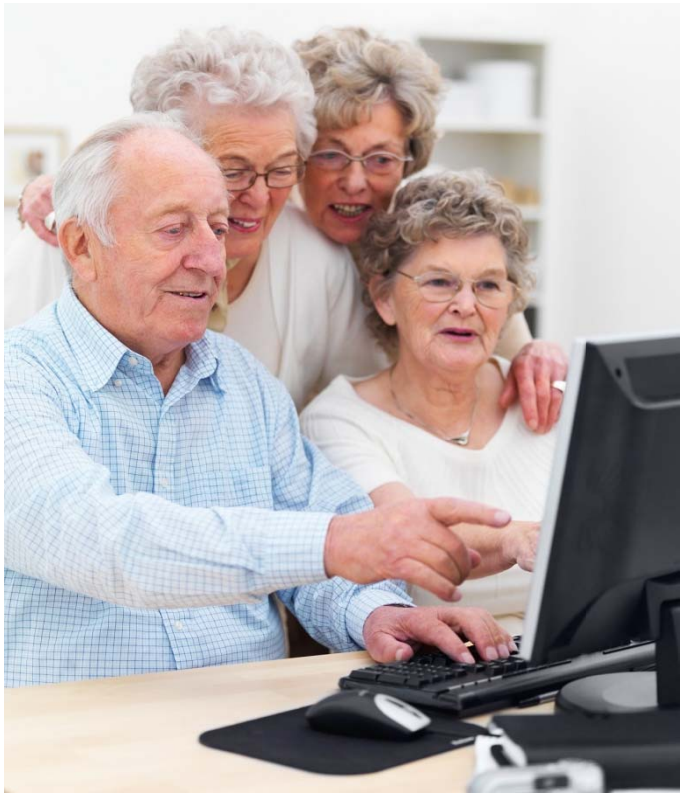
KCE REPORT 176B

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Centre

DÉPISTAGE DU CANCER DU SEIN ENTRE 70 ET 74 ANS



Le Centre fédéral d’expertise des soins de santé

Le Centre fédéral d’expertise des soins de santé est un parastatal, créé par la loi-programme (1) du 24 décembre 2002 (articles 259 à 281), sous tutelle du Ministre de la Santé publique et des Affaires sociales. Il est chargé de réaliser des études éclairant la décision politique dans le domaine des soins de santé et de l’assurance maladie.

Conseil d’administration

	Membres effectifs	Membres suppléants
Président	Pierre Gillet	
Fonctionnaire dirigeant de l'INAMI (vice président)	Jo De Cock	Benoît Collin
Président du SPF Santé publique (vice président)	Dirk Cuypers	Chris Decoster
Président du SPF Sécurité sociale (vice président)	Frank Van Massenhove	Jan Bertels
Administrateur général de l'AFMPS	Xavier De Cuyper	Greet Musch
Représentants du ministre de la Santé publique	Bernard Lange	François Perl
	Marco Schetgen	Annick Poncé
Représentants du ministre des Affaires sociales	Olivier de Stexhe	Karel Vermeyen
	Ri De Ridder	Lambert Stamatakis
Représentants du Conseil des ministres	Jean-Noël Godin	Frédéric Lernoux
	Daniel Devos	Bart Ooghe
Agence intermutualiste	Michiel Callens	Frank De Smet
	Patrick Verertbruggen	Yolande Husden
	Xavier Brenez	Geert Messiaen
Organisations professionnelles représentatives des médecins	Marc Moens	Roland Lemye
	Jean-Pierre Baeyens	Rita Cuypers
Organisations professionnelles représentatives des infirmiers	Michel Foulon	Ludo Meyers
	Myriam Hubinon	Olivier Thonon
Fédérations hospitalières	Johan Pauwels	Katrien Kesteloot
	Jean-Claude Praet	Pierre Smiets
Partenaires sociaux	Rita Thys	Leo Neels
	Paul Palsterman	Celien Van Moerkerke
Chambre des Représentants	Lieve Wierinck	

Contact

Centre Fédéral d'Expertise des Soins de Santé (KCE)
Doorbuilding (10e étage)
Boulevard du Jardin Botanique, 55
B-1000 Bruxelles
Belgique

T +32 [0]2 287 33 88
F +32 [0]2 287 33 85
info@kce.fgov.be
http://www.kce.fgov.be

KCE REPORT 176B
GOOD CLINICAL PRACTICE

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Centre

DEPISTAGE DU CANCER DU SEIN ENTRE 70 ET 74 ANS

FRANÇOISE MAMBOURG, JO ROBAYS, SOPHIE GERKENS

2012 www.kce.fgov.be

COLOPHON

Titre:	Dépistage du cancer du sein entre 70 et 74 ans
Auteurs:	Françoise Mambourg (KCE), Jo Robays (KCE), Sophie Gerkens (KCE)
Relecture:	Frank Hulstaert (KCE), Pascale Jonckheer (KCE), Nancy Thiry (KCE)
Experts externes:	Marc Arbijn (WIV - ISP), Martine Berlière (UCL Saint-Luc), Hilde Bosmans (UZ Leuven), Jean-Benoit Burrion (ASBL Brummammo), Joëlle Desreux (CHU Liège), André-Robert Grivegnée (Institut Jules Bordet), Patrick Neven (UZ Leuven), Myriam Provost (SSMG), Hubert Thierens (UGent), Reinhilde Van Eeckhoudt (WVG), Anne Vandenbroucke (UCL Saint-Luc), Geert Villeirs (UZ Gent).
Validateurs externes:	Philippe Autier (IPRI-Lyon), Geert Page (Jan Yperman Ziekenhuis), Chantal Van Ongeval (KU Leuven)
Conflits d'intérêt:	Aucun conflit déclaré
Layout:	Ine Verhulst

Disclaimer:

- **Les experts externes ont été consultés sur une version (préliminaire) du rapport scientifique. Leurs remarques ont été discutées au cours des réunions. Ils ne sont pas co-auteurs du rapport scientifique et n'étaient pas nécessairement d'accord avec son contenu.**
- **Une version (finale) a ensuite été soumise aux validateurs. La validation du rapport résulte d'un consensus ou d'un vote majoritaire entre les validateurs. Les validateurs ne sont pas co-auteurs du rapport scientifique et ils n'étaient pas nécessairement tous les trois d'accord avec son contenu.**
- **Finalement, ce rapport a été approuvé à l'unanimité par le Conseil d'administration.**
- **Le KCE reste seul responsable des erreurs ou omissions qui pourraient subsister de même que des recommandations faites aux autorités publiques.**

Date de publication:	26 april 2012
Domaine:	Good Clinical Practice (GCP)
MeSH:	Breast Neoplasms ; Mammography ; Mass Screening
Classification NLM:	WP 870 - Breast - Neoplasms
Langue:	français, anglais

Format: Adobe® PDF™ (A4)

Dépot légal: D/2012/10.273/19

Comment citer ce rapport? Mambourg F, Robays J, Gerkens S. Dépistage du cancer du sein entre 70 et 74 ans. Good Clinical Practice (GCP). Bruxelles: Centre Fédérale d'Expertise des Soins de Santé (KCE). 2012. KCE Report 176B. D/2012/10.273/19.

Ce document est disponible en téléchargement sur le site Web du Centre fédéral d'expertise des soins de santé.

PRÉFACE

Faire des choix en matière de soins ressemble vite à de la discrimination, en particulier lorsque ces choix sont basés sur l'âge. Comment par exemple justifier le refus de remboursement d'une intervention coûteuse sur le cœur à une personne, uniquement sur un critère d'âge, même si pour le reste cette personne est encore en bonne forme? De tels raisonnements conduisent systématiquement à des discussions enflammées, nourries à partir de systèmes de valeurs parfois diamétralement opposés.

La présente étude qui pose la question de savoir s'il faut offrir un dépistage organisé du cancer du sein aux femmes âgées de 70 à 74 ans, nous place donc à nouveau sur un sol glissant. Mais il y a encore d'autres raisons d'être particulièrement vigilants sur un tel sujet. Comme dans tout dépistage organisé, on s'adresse en effet à des gens qui ne présentent a priori pas de plainte de santé et qui n'étaient donc pas nécessairement demandeurs d'un tel examen. L'adage *primum non nocere* est donc ici d'autant plus important.

En matière d'argumentation à développer, il y a aussi un défi particulier à relever. Le clinicien est plus coutumier de la logique utilisée pour poser un diagnostic chez une personne qui a une plainte que de celle utilisée en matière de screening. Dans le premier cas, le risque d'un résultat faux positif est non seulement plus petit mais est aussi clairement considéré comme moins important qu'un résultat faux négatif, qui équivaut à louper un diagnostic. Cela explique pourquoi les inconvénients d'un dépistage sont systématiquement sous évalués. De plus, le sujet est loin de laisser l'opinion publique indifférente, un lobbying intense est organisé à son propos, et il est (donc) aussi sensible politiquement.

Même si on mobilise toutes les preuves scientifiques disponibles pour fonder un avis sur la question, on ne peut pas pour autant espérer arrêter la controverse. Nous osons néanmoins espérer que ce rapport apportera tout ce qu'on peut attendre d'un organe d'avis scientifique dans un tel débat.

Jean-Pierre CLOSON
Directeur Général Adjoint

Raf MERTENS
Directeur Général

RÉSUMÉ

INTRODUCTION

Ce travail fait partie d'un projet plus large ayant pour objet la mise à jour du rapport: «Dépistage du cancer du sein», publié en 2005 (rapport KCE n°11). Il concerne plus particulièrement l'extension du dépistage organisé aux femmes âgées de 70 à 74 ans qui ne présentent par ailleurs ni symptôme évocateur, ni facteur de risque particulier.

Le dépistage du cancer du sein est un processus complexe qui comporte des bénéfices et des risques. Les principaux bénéfices attendus du dépistage du cancer du sein sont la diminution de la mortalité et de la morbidité liées à la maladie. La diminution de morbidité implique soit un allègement des traitements, soit une diminution des récidives ou des stades métastatiques de la maladie.

Les risques principaux liés au dépistage concernent la qualité de vie. En effet, un résultat faussement positif, un diagnostic excédentaire (sur-diagnostic) suivi d'un traitement et l'avance au diagnostic (ou *lead time*, qui est la durée par laquelle le diagnostic par dépistage précède le diagnostic clinique) ont des conséquences sur la qualité de vie.

Les résultats faussement positifs ont pour conséquence d'inclure des femmes en bonne santé dans un circuit d'examens diagnostiques anxiogènes voire invasifs (biopsies).

Le sur-diagnostic peut être défini comme la détection de cas de cancers qui n'auraient jamais été perçus cliniquement en l'absence de dépistage. Il s'accroit au fur et à mesure de la diminution de l'espérance de vie de la population dépistée. Le sur-traitement est une conséquence du sur-diagnostic. Vu qu'il est actuellement impossible de prédire si un cancer va se développer, la très grande majorité des cancers diagnostiqués sont traités.

Enfin, le dépistage met les cancers en évidence deux ou trois ans plus tôt que ne le ferait un diagnostic clinique. Ceci implique que la personne devient «malade du cancer» et reçoit des traitements invasifs plus tôt dans le décours de sa vie.

QUESTIONS POSÉES

Ce rapport investigue la question suivante: faut-il étendre le dépistage organisé du cancer du sein aux femmes âgées de 70 à 74 ans?

Si la réponse à cette question est négative, une question subsidiaire se pose: que répondre à la personne de cette tranche d'âge qui demande un dépistage?

MÉTHODOLOGIE

L'étude des bénéfices cliniques du dépistage se base sur une revue de littérature effectuée dans OVID Medline, EMBASE, CDSR et DARE. Cette revue a inclus les articles publiés en Anglais, Allemand, Néerlandais et Français de janvier 2004 à avril 2011.

L'évaluation du rapport bénéfices-risques de ce dépistage se base sur une revue des études de modélisation recherchées dans Medline, Embase, NHS EED et Econlit. Cette revue a inclus les articles publiés en Anglais, Allemand, Néerlandais et Français de janvier 2000 à septembre 2011.

Afin de quantifier le rapport bénéfices-risques dans le contexte belge, un modèle a été construit dans ce but. La construction de ce modèle a nécessité de rechercher dans Medline, Embase, HTA EED et Psycinfo (1950-10/2011) les études relatives à la qualité de vie pendant et après le dépistage et le traitement du cancer du sein. Le modèle contient le maximum de données belges utilisables.

Enfin, des recommandations de bonne pratique ont été rédigées sur la base des éléments de preuve obtenus. Une révision desdites recommandations a été effectuée par les experts externes. Aucun conflit d'intérêts n'a été signalé.

RÉSULTATS ISSUS DE LA LITTÉRATURE

Mortalité

Les résultats des différents essais contrôlés randomisés permettent de mettre en évidence les faits suivants:

- Le dépistage entraîne une diminution de mortalité de 23% sur une période de suivi de 13 ans chez les femmes de plus de 50 ans ayant bénéficié d'un dépistage tous les deux ans.
- Cette diminution de mortalité se manifeste entre 4 et 7 ans après le dépistage. Il convient donc de la mettre en perspective avec l'espérance de vie moyenne de ce groupe d'âge qui est de 16 ans à 70 ans et de 13 ans à 74 ans (données belges de 2009).

Dans l'interprétation des études internationales, il faut tenir compte du petit nombre de participantes âgées de 70 à 74 ans; conséquemment, l'effet sur la mortalité n'a pu être statistiquement démontré pour celles-ci.

Morbidité

Outre le gain en années de vie, le principal avantage attendu du dépistage est de permettre des traitements moins agressifs, vu que le dépistage a pour objectif de mettre en évidence des petites tumeurs. Les données belges dont nous disposons actuellement ne nous permettent pas de valider cette assertion. Les données les plus récentes (rapport KCE 150) font état de 58% de chirurgie conservatrice versus 38% de mastectomies totales dans les stades les moins avancés (Stades I and II). Près de 90% des bénéficiaires de la chirurgie conservatrice reçoivent également un traitement par radiothérapie, 38% d'entre elles reçoivent un traitement de chimiothérapie néo-adjuvante et 41% un traitement hormonal.

D'autre part, les essais contrôlés randomisés n'ont quantifié ni le taux de récidives ni l'évolution vers les stades métastatiques de la maladie. Il est donc impossible d'infirmer ou de confirmer l'hypothèse d'une réduction de morbidité sur cette base. Par contre, la perte de qualité de vie imputable aux métastases est inclue dans le modèle (voir ci-dessous).

ÉTUDES DE MODÉLISATION

Les principales études de modélisation ont été réalisées au sein du projet CISNET (Cancer Intervention and Surveillance Modeling Network). Ces modèles avaient pour objectif d'évaluer la contribution relative du dépistage par mammographie et du traitement adjuvant sur la réduction de la mortalité due au cancer du sein observée aux Etats-Unis de 1975 à 2000, et ils utilisent les données issues du Breast Cancer Screening Consortium.

Les résultats de ces modèles indiquent un gain en années de vie allant de 9 à 22 ans par 1.000 femmes dépistées. D'autres modèles n'utilisant pas la méthodologie CISNET sont également décrits dans le rapport scientifique.

Ces modèles ne sont pas adaptables en tant que tels à la situation belge car il est impossible d'y inclure les données belges. Un nouveau modèle spécifique a donc été construit.

UN MODÈLE DE COHORTE POUR LA BELGIQUE

Méthodologie

Le modèle construit pour ce rapport est un modèle de cohorte qui évolue par cycles annuels. Il compare deux cohortes théoriques de femmes de plus de 70 ans, l'une sans invitation au dépistage (situation actuelle) et l'autre où les femmes continuent à être invitées au dépistage. Le taux de participation et la répartition des cancers détectés par le dépistage versus les cancers d'intervalle sont considérés comme étant les mêmes que dans la tranche d'âge 50-69 ans.

Le dépistage a pour objectif de mettre en évidence les tumeurs à un stade précoce (I et II) afin d'éviter l'évolution vers le stade IV (stade métastatique) qui est incurable. Ce «stage-shift» implique que parmi tous les cancers dépistés, la proportion des stades précoces (I, II) augmente en même temps que la proportion des stades avancés (II et IV) diminue. D'autre part, nous avons émis l'hypothèse selon laquelle la survie et la qualité de vie dépendent de l'âge de la patiente et du stade de la tumeur. Elle ne tient pas compte du fait que le pronostic des cancers détectés par le dépistage est meilleur que celui des cancers détectés cliniquement (cancers d'intervalle et cancers survenant chez les femmes non-participantes).

Paramètres

Ce modèle exploite au maximum les données belges, à savoir : l'espérance de vie moyenne des femmes selon leur âge (2009), les données du registre du cancer (Communauté Flamande), les données issues du programme de dépistage actuel (50-69 ans), le temps nécessaire pour infirmer un diagnostic faussement positif (Agence Intermutualiste, AIM/IMA) et les données de survie à cinq ans en fonction du stade (Registre du Cancer). Les données de la Communauté Flamande ont été privilégiées car elles sont plus complètes et parce que le dépistage opportuniste après 70 ans y est moins fréquent que dans le reste du pays. La durée de l'avance au diagnostic et le pourcentage de sur-diagnostic ont été estimées au départ de l'analyse de la littérature.

Mesure de la qualité de vie

Les données sur la qualité de vie pendant le dépistage et le traitement proviennent de la littérature. L'instrument utilisé pour décrire les états de santé est l'EQ-5D (European Quality of Life-5 Dimensions); ces descriptions ont été valorisées par la population générale anglaise («*UK tariffs*»). Nous ne disposons pas de données relatives à la population belge.

Les variations de la qualité de vie des femmes de plus de 70 ans utilisées dans le modèles sont les suivantes:

- La perte de qualité de vie consécutive à un résultat de dépistage faussement positif est estimée à 16% pendant 45 jours.
- Pour les patientes cancéreuses et pendant la première année qui suit le diagnostic (quel que soit le traitement) la perte de qualité de vie est estimée à 16% pour les stades I, II, III et à 18% pour les stades IV. Pendant les années suivantes, la perte de qualité de vie est estimée à 6% pour les stades I, II, III. Cette perte demeure stationnaire (18%) pour les stades IV.

Plusieurs limitations de cette approche nous obligent à interpréter ces chiffres avec précaution

Résultats

Le scénario de base montre que le dépistage entre 70 et 74 ans permettrait d'éviter 1,3 décès pour 1000 femmes qui y participent, ce qui représente une réduction de 21% des décès. Globalement, le nombre d'années de vie sauvées est estimé à 13,1 et le gain en QALY à 3,9.

Etant donné qu'il existe une incertitude importante (pour les détails, voir les discussions dans le rapport scientifique) au sujet des ces estimations, une analyse de sensibilité du modèle a été réalisée. Cette analyse comprend un scénario pessimiste et un scénario optimiste.

