

Patient follow up screening evaluations. Examples with regard to congenital hip dislocation and congenital heart disease

R E Juttmann, J Hess, G J van Oortmarssen, P J van der Maas

Abstract

Objective—To discuss the merits of the patient follow up study design for the evaluation of some specific mass screening programmes.

Design—Theoretical evaluation illustrated by two examples.

Setting—Department of Public Health Erasmus University Rotterdam.

Main results—The gold standard for evaluation of favourable effects of screening is the randomised controlled trial (RCT). Application of an RCT, however, is often not feasible, in which cases observational studies will have to be relied on. The case-control study design is generally considered to be second best. In some situations, however, a patient follow up study design may be applicable and may have some major advantages. The use of the patient follow up design for screening evaluation will often be very problematic or even unacceptable, particularly as far as screening for cancer is concerned. The most important objections are resulting from lead time bias, length bias, selection bias and over-treatment bias. For the evaluation of screening for congenital heart disease and congenital hip dislocation in Dutch child health care, however, these objections may relatively simply be overcome. Lead time bias will be of little importance, as the ages of onset of these disorders are fixed, namely at birth, and their ultimate outcomes may be expected within relatively short time. Length bias may largely be avoided by correction for severity of the disorder, which can be adequately assessed by modern diagnostic procedures. Selection bias is generally hard to rule out, but in these cases it probably plays a minor part. Over-treatment can be avoided by the policy of “watchful waiting”, which in these disorders can be applied with little risk for fatal outcomes. In principle bias might be avoided more successfully in a case-control screening evaluation than in a patient follow up study. However, the patient follow up study is for both screening programmes discussed here the more feasible design and can provide more supplementary information. The results of two example studies suggest that both screenings probably yield considerable benefits

Conclusion—Under a number of specific conditions a patient follow up study is an

efficient alternative to more customary designs for screening evaluation.

(J Epidemiol Community Health 2001;55:126–131)

Although the basic concept of screening is deceptively simple, there is a general consensus that assessing the favourable effects of any screening programme is fraught with pitfalls and requires a very strict methodology.¹ From a theoretical point of view the most appropriate design for such studies is the randomised controlled trial (RCT), which in principle is the only option to avoid virtually all types of confounding.² Unfortunately, many forms of screening already have an established place in health care without ever having been subjected to a randomised trial. Examples are to be found in occupational health, cervical cancer screening, the periodic examinations during pregnancy and child health care. If a screening programme is already performed on a large scale in a population, it is almost impossible to organise an RCT for the evaluation of such a programme. Definite faith in the benefits of screening among many professionals and the public, will, although based on merely circumstantial evidence, lead to strong resistance, if the screening is withheld from persons who normally would have had the opportunity to participate. When an RCT for assessing favourable effects of an already existing screening programme is not feasible, one has to rely on less decisive observational studies, meticulously considering possible sources of confounding. The most common observational designs for this purpose are the population follow up study and the case-control study. This paper, however, is mainly concerned with the merits and shortcomings of the patient follow up study, which may be considered as an alternative for the case-control study.

For many conditions, especially for cancer, the use of the patient follow up study design for assessing favourable effects of screening is very problematic, or even unacceptable. However, for the evaluation of screening for some disorders with specific characteristics, this design may, under a number of strict conditions, be useful. This may particularly be the case for evaluation of screening for some congenital disorders within the framework of child health care.

After a short description of the patient follow up study design for assessing favourable effects of screening, we will first discuss the most important objections against this kind of study, notably as far as cancer screening is concerned.

Department of Public Health, Erasmus University Rotterdam, the Netherlands
R E Juttmann
G J van Oortmarssen
P J van der Maas

Department of Paediatric Cardiology, the Sophia Children's Hospital and Erasmus University Rotterdam
J Hess

Correspondence to:
Dr Juttmann, Department of Public Health, Erasmus University Rotterdam, Room EE 2008, PO Box 1738, 3000 DR Rotterdam, the Netherlands
(juttmann@mgz.fgg.eur.nl)

Accepted for publication
5 October 2000

Next we will demonstrate why these objections may be less problematic, or may be relatively easily remedied, in assessing the effects of screening for certain conditions in child health care, in particular congenital heart malformations and congenital hip dislocation. Finally, we will review two examples of screening evaluations, in which for these conditions the patient follow up study design was actually applied.

