised CEA the number of potential subgroups tends to infinity, such that: $\lim_{i \to \infty}$. Secondly, the size of subgroups tends to 1, such that: $\lim_{N_i \to 1}$ where $N_i$ is the size of subgroup $i$.

Practically, stratified CEA (in its original form) does not require any major adjustment to the form of standard clinical and economic evaluation. Its principal requirement is that data must be available from a sufficiently large sample to detect differences in costs and outcomes within subgroups, and that those alternative interventions can be tested within subgroups. However, individualised CEA cannot satisfy this requirement because of the two conditions outlined above. One can never observe the counterfactual costs and outcomes where the sample size is equal to one, and therefore cannot estimate incremental effects. Therefore, individualised CEA appears to be impossible in practical terms. Further assumptions will be necessary in order to operationalise an individualised CEA approach.

# Individualisation factors

Here we refer to the top-level factors that determine whether or not an intervention is deemed cost-effective. These are:

1. Probabilities,
2. Costs,
3. Outcomes,
4. Willingness to pay.

Each of these might be heterogeneous across individuals[1]. Individuals are defined by an infinite number of characteristics. Some of these will be observable and will influence the expected costs and outcomes associated with a treatment. The relationship between individual characteristics and the four factors of cost-effectiveness can be estimated and used to predict the expected cost-effectiveness of treatment for an individual. Given a large enough sample, an estimate of the true cost-effectiveness of treatment could be estimated for each individual as shown in Equation 3:

$$NB_i = \lambda_i(E_{1i} - E_{0i}) - (C_{1i} - C_{0i}) . \qquad \text{Equation 3}$$

Where patient level data are available, it is likely that costs and outcomes can be observed for each individual $i$. However, randomised trials are not likely to be large enough to elicit such relationships and it might instead be possible to use aggregate data within a modelling framework. Although the challenge of observing the counterfactual remains, there are now methods by which incremental costs and effects might be estimated for each individual. The analyst might for example use conditional average (CATE) or person-centred (PeT) treatment effects [44]. Existing regression-based approaches to CEA could prove valuable in carrying out individualised CEA.

The use of aggregate data from epidemiological or other observational studies is common in economic evaluation in health care, often being used for parameters in decision analytic models. It would be

[1] Willingness to pay for the outcome of interest is not usually considered a source of heterogeneity and such considerations are often dispatched to the realm of equity analysis. Equation 3 clearly shows that such simple 'equity' concerns can very easily be considered a matter of efficiency analysis.

possible to carry out multifactorial individualised CEA using aggregate level data. For example, multiple trials may have been carried out which might demonstrate a relationship between outcomes and age, and between costs and employment status. These relationships could be incorporated into an individualised CEA.

In some cases it may be desirable to include just one factor in individualised CEA. In this case, only one factor is allowed to alter across individuals and to determine cost-effectiveness at the individual level. In this study we focus on the single factor of individual risk, which is a probability within the estimation of cost-effectiveness. We use the word risk in the absolute sense of an individual's hazard rate or hazard function. That is, an individual with a given set of characteristics, at any moment in time, can be ascribed a hazard rate representing the probability that an event - e.g. disease onset - will occur. Prediction models or 'risk engines' for disease are becoming available in an increasing number of fields [45].

There is a link between individual risk of disease onset and the cost-effectiveness of screening for that disease, which has been implicitly recognised in practice but not explicitly demonstrated in the academic literature. Below we demonstrate the nature of this relationship and present a basis on which to use individualised CEA in risk-based screening.

# Optimising risk-based screening

Within screening programmes, an individual's screening outcome is not random; the probability of screening positive is dependent on a set of (observed or unobserved) risk factors. Through analysis of these risk factors it is now possible, in many cases, to estimate an individual's risk of developing a disease within a given period of time. If we seek to maximise an individual's health, the extent to which a screening programme can be beneficial depends on the effectiveness of treatment or care following a positive screen. If an individual is more likely to screen positive it is more likely that they will receive the intended benefits of screening. As such, it is clear that individuals with different risk levels will experience heterogenous benefits from screening.

The use of risk-based thresholds and patient stratification is well-established [46]. A number of cost-effectiveness analyses of screening programmes that take account of individuals' risk factors have been carried out as illustrated by the following. Using a Markov model to evaluate colorectal cancer screening, Dan et al. found that selective screening, based primarily on an individual's age, was more effective than standardised screening [47]. Similarly, Lansdorp-Vogelaar et al. found a small benefit associated with individualised colonoscopy screening [48]. Also using decision modelling methods, Round et al. found that the most cost-effective screening method for gestational diabetes mellitus depended on a woman's individual risk of disease [49]. Similarly, Aus et al. found that screening intervals for prostate cancer should be individualised based on prostate-specific antigen levels [50].