Le scenario pessimiste fait l'hypothèse d'un excédent de diagnostic de 20%, d'un taux de faux positifs de 10%, entrainant une perte de qualité de vie de 0,19 perdurant pendant 54 jours (temps nécessaire pour infirmer les résultats). La distribution des cancers dépistés par stades observée actuellement dans le cadre du dépistage organisé en Flandre (50-69 ans) a été appliquée au groupe dépisté. Ce scénario pessimiste permet d'estimer un gain de 8,7 années de vie et **une perte** de 3,1 QALY pour 1000 femmes participant au dépistage. Ceci signifie que dans certaines circonstances, au demeurant tout à fait réalistes, le dépistage puisse aboutir à une perte en terme de qualité de vie.

Le scenario optimiste fait l'hypothèse d'un excédent de diagnostic de 3%, d'un taux de faux positifs de 2%, entrainant une perte de qualité de vie de 0,13 perdurant pendant 36 jours. Ce scénario applique au groupe dépisté la distribution par stades observée actuellement dans le cadre du dépistage organisé aux Pays-Bas (70-74 ans). Ce scénario optimiste permet d'estimer **un gain** de 17,0 années de vie et un gain de 16,3 QALY pour 1000 femmes participant au dépistage. Ceci signifie qu'il serait nécessaire d'inviter 67 femmes à participer au dépistage pendant cinq ans pour gagner un QALY.

CONCLUSION

Le dépistage est organisé dans le but d'améliorer le bien-être de la population en évitant notamment des décès prématurés. Il est certain que prolonger le dépistage jusqu'à l'âge de 74 ans devrait permettre de gagner quelques années de vie pour un certain nombre de femmes. Toutefois, l'influence d'un dépistage organisé sur la qualité de vie est nettement plus aléatoire (niveau de preuve très faible car basé sur un modèle). Selon des hypothèses raisonnables, cette intervention pourrait même aboutir à une perte en terme de qualité de vie. Dans ces conditions, il se pourrait que la balance bénéfices-risques de ce dépistage penche du côté d'une perte globale de bien-être de la population. Il n'est donc pas recommandé d'étendre le dépistage organisé du cancer du sein aux femmes âgées de 70 à 74 ans.

RECOMMANDATIONS[a]

- **L'invitation systématique des femmes âgées de 70 à 74 ans à participer au dépistage organisé du cancer du sein n'est pas recommandée.**
- **Si une personne de plus de 70 ans demande une mammographie dans un objectif de dépistage, il importe que le médecin veille à ce qu'elle soit bien informée des avantages et des inconvénients potentiels de celle-ci.**
- **Toute mammographie de dépistage doit répondre aux exigences européennes en matière de qualité, dont notamment : le contrôle de la qualité des installations, la double lecture, l'enregistrement et l'optimisation du taux de rappel. C'est pourquoi, les médecins orienteront la personne qui demande un dépistage vers une structure qui réponde à ces exigences de qualité.**
- **Afin de minimiser le risque de perte de qualité de vie lié aux résultats faussement positifs, il importe que le taux de rappel après mammographie soit le plus bas possible et reste en dessous du seuil défini par les critères européens (<5%).**

[a] Le KCE reste seul responsable des recommandations faites aux autorités publiques

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABBREVIATION	DEFINITION
CPG	Clinical Practice Guideline
CCRT	Cochrane Central Register of Controlled Trials
CI	Confidence Interval
DCIS	Ductal Carcinoma in situ
DET	Data Extraction Table
BCSC	Breast Cancer Surveillance Consortium (USA)
AHRQ	Agency for Health Care Research and Quality
BCR	Belgian Cancer Registry
DNETB	Dutch National Evaluation Team for Breast cancer screening
CISNET	Cancer Intervention and Surveillance Modelling Network
IMA/AIM	Intermutualistic Agency
INAMI/RIZIV	National Institute for Health and Disability Insurance
ICER	Incremental cost-effectiveness ratio
KCE	Belgian Healthcare Knowledge Centre
MST	Mean Sojourn Time
M-A	Meta-analysis
NIS	National Institute for Statistics
NBSS	Canadian National Breast Cancer Screening Study
NBCSP	Norwegian Breast Cancer Screening Programmes
NHS	National Health Service (UK)
NHS EED	NHS Economic Evaluation Database
NCI	National Cancer Institute (USA)
QALY	Quality Adjusted Life Year
QoL	Quality of Life
RCT	Randomized Controlled Trial
RR	Relative Risk

SEER	Surveillance, Epidemiology and End Results (USA)
SR	Systematic Review
ST	Sojourn Time
TTO	Time-trade-off
UK	United Kingdom
USA	Unites States of America
USPSTF	US Preventive Services Task Force

SYNTHÈSE

1. CONTEXTE

Le KCE a déjà publié trois rapports sur le dépistage du cancer du sein. Le rapport de base publié en 2005 (rapport N°11 du KCE) concernait le dépistage du cancer du sein en général, dans la population sans facteurs de risque. Le dépistage du cancer du sein des femmes de la tranche d'âge 40-49 ans a fait l'objet d'une mise à jour partielle publiée en 2010. Dans ce rapport (rapport N° 129 du KCE), le KCE ne recommandait pas le dépistage systématique des femmes de moins de 50 ans. Le troisième rapport (rapport 172 du KCE), publié en 2012, a posé le problème de l'indentification des femmes exposées à un risque accru de cancer du sein. Le rapport actuel pose la question de l'extension du dépistage organisé du cancer du sein aux femmes âgées de 70 à 74 ans

Cette question est régulièrement adressée aux politiciens en raison de l'augmentation régulière de l'espérance de vie de la population féminine. Si la plupart des groupes actifs dans le dépistage demandent cette prolongation, les autorités publiques font preuve de moins d'unanimité. Seuls quatre Etats membres de l'Union européenne ciblent la tranche d'âge des 70-74 (la France, les Pays-Bas, l'Espagne et la Suède)[1]. Les autres pays insistent sur la nécessité d'informer les femmes et de partager avec elles la prise de décision.

2. QUESTIONS POSÉES

Le dépistage organisé du cancer du sein devrait-il être prolongé jusqu'à l'âge de 74 ans? Si la réponse à cette question est négative, que répondre à la personne qui demande ce dépistage?

La première question concerne plus spécifiquement les pouvoirs publics et la seconde, les prestataires de soins.

3. DESCRIPTION DE LA PROBLÉMATIQUE

3.1. Approche intuitive

De façon intuitive, le dépistage du cancer fait sens. Les médias sont généralement enthousiastes à l'égard du dépistage. Cette attitude a été démontrée par Schwartz au début du 21e siècle[2]. Une enquête réalisée aux Etats-Unis a révélé que 87% des adultes considéraient que dépister est une bonne idée. Trois quart des personnes interrogées déclaraient que diagnostiquer un cancer à un stade précoce sauve la vie la plupart du temps. L'enthousiasme des répondants était si fort que pour la majorité d'entre eux le dépistage n'était pas une décision à prendre mais un impératif moral[3].

Cette attitude générale que nous pouvons résumer ainsi "la détection précoce des cancers sauve des vies" peut avoir suscité des attentes irréalistes de la part des femmes. Silverman a réalisé des interviews téléphoniques pour évaluer comment les femmes considèrent le cancer du sein et le bénéfice du dépistage par mammographie[4]. La majorité des répondantes considérait le cancer du sein comme une maladie progressive uniforme et croyait que tous les cancers débutent par une forme curable et silencieuse. En résumé, ces femmes pensaient que si le cancer du sein n'est pas détecté par une mammographie et traité de façon précoce, il grandit, se propage et tue. Fortes de ces croyances, les femmes estimaient que les cancers avancés (et sans doute la plupart des cancers mortels) sont liés à un échec au niveau du dépistage précoce.

Schwartz a souligné que 94% des femmes ne savent pas que le dépistage peut détecter des cancers qui ne vont jamais progresser. De plus, 92% des répondantes sont persuadées du fait que la mammographie ne peut faire de tort à une personne qui n'a pas de cancer du sein[5].

Le corps médical lui-même n'appréhende pas toujours le dépistage de manière adéquate. C'est ainsi que de nombreux cliniciens restent focalisés sur le taux de cancers diagnostiqués (objectif intermédiaire), alors que l'objectif final du dépistage est de diminuer la mortalité. D'autre part, les cliniciens paraissent plus sensibilisés au risque de méconnaître un cancer (faux négatif) qu'aux risques liés aux résultats faussement positifs.

3.2. Approche épidémiologique

Le cancer du sein est le cancer le plus fréquent chez la femme. En Belgique, 10.849 cancers du sein ont été diagnostiqués en 2008. Plus de trois quarts des cancers du sein sont diagnostiqués après l'âge de 50 ans. L'âge moyen au moment du diagnostic est de 62 ans. L'incidence du cancer du sein est de **370,7/100.000** dans le groupe des femmes âgées de 70 à 75 ans[6].

Néanmoins, la part relative de la mortalité due au cancer du sein dans la mortalité totale diffère en fonction de l'âge. En 1999, le cancer du sein était responsable de 18% des décès chez les femmes âgées de 50 à 54 ans, de 13% dans le groupe de 60 à 64 ans et de 6% dans le groupe de 70 à 74 ans (Rapport N°11 du KCE). En 2006, cette proportion était de 14% pour les femmes âgées de 50 à 54, 12% pour le groupe de 60 à 64 ans, 7% pour le groupe des 70 à 74 ans et 5% pour le groupe des 75 à 79 ans[6].

Caractéristiques fondamentales d'un dépistage:

1. Le dépistage s'adresse à des personnes en bonne santé **Contrairement au patient qui consulte son médecin en raison d' une plainte ou d'un symptôme, la personne qui participe à un dépistage est présumée indemne de la maladie recherchée.**
2. Le dépistage a pour objectif à court terme de confirmer l'absence de la maladie.
3. Le dépistage a pour objectif ultime de diminuer la mortalité/morbidité liée à la maladie.
4. Le principe "primum non nocere" est particulièrement d'application en ce qui concerne le dépistage.

Rappelons que pour mille femmes dépistées entre 70 et 74 ans, plus de 990 sont indemnes du cancer du sein.

3.2.1. Objectif à court terme

Le dépistage a pour objectif de confirmer l'absence de la maladie. La personne qui participe au dépistage bénéficie de la "présomption d'innocence" en ce qui concerne le cancer du sein. A l'inverse, la patiente qui consulte son médecin parce qu'elle a une plainte ou parce qu'elle a constaté quelque chose d'inhabituel, devient "suspecte" de maladie. L'objectif du médecin et les moyens à mettre en œuvre dans ces deux situations sont diamétralement opposés. Dans le cas d'une mise au point diagnostique, le médecin a le devoir de tout mettre en œuvre pour trouver une étiologie à la plainte ou au symptôme. A l'inverse, dans le cadre du dépistage, le médecin a le devoir de pratiquer uniquement les examens indispensables. Ceci afin de minimiser les risques et les inconvénients du dépistage pour les 996 femmes (/1.000) qui sont indemnes du cancer du sein.

Le formation des médecins étant essentiellement effectuée en hôpital auprès de malades, ce changement de point de vue est franchement contre-intuitif pour un clinicien.

3.2.2. Objectif ultime

Diagnostiquer les cancers à un stade précoce avant qu'ils ne se développent et essaiment (métastases) est l'hypothèse fondatrice du dépistage du cancer. C'est ainsi que l'on attend du dépistage qu'il diminue la mortalité spécifique à la maladie et conséquemment la mortalité totale. Le fait que la technologie utilisée permette de diagnostiquer des lésions peu avancées et donc potentiellement curables ne représente qu'une étape intermédiaire dans ce processus. Il s'agit d'une condition nécessaire mais qui n'est pas suffisante[7].

On peut également émettre l'hypothèse que le dépistage réduise la morbidité liée à la maladie, en permettant l'utilisation de traitements moins invasifs (mastectomies partielles plutôt que mastectomies totales) et en évitant une partie des évolutions vers les stades métastatiques.

3.2.3. Faux positifs et diagnostics excédentaires

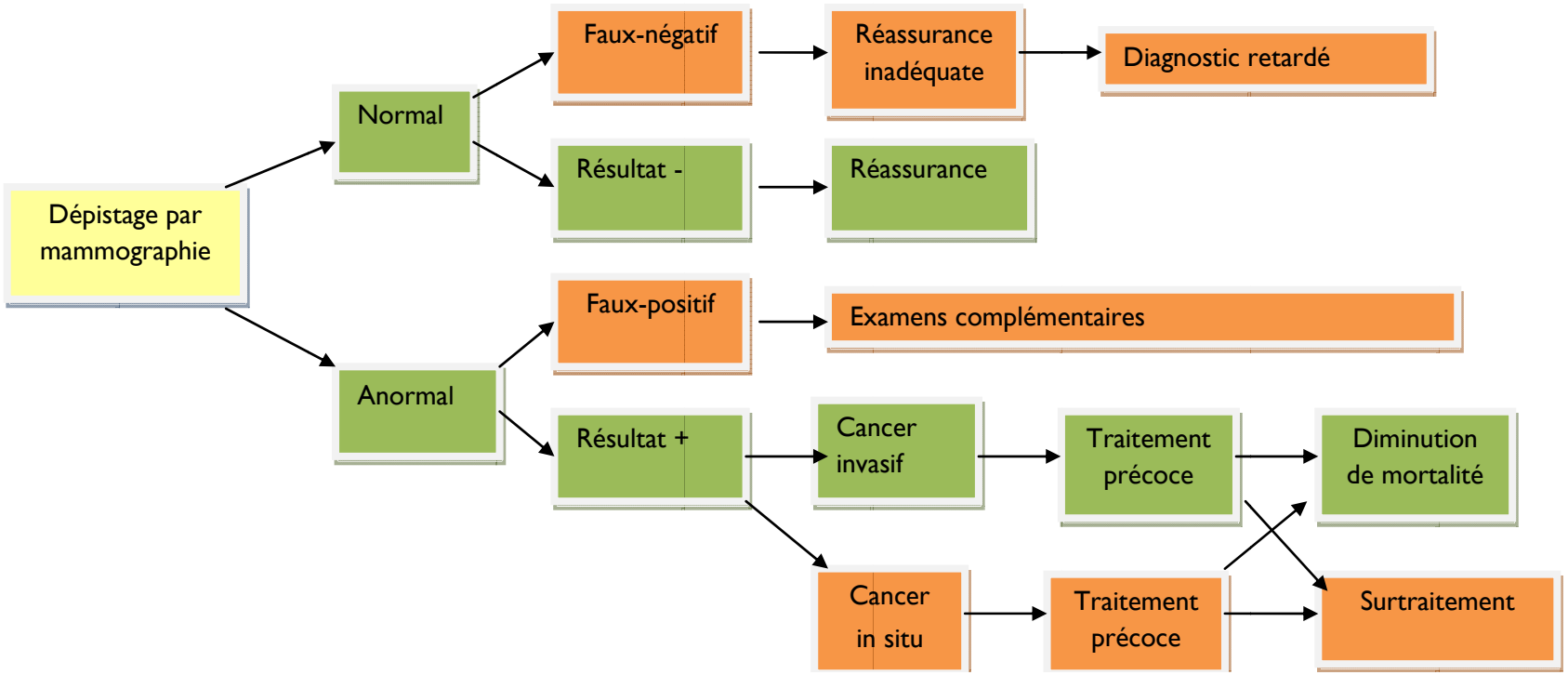
Avant d'instaurer un dépistage organisé, il est nécessaire de s'assurer que la balance avantages/inconvénients du dépistage penche du côté des avantages. Pour ce faire, l'ampleur de la diminution de mortalité doit contrebalancer la perte de qualité de vie consécutive aux inconvénients et aux risques induits par le dépistage.

Les résultats dits: "faux-positifs" (suspicion de lésion cancéreuse en dehors de la présence d'un cancer) sont les effets négatifs indésirables du dépistage du cancer du sein les plus fréquents. Ces résultats faussement positifs créent de hauts niveaux d'anxiété et sont suivis d'examens complémentaires.

Plus encore que les faux–positifs, le risque de diagnostic excédentaire est le risque majeur du dépistage des femmes âgées de 70 à 74 ans. Le diagnostic excédentaire peut se définir comme le fait de diagnostiquer un cancer dont l'évolution est telle qu'il ne se serait jamais manifesté cliniquement en l'absence de dépistage[8]. Ce risque est d'autant plus élevé que le cancer est d'évolution lente et que l'espérance de vie de la personne est faible. Ce risque est particulièrement méconnu dans la population. Très peu de femmes savent en effet que certains cancers évoluent tellement lentement que même s'ils ne sont pas traités ils n'altèreront pas la santé[9].

Ce rapport a pour objectif de quantifier les avantages et les inconvénients (voir Figure 1) de ce dépistage afin de pouvoir les mettre en perspective et s'assurer que les bénéfices l'emportent largement sur les risques de perte de qualité de vie.

Figure 1 mise en perspective des avantages et des inconvénients potentiels du dépistage.



4. MÉTHODOLOGIE

Nous avons recherché des éléments de réponse aux questions précitées, dans la littérature clinique, dans les études de modélisation et dans les données nationales et internationales. Ces recherches ont été menées en suivant les procédures en vigueur au KCE. Elles sont décrites en détails dans le chapitre 2 du rapport scientifique.

4.1. Estimation des bénéfices du dépistage

4.1.1. Diminution de la mortalité

Les principales données probantes relatives au dépistage du cancer du sein, sont issues de huit essais contrôlés randomisés. Sur base de ces essais, on peut retenir deux constats principaux:

1. Le dépistage entraîne une diminution de mortalité de 23% sur une période de suivi de 13 ans pour les femmes de plus de 50 ans ayant bénéficié d'un dépistage tous les deux ans.
2. Cette diminution de mortalité se manifeste principalement entre 4 et 7 ans après le dépistage. Il convient de la mettre en perspective avec l'espérance de vie de la population-cible. L'espérance de vie moyenne de ce groupe d'âge est de 16 ans à 70 ans et de 13 ans à 74 ans (données belges de 2009).

Les données probantes issues de ces essais contrôlés randomisés ne peuvent donner une réponse complète à notre question de base. En effet, un seul essai randomisé, l'étude suédoise dite des "Two County", a inclus des femmes âgées de 70 à 74 ans et le nombre de septuagénaires participant à cet essai était trop faible (10.000 pour les deux groupes) pour pouvoir mettre en évidence un effet statistiquement significatif sur la mortalité. De plus cette étude était entachée de biais méthodologiques.

4.1.2. Amélioration de la qualité de vie des patientes

Le dépistage ayant pour objectif de mettre en évidence des tumeurs de petite taille, un des avantages attendus est de permettre des traitements moins agressifs. Ni les données issues de essais contrôlés randomisés, ni les données factuelles recueillies en Belgique, ne permettent de confirmer cette attente.

Les essais contrôlés randomisés n'ont quantifié ni le taux de récidives ni l'évolution vers les stades métastatiques de la maladie. Il est donc impossible d'infirmer ou de confirmer l'hypothèse d'une réduction de morbidité sur cette base. Par contre, la perte de qualité de vie imputable aux métastases est inclue dans le modèle décrit ci-dessous.