The design

In a screening evaluation according to the patient follow up study design, a representative group of patients is followed up from diagnosis until the outcome of treatment can be established. Retrospectively the exposure to screening will be established. If the odds for reaching an adverse outcome are more favourable for well screened patients (the screening group) than for patients not or scarcely exposed to screening (the non-screening group), the screening is considered to be effective.

Objections against patient follow up studies in cancer screening

There are four evident objections against patient follow up evaluations of cancer screening:

LEAD TIME BIAS

Considering the nature of most cancers, the follow up time after diagnosis, in which the outcome of the treatment is supposed to be established in patient follow up screening evaluations, is not fixed. Because for most cancers the adverse outcome to be prevented is death, the usual outcome measure is whether or not patients will survive a fixed period after diagnosis of for example five years. As the age of onset of cancer in relation to the age of screening may vary considerably, the screening group may contain a large proportion of patients with disorders in an early stage. Many of these patients will, regardless of being treated or not, probably not die within five years, but may despite receiving treatment still die from cancer after five years. In the non-screening group such patients will be virtually absent. This will lead to an overestimation of the proportion of patients with favourable outcomes in the screening group and consequently to overestimation of the potential favourable effects of screening.

LENGTH BIAS

Patients with rapidly progressing cancers are more likely to reach an adverse outcome than patients with slowly progressing diseases, while they have also less chance of being screened during the short preclinical detectable phase. In a comparison of screen detected with non-screen detected patients this will lead to an overrepresentation of such patients in the non-screen detected group and consequently to overestimation of the favourable effects of screening. This kind of contamination is called "length bias". In a patient follow up study in which exposure to screening rather than detection by screening is the subject of the compari-

son, length bias is not necessarily a problem. However, when rapid progression is somehow connected with a decreased participation in the screening programme, overestimation of favourable effects of screening as a result of differences in natural course will also occur in this design.

SELECTION BIAS

People who have a better chance of being diagnosed in time and treated with favourable outcomes without screening anyway, may also be the better screened. For example, people who are watchful as far as their health is concerned and who are assertive in acquiring treatment, probably scrupulously attend the screening programme. This will also lead to an overestimation of the effectiveness of screening.

OVER-TREATMENT

Another problem in the patient follow up study design is the potential presence among those who were apparently treated successfully, of persons who were wrongly indicated as a patient by screening in the first place. This is a general problem in cancer screening: the natural course of anomalies found by cancer screening may vary considerably. Conditions may be regressive and resolve spontaneously or may be slowly progressive and never become a real threat. Generally speaking it would be a sensible policy to postpone treatment of screen detected disorders until the disease has almost progressed up to a stage in which a favourable outcome can no longer be expected. This may be feasible for disorders of which the prognosis is relatively easy to establish, yet in many cancers the natural course is quite unpredictable and a policy in which treatment is postponed may easily lead to fatal outcomes. In these cases therapeutic interventions will therefore be applied without delay. Hence over-treatment must be considered as an inevitable consequence of screening for many cancers, of which cervical cancer and prostate cancer are obvious examples.^{3 4} In the light of the possible benefits of screening such a disadvantage may be acceptable. However, in a patient follow up cancer screening evaluation, this may also lead to overestimation of the favourable effects of screening as in such a study favourable outcomes in over-treated screening participants will contribute to the observed positive effect of the screening.

Status of these objections in screening for congenital heart malformations and congenital hip dislocation

LEAD TIME

The most important differences of these conditions compared with cancer are that their ages of onset are fixed, namely at birth, and that their ultimate outcomes may be expected within a relatively short time.