The expected incremental costs and outcomes of screening, at the individual level, are dependent on at least two binary probabilities; whether the individual has a given disease and whether they are screened positive or negative. Accounting for the specificity (false positive rate) and sensitivity (false negative rate) of the screening test, this means there will be four possible outcomes for an individual who is screened and two for an individual who is not. Figure 1 shows these six possible screening pathways, which are widely applicable to screening programmes generally.

Figure 1: Screening pathways

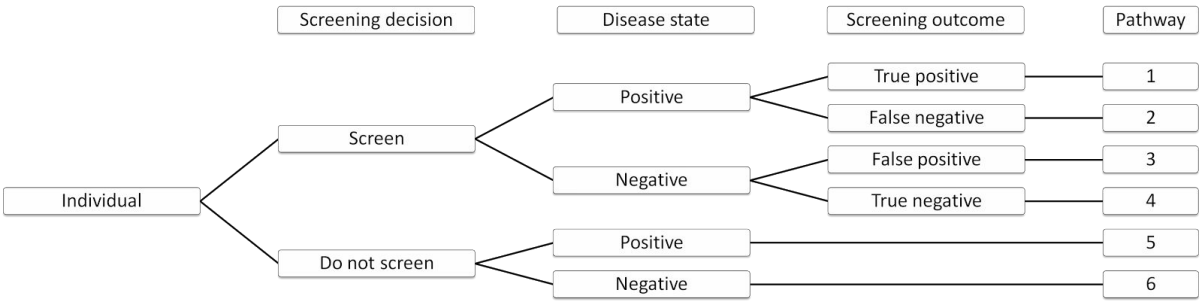



Table 1: Expected costs and outcomes

| Pathway | Probability | Cost | Expected cost | Outcome | Expected outcome |
|---|---|---|---|---|---|
| Screening | | | | | |
| 1 | $r(1-\beta)$ | $C_S$ | $r(1-\beta)C_S$ | $E_S$ | $r(1-\beta)E_S$ |
| 2 | $r(\beta)$ | $C_S$ | $r(\beta)C_S$ | $0$ | $r(\beta)0$ |
| 3 | $(1-r)(\alpha)$ | $C_S$ | $(1-r)(\alpha)C_S$ | $0$ | $(1-r)(\alpha)0$ |
| 4 | $(1-r)(1-\alpha)$ | $C_S$ | $(1-r)(1-\alpha)C_S$ | $0$ | $(1-r)(1-\alpha)0$ |
| Total | $1$ | | $C_1$ | | $E_1$ |
| No screening | | | | | |
| 5 | $r$ | $0$ | $(r)0$ | $0$ | $(r)0$ |
| 6 | $1-r$ | $0$ | $(1-r)0$ | $0$ | $(1-r)0$ |
| Total | $1$ | | $C_0$ | | $E_0$ |

In demonstrating the relationship between individual risk and the cost-effectiveness of screening, we adopt four simplifying assumptions:

1. A true positive outcome (i.e. pathway 1 in Figure 1) is the only outcome of value
2. The value of a true positive outcome is constant across cohorts

3. The incremental cost of screening is constant across cohorts
4. Individual risk is equal to the true probability that disease is present.

Table 1 shows the expected costs and outcomes associated with each pathway under these assumptions. The expected incremental cost and effect of screening is shown by

$$C_1 - C_0 = \sum_{k=1}^{4} (Prob(k) \times C(k)) - \sum_{k=5}^{6} (Prob(k) \times C(k)) \quad \text{Equation 4}$$

$$E_1 - E_0 = \sum_{k=1}^{4} (Prob(k) \times E(k)) - \sum_{k=5}^{6} (Prob(k) \times E(k)) \quad \text{Equation 5}$$

where $k = 1, 2...6$ refers to the possible pathways shown in Figure 1. Simplified, these expressions allow for an ICER to be estimated for an individual as

$$ICER_i = \frac{C_S}{(1-\beta) r_i E_S} \quad \text{Equation 6}$$

where $C_S$ is the cost of screening, $\beta$ is the false negative rate of the screening test, $r_i$ is the risk of disease onset for individual $i$ and $E_S$ represents the value (here assumed equal to 1) of a true positive screening outcome. This expression can be presented in terms of net monetary benefit as

$$NB_i = \lambda(1-\beta) r_i E_S - C_S . \quad \text{Equation 7}$$

Here we assume a true positive screen ($E_S$) to be beneficial and fixed across risk levels. $C_S$ only includes the cost of screening; the sunk cost associated with risk estimation and the cost of treatment are not included. Under our assumptions the cost and outcomes in pathways 3 and 4 are equivalent, meaning that the false positive rate ($\alpha$) does not influence the ICER. This means that the relationship between individual risk ($r$) and the expected incremental benefit of screening is positive and linear. Similarly, if we assume that the cost of screening is positive and constant across risk levels, a higher level of risk (*ceteris paribus*) will be associated with a lower ICER. As such, the relationship between an individual's risk of developing a disease and the expected cost-effectiveness of screening them must be positive.