Les données belges dont nous disposons actuellement ne nous permettent pas de valider cette assertion. Les données les plus récentes (rapport KCE 150) font état de 58% de chirurgie conservatrice versus 38% de mastectomies totales dans les stades les moins avancés (Stades I et II). Près de 90% des bénéficiaires de la chirurgie conservatrice reçoivent également un traitement par radiothérapie, 38% d'entre elles reçoivent un traitement de chimiothérapie néo-adjuvante et 41% un traitement hormonal.

4.2. Estimation des inconvénients du dépistage

4.2.1. Diminution de la qualité de vie des participantes

Le dépistage provoque une diminution de la qualité de vie d'une partie des personnes dépistées. Ceci s'explique par une série de facteurs:

1. Les résultats faussement positifs du dépistage sont perçus par les patientes comme de vrais positifs, aussi longtemps que les examens complémentaires n'ont pas permis de les infirmer. Ils provoquent de l'inquiétude par rapport au cancer du sein et aux procédures invasives telles que les ponctions mammaires.
2. Les diagnostics excédentaires et les traitements qui les suivent (over-diagnosis and over-treatment, pour plus de détails, voir le rapport scientifique) conduisent à des inquiétudes graves et à des traitements lourds dont des amputations mammaires qui n'ont pas d'influence sur la survie de la personne.

3. L'avance au diagnostic peut entrainer une perte de plusieurs années de vie en bonne santé. Le dépistage a pour objectif de diagnostiquer le cancer plus précocement que ne le ferait un diagnostic clinique. La patiente devient de ce fait malade du cancer plus tôt dans le décours de sa vie. Toutefois, si cette patiente décède d'une cause indépendante de son cancer avant que celui-ci n'ait eu le temps d'évoluer, elle aura été "malade du cancer" quelques années trop tôt sans que cette avance au diagnostic et au traitement n'aient pu influencer son espérance de vie[10].

4.3. Approche par modélisation

Les revues de littérature précitées ne nous ayant pas permis de quantifier le poids des bénéfices et des risques, nous avons construit un modèle spécifique dans ce but. La construction de ce modèle a nécessité de rechercher les études relatives à la qualité de vie des femmes pendant le dépistage et à la qualité de vie des patientes au cours de leur maladie.

4.3.1. Mesures de la qualité de la vie

Différents instruments sont disponibles pour mesurer la qualité de vie. Certains instruments sont spécifiquement adaptés à la maladie, comme par exemple, le questionnaire relatif à la qualité de vie des patientes atteintes d'un cancer du sein de l'European Organization for Research and Treatment of Cancer (EORTC). Ces outils évaluent l'image du corps, le fonctionnement physiologique, la peur de la récidive… Toutefois, il n'est pas possible de prendre en compte ces données de santé multi-dimensionnelles dans un modèle. Elles doivent être converties en un indice global de qualité de vie, à savoir, le Quality-Adjusted Life-Year (QALY). Les QALYs sont le nombre d'années de vie ajustées à la qualité de vie.

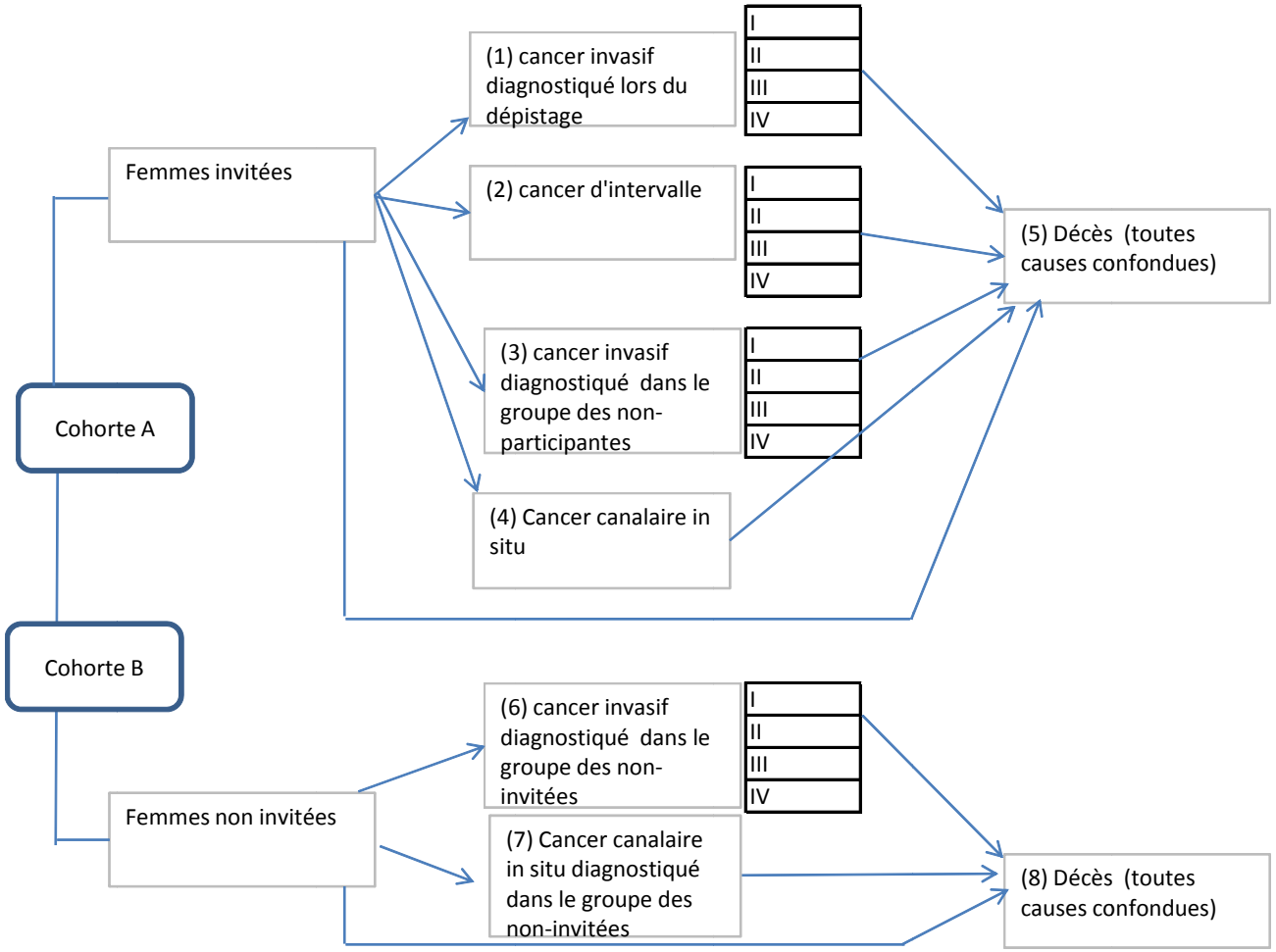
Les recommandations pharmaco-économiques du KCE, considèrent que le questionnaire appelé EQ-5D (European Quality of Life-5 Dimensions) est un des meilleurs instruments disponibles pour évaluer les QALYs. Avec cet instrument, la qualité de la vie liée à l'état de santé est mesurée en prenant en compte cinq dimensions: la mobilité, l'autonomie de la personne, les activités courantes, la douleur/la gêne, l'anxiété/la dépression. Pour chacune de ces dimensions, plusieurs réponses sont possibles. Celles-ci reflètent le niveau de sévérité du problème (aucun problème, quelques problèmes, des problèmes modérés, ou des problèmes graves) Ce questionnaire est soumis à la population concernée, soit en ce qui concerne le dépistage, une population de femmes indemnes du cancer du sein et en ce qui concerne la maladie, une population de personnes atteintes de ce cancer. La revue de la littérature a permis d'identifier trois études qui correspondaient à nos critères d'inclusion. Sur base de ces études, les variations de la qualité de vie des septuagénaires sont estimées comme suit:

1. La perte de qualité de vie consécutive à un résultat de dépistage faussement positif est estimée à 16% pendant la période nécessaire pour infirmer ce faux positif. En Belgique, cette période dure en moyenne 45 jours (minimum 36, maximum 54 jours) selon les données AIM (Agence Intermutualiste)
2. Pour les patientes cancéreuses et pendant la première année qui suit le diagnostic (quel que soit le traitement), la perte de qualité de vie est estimée à 16% pour les stades I, II, III et à 18% pour les stades IV. Pendant les années suivantes, la perte de qualité de vie est estimée à 6% pour les stades I,II,III. Cette perte demeure stationnaire (18%) pour les stades IV.

Plusieurs limitations de cette approche nous obligent à interpréter ces chiffres avec précaution. Il s'agit de résultats provenant de pays anglo-saxons. Le questionnaire utilisé, à savoir, l'EQ-5D mesure les dimensions sanitaires générales et non les dimensions spécifiques au cancer du sein. Les mesures concernant les patientes ne prennent que sommairement en compte l'impact à court terme du diagnostic et de la chirurgie. Ce questionnaire ayant été utilisé lors des consultations ambulatoires; ses résultats ne reflètent pas la qualité de vie des patientes gravement malades ne pouvant plus se déplacer. Les particularités de l'étude utilisée pourraient expliquer le faible changement de qualité de vie constaté entre les patientes ayant un cancer du sein et la population générale ou entre les patientes ayant développé des métastases et celles qui n'en ont pas.

4.3.2. Description du modèle

Le modèle compare deux cohortes théoriques. Ces deux cohortes sont constituées de 100.000 femmes dont l'évolution est suivie jusqu'à la mort. Le schéma ci-dessous représente cette évolution:


Cohorte A
Femmes invitées
(1) cancer invasif diagnostiqué lors du dépistage
I
II
III
IV
(2) cancer d'intervalle
I
II
III
IV
(3) cancer invasif diagnostiqué dans le groupe des non-participantes
I
II
III
IV
(4) Cancer canalaire in situ
(5) Décès (toutes causes confondues)
Cohorte B
Femmes non invitées
(6) cancer invasif diagnostiqué dans le groupe des non-invitées
I
II
III
IV
(7) Cancer canalaire in situ diagnostiqué dans le groupe des non-invitées
(8) Décès (toutes causes confondues)

La cohorte A illustre l'hypothèse d'une prolongation du dépistage organisé jusqu'à 74 ans. Elle est constituée des femmes invitées au dépistage. Parmi celles-ci, certaines participent au dépistage (invitées/participantes) et d'autres non (invitées/non-participantes). Les cancers survenant dans la cohorte sont répertoriés. Il s'agit soit de cancers diagnostiqués lors du dépistage (1), soit de cancers diagnostiqués dans l'intervalle entre deux sessions de dépistage (2), soit de cancer diagnostiqués dans le groupe des invitées/non-participantes (3). Enfin, les cancers canalaires in situ peuvent survenir dans le groupe des invitées/participantes tout comme dans le groupe des invitées/non-participantes (4). La très grande majorité des femmes constituant cette cohorte décèdera d'une autre affection que le cancer du sein (5).

La cohorte B (cohorte de contrôle) correspond à la situation actuelle. Les membres de cette cohorte ne sont pas invitées au dépistage. Certaines femmes seront atteintes d'un cancer invasif (6), d'autres d'un cancer canalaire in situ (7). La très grande majorité des femmes constituant cette cohorte décèdera d'une autre affection que le cancer du sein (8).

Le cancer du sein évolue en quatre stades (I, II, III, IV). Le stade I est le stade le moins avancé. La survie est d'autant moins bonne et le traitement d'autant plus lourd et plus invasif que le stade est avancé au moment du diagnostic.

4.3.3. Hypothèses de base

L'hypothèse de base est la suivante: parmi les cancers détectés par le dépistage, la proportion de stades peu avancés (I et II) est plus importante que parmi les cancers diagnostiqués sur base de la clinique. Tout le bénéfice du dépistage provient des différences dans la répartition des stades (stage-shift) consécutive au dépistage.

L'autre hypothèse retenue est que la survie et la qualité de vie des femmes dépendent uniquement du stade de la tumeur et de l'âge de la femme au moment du diagnostic, que celui-ci soit consécutif ou non au dépistage.

Les cohortes sont suivies d'année en année et évoluent en fonction de paramètres de transition tels le nombre de femmes atteintes chaque année (incidence) et le taux de survie en fonction du stade du cancer.

4.3.4. Alimentation du modèle

Pour réaliser cet exercice, nous avons autant que faire se peut, alimenté notre modèle avec des données belges. Ces paramètres sont décrits en détails dans le chapitre 3.3. du rapport.

L'espérance de vie de la population étudiée provient des tables de survie de la population féminine belge du même âge. L'incidence du cancer en fonction de l'âge et des stades de la maladie provient du registre belge du cancer (Communauté flamande). Les données relatives au dépistage sont issues des programme actuels (femmes de 50-69 ans en Wallonie, à Bruxelles et en Communauté Flamande).

Une mesure de qualité de vie a été appliquée à chaque compartiment du modèle. Le modèle contient un cas de base (base case) qui correspond à la situation la plus vraisemblable.

"Par essence, tous les modèles sont faux mais certains sont utiles"[a]

4.3.5. Analyse de sensibilité

Dans notre modèle, nous avons émis un certain nombre d'hypothèses simplificatrices, en raison des données dont nous disposions et de la nécessité d'éviter l'utilisation d'un modèle trop complexe. Ce choix conduit à une incertitude liée à la structure du modèle, au bon choix des paramètres et de la source des informations. Pour faire face à ces différents types d'incertitude, nous avons réalisé une analyse de sensibilité approfondie utilisant différents scénarios. Ces différents scénarios sont décrits en détails dans la table 3.2 du rapport scientifique.

[a] citation attribuée au statisticien George Box.

5. RÉSULTATS

Le scénario de base montre que la prolongation du dépistage jusqu'à 74 ans permettrait d'éviter 1,3 décès pour 1000 femmes participantes, ce qui représente une réduction de 21% des décès. Globalement, le nombre d'années de vie sauvées est estimé à 13,1 et le gain en QALY à 3,9.

L'analyse de sensibilité du modèle comprend un scénario pessimiste et un scénario optimiste .

Le scenario pessimiste fait l'hypothèse d'un excédent de diagnostic de 20%, d'un taux de faux positifs de 10%, entrainant une perte de qualité de vie de 0,19 perdurant pendant 54 jours (temps nécessaire pour infirmer les résultats). La distribution des cancers dépistés par stades observée actuellement dans le cadre du dépistage organisé en Flandre (50-69 ans) a été appliquée au groupe dépisté. Ce scénario pessimiste permet d'estimer un gain de 8,7 années de vie mais une **perte de 3,1 QALYs** pour 1000 femmes participant au dépistage. Ceci signifie que dans certaines circonstances, au demeurant tout à fait réalistes, le dépistage peut aboutir à une perte en terme de qualité de vie.

Le scenario optimiste fait l'hypothèse d'un excédent de diagnostic de 3%, d'un taux de faux positifs de 2%, entrainant une perte de qualité de vie de 0,13 perdurant pendant 36 jours. Ce scénario applique au groupe dépisté la distribution par stades observée actuellement dans le cadre du dépistage organisé aux Pays-Bas (70-74 ans). Ce scénario optimiste permet d'estimer un gain de 17,0 années de vie et un gain de 16,2 QALY pour 1000 femmes participant au dépistage. Ceci signifie qu'il serait nécessaire d'inviter 62 femmes à participer au dépistage pendant cinq ans pour gagner un QALY.

6. DISCUSSION

Les résultats du modèle décrit ci-dessus indiquent que en ce qui concerne la situation de base, le gain en années de vie est de 13 ans pour 1000 femmes dépistées. Ce résultat reste fiable tout au long de l'analyse de sensibilité. A l'inverse, les QALYs varient substantiellement en fonction des hypothèses choisies, allant d'un gain relativement faible à, selon certaines hypothèses plausibles, une perte en qualité de vie.

6.1. Ajouter des années à la vie ?

L'augmentation de l'espérance de vie de la femme est un des arguments utilisés pour justifier de poursuivre le dépistage du cancer du sein chez la femme âgée de plus de 69 ans. Cet argument présuppose que la population des septuagénaires a les mêmes caractéristiques que la population des sexagénaires. Il n'en est rien en ce qui concerne la fréquence et les causes de décès.

Le nombre de décès observé dans la tranche d'âge des 70-79 ans est deux fois et demi plus élevé que celui de la tranche d'âge des 60-69 ans. En fait, la population féminine belge perd 4% de ses effectifs entre 50 à 59 ans, 8% entre 60 à 69 ans et 20% entre 70 à 79 ans (Belgian life table 2009).

Les causes de décès varient également. En Belgique, la proportion des décès dus au cancer du sein passe de 13% entre 60 et 64 ans à 6% de tous les décès entre 70 et 75 ans. A cet âge, la mortalité par cancer tout comme la mortalité cardiovasculaire sont pratiquement équivalentes et responsables chacune d'un peu plus d'un tiers des décès. Parmi tous les décès, la part de décès consécutifs au cancer du sein diminue donc avec l'âge (KCE report 11).

6.2. Ajouter de la (qualité de) vie aux années ?

6.2.1. Traitements moins agressifs?

Outre le gain en années de vie, le principal avantage attendu du dépistage est de permettre des traitements moins agressifs. Toutefois, ni les données issues de essais contrôlés randomisés, ni les données factuelles recueillies en Belgique, ne permettent de confirmer cette attente.

6.2.2. Faux positifs

Dans notre modèle, les diagnostics ”faussement positifs” représentent une source importante de perte de qualité de vie. Un taux élevé de résultats faussement positifs (pouvant aller jusqu'à 10%) conjugué à un délai d'attente relativement élevé (45 jours en moyenne) pour les examens complémentaires peut amener à un résultat total du dépistage négatif en termes de QALY. Si on parvient à garder ce taux dans les normes européennes (3,5%) comme c'est le cas dans une région du pays (en Flandre), le gain en QALY est de 3 pour 1000 femmes.

6.2.3. Excès de diagnostics et de traitements

Le risque de diagnostic excédentaire est le risque majeur de ce dépistage pour les septuagénaires. Si nous appliquons un taux de surdiagnostic de 3 %, on peut s'attendre à ce que dans chaque cohorte de 100.000 femmes, 108 femmes supplémentaires auront un diagnostic de cancer et subiront très vraisemblablement un traitement. Si nous appliquons un taux de surdiagnostic de 10 %, ce nombre monte à 367.

D'autre part, toutes les femmes dont le cancer est diagnostiqué par screening deviennent malades du cancer deux ou trois ans plus tôt qu'en cas de diagnostic clinique. Ceci a un impact négatif sur la qualité des années de vie qui leur restent.