Both congenital hip dislocation and congenital heart malformations are in principle present at birth. A persistent untreated dislocation will almost inevitably reach the adverse outcome—that is, limping—shortly after the age of 1 year. In most untreated clinically significant

congenital heart malformations the adverse outcomes—that is, haemodynamic complications such as heart failure and hypoxemia—will occur even before the age of 1 year. As a result lead time bias will not occur in evaluating the screening programmes for these congenital conditions with a patient follow up design, in contrast with the situation in cancer screening.

LENGTH BIAS

In patient follow up studies for congenital heart malformations length bias may not be ruled out. Screening examinations for this condition are scheduled relatively close together in the first months of life. Patients with severe disorders may deteriorate rapidly and consequently fail to attend the screening examination. Subsequently after diagnosis they may be included in the non-screening arm of a patient follow up study and may be considered to be diagnosed after the adverse outcome was reached. Such a sequence of events is less likely in patients with moderate disorders, so such patients may be over-represented in the screening group and the non-adverse outcome group, which will result in an overestimation of the favourable effects of screening. This might be neutralised by correcting the analysis for severity of the disorder. Reliable assessment of severity—that is, the tendency to a rapid deterioration—is obviously a requirement for such a solution. In contrast with cancer, which often is a condition developing, as it were, in disguise, in congenital heart disease, current diagnostic procedures make it possible to visualise the malformation completely and thus establish its severity reliably.⁵

SELECTION BIAS

As in all observational designs stringently ruling out selection bias is one of the most arduous problems in patient follow up screening evaluations. Nevertheless, in an observational evaluation of child health care screening, selection bias may play a somewhat less important part than in similar evaluations of cancer screening. Selection bias is related to the extent to which, and the reason why, people comply with the invitation for the screening. However, in the screening programmes under discussion, exposure to screening generally does not depend on the compliance of the parents (which is usually very high), but above all on whether or not screening is well performed by the child health centre physician. Variation in this performance will not automatically lead to selection bias.

OVER-TREATMENT

As clarified above, postponing treatment until the disorder almost reaches a stage in which spontaneous regression is judged to be impossible is the obvious strategy to prevent over-treatment. This policy of “watchful waiting” enables researchers to exclude screening participants wrongly picked out as “patients” from a patient follow up screening evaluation. In contrast with many forms of cancer, in which predicting the natural course is very hazardous, this is relatively straightforward for

KEY POINTS

- The gold standard for evaluation of favourable effects of screening is the randomised controlled trial (RCT).
- Application of an RCT is often not feasible, in which cases observational designs will have to be relied on of which the case-control study design is generally considered to be second best option.
- The most important objections against patient follow up screening evaluations are related to lead time bias, length bias, selection bias and over-treatment bias. For the evaluation of screening for congenital heart and hip disorders, however, these objections may relatively simply be overcome.
- Although in principle in a case-control screening evaluation selection bias might be avoided more successfully than in a patient follow up study, for several screening programmes in child health care the latter is the more feasible design and can provide more supplementary information.

the congenital conditions under discussion. While many cancers develop on a cellular level, these conditions are generally relatively large anatomical malformations. Congenital heart and hip disorders may nowadays also be completely visualised, with the help of for example ultrasound technology.⁶ Consequently the natural course of both conditions can be monitored adequately and interventions can be postponed until deterioration can be accurately foreseen.

Examples of patient follow up screening evaluations in child health care

During the preparation of our child health care evaluation programme we perceived that congenital heart disease and congenital dislocation of the hip may have the specific characteristics that allow applying a patient follow up screening evaluation. Thus we decided to conduct two studies with such a design on these diseases.