Note that equations 6 and 7 are not necessarily generalisable, and could simplify to different expressions if the cells in Table 1 were changed. It is the process of defining them that is generalisable. They are presented to illustrate the relationships between the variables under our assumptions.

These expressions can be used to stratify screening in at least two ways: i) by setting an optimal threshold for screening; above which an individual is offered screening and below which they are not

and ii) by setting an optimal recall period for each individual. Later we will see that these approaches are equivalent in theory, if not in practice.

We can use the net benefit approach to estimate whether, at a given level of willingness to pay per true positive screen, it is cost-effective to screen an individual with a given level of risk. In order to do this, we need simply solve equation 7 for $r$ where net monetary benefit equals zero, such that:

$r = \frac{-C_S}{\lambda(\beta-1)E_S}$. Equation 8

This indicates the minimum level of risk at which individuals should be screened, and thus defines a threshold.

In an optimised programme, the time to next screen would be decided following each negative screening outcome, rather than a screening frequency being defined at an arbitrary time point. The incremental cost-effectiveness of subsequent screens can be characterised in the same way as the first; the decision process is the same, but the inputs may have changed. The relationship rationalises to that already discussed.

For some programmes it will be more useful to estimate an optimal recall period for an individual at a given point in time, rather than to define a minimum level of risk for population screening. In order to achieve this, it is first necessary to decide on the lowest practical screening interval; a clinic may have the administrative capacity to recall individuals on a weekly, monthly or yearly basis. Maintaining our earlier assumptions, and assuming that individuals' hazard ratios are estimated based on a time at risk ( $T$ ) of 1 year and that screening clinics are capable of recalling individuals for screening on a weekly basis, we can estimate the recall time at which net monetary benefit becomes positive. This is simply a matter of estimating the cost-effectiveness of screening an individual at various periods of recall, accounting for the fact that risk increases in time. The optimal recall period, in terms of cost-effectiveness, can be found by solving for $NMB = 0$ at any given level of individual risk. The optimal recall period for an individual can be established in 3 steps:

1. assuming a constant average rate of disease onset, convert the individual's estimated level of risk to an instant rate using the formula $= -\frac{ln(1-r)}{T}$,
2. convert the individual's instant rate back to the risk associated with the minimum feasible interval ($f$) using the formula $\bar{r} = 1 - exp^{-r(\frac{p}{(T/f)})}$,
3. select the recall period ($p$) such that net monetary benefit is at its lowest possible positive value, where $r$ is replaced with $\bar{r}$.

Because $r \sim \bar{r}$, the risk threshold associated with the optimal recall period will be approximate to that found at the programme level.

# An example of individualised CEA in risk-based screening

We present a stylised application of our framework for risk-based screening using aggregate data relating to the NHSDESP. We use reported figures from a CEA of systematic photographic screening compared with opportunistic screening [51], though we assume that the study compared systematic screening with no screening for the sake of the demonstration. The authors estimate prevalence to be 14.1%, but that only 4 in 5 people attend screening. Here we assume that attendance in the study was 100%. The primary outcome is cost per true positive screen. These examples are simplified, and as such the findings are not designed to inform policy.

Table 2: Model parameters

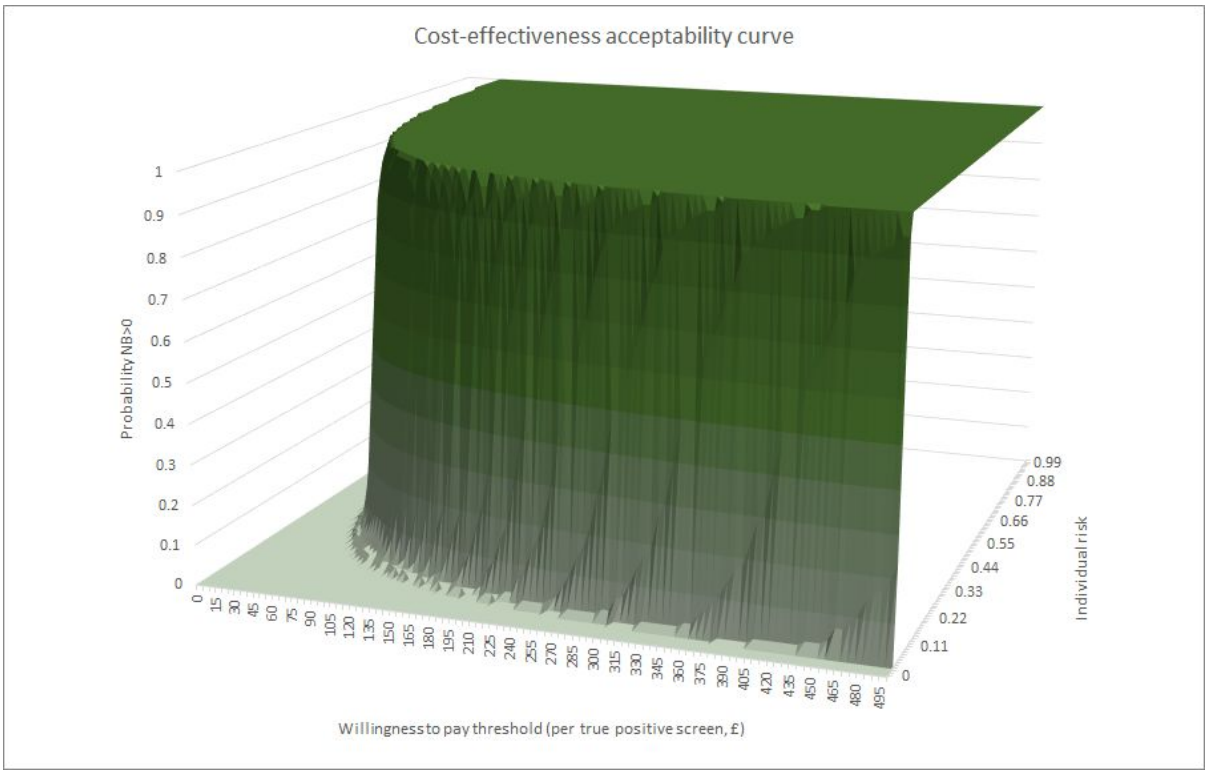
| | mean | standard error |
|---|---|---|
| $\alpha$ | 0.49 | 0.016 |
| $\beta$ | 0.02 | 0.003 |
| $C_S$ | £21 | 2[+] |
| $E_S$ | 1* | N/A |
| $ICER$ | £209 | N/A |