7. CONCLUSIONS

7.1. Faut-il prolonger le dépistage jusqu'à l'âge de 74 ans ?

La conclusion de cette étude est que la réponse à cette question est non. Cette affirmation est basée, d'une part, sur les résultats du modèle et d'autre part sur le contexte spécifique de cette question. Les résultats du modèle démontrent un gain de 13 années de vie pour 1000 femmes dépistées. Toutefois, certaines hypothèses qui sont loin d'être irréalistes, indiquent que le résultat net du prolongement du dépistage pourrait résulter en une perte globale en qualité de vie. Ces résultats ne sont donc pas décisifs en tant que tels et doivent être interprétés dans le contexte particulier d'un dépistage organisé. Le dépistage organisé s'adresse par définition à un individu qui n'exprime ni plainte ni demande. Cette spécificité implique d'être d'autant plus vigilant aux principes éthiques[11]. Les trois principes éthiques de base applicables notamment au dépistage sont: les principes de bienfaisance ou de non malfaisance, le principe de justice ou d'équité et le principe d'autonomie[12].

Les principes de bienfaisance ou de non malfaisance sont définis comme suit: “Ne pas faire de mal (primum non nocere) est le premier. Il doit se doubler d'un devoir de bienfaisance qui va de pair avec une attitude de bienveillance”. Le principe de justice ou d'équité est: “cette préoccupation qui fait intervenir la dimension collective des problèmes de santé, dans le sens d'une préférence pour les plus faibles, les plus démunis”[12].

Le dépistage est organisé dans le but d'améliorer le bien-être de la population en évitant notamment des décès prématurés. Cependant, les résultats obtenus par le modèle ne permettent pas d'exclure que dans certaines situations, le dépistage puisse affecter négativement la qualité de vie dans la tranche d'âge étudiée. Dans ces conditions, il y a risque de violation du principe de base ”primum non nocere”(ne pas faire de mal).

D'autre part, le dépistage est nettement moins efficace pour les femmes dont l'espérance de vie est la plus basse. Cette différence d'efficacité existe certes dans les autres tranches d'âge mais elle y est moins prononcée. Le respect du principe de justice ou d'équité se révèle donc être une raison supplémentaire de répondre par la négative à la question posée.

7.2. Que répondre à la personne qui demande un dépistage?

Le contexte de cette question diverge de celui de la question précédente en deux points: l'individu est demandeur et le problème doit être évalué sur un plan individuel. Le principe d'autonomie s'applique particulièrement bien à cette situation. Ce principe est défini comme suit: "le respect de la personne est le principe de base, le respect de l'autonomie de cette personne en découle ; Il s'agit de reconnaître la capacité de l'individu à faire des choix pour lui-même (autodétermination et libre choix) et à régir sa conduite (autogestion)"[12]. Pour que la personne puisse faire un libre choix, il importe qu'elle soit clairement et correctement informée des avantages et des inconvénients du dépistage dans sa situation personnelle. Le droit d'être informé (Article 7) et le droit au consentement éclairé sont décrits dans la loi belge relative aux droits des patients. Le consentement éclairé de la patiente ne peut être obtenu uniquement sur base de la lecture d'un document d'information. Il s'agit d'un processus qui devrait idéalement inclure un échange d'idées avec le praticien.

Il convient également que le médecin développe pour sa patiente qui demande le dépistage, une stratégie qui en minimise les inconvénients[13]. Ainsi, une attitude articulée en trois étapes peut être recommandée:

- Information spécifique à la tranche d'âge
- Prise de décision en fonction de l'appréciation personnelle de la patiente[14].
- Orientation de la personne qui le souhaite vers un dépistage dont les modalités minimisent les inconvénients.

Les critères définis dans le cadre du programme européen prévoient notamment la surveillance de la qualité technique des équipements utilisés, la double lecture des mammographies et l'optimisation du taux de rappel[1]. En Belgique, les unités de mammographie agréées répondant aux critères définis dans le cadre du programme européen, il est donc logique d'orienter les femmes qui demandent explicitement un dépistage vers ces structures.

7.3. Message clé

Le dépistage est organisé dans le but d'améliorer le bien-être de la population en évitant notamment des décès prématurés. Il est certain que prolonger le dépistage jusqu'à l'âge de 74 ans devrait permettre de gagner quelques années de vie. Toutefois, l'influence de cette mesure sur la qualité de vie est nettement plus aléatoire. Selon des hypothèses raisonnables, cette intervention pourrait même aboutir à une perte en terme de qualité de vie. Dans ces conditions, il se pourrait que la balance bénéfices-risques de ce dépistage penche du côté d'une perte globale de bien-être de la population.

8. RÉFÉRENCES

1. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition--summary document. Ann Oncol. 2008;19(4):614-22.
2. Schwartz LM, Woloshin S. News media coverage of screening mammography for women in their 40s and tamoxifen for primary prevention of breast cancer. JAMA. 2002;287(23):3136-42.
3. Schwartz LM, Woloshin S, Fowler FJ, Jr., Welch HG. Enthusiasm for cancer screening in the United States. JAMA. 2004;291(1):71-8.
4. Silverman E, Woloshin S, Schwartz LM, Byram SJ, Welch HG, Fischhoff B. Women's views on breast cancer risk and screening mammography: a qualitative interview study. Med Decis Making. 2001;21(3):231-40.
5. Schwartz LM, Woloshin S, Sox HC, Fischhoff B, Welch HG. US women's attitudes to false positive mammography results and detection of ductal carcinoma in situ: cross sectional survey. BMJ. 2000;320(7250):1635-40.
6. Belgian Cancer Registry, editor. Cancer incidence in Belgium, 2004-2005. Brussels; 2008.
7. Paulus D, Mambourg F, Bonneux L. [Breast cancer screening]. Good Clinical Practice (GCP). Brussels: Belgian Health Care Knowledge Centre (KCE); 2005 02/05/2005. KCE Reports 11 Available from: http://kce.fgov.be/index_en.aspx?SGREF=5221&CREF=9348
8. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms.[Erratum appears in Ann Intern Med. 2010 Jan 19;152(2):136]. Ann Intern Med. 2009;151(10):738-47.
9. Woloshin S, Schwartz LM, Byram SJ, Sox HC, Fischhoff B, Welch HG. Women's understanding of the mammography screening debate. Arch Intern Med. 2000;160(10):1434-40.
10. Mandelblatt JS, Silliman R. Hanging in the balance: making decisions about the benefits and harms of breast cancer screening among the oldest old without a safety net of scientific evidence. J Clin Oncol. 2009;27(4):487-90.
11. Doumont D, Verstraeten K. Enjeux éthiques du dépistage organisé. Santé en Communauté Française. 2012(7):3-7.
12. Gallois. Dépister les cancers, mais à quelle condition In: UNAFORMEC, editor. Médecine. Paris; 2005. p. 72-7.
13. USPSTF. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement Annals of Internal Medicine 2011(151):716-26.
14. Woloshin S, Schwartz LM. The benefits and harms of mammography screening: understanding the trade-offs. JAMA. 2010;303(2):164-5.

SCIENTIFIC REPORT

1. INTRODUCTION

1.1. Context of this report

This report is a partial update of the clinical practice guideline (CPG) on breast cancer screening published in 2005[1]. Therefore, the KCE experts made a list of clinical questions related to breast cancer screening. Representatives of stakeholders' organizations were then invited to review the choice and the wording of the questions, to highlight the main problems related to each question and to score the relevance of clinical questions(see KCE report 172)[2]. Selected questions were then divided over three KCE reports. A first KCE report published in 2010 is focused on breast cancer screening with mammography for women in the age group of 40-49 years (KCE report 129)[3]. The second is focused on identification of women at risk for breast cancer and technical methods for breast cancer screening (KCE report 172)[2].

1.2. Scope of this report

This report focuses on the extension of organized breast cancer screening with mammography to older women. Eligible population is defined as women between 70-74 years of age with average risk of breast cancer.

1.3. Breast cancer screening in Belgium

The Belgian federal and regional governments signed a protocol agreement in 2001 for an organized screening programme for women aged 50-69 years, to be organized by the regional governments with appropriate financial resources supplied by the federal government. Since 2001, Flanders, the Walloon region and the Brussels capital region have each introduced an organized screening programme within their specific context of already existing practices. Indeed, opportunistic screening remains quite frequent in the Walloon and Brussels region among women in the age-group 50-69, but also among younger (40-49 years of age) or older women (>70 years). In Flanders, screening mammographies are dominant in the age-group 50-69. In the age-group 70-79 overall coverage drops, mainly because organized screening stops at age 69. The coverage by means of diagnostic mammography decreases also with 3%, indicating that substitution of screening mammography by opportunistic screening at

the age of 70 is not frequent in Flanders. At this age, total coverage (including both diagnostic or follow up mammographies and opportunistic screening) remains at 18% in Flanders, 33% in Brussels and 30% in the Walloon region (KCE report 172)[2].

1.4. Clinical questions

This specific report addresses the following questions:

1. What are clinical benefits of an extension of breast cancer organized screening in women between 70 and 74 years?
   - 1.1. What is the effect of an extension (70-74 years) of breast cancer organized screening on the breast cancer related mortality?
   - 1.2. How long is the delay between the screening and the associated breast cancer related mortality reduction?
   - 1.3. What is the effect of an extension (70-74 years) of breast cancer organized screening on the overal mortality?
   - 1.4. What is the effect of an extension (70-74 years) of breast cancer organized screening on morbidity?
2. What are the specific harms of an extension of breast cancer organized screening in women between 70 and 74 years?Harms in terms of false positive or false negative results?
   - 2.2. Harms in terms of additional diagnostic tests?
   - 2.3. Harms in terms of over-diagnosis?
   - 2.4. Harms in terms of overtreatment?
3. What attitude should be recommended for women in case of self referral?

1.5. Scientific approach

For each clinical question, a systematic search of the literature was performed and discussed with the support of external experts chosen for their scientific competency in several fields: gynaecology, radiology, epidemiology, or health economics. For question 3, we searched for models. To quantify what the implications of our findings are on the Belgian situation we applied data from the Intermutualistic Agency (IMA/AIM), cancer registry and data from the literature on the Belgian life tables and constructed a simple time dependent Markov chain with annual cycles.

The methodology used and the results are described in each chapter.

2. LITERATURE REVIEWS

2.1. Review of clinical studies

2.1.1. Methodology

2.1.1.1. Sources

A broad search of the electronic databases OVID Medline, EMBASE, CDSR and DARE was conducted in April 2011. Search was conducted first for systematic reviews (SR) and meta-analysis (M-A).

2.1.1.2. Search terms

For searching on Medline database, the following MeSH terms were used in combination with usual language: Breast neoplasms (MESH) and mass screening (or early detection) (MESH) and mammography (MESH). For EMBASE, the following Emtree terms were used: 'cancer screening', 'breast cancer' and 'mammography'. These MESH and Emtree terms were combined with a standard search strategy to identify systematic reviews (SR) or meta-analysis (M-A).

2.1.1.3. In- and exclusion criteria

Databases were searched for SR and M-A in English, French, Dutch or German. This report is a update of previous KCE report[1] (search made in 2004), thus we used a date restriction (2004-2011) and a language restriction (English, Dutch, French and German). Inclusion criteria used for selection based on title, abstract or full text were: population (women without breast cancer and without particular breast cancer risk), intervention (mammography), outcome (mortality, morbidity, additional diagnosis tests, over diagnosis and over treatment), design (SR or meta-analysis or RCT), key question (screening), age of population (>70 and <75 years), and original publication. Relevant publications were selected independently by 2 reviewers (FM, JR).

2.1.1.4. Additional evidence

We identified two SR[4, 5] as the more extensive source for the research question 2. Therefore the evidence-identified through those SR-was updated by searching Medline and the Cochrane Database of Systematic Reviews from the search date of the two SR's on (search date Nov-Dec 2008). Additional hand searching of reference lists was also undertaken to ensure that no potentially relevant studies were missed. We also scanned reference lists of SR and of our previous report on breast cancer screening[3].

The identified studies were selected based on title and abstract. For all eligible studies, the full-text was retrieved. In case no full-text was available, the study was not taken into account.

The description and results of the literature searches and flow of studies search are in Appendix 1.1.

2.1.1.5. Quality appraisal

The methodological quality of systematic reviews and associated risk of bias were rated using the checklists of the Dutch Cochrane Centre (www.cochrane.nl). The assessment of the risk of bias in the included SR was conducted by a team of two reviewers (FM, JR).

The methodological quality of selected additional evidence was also rated using the adequate checklists of the Dutch Cochrane Centre (www.cochrane.nl).

The results of the quality appraisal are in Appendix 1.5.

2.1.1.6. Identified systematic reviews

In the systematic search for literature reviews, 53 citations on the topic were identified in database searches. The majority of citations were excluded on the basis of title and abstract; 10 citations were retrieved in full and reviewed in more detail. On the basis of the full text, 5 reviews were included[4-8].

The reviews written by Götzsche and Nelson[4, 5] are mainly focused on mortality as outcome, those from Biesheuvel and Jorgensen[6, 7] on over-diagnosis and the review of Virnig[8] on ductal carcinoma in situ (DCIS).

As a first step, a quality appraisal of all the reviews was carried out to determine their suitability for inclusion. Götzsche and Nelson SR[4, 5] were judged to be of high quality with a low risk of bias. Nelson review[5] was an update of one other review performed by Humphrey for the U.S. Preventive Task Force[9]. Humphrey review was also judged to be of high quality and used here as complementary information source.

The review written by Biesheuvel[6] was judged to be of good quality (quality appraisal of selected trials not sufficiently described) and those written by Jorgensen[7] was judged to be of high quality. The review written by Virnig was also judged to be of high quality.

2.1.1.7. Identified RCT

The evidence was updated using the key words reported in Nelson SR[5] by searching Medline and the Cochrane Database of Systematic Reviews from the search date of this SR on (search date Nov 2008). The literature search for relevant RCTs carried out in Medline, EMBASE and CCRT (in April 2011) identified 432 citations. The majority of citations were excluded on the basis of title and abstract; the other papers (n=8) were retrieved in full and reviewed in more detail. On the basis of the full text, all eight studies were excluded because of the study design (not an RCT) shows the flow of randomized controlled trials from selection to in-or exclusion.

By hand searching of reference lists of Götzsche and Nelson[4, 5], the Swedish RCT's were identified. Among those, the Two County trials is the only RCT that includes women aged 70-74 years at the time of randomization. Quality appraisal of this RCT was carried out to determine their suitability for inclusion. The Two County trials was judged to be of fair quality by Nelson and of low quality by Götzsche and included for further analysis[10].

2.1.1.8. Identified additional evidence

For diagnostic errors and over-diagnosis, this update was carried out in July 2011 identifying 10 citations. The majority of citations were excluded on the basis of title and abstract; 2 papers on diagnostic errors were retrieved in full and reviewed in more detail. On the basis of the full text, those two studies were excluded. Most papers are discussions and comments on the two main SR[4, 5].

For DCIS, this update was carried out in July 2011 identifying 7 citations. All citations were excluded on the basis of title and abstract.

For overtreatment, this update was carried out in July 2011 identifying 19 citations on Medline and 7 citations on the Cochrane Library. The majority of citations were excluded on the basis of title and abstract; 2 papers on overtreatment were retrieved in full and reviewed in more detail. On the basis of the full text, we retrieved again the SR written by Götzsche[4] and selected one publication presenting data issued from the UK Breast Screening Programme[11]. The description and results of those updates are in Appendix 1.4.

2.1.1.9. Ongoing clinical trials

In addition to the database searches, the ClinicalTrials.gov website was searched for clinical trials. The search terms 'breast neoplasm' as well as 'screening' and 'mammography' were used to search for studies. The majority of search results (n=135) were ongoing trials. Two potentially relevant trials (NCT00963911, NCT00247442) were identified but were considered as out of scope after receiving more information on the full protocol.

2.1.1.10. Data extraction

Data from systematic reviews and from trials were extracted into a data extraction table (DET) summarizing key design features and results. All data extraction table are in Appendix 1.6.

2.1.2. Description of screening benefit

2.1.2.1. Sources

In the years 1960-1980, USA, Sweden, Canada and United Kingdom conducted randomized controlled trials of mammography screening. In US, the HIP trial (N = 60 995) started in 1963. In Sweden, the Malmö trial (phase I and II, N = 60 076) started in 1976 and 1978, the Two county Trial (Kopparberg and Ostergötland, N= 133 065) in 1977-78, the Stockholm trial (N= 60 117) in 1980 and finally the Göteborg trial (N= 51 611) in 1981. In Canada, the National Breast Screening Trials (NBSS-1 and 2, N = 89 835) were initiated in 1980. In United Kingdom, the Edinburgh, trial started in 1979 in 1980 (N=44 268) and the UK Age Trial in 1991 (trial limited to

women aged 40-49 years)[12]. Numerous publications and some SR summarizing theirs results are now available.

This part is based on the SR (2002) commissioned to assist the US Preventive Services Task Force (USPTSF) and its update of 2009[5,9] and on the Cochrane SR [4]. We analysed more in detail one RCT named the Swedish Two-County trial (Ostergötland) which was included by both SR[10,13,14].

Both reviews included the same trials in their meta-analysis: the HIP trial, Malmö I and II, the Two county trial, the NBSS trials (1 and 2), the Stockholm trial, the Göteborg trial and the UK Age Trial. The Edinburgh study was rated as poor quality by both authors and excluded therefore[4, 9].

Nelson updated the meta-analysis from Humphrey[9] to include new findings about younger women (40-49 years of age). Therefore, we refer to the Humphrey publication for mortality analysis performed on women aged from 50 to 74 years. Götzsche[4] performed first a meta-analysis among women 39 to 74 years of age. Then he did a separate analysis for women younger than 50 years and for women older than 50 years.

Two-County trial

We analysed the Swedish Two-County trial in order to find more information on our specific population (70-74 years). The Swedish Two-County trial is the largest of the first eight randomized trials. We used therefore three publications that describe this study. We used the first publication of the initiator[10], the publication of Nyström[13] who was selected by Nelson and the last publication of Tabar published in July 2011[14]. The Two-County trial was commissioned by the Swedish National Board of Health and Welfare and included women in two Swedish counties: Kopparberg and Östergötland. In 1977-78, 134 867 women aged 40 to 74 years were cluster-randomized by geographic area. They were also stratified by socioeconomic status, urban or rural residency, and size of cluster. Finally, 78 085 women were invited to the screening. Among those, they were 10 568 women aged 70-74 years in the screening group and 7 462 in the control group. At this age, women were invited to two screening rounds with a screening interval of 33 months. The trial was closed in 1984 after approximately 7 years of screening[10].

In 2002, Nyström performed one review of the Swedish RCT's including the Malmö, Ostergötland, Stockholm, Göteborg trials. Results of the Kopparberg trial were not available at this time. This publication assessed the age-dependency of the effect of screening. The author calculated mortality relative risks for consecutive 5-years age group based on results from the Ostergötland trial. The median follow-up time was 17.9 years. Unfortunately, without the Kopparberg part of the Swedish trial, the number of women 70 to 74 years of aged enrolled was low (approximately 5000 in each group)[13].

Finally, we found one publication summarizing long term data (29 years) on mammographic screening effect on mortality[14].