CONGENITAL HEART DISEASE

About 0.8% of all children are born with congenital heart disease. Large anatomical anomalies, in which spontaneous recovery is inconceivable, are treated immediately after diagnosis by medication, catheterisation or surgery. In all other cases, the natural course of the disorders is surveyed until spontaneous regression occurs—that is, the disorder disappears or proves to be haemodynamically insignificant, or until the adverse outcome of the disease (heart failure or hypoxemia) is judged to be inevitable unless treatment is started.⁷

After the neonatal check up by the doctor or midwife who assisted birth, which is not standardised in the Netherlands, the cardiovascular system of children in the Netherlands is screened during recurrent physical examinations in the child health care programme.

Table 1 2 × 2 table for analysis of a pilot patient follow up study for the evaluation of child health care screening for congenital heart malformations

	"Too late"	"In time"	Total
Adequately screened	2	10	12
Inadequately screened	37	33	70
Total	39	43	82

Odds ratio for being "too late" depending on whether or not adequately screened:

2:10/37:33 = 0.18 (95% CI: 0.04, 0.87)

= 0.20 (95% CI: 0.04, 1.05) corrected for severity

From 1994 until 1996 in the Sophia Children Hospital in Rotterdam a patient follow up study was carried out, comprising 82 children with congenital heart disease.⁸ Children were classified in the adverse outcome category (diagnosed "too late") if heart failure or hypoxemia had occurred before diagnosis. Children who could be treated before the onset of heart failure or hypoxemia was judged to be inevitable were classified in the non-adverse outcome category (diagnosed "in time"). Children were classified as "adequately screened" if they had at least been exposed to all scheduled screening tests in the child health care programme until being diagnosed and if all these screening tests had been performed adequately—that is, in accordance with the guidelines of the child health care authorities. All others were classified as "inadequately screened". The number of children in each of the four categories and the odds ratio for reaching the adverse outcome depending on whether or not being adequately screened are presented in table 1. As the disorders were also classified as moderate, severe or very severe by paediatric cardiologists, analysis could be corrected for severity.

The results of this study show that systematic screening in child health care can prevent episodes of heart failure and hypoxemia in children with congenital heart disease, although after correction for severity the confidence intervals for the odds ratio just includes 1.

CONGENITAL HIP DISLOCATION

About 1% of all children are born with a hip dislocation or a dislocatable hip. In the absence of (early) intervention this disorder will develop into a permanent anomaly, which finally results in limping, in only 0.08–0.16 % of all children.⁹ Neonatal screening by the Barlow and Ortolani methods is applied, for instance in the United Kingdom and Scandinavia.¹⁰ Splinting for four to six weeks starting as soon as possible after birth is often considered an effective and little taxing intervention. However, as in 90% of the cases the disorder will be regressive, the number of children apparently treated successfully, but actually wrongly picked out as a patient by screening will be substantial. Therefore at first sight this programme seems to be a poor candidate for evaluation with a patient follow up study design.

Some authors, however, advise to postpone intervention until approximately the fifth month.⁹ If the disorder still exists at that age it can safely be considered as an anomaly that will

not recover without treatment. Splinting at that age will generally still be successful, though more taxing than in the very young children. Surgery however will be avoided, as will over-treatment.

As in Great Britain screening and intervention are applied on a large scale soon after birth, evaluation by an RCT is problematic.¹⁰ As in Great Britain the mere existence of congenital dislocation after the age of 1 month is considered the adverse outcome to be avoided, in that country a patient follow up study, as presented in this paper will be of no use. If, for instance, the need for surgery would be considered as the adverse outcome to be avoided, such study would become a possibility.

In the Netherlands there is no neonatal screening programme for congenital hip dislocation. Instead, children are screened much later during periodic physical examinations in the child health care programme, which include assessment of the abduction range of the hips and the length of the legs.¹¹ In the Sophia Childrens Hospital in Rotterdam, treatment policy is expectant: splinting is postponed until spontaneous recovery has become very unlikely. Under these circumstances the need for surgery is a useful definition of the adverse outcome, and the patient follow up design a possible option for screening evaluation.