[+]Assumed.
*the outcome is true positive screening result

Table 2 shows figures provided by James et al. for screening for diabetic retinopathy, which we use to populate Equation 7. The CEA is individualised based on the single factor of individual risk of disease onset, in the way described above. In order to demonstrate the individualised CEA approach - including probabilistic sensitivity analysis - we use the parameters in Table 2 to simulate 100 sets of net benefit estimates for individual risk values from 0 to 1 and willingness to pay values from £0-500.

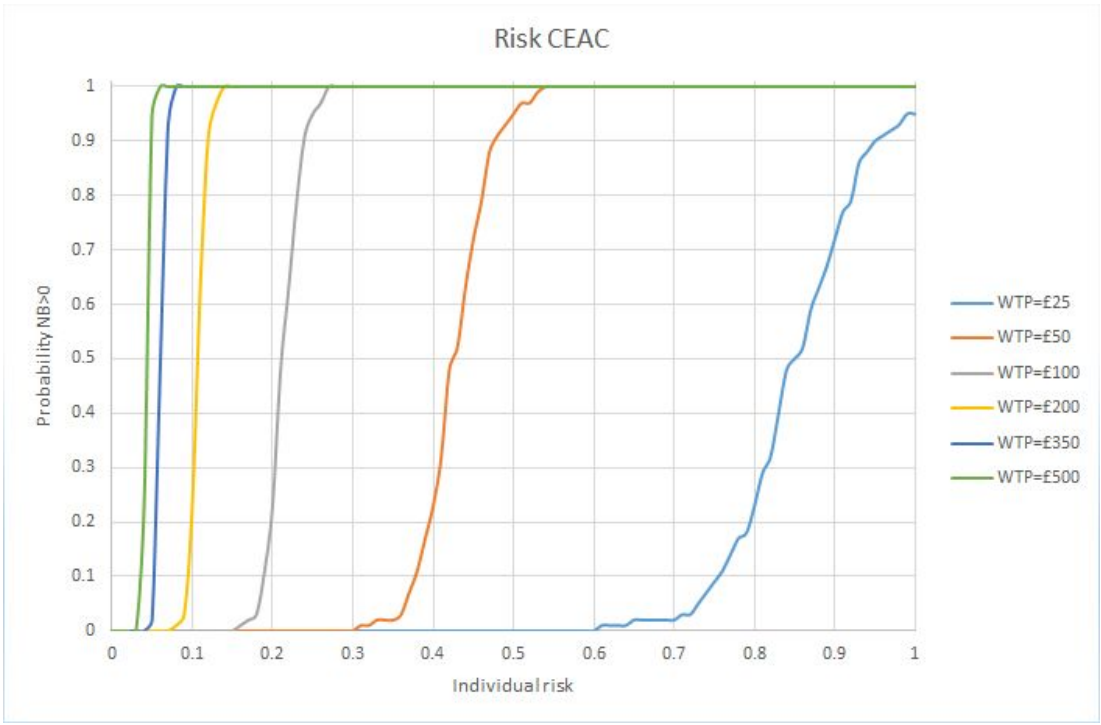
It is possible to demonstrate a relationship between individual risk, willingness to pay per true positive screening outcome, and the probability of cost-effectiveness. The resulting 3-dimensional cost-effectiveness acceptability surface is shown in Figure 2.