Trials quality and bias

All studies included by Humphrey in 2002 and later by Nelson were rated as fair[5, 9]. Götzsche assessed the randomization quality. This author divided his results on results based on adequately randomized control trial and results based on suboptimally randomized control trial[4].

Nevertheless, the third meta-analyses were judged of high quality with a low risk of bias (see Appendix 1.5.1).

Some publications based on The Swedish Two-County reported varying numbers of women enrolled. To explain this variation, Nyström replied that some studies analysed results by year-of-birth while some others used exact age at randomization[13]. Nevertheless, Götzsche assessed this trial as suboptimally randomized and likely to be biased. He argued that for Ostergötland, a public notary allocated the clusters by tossing a coin while witnesses were present. Breast cancer mortality in the control group was almost twice as high in Kopparberg compared to Ostergötland (0.0021 versus 0.0012, p = 0.02). The autopsy rate was 36% for all the Two-County trial and cause-of-death assessments were not blinded[4]. According to that the validity of local end point committee data was criticized, a third committee (named consensus committee) reviewed the records containing a doubtful cause of death[14].

2.1.2.2. Breast cancer related mortality reduction

For women aged 39 to 74 years and at approximately 13 years of follow-up, the Humphrey meta-analysis (M-A)[9] and the Cochrane review[4] showed a significant reduction in breast cancer mortality of 16% (Relative Risk (RR) 0.84, 95% confidence interval (CI) 0.77 to 0.91) and 19% (Relative Risk (RR) 0.81, 95% (CI) (0.74, 0.87) respectively.

For women aged 39 to 74 years, the Review of Swedish randomized control trial showed a significant reduction in breast cancer mortality at 15.8 years (median follow up) of 21% (Relative Risk (RR) 0.79, 95% (CI) 0.70 to 0.89). This study showed that the effect of breast cancer screening in terms of breast cancer mortality reduction varies according to age range[13].

For women aged at least 50 y at randomization, three trials with adequate randomization did not show a significant reduction in breast cancer mortality at 13 years (Relative Risk (RR) 0.94, 95% (CI) 0.77 to 1.15). Four trials with suboptimal randomization showed a significant reduction in breast cancer mortality (RR of 0.77 (95% CI 0.67 to 0.83)). The RR for all seven trials combined was 0.77 (95% CI 0.69 to 0.86)[4].

The review of Swedish randomized control trial, applying a more conservative determination of cause of death for women aged at least 70 y at randomization, did not show a significant reduction in breast cancer mortality at 17.4 years ((RR) 1.12, 95% (CI) 0.73 to 1.72)). Unfortunately, this age group was relatively small (approximately 5000 women in each group) and this study is underpowered[13]. Consequently, we must conclude together with Nelson that data are insufficient for this age group[5].

2.1.2.3. Delay between screening and specific mortality reduction

Tabar published in July 2011 the last follow-up result (29-year) of the Swedish Two-County Trial[14]. This publication modulated breast cancer mortality reduction in function of length of follow up. In this report both data issued from local end point committees and consensus-based data were presented. The validity of local end point committee data was criticized, we present here consensus data. For women aged 39 to 74 years, this publication showed specific mortality reductions of 20% ((RR) 0.80, 95% (CI) 0.62 to 1.05), 27% ((RR) 0.73, 95% (CI) 0.59 to 0.92), and 27% at respectively 10, 15 and 20 to 29 years of follow up. In the same time, deaths from breast cancer prevented in the study group increased along length of follow-up. They were respectively 50, 99, 114, 122 and 126 deaths prevented at 10, 15, 20, 25 and 29 years of follow-up for all women included in this study. Author emphasized that breast cancer screening prevents deaths more in the medium to long term than in the immediate future. So most of the breast cancer deaths would have occurred (in the absence of screening) more than 10 years after randomization.

Authors did not calculate mortality relative risks for each age group separately. Results presented are based on 133 065 women aged 40-74 (77 080 in the screening group and 55 985 in the control group), while they were 10 568 women aged 70-74 years in the screening group and 7 462 in the control group. In Kopparberg, cancers diagnosed after the two screening rounds in women aged 70-74 years and breast cancer deaths from these cases were still included in the results[14].

As cited on previous point, the group of women aged 70-74 years included in the Swedish Two-County Trial was relatively small (approximately 5000 women in each group) and this study is underpowered[13].

2.1.2.4. All-cause mortality

The Cochrane SR has reported data on all-cause mortality. For women aged at least 50 y at randomization, two trials with adequate randomization (n=73654) did not show a significant reduction in all-cause mortality at 13 years (Relative Risk (RR) 1.00, 95% (CI) 0.95 to 1.04). The two trials with suboptimal randomization (n=98261) also did not show a significant reduction in all-cause breast cancer mortality (RR of 0.99 (95% CI 0.97 to 1.02))[4].

Unfortunately studies did not have statistical power to detect an all-cause mortality reduction. According to that disease specific mortality is a small fraction of all-cause mortality in cancer screening trials, detect a mortality reduction would require inclusion of millions of subject.

2.1.2.5. Morbidity reduction

We found no data related to the cancer related morbidity in our selected sources. In other words, we do not accept or reject the hypothesis that screening reduces the morbidity of the breast cancer disease.

2.1.3. Description of screening harms

2.1.3.1. Sources

This part is based on the 5 SR selected in our main search[4-8]. As explained in part 2.3.5, we updated those in July 2011 starting from the last literature search date. See more details in appendix 1.4.

2.1.3.2. Study description

SR written by Götzsche and Nelson are described in point 3.1.1. The reviews written by Biesheuvel and Jorgensen[6, 7] were focused on over-diagnosis and subsequently on overtreatment. Each author used very different methods to address this issue. Biesheuvel analysed reports issued from the first RCTs while Jorgensen analysed data issued from publicly organized screening programmes. The review written by Virnig was focused on ductal carcinoma in situ (DCIS)[8].

2.1.3.3. Performance of mammography

The sensitivity of first mammography for women aged 70-74 years was 81% in the Two County trial. This includes over-diagnosis and may be difficult to interpret. This data cannot be applied to individual patients because they are not adjusted for patient factors (use of hormone replacement therapy, mammographic breast density), technical factors (quality of mammography, number of mammographic views) or provider factors (the experience of radiologists and their propensity to label the results of an examination abnormal)[9]. Provider factors may explain that sensibility may vary between countries[4]. In the Two County trial, the specificity of a single mammographic examination was 95.6% for women aged 40-74 years. This indicates that 4% of women who did not have cancer underwent further diagnostic evaluation. The positive predictive value of one-time mammography was 12% for abnormal results requiring further evaluation and from 50% to 75% for abnormal results requiring biopsy. Positive predictive value increases with age and ranged from 18% to 20% among women 70 years of age or older[9].

Nelson reported data from the Breast Cancer Surveillance Consortium (USA) BCSC for regularly screened women that are based on results from a single screening round. False-positive mammography results are less common among women aged 70-79 years (68.8 per 1000 women per screening round). Conversely, false-negative mammography results are a little more common among women aged 70 to 79 years (1.5 per 1000 women per screening round)[5].

2.1.3.4. Additional diagnostic tests

Rates of additional imaging are relatively low among women aged 70 to 79 years (64.03 per 1000 women per screening round). Biopsy rates are higher among women aged 70 to 79 years (12.2 per 1000 women per screening round) than among younger women. As expected, the number of screen detected cancer is highest in this age group. Results indicate 6.5 screen-detected invasive cancer and 1.4 screen-detected DCIS per 1000 women per screening round. The BCSC results indicate that for every case of invasive breast cancer detected by mammography screening in women aged 70 to 79 years, 154 women have additional mammography, 10 have other imaging test, and 2 have biopsies[5].

2.1.3.5. Over-diagnosis

Over-diagnosis of breast cancer at screening may be defined as the detection with screening of cancer that would not have presented clinically during the woman's lifetime (and therefore would not be diagnosed in the absence of screening)[6].

Nelson reported rates of over-diagnosis varying from less than 1% to 30% with most from 1% to 10%. She explained variations by inclusion or exclusion of DCIS cases, by whether cases are incident or prevalent, and by age. She concluded that the studies are too heterogeneous to combine statistically[5].

Götzsche reported that the level of over-diagnosis was about 30% in the RCT's that did not introduce early screening in the control group, and somewhat larger in the sub optimally randomized trials before screening of the control group. He found also a 40% to 60% increase in incidence of breast cancer in observational studies performed in Australia, Europe and USA after beginning of the screening[4].

Biesheuvel analysed publications issued from the first RCTs (New York/HIP, Malm II, Two County, Canada a and b, Stockholm, Göteborg, Edinburgh) and from four population-based programme (Sweden, Norway, Netherlands and Italy). He selected papers that attempted to estimate over-detection of invasive breast cancer by mammography screening.

Note that he did not include DCIS. He excluded potentially biased publications. Bias were described as: different breast cancer risk in screened and unscreened population, low participation in screening group and high participation in non-screening group, offering screening to the control group before or during follow up, inappropriate adjustment for lead time. After exclusion, he selected 22 estimates of over-detection from several (some overlapping) sources. Publications were categorized as being based on cumulative-incidence or incidence-rate methods (definitions of terms are in appendix 1.4.3). Excluding biased studies as described before, he selected the least biased over-detection estimates. Excluding DCIS cases, over-detection ranged from 7% to 21% for women aged 60–69 years[6].

Jorgensen analysed data issued from publicly organized screening programmes. He selected papers that published trends in incidence of breast cancer before and after the introduction of mammography screening. Note that when data were present, DCIS were included. If not, he estimated that they would contribute to 10% of the diagnoses in a screened population. After exclusion of the implementation phase of the screening, he compared data covering at least seven years before screening with data covering at least seven years after screening in screened and non screened age groups. The most common age-range for mammography screening programmes was 50-69 years. No data specific for women aged 70 to 79 years are available. The increase in incidence of breast cancer was closely related to the introduction of screening. Surprisingly, little of this increase was compensated for by a drop in incidence of breast cancer in women older than 70 years. Jorgensen calculated that over-diagnosis for invasive cancer was 35%. The rate of over-diagnosis including DCIS cases was 52% in this meta-analysis (95% CI 46% to 58%)[7].

Discrepancies between results reported by Biesheuvel or by Jorgensen have led to a lot of controversial discussions.

The approach Biesheuvel et al. to adjusting for lead time was contested by Zahl, Jorgensen and Götzche (2008), who stated that their estimations were substantially downwardly biased, due to over-adjustment, use of a hypothetical increase in incidence based on theoretical models and use of long term follow up data that are considerably diluted. They also considered estimation unhelpfully wide.

Jorgensen & Götzsche used linear regression to compare observed incidence with a in an (hypothetical) population that did not undergo screening. They assume a linear increase extrapolated from prescreening trends, following the same pattern as the linear trend observed in women too young to be screened. It is difficult to judge if this assumption holds or not, the graphs the authors present show non-linear increases in incidences before screening was introduced in the UK and Norway for whom no explanation was given.

2.1.3.6. DCIS

Historically, DCIS was rare and diagnosed by surgical removal of a suspicious breast mass. Since the wide use of mammography, a increasing numbers of patients were diagnosed with DCIS. The prognosis of the disease is excellent. Maass reported data issued from the SEER database (Surveillance, Epidemiology and End Results database of the United States National Cancer Institute). Those data showed a 10-year survival rate of 96.6% for cases between 1978 and 1983, when no screening was performed. The rate was 98.1% between 1984 and 1989, when screening was performed[3, 15].

Recent changes in DCIS incidence in USA were emphasized by Virnig. This author performed a SR on incidence, treatment and outcomes of DCIS in name of Agency for Healthcare Research and Quality (AHRQ). She included 63 publications addressing incidence for analysis. She compared data obtained before the screening (1973-1975) with current century data collected in US where screening is common. DCIS incidence rose there from 1.87 per 100 000 in 1973–1975 to 32.5 per 100 000 in 2004. Incidence increased most in women older than 50 years. Increased use of mammography may explain some but not all of this increased incidence[8].

2.1.3.7. Overtreatment

Götzsche reported that the number of mastectomies and lumpectomies was significantly larger in the screened groups. Three trials with adequate randomization showed a significant increase in mastectomies and lumpectomies (Relative Risk (RR) 1.31, 95% (CI) 1.22 to 1.42). Two trials with suboptimal randomization showed the same increase in interventions (RR of 1.42 (95% CI 1.26 to 1.61)). The RR for all five trials combined was 1.35 (95% CI 1.26 to 1.44)[4].

Based on recent data from the UK Breast Screening Programme, Dixon emphasized the increasing numbers of patients with DCIS. In 1998/99 there were approximately 1500 cases, but in 2007/08 there were close to 3500 cases. Although, most DCIS cases may be treat by breast-conserving surgery, the percentage of patients being treated with this method has remained constant at 30% during this period. Because of the increasing incidence of DCIS treatments, the absolute numbers of women having mastectomies has increased from just under 500 in 1998/99 to over 900 in 2007/08[11].

2.1.4. Screening conditions

The sojourn time (ST) is the average duration of the preclinical screen-detectable phase. Estimation of sojourn time can be performed by from simple mathematical estimates or using microsimulation techniques (mainly Markov Models)[12]. Sojourn time provides an absolute upper limit to the lead time obtainable. If the sojourn time is long, the maximum attainable lead time is corresponding long[16]. A longer sojourn time results in higher number of additional breast cancer detected, more life-years gained and higher number of years with cancer due to lead-time[17].

2.1.4.1. Literature search

In a first stage, studies assessing sojourn time were searched. Ovid Medline was consulted from 1948 to October Week 1 2011. The main search terms (MESH) were: Breast Neoplasms/ Mass Screening/ or Mammography/. Sojourn time was included in free text. The search was limited to papers written in English, Dutch, French, or German. Reference lists of the selected studies were checked for additional relevant citations. See more details in appendix 1.4.4.

Selection criteria

All retrieved references were assessed against pre-defined inclusion criteria (in terms of population, intervention, outcomes, and design-Table 1) in a two-step procedure: initial assessment of the title, abstract and keywords; followed by full-text assessment of the selected references. Estimation of sojourn time not based on data were excluded. After excluding of 3 duplicates, 40 unique citations were identified from the databases. Of this total of 40 references, 23 did not meet the inclusion criteria based on title and abstract evaluation. Among the 17 citations retained for full-text assessment, 6 did not fulfill the population criteria[18-23] and 1 did not fulfill the outcome criteria[24]. Finally, 10 studies were retained[16, 17, 25-32]

Our search was based on ST duration estimations as search. We found several publications were ST estimations were issued from others studies cited as references by the author. For example, Zappa in 2003[32]referred to data published by Tabar in 1995[29]. Duffy in 2005[26] referred also to those data[29]. Therefore, we used original publications. If one author published two or more articles based on the same data, we choose the most accurate for our study[30, 31]. Finally, 7 publications are summarized in data extraction table (see appendix 1.6.7).

2.1.4.2. Results

Sojourn times calculated on RCT's data

We found 4 studies based on the results of the Two- County Trial. The Two- County Trial is described in chapter 2 (point 2.2.1.1.)[10, 13]. First estimates of sojourn time published by Tabar and Duffy were based on approximately the same data. Both authors used the same Markov chain model, but results were not the same. Shen underlined that the difference in estimates published by the two authors[25, 29] may be caused by different statistical methods or by discrepancy in the data. Shen applied his recently developed statistical methods based on the maximum likelihood estimates to data from the Two County Trial. Authors estimated the sensitivities of early detection modalities as 0.92 (SD, 0.09) and the mean ST as 4.4 years (SD, 0.76)[27].

Sojourn times calculated on screening programmes data

Spratt estimated the duration of breast cancer before detection by dividing prevalence rates at first screening round by incidence rates in the following years. Therefore, he used data from 10 000 women aged 35 to 70y at start included in the Breast Cancer Detection and Demonstration Project (Louisville). For women aged 70-74, he estimated that sojourn time ranged between 2.5 y to 3.8 y[28].

Fracheboud compared the results of the Dutch breast cancer screening programme for women aged 70-75 with the hypothesis developed by Boer in 1995. Boer had described optimistic and pessimistic assumptions for use in his MISCAN model. Optimistic assumption assumed no further increase in preclinical duration of breast cancer after 65years of age although pessimistic assumption assumed a further increase in preclinical duration with age[33]. Based on 187 207 screening examinations (women aged 70-74 years), Fracheboud found that detection rates in both initial and subsequent screens increased steadily with age and got close to assumption which assume a continuously increasing sojourn time beyond the age of 69. This increasing sojourn time of breast tumours lead to a strong increase in detection of cancers, but also to more life- years in lead time[17].

Weedon constructed one inventive solution for screening programme who do not have full registration of interval cancers or where opportunistic screening is common. Although Norwegian registration is of very high quality, incidence data from the first screening round, interval between screening examination or registration of interval cancer may be insufficient. Therefore, he replaced data lacking by data issued from questionnaire send to 336 533 women in the Norwegian Breast Cancer Screening Programme (NBCSP). This new approach gave estimation of MST to 6.9 years for women aged 60-69 years, although STS was estimated to 60%[30].

2.1.4.3. Discussion

Most estimates of sojourn time have been based on Models (mainly Markov chain models). Such models assume a chronological stepwise growth of cancer. Unfortunately, it remains unknown whether cancer really develop according to a chronological sequence. Estimations of sojourn time must consequently be interpreted with caution.

2.1.5. Key data

Data issued from literature search are summarized in table 2.1.

Table 2.1: Data issued from clinical literature review

Question 1: Should breast cancer organized screening extended in women between 70 and 74 years?	
Population	Women between 70-74 years of age without breast cancer and without particular risk of breast cancer.
Intervention	Organized screening with mammography
Comparison	No organized screening
Outcomes:	
Mortality (specific)	For women >50 y at randomization, the specific mortality reduction after a follow-up of 13 years is 23% (RR: 0.77, (CI) 0.69 to 0.86). In the Two County trial, specific mortality reduction reach at significant reduction of 27% (RR: 0.73, (CI) 0.59 to 0.92) at 15 years of follow up and increases afterwards.
Mortality (all cause)	Studies did not have statistical power to detect an all-cause mortality reduction.
FP	68.8 per 1000 women aged 70 to 79 years per screening round (BCSC-USA)
FN	1.5 per 1000 women aged 70 to 79 years per screening round (BCSC-USA)
Additional imaging	64.03 per 1000 women aged 70 to 79 years per screening round (BCSC-USA)
Biopsy	12.2 per 1000 women aged 70 to 79 years per screening round (BCSC-USA)

DCIS	1.4 screen-detected DCIS per 1000 women aged 70 to 79 years per screening round (BCSC-USA)
Over-diagnosis	Over-detection (excluding DCIS cases), ranged from (7% to 21%) to 35% (no data specific for women aged 70 to 79 years are available).
Over-treatment	The number of mastectomies and lumpectomies was significantly larger in the screened groups (RR:1.35 (95% CI 1.26 to 1.44).