In 1992 a study was carried out in the Sophia Children Hospital, comprising 60 children with a congenital hip dislocation.¹² All these children had progressive disease: pathological changes progressed up to a stage in which spontaneous recovery of the dislocation was judged to be impossible, before treatment was applied. Children were classified in the adverse outcome category if surgery was needed and in the no adverse outcome category if they were successfully treated by non-invasive methods. Children were classified as "adequately screened" if until the definite diagnosis they had been at least exposed to all scheduled screening tests in the child health care programme and if all these screening tests were performed adequately—that is, in accordance with the child health care professional guidelines. All other children were classified as "inadequately screened". The results and the estimated odds ratio are presented in table 2.

The results of this study support the idea that systematic screening in child health care can prevent the necessity of surgical intervention for congenital hip dislocation.

Table 2 2 × 2 table for analysis of a pilot patient follow up study for the evaluation of child health care screening for congenital hip dislocation

	Necessity for surgery	Conservative intervention possible	Total
Adequately screened	6	13	19
Inadequately screened	26	15	41
Total	32	28	60

Odds ratio for necessity for surgery depending on whether or not adequately screened.

6:13/26:15 = 0.27 (CI: 0.08, 0.85)

Discussion

The first question to be answered here is whether it is worthwhile to apply second best evaluation designs in situations where an RCT is not feasible because of established practice and circumstantial evidence for at least some effectiveness of the intervention. In our opinion the answer depends on the quality of the circumstantial evidence and the chance of further improving the intervention. As review of the available literature reveals large gaps in our knowledge of the effectiveness of screening protocols in child health care,^{13 14} we believe that observational studies for evaluation of the benefits of these screenings are justified.

We believe that for at least congenital hip dislocation and congenital heart disease a partly retrospective partly prospective patient follow up study can be an efficient alternative for more customary designs for screening evaluation like the population follow up study and the case-control study.

In terms of efficiency a patient follow up study offers the advantage of the availability of a study group directly from the patient population of for instance an academic hospital providing specialised medical care to a large area. Thus the laborious collection of data in the general population, necessary in a population follow up study, can be avoided.

Case-control studies provide another relatively efficient alternative for assessing favourable effects of screening. In such a study the case group consists of patients who have reached the adverse outcome of the condition. The exposure to screening in this group is retrospectively compared with that in a control group. To minimise bias, this control group should be sampled directly from the total population that generated the cases: the source population.^{2 15} In practice complying with this condition may not be easy. To form a sample from the source population, reliably presenting the average exposure to screening in that population may, especially if data concerning this population are not easily available, be more problematic than gathering complete data of relevant patients from a circumscribed area in a well defined time window, as required in a patient follow up study.

In screening evaluations by case-control studies one has to deal with the so called "healthy-screenee-bias". Once a disease is diagnosed the patient is thereafter no longer screened, while "healthy" non-patients are screened again and again. As the control group (almost) exclusively consists of "healthy" non-patients this phenomenon will, if the total number of applied screening tests per study participant is used to quantify the screening history, lead to an overestimation of the effectiveness of screening. To avoid this bias controls and cases are matched for age, and assessment of exposure to screening in a control is exclusively aimed at the period up to the age that the matched case was diagnosed. Matching for other variables may be hazardous, as it may introduce new sources of confounding. In Dutch child health care, for instance, differences in exposure to screening

within the population are probably strongly connected with differences in screening performance in child health centres. Matching, for example, for living area may easily lead to matching for child health centre as well. Should there be a real effect of screening, this would lead to an underestimation of the average exposure in the source population, and consequently to an underestimation of the favourable effect of screening.

An advantage of patient follow up studies is that healthy-screenee-bias is ruled out, although a similar problem remains. The total number of applied screening tests per individual cannot be used to establish the screening history in patient follow up studies in child health care either. Cases diagnosed early in life may have a better chance of being successfully treated, but will be exposed to fewer screening tests than those who are diagnosed later. Therefore using the total number of applied screening tests as a measure for screening history would lead to underestimation of the effectiveness of screening. To avoid this bias in establishing the screening history the proportion of scheduled screening examinations that have actually been carried out until the age of the definite diagnosis should be used.