Figure 2: Risk/cost-effectiveness acceptability curve




For any given willingness to pay value, a risk-based CEAC can be created. Figure 3 shows the probability that screening is cost-effective across risk levels for a selection of willingness to pay values.

Figure 3: Risk-acceptability curves




Equivalent curves can be estimated for recall periods (recall-acceptability curves) for any given level of risk and willingness to pay threshold.

# Individual risk as an approximation of need

Ethical concerns might represent a barrier to accounting for heterogeneity in resource allocation decisions [4]. These concerns can be addressed in the case of risk-based screening, because individual risk should be understood as an approximation of need and therefore an appropriate basis for rationing.

In order to ensure fairness in risk-based screening, we propose two requirements, which are demonstrated in the examples above: i) individualised CEA should use average costs and benefits from the population and ii) screening should be personalised only on the basis of individual risk.

# Defining need for screening

A popular interpretation of health care need in resource allocation is as 'capacity to benefit from cost-effective care' [52]. This interpretation is particularly pertinent to screening, for which current ill health cannot be used to define need. Furthermore, an interpretation that allows for some needs to go unmet within an equitable system is necessary because screening is delivered to asymptomatic individuals. It could prove infeasible to meet the needs of many who may stand to gain extremely small potential benefits from screening.

The World Health Organisation definition of screening incorporates the requirement that effective treatment be available for those screening positive [53]. For many diseases, the majority of the population will have a probability of screening positive that is greater than zero. From the perspective of the analyst, if the distribution of risk in the population is unknown, the entire asymptomatic population has a capacity to benefit from screening. Many individuals might have a negligible probability of screening positive and therefore a small capacity to benefit from screening. To understand such low-risk people as having a 'need' for screening - even if universal screening is affordable - is not a useful interpretation and cannot inform a fair or efficient allocation of resources. Instead, for the case of screening, we should consider two levels of need.

The first level of need we refer to as 'treatment need'. This is the definition of need with which we are familiar; the capacity to benefit from cost-effective treatment for those with the disease. It can be measured in terms of health benefits, such as quality-adjusted life years. The purpose of screening is not to satisfy this need but to assess it. Screening is a means of obtaining more information about a person's health state; analogous to a physician consultation or a diagnostic test.

The second level of need we refer to as 'screening need', which can be interpreted as an intermediate or instrumental capacity to benefit. Screening tests are not designed to confer direct causal health benefit.

Although the results of screening tests might influence quality of life, they are not usually valued in these terms. Rather, their purpose is to identify people who have a capacity to benefit (treatment need) and this is the information provided by a true positive screening outcome. An individual's intermediate capacity to benefit from screening should therefore be understood as the probability that they will have a true positive screening outcome following a screening test. Individual risk of disease onset - as estimated by a risk calculation engine - should be used as an approximation of an individual's screening need.

Though we present two definitions of need, screening should still be evaluated in terms of overall capacity for health benefit. This is because - following our first requirement - treatment need is assumed to be fixed across risk levels, while screening need is a function of risk. Therefore, overall need ought to be estimated by multiplying treatment need (measured as health benefit) and screening need (a probability). This function is vital if screening programmes are to be evaluated against other possible spending options.

## Risk discrimination is non-discriminatory

Only the probabilities of different screening outcomes should differ across individuals when carrying out individualised cost-effectiveness analysis. In the allocation of resources, it is important to maintain a non-discriminatory approach. Generally, this approach is adopted by the National Institute for Health and Care Excellence in the UK [54]. Yet screening has to some extent been exempt from fairness concerns; many programmes discriminate based on age and sex. This is likely because rationing in screening is seen as necessary, even though adequate need-based criteria for rationing have not been developed.

Even if non-discrimination ought to apply to treatment need, it does not follow that all people should be eligible for screening. Rather, rationing on the grounds of an individual's screening need should take place. Individual risk is a supra-personal characteristic that should not in itself hold moral weight. To distinguish between individuals based on their risk levels is not equivalent to doing so based on age, sex or race[2]. An individual's risk of disease onset is determined by population-level dynamics; hazard rates are elicited from large samples. Estimates are based on the optimal statistical specification and should not be influenced by human bias and normative assumptions.

Decision-makers should not bear the same moral responsibility to provide screening to all those with a capacity to benefit, because the role of screening is to provide information rather than health improvement.

[2] Though it should be noted that existing screening programmes do discriminate on such characteristics and therefore do not satisfy usual non-discrimination rules.

If screening is personalised exclusively on the non-discriminatory basis of individual risk (our second requirement), proportionality is maintained in the distribution of expected health care. That is, the expected immediate benefit of screening is equal across all people. If the expected benefit of a true positive screening outcome is held constant across all levels of risk (our first requirement), we can assert that capacity to benefit increases linearly with risk. Therefore, the capacity to benefit from screening (screening need) for an individual with a 5% risk of disease onset is double that of an individual with a 2.5% risk of disease onset.