2.1.6. Conclusion

At this age group, performance of mammography is high and rates of additional imaging are relatively low. Breast cancer screening achieves a specific mortality reduction of 23% to 27% according to authors. This mortality reduction did not appear in the first years after screening. The specific mortality reduction is not statistically significant before 10 years after screening ((RR) 0.80, (CI) 0.62 to 1.05). Breast cancer mortality reduction must be put in perspective with life-expectancy for this age-group in our country.

On the other hand, aspects related to quality of life raises questions pertinent to discussion of the benefit and harms of breast cancer screening in this age-group. First, over-diagnosis being an inevitable consequence of cancer screening, the risk of overtreatment persists. Secondly, the lead time bias although difficult to estimate, may be crucial for older women. Screening diagnosed breast cancer and consecutive treatment may mean the end of “the life in good health condition” some years earlier than clinical diagnosed breast cancer[34].

2.2. Review of modeling studies

2.2.1. Literature search strategy

In a first stage, randomized clinical trials analysing the impact of screening on morbidity and mortality were searched (see above). Then, because the effectiveness of screening require a lot of information from a wide range of sources to correctly inform decision makers, modeling studies were searched[35].

Medline, Embase, NHS EED and Econlit databases were consulted from January 2000 up to September 2011 (see appendix 2.1). The search was limited to papers written in English, Dutch, French, Spanish, or German. Reference lists of the selected studies were checked for additional relevant citations.

The keywords used and the results are detailed in appendix 2.1. The main search terms (MESH) were:

- Breast Neoplasms; and
- Mass Screening or Early Detection of Cancer ; and
- Mammography; and

2.2.2. Selection criteria

All retrieved references were assessed against pre-defined selection criteria (in terms of population, intervention, outcomes, and design - Table 2.2.) in a two-step procedure: initial assessment of the title, abstract and keywords; followed by full-text assessment of the selected references. It should be noted that studies assessing screening techniques (such as digital mammography) were excluded because such topic was investigated in KCE report 172[2].

Table 2.2: Selection criteria

	Inclusion criteria	Exclusion criteria
Population	caucasian women without breast cancer and without particular risk	Other (e.g. woman at risk, Asian women, etc.)
Intervention	Screening mammography	Other, including mammography techniques (e.g. digital mammography)
Outcomes	Morbidity and Mortality (e.g. LYG and QALYs)	Other outcomes (e.g. over diagnosis)
Design	Modeling studies	Other designs

LYG: life-year gained; QALY: Quality-adjusted life-year gained

2.2.3. Quantity of research available

After excluding 195 duplicates, 1058 unique citations were identified from the databases. Hand searching did not allow us to identify additional citations. Of this total of 1058 references, 1016 did not meet the inclusion criteria based on title and abstract evaluation. Among the 42 citations retained for full-text assessment, 2 did not fulfill the population criteria and 15 did not fulfill the design criteria. Finally, 25 modeling publications were retained, concerning 6 models developed by modeling groups involved in CISNET, 2 applications of these models on different context and 7 models developed by other groups or authors, as some models have several publications[17, 36-59]. The flow chart of this selection is presented in the appendix 2.2.

Table 2.3: Modeling studies excluded after full-text assessment

Exclusion criteria	Studies
Population	Messecar 2000; Wen 2005[60, 61].
Intervention	
Outcome	
Design	Advisory Committee on Breast Cancer 2006; Anonymous 2000; Barratt 2002a; Barratt 2002b; Bonneux 2009; Caplan 2001; Carney 2007; De Koning 2000; Feuer 2004; Grivegnee 2001; Habbema 2006; Mandelblatt 2003; Prevost 2000; Rautenstrauch 2000, Xu 2000[62-76].

2.2.4. Selected studies

2.2.4.1. The CISNET Project

The Cancer Intervention and Surveillance Modeling Network (CISNET) (http://cisnet.cancer.gov) is a consortium of National Cancer Institute (NCI)-sponsored investigators whose focus is modeling the impact of cancer control interventions on population trends in incidence and mortality for breast cancer. These models are also used to project future trends and to help determine optimal cancer control strategies[40]. Seven groups developed their own breast cancer models spanning a wide range of modeling philosophies: The University of Texas M. D. Anderson Cancer Center model[37], University of Wisconsin model, Georgetown[47], Erasmus (MISCAN) model[57], Dana-Farber model[44] University of Rochester model[43] and Stanford model[51]).

The seven models were first used to assess the relative and absolute contributions of screening mammography and adjuvant treatment to the reduction in breast-cancer mortality in the United States from 1975 to 2000[77]. Mandelblatt et all[48] used 6 of those CISNET models to provide estimates of potential benefits and harms of mammography screening under different screening schedules. One of the 7 models, the University of

Texas M. D. Anderson Cancer Center model[37] was not used as it was purely descriptive.

The models were developed by different groups but not independently, they were compared, discussed and adapted during the development process, they also used a common set of variables and inputs, based on US datasets BCSC (Breast Cancer Surveillance Consortium), SEER 9 (Surveillance, epidemiology and end results), Connecticut Tumor registry and the Berkeley mortality Database.

A detailed discussion of each of this models can be found in the publications and on the CISNET website, we will not discuss each model in detail, but summarize the pooled comparison of Mandelblatt et al.[48] and discuss the main limitations and implications for our research question. The models estimated a large number of scenarios, but we will only present the results of the part relevant to our research question, the comparison of a screening policy screening age 50-69 to a screening policy screening age 50-74.

Feuer et al.[40] identifies two dimensions to characterize the types of surveillance models used here. The first dimension incorporates micro simulation models at one end of the spectrum, where individuals are run through the model one at a time, where at each transition a random number is generated and individual life histories are generated, to mechanistic or analytic models, where a set of analytically derived equations describe the relationships between key health states and/or tumor growth and metastasis. The University of Texas M. D. Anderson Cancer Center, University of Wisconsin, Georgetown, and Erasmus models could be characterized as micro simulation models; the Dana-Farber model could be characterized as analytic; and the remaining two models (University of Rochester and Stanford) could be described as having some aspects of each. The second dimension of model characterization runs from biologic, where the model goes beyond observable quantities to model the underlying disease onset, growth, and progression of disease, to epidemiologic, where only a portion of the disease process is modeled (usually the observable portion).

The models start with estimates of breast cancer incidence and mortality trends without screening and treatment and then look at the effect of screening use and improvements in survival associated with treatment.

Breast cancer is assumed to have a preclinical, screening-detectable period (sojourn time) and a clinical detection point. On the basis of mammography sensitivity (or thresholds of detection), screening identifies disease in the preclinical screening-detection period and results in the identification of earlier-stage or smaller tumors than might be identified by clinical detection, resulting in reduction in breast cancer mortality. Age, estrogen receptor status, and tumor size– or stage–specific treatment have independent effects on mortality. Women can die of breast cancer or of other causes. as mentioned before, the 6 models use a common set of age-specific variables for breast cancer incidence, mammography test characteristics, treatment algorithms and effects, and non-breast cancer competing causes of death. On the other hand, unobserved variables such as preclinical detectable times (sojourn time), lead time, dwell time within stages of disease, were in these models estimated intermediate outputs that followed from the model structure and assumptions concerning tumor growth.

The stage distributions in unscreened versus screened women in these models were also intermediate outcomes, this in contrast to some other models that use this observable variable as input. As end output from the model reductions in mortality, life years gained were calculated, no QALYs were used. The harmful effects false positive mammograms, unnecessary biopsies and over diagnosis followed from the model, also here no direct observed input was used, no attempt was made to quantify those harms in terms of QALYs. Morbidity associated with surgery for screening-detected disease or decrements in quality of life associated with false-positive results living with earlier knowledge of a cancer diagnosis or over diagnosis was not considered, which makes the models less useful for our purposes.

Table 2.4 gives the results of the different models in terms of mortality reduction and years of life gained for the different models. Gains are fairly limited and there is some variability between models, with number years gained per 1000 women screened ranging from 9 to 17 and number of deaths averted ranging from 4 to 6.

This class of models relies heavily on unobservable variables, and as most models are individual bases they are not always very transparent. Independent validation was made difficult because results from trials and

the main US breast cancer registries were used to parameterize or calibrate the model. Model outputs are similar to the results from RCT's and some observational studies, but this does not say much about the validity of the model as data from those studies were partly used to calibrate the model.

Table 2.4: results of the different models in terms of mortality reduction and years of life gained per 1000 women screened for the different models

Model	D	E	G	M	S	W
Mortality (specific) reduction over the whole period in %						
Screening in agegroup 50-69	16	23	17	16	15	23
Screening in agegroup 50-74	22	27	21	21	20	28
Incremental mortality reduction screening in agegroup 50-74 compared to screening in agegroup 50-69	6	4	4	5	5	5
Years of life Gained per 1000 women screened						
Screening in agegroup 50-69	88	107	111	82	99	84
Screening in agegroup 50-74	106	116	128	96	121	95
Incremental years of life gained screening in agegroup 50-74 compared to screening in agegroup 50-69	18	9	17	14	22	11
Incremental days of life gained per women screened						
Screening in agegroup 50-74 compared to screening in agegroup 50-69 women screened	6,6	3,3	6,2	5,1	8,0	4,0

Model group abbreviations: D _ Dana-Farber Cancer Institute; E _ Erasmus Medical Center; G _ Georgetown University; M _ M.D. Anderson Cancer Center; S _Stanford University; W _ University of Wisconsin/Harvard

Stout et al 2006[56] used the Wisconsin model to do a cost effectiveness analysis, including the use of QALYs, but comparisons of the age groups 50-74 with age groups 50 - 69 were not made.

Rue et al.[55] adapted de Dana-Farber Cancer Institute model of Lee and Zeelen[44] to data in Catalonia. Because there was insufficient information on Catalan survival they combined the survival data from the SEER in the US with Catalan data in a previous publication of Vilaprinyo, 2009[58]. Obtained results were very similar to the ones Lee & Zeelen originally found, a mortality reduction of 21% and 131 life years gained per 1000 women screened for a biannual screening in the age group 50-74, with an incremental benefit for biannual screening 50- 74 of 1.7% in terms of mortality and 2 life years gained per 1000 women screened. As authors had no choice than to use US data for most key variables one can question in what degree this can really be called an adaptation to the Catalan context. Carles et al, 2011[38] finally used the results of Rue et al and Vilaprinyo et al[58] to do a cost effectiveness analysis, including QALYs. They found 3990 life years gained for a cohort of 100 000 women for a biannual screening in the age group 50-74, with an incremental benefit for biannual screening 50- 74 of 299 life years gained per 100 000 women. They found 3891 QALYs gained per 100 000 women screened with a biannual screening in the age group 50-74, with an incremental benefit for biannual screening 50- 74 of 277 life years gained per 100 000 women compared to a schedule 50-69. They did not report the QALYs gained with extending the screening to 50-74 from 50 -69, as it was dominated by screening from 45- 69, but reported that 186 QALYs per 100 000 were gained by extending the screening to 45-74 from 45 -69. Interestingly, they did not incorporate the results of Vilaprinyo et al,2009[58] into their calculations, but used US survival data. They did not take into account over diagnosis.

2.2.4.2. Models not related or not using CISNET methodology

Carter et al, 2005[39, 78] developed a micro simulation model based on tumor growth using mainly SEER data. The model lacks credibility though mainly because of unrealistic assumptions concerning stage specific survival, as they assume a fixed survival of 2 years for stage 4 and complete cure for stages 1, 2 and 3. This leads to considerably higher years of life gained for screening than other models but is in absolute contradiction to what we know about stage specific survival.

Rojnik et al, 2008 produced a time dependent Markov model with 4 stages, DCIS, local, regional and distant. Overall model structure was described but details on how the model was parameterized are lacking so we cannot judge how this was done or if assumptions were reasonable. They only report ICERs so we have no information on assumed gains in Life Years Gained and QALYs.

Neeser et al developed a simple Markov model comparing organizes screening with a coverage of 70% with opportunistic screening with a coverage of 20%. They assumed that the organized screening reduces breast cancer mortality with 15% based on the IARC handbook, but it is unclear how they come to this figure as IARC postulates a reduction ranging from 5 to 20%. They calculated the years of life gained for a 10 years screening beginning at 70 (they evaluated other schedules not relevant for our research question as well. They found that organized screening would save 41 lives per 100 000 and add 0.008 life years (2.9 days) per women screened for 10 years. The model is somewhat overly simplistic by not taking into account lead time but applying assumed reductions immediately. No QALYs were used and effects on morbidity was not taken into account.

Rauner et al, 2010 developed an ant colonization optimization model but only evaluated the effect of screening amongst women 50-70 and their rather experimental model is not useful for our purposes. It is also unclear how they actually modeled stage specific survival.

Mahnken et al, 2008[46] developed a method to adjust for lead time bias, length bias and over-detection and applied this to SEER data, but provided only adjusted Hazard ratio's.

Rijnsburger et al, 2004[53] used the MISCAN micro-simulation model developed by the Rotterdam[57] (see above) to replicate the data of the Canadian CNBSS-2 trial on breast cancer screening among women aged 50–59, so their findings are not really useful for our purposes.

Barratt et al 2005[36] constructed a Markov model for two hypothetical cohorts, with one cohort women undergoing biennial screening and the other not, assuming 100% participation. Within this model, they evaluated the outcomes of women over 70 years old undergoing 10 years of biennial screening. They assume a 37% mortality reduction, adjusting the 25% reduction from for non compliance, and assume that benefit accrues linearly to maximal level over first five years after starting screening and that benefit declines linearly to nothing over five years after stopping screening. For women who continue screening for 10 years after the age of 70s, two fewer women per thousand die from breast cancer than in women who stop screening (six *v* eight deaths from breast cancer). The number of diagnoses of breast cancer in screened women is about 41 and the number in unscreened women about 26. assuming a risk reduction of 50% brings the number of deaths in the screened group down 6.2 to 5.1. This simple model has the advantage of transparency, but does not take into account the effects of lead time and stage-shifts on morbidity.

2.2.5. Conclusion

Models described are give useful insights and elements but it is difficult to adapt them to the Belgian situation as we do not have the necessary data to parameterize them. The CISNET models give a modest gain in year of life between 9 an 22 years per 1000 women screened.

2.3. Review of quality of life studies

Because breast cancer screening programs are expected to have an impact on the quality of life (QoL) of the patients, models with a one-dimensional health-outcome measure in terms of survival are not enough informative. It is important to take into account all the multidimensional health outcomes in the assessment of breast cancer screening programs. To value these multidimensional outcomes into a single measure, quality-adjusted life-year (QALY) must be used. QALYs permit to adjust the expected length of life by the health-related quality of life. These adjustments are made using utilities derived from individuals' preference for different health states.

Determination of utility values, needed for the calculation of QALYs, requires two steps:

1. The health state description. According to the pharmaco-economic guidelines of the Belgian Health Care Knowledge Centre (KCE), health states should be described on a standardized descriptive system. Ideally, the description should be done by Belgian patients using a generic descriptive system, such as the EQ-5D. If health states descriptions from Belgian patients are not available, descriptions from similar patients in other countries may be used79.

2. The valuation of these health states. According to the pharmaco-economic guidelines of the KCE, health state values should be valued on a 0 (=value for death) to 1 (=value for perfect health) scale by a representative sample of the general public. Ideally, they should be valued by the Belgian population but if no original Belgian data are collected, valuations from other countries can be used and discussed79.

In this section, the availability and the quality of published utility values describing the burden of disease due to breast cancer (screening and treatment) is assessed.

2.3.1. Methods

2.3.1.1. Literature search strategy

Electronic databases were consulted for original publications on utility estimates for different health states associated with breast cancer screening and treatment. Systematic searches were carried out up to the end of October 2011 in the following databases: Medline (via OVID), Embase (via Embase.com), HTA and EED (via CRD NHS) and Psycinfo (via OVID).

Searches using various qualifiers for “quality of life” were used as Subject heading or text word. See appendix 3.1 for an overview of the search strategies and terms used.

2.3.1.2. Selection criteria

Identified references were assessed against pre-defined selection criteria (in terms of population, intervention, outcome and design –Table 2.5) in a two-step procedure: initial assessment of the title, abstract and keywords; followed by full-text assessment of the selected references. When no abstract was available and the citation was unclear or ambiguous, the citation was assessed based on keywords and full-text. Reference lists of the selected studies were scrutinized for additional relevant citations.

Table 2.5: Article selection criteria

	Inclusion criteria	Exclusion criteria
Population	Screened or treated patients for breast cancer, with a Caucasian origin and without high risk factors	Other diseases, non Caucasian, high risk women
Intervention	Any intervention relevant to the Belgian settings	Interventions not used in Belgium
Outcome	Unique QoL weights allowing to derive QALYs (=utilities)	Multi-dimension HRQoL scores, DALYs, HYEs, …
Design	Direct (TTO, PTO, SG) or indirect (EQ-5D, SF-6D, HUI, QWB) valuation methods in primary studies	Letters, secondary studies, CUA with QALYs derived from the literature, … Direct valuations using VAS (not recommended in the KCE pharmaco-economic guidelines)[79].

QoL: Quality of Life. QALY: Quality adjusted life year. HRQoL: Health-Related Quality of Life. DALY: Disability-Adjusted Life-Years. HYE: healthy-years-equivalent; TTO: Time-Trade-Off. PTO: Person Trade-Off. SG: Standard-Gamble. HUI: Health Utility Index. QWB: Quality of Well Being scale. CUA: cost-utility analysis. VAS: visual analogue scale

2.3.1.3. Selection process

The flowchart of the selection process is presented in appendix 3.2. The searches on the databases returned 524 citations. After exclusion of 172 duplicates, 352 unique citations were left (see also appendix 3.2). Hand searching allowed us to identify 3 additional citations. Two-hundred and ninety (290) references were discarded based on title and abstract, leaving 65 references for full-text evaluation. Another 49 references were excluded at this stage, mostly because of the unmet design and population criteria. Overall, we selected 16 primary studies (see appendix 3.2).

2.3.2. Results

A summary of the selected studies can be found in appendix 3.3. It should be noted that this summary only report methods used to derive utility values and their results. If other parameters were measured, they were not reported in the summary.