Patient follow up studies (as well as case-control studies) aim exclusively at estimating the favourable effects of screening and not at weighing advantages and disadvantages (for example, arising from false positive and false negative tests). This requires additional data collection.

A remarkable characteristic of case-control screening evaluations is the fact that although such studies are aimed at assessing benefits of a screening not a single person who actually benefits from screening may be included in the study. The case group consists of people with adverse outcomes of the condition, the control group virtually always consists of people who do not have the condition at all. As a result, in case-control screening evaluations, only the prevention process as a whole starting from exposure to screening can be evaluated. The contribution of separate factors, such as the influence of delay between a positive screening test and adequate diagnosis and intervention, which may be of crucial importance for the effectiveness of the prevention programme, cannot be evaluated. In a patient follow up study, however, this is very well possible.

If in a case-control screening evaluation a truly representative sample can be taken from the source population, selection bias will probably be avoided more successfully than in a patient follow up study in which correction for such contamination is not straightforward. The patient follow up study, however, is for both screening programmes discussed here the more feasible design and will provide more supplementary information. The results of the trials presented in this study indicate that both screenings might yield considerable benefits. The odds ratios presented are very low. One is tempted to conclude that, even if selection bias would play a part, there must be also a real effect of screening.

The final conclusion is that under a number of specific conditions a patient follow up study is an efficient alternative to more customary designs for screening evaluation.

Funding: this study was supported by a grant of the Netherlands Heart Foundation.

Conflicts of interest: none.

- 1 Cochrane A. Screening: the rules of the game. *Tijdschrift voor Sociale Geneeskunde* 1978;**56**:6–8.
- 2 Morrison AS. *Screening in chronic disease*. New York: Oxford University Press, 1992.
- 3 Ballegooijen M van. *Effects and costs of cervical cancer screening*. PhD thesis. Rotterdam: Erasmus University, 1998.
- 4 Schroder F, Damhuis R, Kirkels W, et al. European randomized study for screening prostate cancer, the Rotterdam pilot studies. *Int J Cancer* 1996;**65**:145–51.
- 5 Park M. *Pediatric cardiology for practitioners*. St Louis: Mosby, 1995.
- 6 Boere-Boonekamp M. *Screening for developmental dysplasia of the hip*. PhD thesis. Enschede: Twente University, 1996.
- 7 Danford A, McNamara D. Infants with congenital heart disease in the first year of life. In: Garson A, Bricker J, McNamara D, eds. *The science and practice of paediatric cardiology*. Philadelphia: Lea and Febiger, 1990:1959–72.
- 8 Juttman R, Hess J, Looman C, et al. Screening for congenital heart malformation in child health centres. *Int J Epidemiol* 1998;**27**:989–94.
- 9 Burger B, Burger J, Bos C, et al. Neonatal screening and staggered early treatment for congenital dislocation and dysplasia of the hip. *Lancet* 1990;**336**:1549–53.
- 10 Dezateux C, Godward S. Evaluating the national screening programme for congenital dislocation of the hip. *J Med Screen* 1995;**2**:200–6.
- 11 Boere-Boonekamp MM, Kerckhoff TH, Schuil PB, et al. Early detection of developmental dysplasia of the hip in The Netherlands: the validity of a standardized assessment protocol in infants. *Am J Public Health* 1998;**88**:285–8.
- 12 Juttman RE. Unnecessarily taxing treatment of congenital hip dislocation to be avoided by timely diagnosis at the well-child center. *Ned Tijdschr Geneesk* 1992;**136**:1467–71.
- 13 Hall D. *Health for all children. A program for child health surveillance*. Oxford: Oxford University Press, 1996.
- 14 Winter M de, Balledux JM, Burgmijer R. *Screening in child health care. Report of the Dutch Working Party on Child Health Care*. Oxford: Radcliffe Medical Press, 1995.
- 15 Weiss NS. Case-control study of screening. *Epidemiol Rev* 1994;**16**:102–8.