A risk-based screening programme of the type described above allocates screening to all people with capacity to benefit from cost-effective care. Consequently, the threshold defined by our approach can be reinterpreted as a need-based threshold. If individual risk is understood to be an approximation of screening need, risk-based screening increases both horizontal and vertical equity. People with equivalent screening need receive equivalent screening. Where screening need differs between people, the frequency of screening differs proportionally.

Individualised risk-based screening could be more efficient and more equitable than either a standardised programme or a stratified programme. This can be illustrated by an example contrasting a standardised (current) approach to screening with a risk-based screening programme. Consider a standardised screening programme for which all men over 65 are eligible[3]. Whatever the disease, it is highly unlikely that age is the only risk factor. It is probable that there are many identifiable pairs of 60 year olds and 70 year olds who have the same level of risk. Within these pairs, the 60 year olds likely have a greater life expectancy and therefore a greater capacity to benefit from treatment (treatment need). Non-discrimination implies that we ought not allocate resources on such criteria and the 60 year olds would not be judged to have a greater right to treatment than the 70 year olds. There seem to be no grounds in terms of either efficiency or fairness on which to offer screening to the 70 year olds but not the 60 year olds. Yet this is what would happen within the standardised screening programme described. Within a risk-based screening programme, people with equivalent risk would have equivalent eligibility. Inefficient and inequitable outcomes would only result from imprecision in the risk calculation engine.

# Discussion

Cost-effectiveness analysis has to date been based only on costs and outcomes at the aggregate level. This is ethically convenient as it prevents discrimination against those who experience either higher costs or lower benefits from treatment because of their personal characteristics. Screening programmes do, however, discriminate on such characteristics.

[3] This example corresponds to the NHS abdominal aortic aneurysm screening programme.

In this study we have presented a means of estimating cost-effectiveness at the individual level based on particular factors. Such an approach to cost-effectiveness analysis can be used to guide the allocation of resources in risk-based screening and to achieve optimality in terms of both efficiency and equity.

It is likely to be important to relax the assumptions we set out above. Our first assumption - that a true positive screening is the only outcome of value - is not likely to hold true. For example, it is likely that a true negative outcome will be preferred to a false positive. There are also likely to be costs associated with a false positive that are not incurred with a true negative. A more realistic approach would be to associate different costs and outcomes (ideally measured as quality-adjusted life years) with each screening pathway. As such, the inputs to the ICER calculation (i.e. the figures shown in Table 1) would be adjusted. Taking such an approach will be crucial, particularly given the detrimental effects of overdiagnosis.

Our second and third assumptions relate to the consistency of costs and outcomes across risk levels. For the reasons outlined above, we do not believe that differential costs and outcomes ought to be taken into account where these costs and outcomes relate to treatment following screening. Nevertheless, relaxing this assumption would be possible by determining relationships between the inputs to the ICER calculation.

Regarding our fourth assumption, a risk engine is likely to produce an estimate of probability within a given margin of error. Furthermore, accuracy of the risk engine might vary by risk level, and this might also need to be accounted for in the model.

Individualised CEA can be computationally expensive due to the increase in dimensionality. Even our simple example of unifactorial individualised CEA required the generation of 1,000,000 data points.

Calculating individual risk enables better definition of the basis on which to distribute care by giving us a more accurate understanding of individual need. Risk-based screening can improve both vertical and horizontal equity; a programme in which people of equivalent risk receive equivalent screening has been described as being in line with expectations of justice [55]. A risk-based screening programme that satisfies our requirements can be interpreted as egalitarian because everybody with a given level of need is treated equally.

Risk-based screening programmes that use risk calculation engines to differentiate screening recall periods and to determine limited use criteria are feasible in practice, as being demonstrated in ongoing research. However, decision rules for determining treatment thresholds and optimal regimes are not clearly defined. Here we have introduced an approach to optimising risk-based screening on the basis of

cost-effectiveness, and have demonstrated that such an approach can be both more effective and more equitable than standardised screening programmes.

# Acknowledgments

We would like to thank attendees at the Personalised Medicine and Resource Allocation conference at University of Oxford, where some of this work was discussed, for their helpful comments. Financial support for this study was provided in part by a grant from the National Institute for Health Research. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

# References

1. Kravitz RL, Duan N, Braslow J (2004) Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. Milbank Q 82:661–687

2. Stevens W, Normand C (2004) Optimisation versus certainty: understanding the issue of heterogeneity in economic evaluation. Soc Sci Med 58:315–320

3. Sculpher MJ (2008) Subgroups and heterogeneity in cost-effectiveness analysis. Pharmacoeconomics 26:799–806

4. Grutters JPC, Sculpher M, Briggs AH, Severens JL, Candel MJ, Stahl JE, De Ruysscher D, Boer A, Ramaekers BLT, Joore MA (2013) Acknowledging patient heterogeneity in economic evaluation : a systematic literature review. Pharmacoeconomics 31:111–123