The selection of utilities was done according to the following stages:

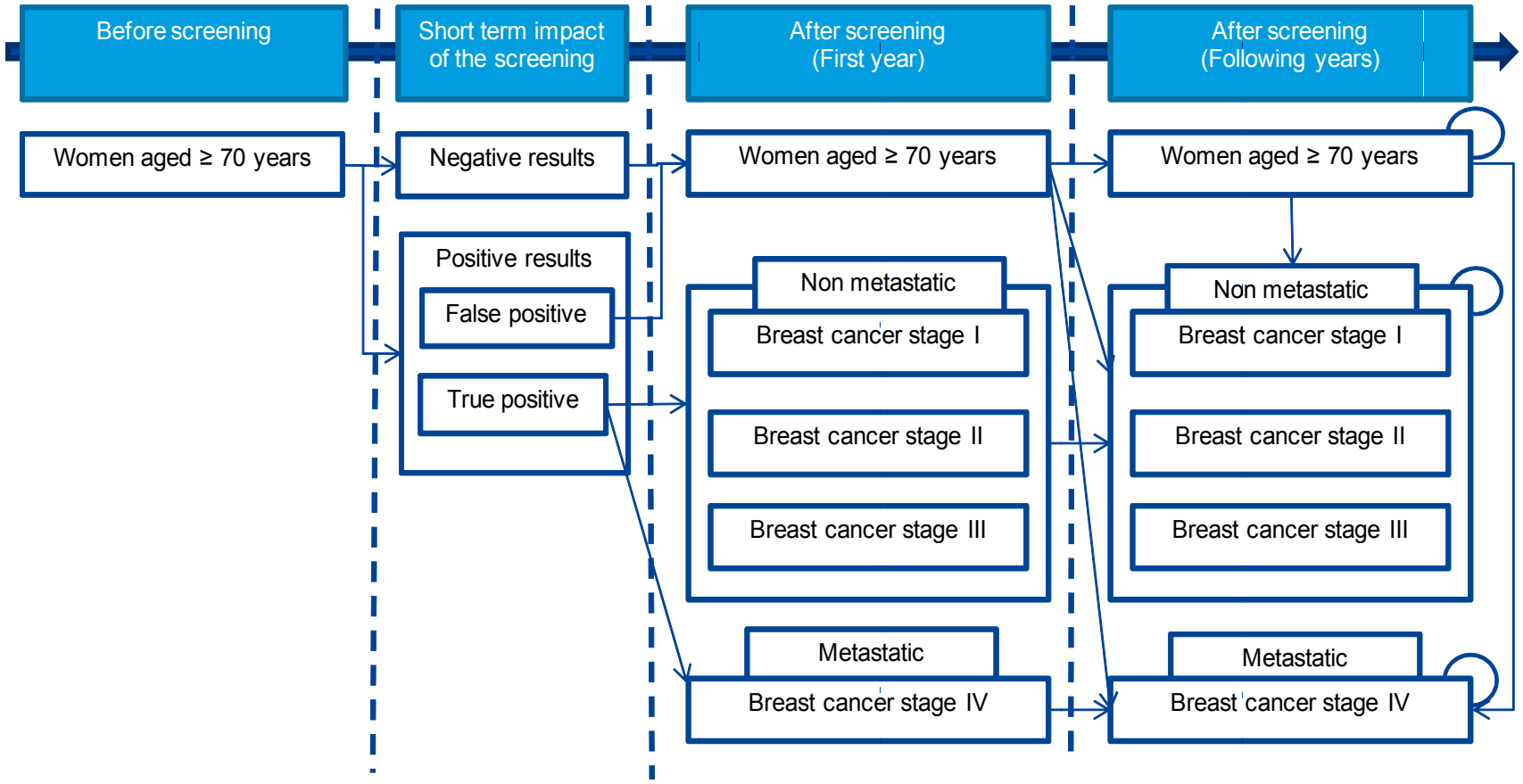
- Determination of health states for which utilities were needed
- Selection of utilities
  Selection of a basecase study
  Selection of other studies
- Pooling of selected utilities and calculation of percentage changes

2.3.2.1. Determination of health states

Health states for which utility values are needed are listed in Figure 2.1.

It should be noted that this figure is a schematic representation of the reflection process but not the model itself (described in section 3.2).

Figure 2.1: Health states for which utilities are needed (reflection process)



2.3.2.2. Selection of utilities

To select utility values, we first tried to find Belgian data as recommended by the pharmaco-economic guidelines of the KCE[79]. However, no Belgian data could be found.

Then, we tried to find the most complete study which best fit with our model. The aim was to avoid as much as possible the use of multiple instruments and multiple populations to derive them. Indeed, according to the pharmaco-economic guidelines of the KCE, it is strongly recommended to use the same descriptive instrument and the same set of values for quality of life weights coming from different studies[79].

However, no study assessing all of the heath states described in section 2.3.2.1 with the same design was found. We therefore tried to find the study with the greatest number of health states corresponding to our model and to use it at the starting point of the selection process.

Selection of the base case study

We found only one study having assessed utility values for both non-metastatic and metastatic patients, i.e. the study of Lidgren et al[80]. We had the chance that this study also used the better available instrument according to the pharmaco-economic guidelines of the KCE, i.e. the EQ-5D[79]. This study was therefore the starting point of our selection process.

Utility values in the study of Lidgren et al.[80] were derived from two methods, i.e. a direct valuation method (i.e. the time-trade-off (TTO) technique) by Swedish patients and an indirect valuation method using a generic instrument (i.e. the EQ-5D instrument). Because pharmaco-economic guidelines of the KCE recommend the use of the EQ-5D, only these valuations were retained (i.e. utility values from EQ-5D and not from TTO). In this study, health states were described by Swedish patients and valued using UK tariffs (because no tariffs from the Swedish population are available). Health states descriptions can be found in Table 2.6. Utility values for non-metastatic patients in the first year (i.e. the year of the treatment) and the following years as well as utility values for metastatic patients are described in Table 2.8.

This study had the following limitations:

- Utility values were measuring during out-patient visits at a breast cancer outpatient clinic (Karolinska University hospital), implying the following limitations:

  Utility values for non-metastatic patients did not fully take into account the short term impact of surgery. However, on an annual basis, this short term impact was expected to cover a limited length of time and was therefore not included in the model.

  Utility values for metastatic patients did not represent patients in palliative care. It was therefore assumed that these utility values only reflected the quality of life of metastatic patients during the first year of diagnosis.

  The short term impact of diagnosis is also not fully taken into account (not measured at the moment of the diagnosis), even if authors of this study reported that this impact was expected to be included in the valuation (measured the year of diagnosis).

- Non-metastatic patients were divided in only two groups, i.e. the first year of diagnosis and the following years. It was therefore assumed that after the year of treatment, utility values remained constant. This assumption is supported by an US study where no significant difference in utility values (from EQ-5D using US tariffs) was found at year 5, 10 and 15[81]. This US study is described in the appendix 3.3.

- It should also be noted that utility estimates for non metastatic patients (primary breast cancer in the first year of diagnosis) and metastatic patients were similar (0.696 and 0.685 respectivelly). This inconstistency may be due to the following reasons:

  Metastatic patients only include patients going in out-patient consultations (best cases).

  Generic instruments such as the EQ-5D are less sensitive to capture relevant changes in health in a specific disease than disease-specific instruments. However, diease-specific instruments can only be used if validated mapping functions to derive utilities from these instruments are available, which was not the case[79].

Table 2.6: Health states descriptions for the study of Lidgren et al.

Primary breast cancer (year 0-1)	Patients who had primary diagnosis breast of cancer within 1 year or less prior to answering the questionnaire, no recurrence and no metastatic disease
Recurrence (year 0-1)	Patients who had at least one recurrence (loco-regional and/or contra-lateral) within 1 year or less prior to answering the questionnaire, and no metastatic disease.
Primary breast cancer and recurrence following years	Patients who had been diagnosed with a primary breast cancer or their last recurrence more than 1 year prior to answering the questionnaire, and no metastatic disease.
Metastatic patients	Patients who had metastatic disease

Selection of other studies

For other health states, we tried to find studies having used similar instruments for the same population. The study of Lidgren et al.[80] allowed us to identify a study assessing utility values for the general Swedish population stratified by age and gender using the same instrument (EQ-5D with UK tariffs), i.e. the study of Burström et al.[82] These utility values were therefore used for women aged 70 and over (see Table 2.8).

For the short term impact of positive results after screening, one study using the EQ-5D instrument was identified (Gerard et al.)[83]. This study assessed utility values for false positive, true positive, false negative and true negative. Health states were described by the UK population (and not the Swedish population) but UK tariffs were used to valuate these health states (as in the other selected studies). A description of the "false positive" state is given in Table 2.7.

As showed in this table, the description include the following stage: being invited for screening, having a breast screen, waiting for results, being recalled for further examinations, having further examinations and obtention of a diagnosis, i.e. no evidence of breast cancer. For the assessment, only three of the five EQ-5D dimensions were used, i.e. usual activity; pain/discomfort; and anxiety/distress and it was assumed that the remaining two dimensions (i.e. mobility and ability of self-care) were unaffected. The quality of life effects associated with true negatives and false positives lasted 12 months while true positive and false negative were measured for the remaining life expectancy. These values can therefore not be used to measure the short term impact of screening. We decided to make the following assumptions:

- True negative patients have utility values equal to the general population.
- The short term impact of positive results at screening is measured by the percentage change between true negative and false positive.
- This impact is present until the diagnosis, i.e. on average 45 days after screening according to IMA data. After, either valuations of Burström et al.[82] (general population for false positive) or valuations of Lidgren et al.[80] (non metastatic or metastatic disease year 1 for true postive) were used.

Utility values for false positive and true negative can be found in Table 2.8. It should be noted that the study of Domeyer et al.[84] described in appendix 3.3 assessed the short term impact of biopsy. However, to avoid model complexity and because the biopsy is included in the description of a "false positive" in the study of Gerard et al.[83], we decided to not take the study of Domeyer et al.[84] into account.

Concerning the evolution of utility values for patients with metastatic breast cancer in the long term, no study was found.

-

Table 2.7: Description of a "false positive" state (Gerard et al)[83]

Routine breast screen	• She is invited by letter for routine breast screening. • The appointment is about 2 weeks from receiving the invitation. • The visit at the breast screening centre takes about half an hour, which may include waiting time. • A female radiographer asks about any symptoms or history of breast disease and explains what will happen. • To take the X-ray she is asked to undress to the waist. Each breast is placed in turn between two special X-ray plates and compressed to get the best possible picture.
Further tests	• She is asked by letter to go to the breast screening centre the following week. • Other tests are needed because the breast X-ray result is not clear. • This visit may take up to half a day. • The breast X-ray is repeated. • The doctor examines her breasts. • The doctor may carry out an ultrasound examination. • Fluid from the affected area is taken for laboratory analysis using a fine needle to do this her breast may again be compressed between the X-ray plates.
The results of the tests are ready within the week	• The tests show no evidence of breast cancer.
Quality of life effects of routine breast screening (short term)	The QoL of some women is affected by the experience of routine breast screening and breast cancer diagnosis. The effects may continue for some time.
Receiving the invitation	• Most women are pleased to receive the invitation. • Some women are made nervous, anxious or depressed, and are worried about having breast cancer.
Waiting for the day of the appointment	• Most women carry on with their usual activities and interests. • Some women are anxious and depressed, unable to concentrate, sleep badly and are moody and irritable. They are unable to carry on with their usual activities and interests. • Personal and sexual relationships may be affected.

At the breast screening clinic	• Most women are nervous, but are not anxious or depressed. • Most women are not embarrassed by the screening procedure. • Most women are not unduly worried about breast cancer developing. • Most women find the breast X-ray is uncomfortable and slightly painful, but this is short lived. • Some women find the breast X-ray very uncomfortable and painful.
Waiting for the results	• Most women carry on with their usual activities and interests. • Some women are anxious and depressed, un-able to concentrate, sleep badly and are moody and irritable. They are unable to carry on with their usual activities and interests. • Personal and sexual relationships of some women may be affected. If recalled for further tests: • Most women are very anxious at being recalled for further tests. • One of the tests, where the doctor removes fluid from the affected area, is painful.
Clear results after the test	• Most women are reassured by the clear results. • Some women remain anxious for up to a year before they are back to their usual self.

Table 2.8: Description of the selected utilities

Author (year)	Instrument	Population for health state description	Population for valuation	Health state	Mean value
Lidgren et al. (2007)[80]	EQ-5D	Sweden patients (Mean age: see health states)	UK tariffs (general population)	Primary breast cancer (Year 0-1); Mean age: 56	0.696
			UK tariffs (general population)	Breast cancer (following years); Mean age: 58	0.779
			UK tariffs (general population)	Metastatic patients; Mean age 56	0.685
Burström et al. (2001)[82]	EQ-5D	Sweden patients (Mean age: see health states)	UK tariffs (general population)	Women aged 50-59	0.833
				Women aged 70-79	0.792
				Women aged 80-88	0.740
Gerard et al. (1999)[83]	EQ-5D	Women from UK aged 40-64 years (eligible for screening)	UK tariffs (general population)	True negative	0.940

2.3.2.3. *Pooling of selected studies and calculation of percentage changes*

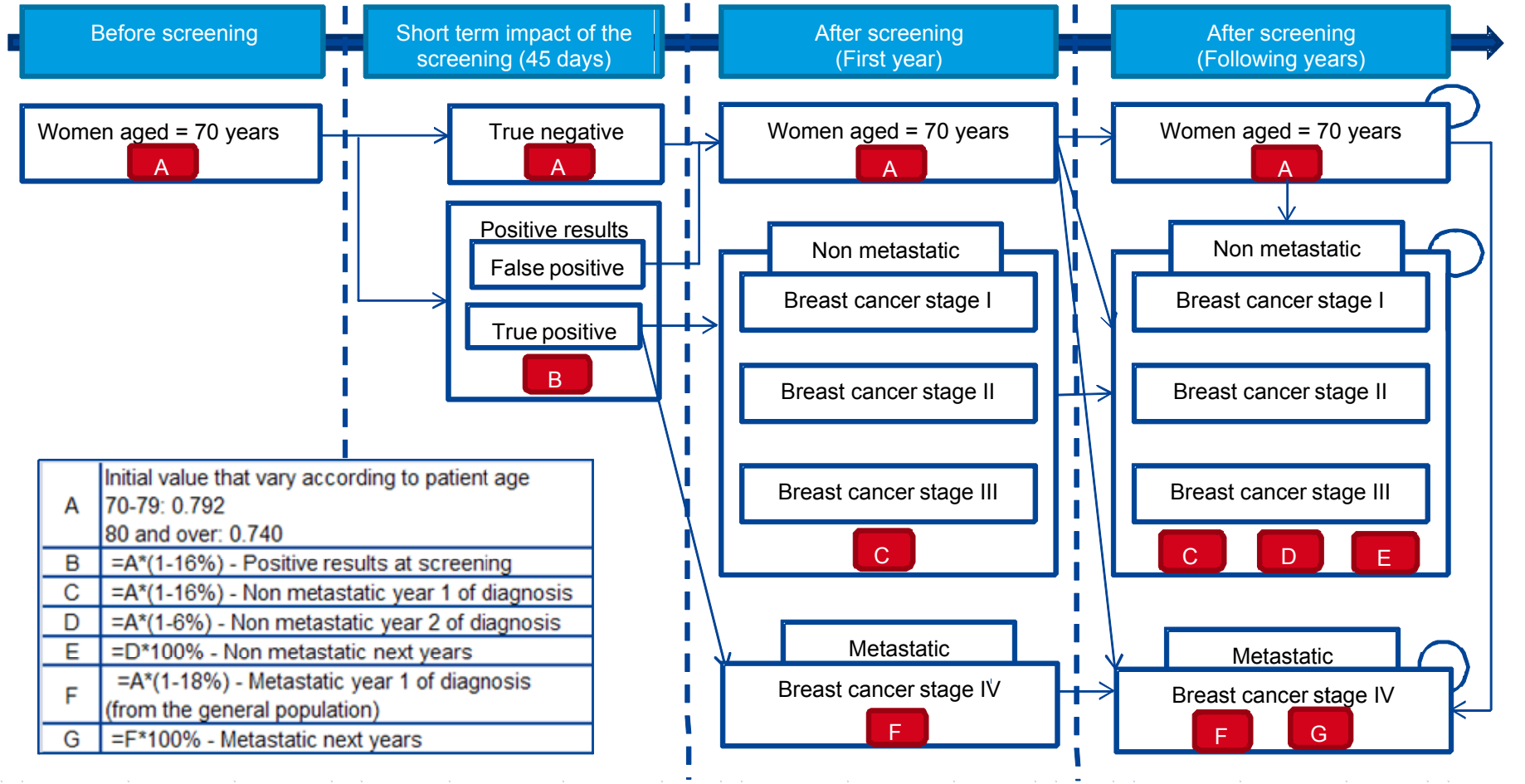
A summary of selected utilities and of calculation of percentage changes can be found in Figure 2.2. The utility values of the study of Burström et al.[82] were chosen as the initial values of the model (first state of the model (A)). These values varied according to women age. Then, the percentage change relative to these values was applied. It was assumed that these percentage changes did not vary according to the women age (no data).

The next stage concerns the short term impact of the screening. It was assumed that utility values for true negative women were equal to utility values in the general population (A). Then, the percentage decrease in utilities between true negative women and false positive women was calculated, i.e. -16% (B). Initial values were thus maintained for true negative women and decreased by 16% for women with a positive result after screening (false or true positive). As mentioned in the section 0, these utilities will be maintained for 45 days.

For the first year of screening, women without breast cancer had utility values equal to the general population (A). For true positive, utility values of the study of Lidgren et al. were used.[80] To make the link between the study of Lidgren et al.[80] and the study of Burström et al.[82], percentage changes between values for the same population were used, i.e. Swedish women aged 50-59 (and UK tariffs). Utility values were therefore reduced by 16% ((0.696-0.833)/0.833) for non metastatic patients (C) and by 18% ((0.685-0.833)/0.833) for metastatic patients (E).

For the next years, people from the general population who developed non-metastatic or metastatic breast cancer had utility values reduced by 16% (C) and 18% (E) respectively (as calculated above). Non-metastatic patients who stayed in this stage had their utility decreased by 6% compared to the general population ((0.779-0.833)/0.833) and maintained the years after (D). Metastatic patients maintained their utility until death (G).

Figure 2.2: Percentage change in utilities



A	Initial value that vary according to patient age 70-79: 0.792 80 and over: 0.740
B	=A*(1-16%) - Positive results at screening
C	=A*(1-16%) - Non metastatic year 1 of diagnosis
D	=A*(1-6%) - Non metastatic year 2 of diagnosis
E	=D*100% - Non metastatic next years
F	=A*(1-18%) - Metastatic year 1 of diagnosis (from the general population)
G	=F*100% - Metastatic next years

2.3.3. Discussion

To include the quality of life impact of screening in the analysis, utility values for each health state of the model had to be identified. The aim of this chapter was therefore to select these values. The method was based on the KCE pharmaco-economic guidelines[79]. We tried to avoid the use of multiple instruments and multiple valuations and focused on utility values derived from the EQ-5D instrument.