5. Bradbury A, Olopade OI (2006) The case for individualized screening recommendations for breast cancer. J Clin Oncol 24:3328–3330

6. Brawley OW (2012) Risk-based mammography screening: an effort to maximize the benefits and minimize the harms. Ann Intern Med 156:662–663

7. Burton H, Chowdhury S, Dent T, Hall A, Pashayan N, Pharoah P (2013) Public health implications from COGS and potential for risk stratification and screening. Nat Genet 45:349–351

8. Glass AS, Cary KC, Cooperberg MR (2013) Risk-based prostate cancer screening: who and how? Curr Urol Rep 14:192–198

9. Zhu X, Albertsen PC, Andriole GL, Roobol MJ, Schröder FH, Vickers AJ (2012) Risk-based prostate cancer screening. Eur Urol 61:652–661

10. Burton H, Sagoo GS, Pharoah P (2012) Time to revisit Geoffrey Rose: strategies for prevention in the genomic era? Ital J Public Health. doi: 10.2427/8665

11. Roobol MJ, Carlsson SV (2013) Risk stratification in prostate cancer screening. Nat Rev Urol 10:38–48

12. Pashayan N, Pharoah P (2012) Population-based screening in the era of genomics. Per Med 9:451–455

13. Pharoah PDP, Antoniou AC, Easton DF, Ponder BAJ (2008) Polygenes, risk prediction, and targeted prevention of breast cancer. N Engl J Med 358:2796–2803

14. Olafsdóttir E, Stefánsson E (2007) Biennial eye screening in patients with diabetes without retinopathy: 10-year experience. Br J Ophthalmol 91:1599–1601

15. Dominitz JA, Robertson DJ (2014) Tailoring colonoscopic screening to individual risk. Gastroenterology 147:264–266

16. Harding SP (2013) RP-PG-1210-12016 - Introducing personalised risk based intervals in screening for diabetic retinopathy: development, implementation and assessment of safety, cost-effectiveness and patient experience. National Institute for Health Research Programme Grants for Applied Research

17. Broadbent D, Harding S, García-Fiñana M, Fisher A, Bennett A, Appelbe D, Wang A (2015) Introducing personalised risk based intervals in screening for diabetic retinopathy: the ISDR study. Eur J Ophthalmol 25:e20–e21

18. Sampson C, James M, Fisher AC, Harding SP (2015) Cost-effectiveness of a risk-based screening programme for diabetic retinopathy: a modelling approach. Eur J Ophthalmol 25:e13

19. Bunce C, Wormald R (2006) Leading causes of certification for blindness and partial sight in England & Wales. BMC Public Health 6:58

20. Poku E, Brazier J, Carlton J, Ferreira A (2013) Health state utilities in patients with diabetic retinopathy, diabetic macular oedema and age-related macular degeneration: a systematic review. BMC Ophthalmol 13:74

21. Lloyd A, Nafees B, Gavriel S, Rousculp MD, Boye KS, Ahmad A (2008) Health utility values associated with diabetic retinopathy. Diabet Med 25:618–624

22. Sampson C, Tosh J, Cheyne C, Broadbent D, James M (2015) Health state utility values for diabetic retinopathy: protocol for a systematic review and meta-analysis. Syst Rev 4:15

23. Jones S, Edwards RT (2010) Diabetic retinopathy screening: a systematic review of the economic evidence. Diabet Med 27:249–256

24. Scanlon PH, Aldington SJ, Leal J, Luengo-Fernandez R, Oke J, Sivaprasad S, Gazis A, Stratton IM (2015) Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. Health Technol Assess 19:1–116

25. Stratton IM, Aldington SJ, Taylor DJ, Adler AI, Scanlon PH (2013) A simple risk stratification for time to development of sight-threatening diabetic retinopathy. Diabetes Care 36:580–585

26. Looker HC, Nyangoma SO, Cromie DT, et al (2013) Predicted impact of extending the screening interval for diabetic retinopathy: the Scottish Diabetic Retinopathy Screening programme. Diabetologia 56:1716–1725

27. Porta M, Maurino M, Severini S, et al (2013) Clinical characteristics influence screening

intervals for diabetic retinopathy. Diabetologia 56:2147–2152

28. Echouffo-Tcheugui JB, Ali MK, Roglic G, Hayward RA, Narayan KM (2013) Screening intervals for diabetic retinopathy and incidence of visual loss: a systematic review. Diabet Med 30:1272–1292

29. Thomas RL, Dunstan F, Luzio SD, Roy Chowdury S, Hale SL, North RV, Gibbins RL, Owens DR (2012) Incidence of diabetic retinopathy in people with type 2 diabetes mellitus attending the Diabetic Retinopathy Screening Service for Wales: retrospective analysis. BMJ 344:e874

30. Leese GP (2013) Should diabetes retinal screening intervals change? Diabet Med 30:43–45

31. Agardh E, Tababat-Khani P (2011) Adopting 3-year screening intervals for sight-threatening retinal vascular lesions in type 2 diabetic subjects without retinopathy. Diabetes Care 34:1318–1319

32. Rein DB, Wittenborn JS, Zhang X, Allaire BA, Song MS, Klein R, Saaddine JB, Vision Cost-Effectiveness Study Group (2011) The cost-effectiveness of three screening alternatives for people with diabetes with no or early diabetic retinopathy. Health Serv Res 46:1534–1561

33. Misra A, Bachmann MO, Greenwood RH, Jenkins C, Shaw A, Barakat O, Flatman M, Jones CD (2009) Trends in yield and effects of screening intervals during 17 years of a large UK community-based diabetic retinopathy screening programme. Diabet Med 26:1040–1047

34. Brailsford SC, Gutjahr WJ, Rauner MS, Zeppelzauer W (2007) Combined Discrete-event Simulation and Ant Colony Optimisation Approach for Selecting Optimal Screening Policies for Diabetic Retinopathy. CMS Books Math/Ouvrages Math SMC 4:59–83

35. Vijan S, Hofer TP, Hayward RA (2000) Cost-utility analysis of screening intervals for diabetic retinopathy in patients with type 2 diabetes mellitus. JAMA 283:889–896

36. Taylor-Phillips S, Mistry H, Leslie R, Todkill D, Tsertsvadze A, Connock M, Clarke A (2015) Extending the diabetic retinopathy screening interval beyond 1 year: systematic review. Br J Ophthalmol. doi: 10.1136/bjophthalmol-2014-305938

37. Aspelund T, Thornórisdóttir O, Olafsdottir E, et al (2011) Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. Diabetologia 54:2525–2532

38. Van der Heijden AAWA, Walraven I, van ’t Riet E, et al (2014) Validation of a model to estimate personalised screening frequency to monitor diabetic retinopathy. Diabetologia 57:1332–1338

39. Mehlsen J, Erlandsen M, Poulsen PL, Bek T (2012) Individualized optimization of the screening interval for diabetic retinopathy: a new model. Acta Ophthalmol 90:109–114

40. Day TE, Ravi N, Xian H, Brugh A (2014) Sensitivity of diabetic retinopathy associated vision loss to screening interval in an agent-based/discrete event simulation model.

Comput Biol Med 47:7–12

41. Chalk D, Pitt M, Vaidya B, Stein K (2012) Can the retinal screening interval be safely increased to 2 years for type 2 diabetic patients without retinopathy? Diabetes Care 35:1663–1668

42. Davies R, Roderick P, Canning C, Brailsford S (2002) The evaluation of screening policies for diabetic retinopathy using simulation. Diabet Med 19:762–770

43. Coyle D, Buxton MJ, O'Brien BJ (2003) Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria. Health Econ 12:421–427

44. Basu A (2011) Economics of individualization in comparative effectiveness research and a basis for a patient-centered health care. J Health Econ 30:549–559

45. Bouwmeester W, Zuithoff NPA, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, Altman DG, Moons KGM (2012) Reporting and methods in clinical prediction research: a systematic review. PLoS Med 9:1–12

46. Pauker SG, Kassirer JP (1980) The threshold approach to clinical decision making. N Engl J Med 302:1109–1117

47. Dan YY, Chuah BYS, Koh DCS, Yeoh KG (2012) Screening based on risk for colorectal cancer is the most cost-effective approach. Clin Gastroenterol Hepatol 10:266–71.e1–6

48. Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Boer R, Wilschut J, Winawer SJ, Habbema JDF (2009) Individualizing colonoscopy screening by sex and race. Gastrointest Endosc 70:96–108, 108.e1–24

49. Round JA, Jacklin P, Fraser RB, Hughes RG, Mugglestone MA, Holt RIG (2011) Screening for gestational diabetes mellitus: cost-utility of different screening strategies based on a woman's individual risk of disease. Diabetologia 54:256–263

50. Aus G, Damber J-E, Khatami A, Lilja H, Stranne J, Hugosson J (2005) Individualized screening interval for prostate cancer based on prostate-specific antigen level: results of a prospective, randomized, population-based study. Arch Intern Med 165:1857–1861

51. James M, Turner DA, Broadbent DM, Vora J, Harding SP (2000) Cost effectiveness analysis of screening for sight threatening diabetic eye disease. BMJ 320:1627–1631

52. Culyer A (1998) Need--is a consensus possible? J Med Ethics 24:77–80

53. Wilson JMG, Jungner YG (1968) Principles and practice of mass screening for disease.

54. Cookson R (2015) Justice and the NICE approach. J Med Ethics 41:99–102

55. Shickle D, Chadwick R (1994) The ethics of screening: is "screeningitis" an incurable disease? J Med Ethics 20:12–18