- The analysis had the following limitations:
- No Belgian data were available and a transferability analysis was not possible (no access to primary data). Even if we expected that using UK tariffs instead of Belgian valuations would not greatly influence results, Belgian data would be interesting for future models.
- The short term impact of surgery and of diagnosis was not taken into account because no valid data were available.
- Even if the EQ-5D is one of the best available instrument to assess these utilities (according to the KCE pharmaco-economic guidelines), this instrument is less sensitive than disease specific instruments. Consequently, it can be expected that the impact of some conditions such as a mastectomy (partial or total) would have been more important if a disease specific instrument instead of a generic instrument had been used. This lack of sensitivity could explain the low percentage change between patients with breast cancer and women in the general population or between metastatic and non metastatic patients. The assessment of the quality of life from disease specific instruments was nevertheless not investigated in this chapter because these instruments do not permit to derive QALYs.
- Finally, the review of the literature showed an important variability between reported utility estimates for breast cancer health states (see appendix 3.3), revealing a high level of uncertainty around these parameters. Because of this uncertainty, a sensitivity analysis on these parameters should be done in the chapter on model results.

3. DECISION ANALYSIS

To quantify what the implications of our findings are on the Belgian situation we constructed a decision analysis model using two different approaches. For the first simple approach, we applied data from IMA, cancer registry and data from the literature on the Belgian life tables (see below). For the second, we constructed a simple time dependent cohort with annual cycles.

We consider performing one Belgian decision analysis a better approach than trying to adapt the models discussed in chapter 2 to the Belgian situation. Indeed, Belgian data needed to parameterize these models are not available and we would merely reproduce the already published findings of these models, as we would be obliged to use the same (mainly US) data.

We look at the effect of introducing mammography screening in addition to the currently existing situation with the opportunistic screening going on at the current level. This has the advantage that we can use Belgian data as baseline without having to modify then, as this can only be done making use of an additional number of non verifiable assumptions. We describe here:

- Available data used in this decision analysis;
- Additional literature review focused on quality of life related to the screening and to the breast cancer as such;
- The model used for this decision analysis.

3.1. Data sources

Belgian life table (2009)

Overall survival was taken from the Belgian life table of 2009 from be.STAT (http://statbel.fgov.be)

Belgian Cancer Registry (BCR)

The Belgian Cancer Registry Foundation is an public institution which collects data concerning new cancer cases in Belgium and makes up statistics from these data (http://www.kankerregister.org/).

Belgian organized screening

As recommended by European Commission, Belgium started a national organized screening programme. The target age groups as defined by the program are women aged 50 to 69 years. Belgian breast cancer screening programs are organized by: Brumammo (Bruxelles, http://www.brumammo.be/), Centre Communautaire de Référence pour le dépistage des cancers (CCRef: http://www.ccref.org/) (Communauté Française) and BorstKankerOpsporing (BKO) (Vlaamse Gemeenschap: http://www.zorg-en-gezondheid.be/).

Intermutualistic Agency (IMA)

The Intermutualistic Agency (IMA) centralises data coming from all Belgian sickness funds. IMA compiled and published several reports on the national screening program containing data on the target age groups as defined by the program (50-69 years). IMA complemented this with information on persons outside the target age-group, with a particular focus on the tests used, delays between screening tests and possible confirmation and treatments following testing (http://www.nic-ima.be/).

Dutch National Evaluation Team for Breast cancer screening (DNETB)[85].

The Dutch National Evaluation Team for Breast cancer screening published a report with their findings covering the period 1990-2007 containing information on age specific stage distributions in the screened population.

SEER database

The Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute works to provide information on cancer statistics in an effort to reduce the burden of cancer among the U.S. Population (http://seer.cancer.gov/). SEER collects data on cancer cases from various locations and sources throughout the United States. Data collection began in 1973. As they used an outdated distribution we could not incorporate these in the model.

3.2. Model description

In a first simple approach we applied the 22% reduction in breast cancer mortality caused by screening coming from RCT and its range, resulting from the results of the meta-analysis of Gøtzsche et al, 2008[4] on the Belgian life table. We assume here that the reduction in women aged 70-74 is similar to the reduction in other age groups. We also assume, following Barratt et al 2005[36] that benefit accrues linearly to a maximal level over first five years after starting screening and that benefit declines linearly to nothing over five years after stopping screening. Life years saved can then be derived from the life table. However, effects of harms and effects on quality of life resulting from earlier diagnosis, over-diagnosis and stage-shift is more difficult to assess in this approach. Therefore this approach was only used for cross validation by comparing it with a more complex approach that makes use of the stage-shift caused by screening.

The second approach makes use of the Belgian Cancer Registry (BCR) data on incidence of invasive cancer and DCIS for the construction of a time dependent state transition cohort model with annual cycles.

The model compares 2 cohorts:

- A cohort of women starting at age 70 where screening is extended to the population in the age group 70-74, where a part of the women participates in the screening and where a part of the cancers is found by screening, depending on participation rate and sensitivity of the screening. There is a mix of screen detected and not screen detected cases (interval cancers and cancers amongst unscreend womens). The screen detected cancers will have a different stage distribution than the cancers not detected by screening.

- A cohort of women starting at age 70 where the screening is not extended beyond the age of 69 years. For this cohort all women have the stage distribution of the non screened.

All women are followed to death. The cumulative number of life years, the number of QALYs and deaths to breast cancer of the two cohorts are compared. Overall mortality is not compared as in the end everybody dies.

We assume that:

- Survival and quality of life of the women depends only on the stage of the tumor at the moment the tumor is detected and the age of the women, and not on the presence or absence of screening;
- All benefit of the screening results from the stage-shift, the differences in stage-distribution caused by the screening.

Harm caused by false positives at the moment of the screening is accounted for separately, by assuming 3 screening rounds with a 2 years interval in the participation women and applying recall rates at the proportion women that are alive and without breast cancer at the moment the screening round actually takes place.

Figure 3.1 shows the different compartments in the two cohorts and the transitions between them.

In the unscreened cohort, transitions between compartments from year to year are determined by:

- Incidence of breast cancer;
- Stage distribution of unscreened cancers;
- Stage specific survival; and
- Age specific overall mortality due to other causes.

On top of that, for the cohort where screening takes place, transition is also determined by some aspects of the screening:

- Lead time as part of the cancers will be found earlier;
- The proportion of cancers found by screening and proportion not found by screening and their respective stage distributions.

As survival and quality of life depends on both age and time since diagnosis in the model, a separate compartment is made for each age of diagnosis and stage, and stage specific survival is than applied. As screening is applied during 5 years and there is an assumed lead time of 2 years (or 3 in sensitivity analysis) the number of compartments remained manageable.

Transitions between stages are not included as stages are assessed at the moment the diagnosis is made followed by treatment. Even if the cancer evolves after treatment it does not necessarily go through the 4 stages anymore.

Figure 3.1: Comparison of the two cohorts with and without a screening program

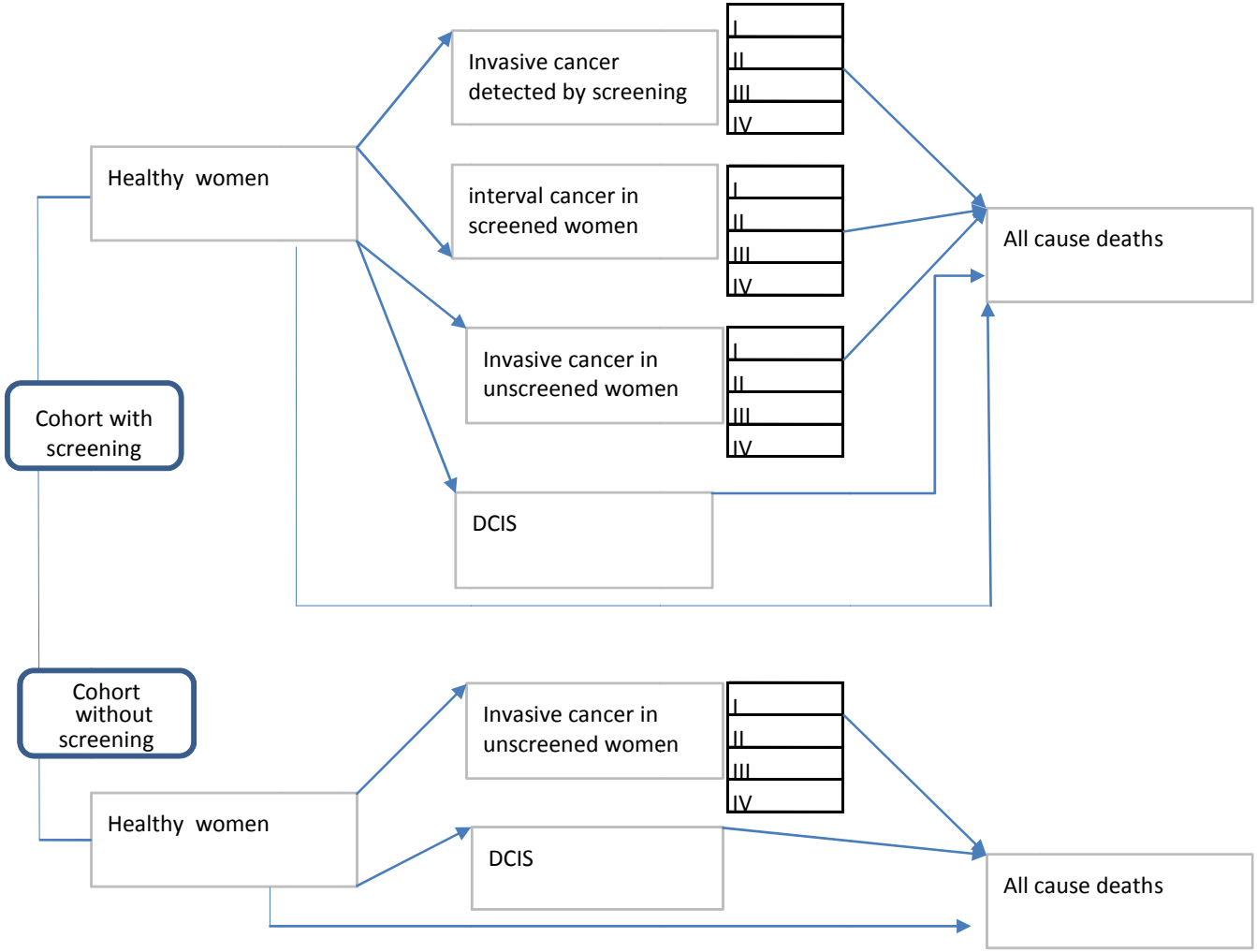
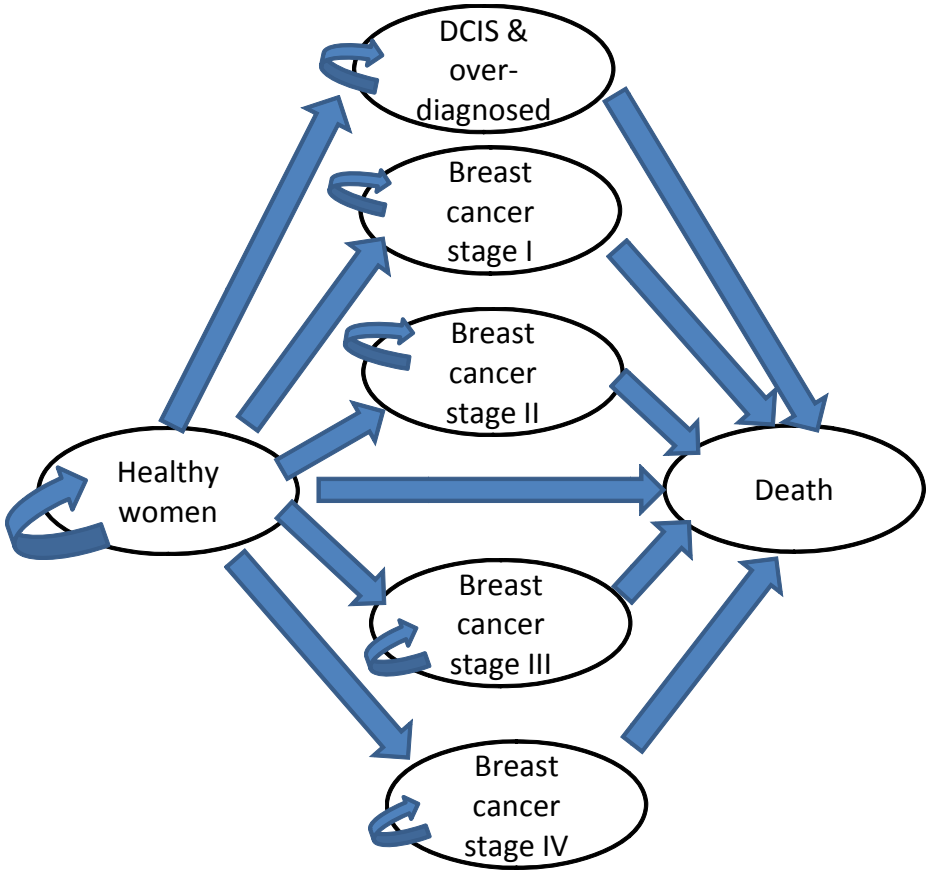


Figure 3.2: Compartments in the two cohorts and the transitions between them



3.3. Description of the parameters

3.3.1. Age specific overall survival

Overall survival was taken from the Belgian life table of 2009 from be.STAT (http://statbel.fgov.be) after adjusting for breast cancer specific mortality based on data from the Belgian Cancer register.

3.3.2. Breast cancer incidence

For the baseline without screening the BCR data on incidence of DCIS and the 4 stages for invasive for the age group 70-74 of the period 2004-2008 were used. There is some opportunistic screening in that age group. From the IMA data we infer that in Flanders the coverage with at least one mammography in the past 2 years is 18% (for details see the KCE report 172 on breast cancer screening in risk groups)[2]. Given that we can assume that an important part of this is also for diagnostic and follows up purposes, so we choose to use data issued from Flanders because they are less contaminated by opportunistic screening.

For the situation where screening takes place, incidence in the 70-74 group will increase with a number of cancers coming from two sources:

- **Lead time**, cancers that would have appeared later but are found now because of lead time. This will lead to a compensatory decrease in number of cases in the following years. The moment and degree of this shift depends on the assumed lead time (see point 3.2.2). We used 2 years lead time in the baseline and 3 years in the sensitivity analysis.
- **Over diagnosis invasive cancer**, we modeled the over diagnosis based on the findings in the literature as described in the literature review above under 2.1.3.5. We assume a range of 2 to 30% for over diagnosis excluding DCIS.
- **Over diagnosis DCIS**, we model the over diagnosis of DCIS in a different way: we use the observation that in Flanders the incidence DCIS per 100 000 is twice in the group 60-69 where screening takes place compared to the age-groups 70-74 and 75-79 where only a limited amount of opportunistic screening takes place. This is in contrast with the Brussels capital region and Walloon region where the drop in DCIS is much less pronounced. So we take as an estimation of over diagnosis the difference in DCIS incidence in Flanders between the age-groups 60-69 and 70-74, augmented by 1.5 to adjust for the fact that screening coverage is only 60% as a proxy for overdiagnosed DCIS. This brings us to an over diagnosis of DCIS of 40 per 100 000 women per year.

3.3.3. Participation rate

We used a 70% participation (plausible range 60% to 80%) as baseline.

3.3.4. Proportion of screen detected breast cancers

The data of the Belgian screening program show that in the age group 50-69, 49% of the cases are found by screening, and the rest is either interval cancer or not participating in the screening. Among the screened women, 75% of the found cancers are screen-detected and 25% is interval cancer. We used a proportion of cancers found among the women participating in screening of 70% (plausible range 60% to 80%).

3.3.5. Recall rate

We assume a recall rate of 3.5% based on the data from the Flemish screening program concerning follow up rounds (as the screening would be an extension of the screening among women aged 50-69. For the sensitivity analysis we used 2% in an optimistic scenario and 5 and 10% in the more pessimistic scenario (10% recall rates are observed for the moment in certain regions).

As a baseline we assume a delay of 45 days, based on IMA data, with a plausible range for the sensitivity analysis of 36 and 45 days (subtracting and adding 20%).

The short term impact of positive results at screening were measured by the percentage change in utility values between true negative and false positive results.

3.3.6. *Stage distribution and stage shift*

We take estimations of the stage distribution for breast cancer amongst screened and unscreened from the BCR data on the Flanders and data provided by the Flemish screening program.

For the stage distribution in the unscreened women, we can consider the stage distribution amongst women in the group 70-74 in Flanders for the years 2004 - 2008 a good estimation. The stage shift will be slightly underestimated as there is some opportunistic screening in that group going on, see above.

For stage distribution in the screened population, the base case estimation is based on the data from the Dutch National Evaluation Team for Breast cancer screening report of 2009 (DNETB)[85] who provide data specifically for the age group 70-74 from 1998-2007. Although using 2 stage distributions from different sources is a suboptimal way of modeling a stage shift we think this approximates best the Belgian situation, as we do not have date on screen detected cancer in this age group. We assume stage distribution of cases among the non screened and interval cancer to be the same, based on the data from the Flemish screening program.

The Flemish screening program provided data on the stages among screen detected cancers, interval cancers and cancers amongst non participants, collected amongst women who gave their consent in the period 2001-2006. Stage distribution of interval cancers and cancer among non participants is very similar.

Table 3.1: Stage distribution among screen detected breast cancers, interval cancers and cancers among non participants, age 50-69, Flemish screening program 2001-2006.

	Screen detected cancers		Interval cancers		Cancers amongst non participants	
Stage	**n**	**%**	**n**	**%**	**n**	**%**
I	**2586**	62.5%	624	41.5%	1454	41.8%
II	**1306**	31.6%	656	43.6%	1460	42.0%
III	**232**	5.6%	200	13.3%	493	14.2%
IV	**15**	0.4%	24	1.6%	71	2.0%
TOTAL	**4139**	**100%**	**1504**	**100%**	**3478**	**100%**

This baseline stage shift we call Scenario 1:

Stage distribution of cancers not found by screening

BCR data (Flemish population,70-74y, 2004-2008)

Stage	%
I	31.6%
II	42.3%
III	16.6%
IV	9.5%

Stage distribution of cancers found by screening

Data of the DNETB screening report 2009

Stage	%
I	80%
II	18.7%
III	0.8%
IV	0.5%

Important remark: It is important to note that this shift concerns **only screen detected cancers** and that in the cohort with screening interval cancers and nonparticipants keep the stage distribution of cancer not found by screening. In most cases this is around 50% but depends on the other parameter values of the screening and varies in time.

For the sensitivity analysis we use 2 supplementary scenario's:

As the stage distribution from the Dutch National Evaluation Team for Breast cancer screening report of 2009 may be more favorable than what can be achieved in the Belgian context, we used as an alternative scenario the stage-distribution for screen detected patients of the age group 50-69 from the Flemish cancer screening program.

This we call Scenario 2:

Stage distribution of cancers not found by screening

BCR data (Flemish population,70-74 years, 2004-2008)

Stage	%
I	31.6%
II	42.3%
III	16.6%
IV	9.5%

Stage distribution of cancers found by screening (Flemish screening programme (50-69 years)

Stage	%
I	62.5%
II	31.6%
III	5.6%
IV	0.4%

In a third scenario we use a slightly different modeling approach.

Instead of using stage distributions amongst screened and unscreened women, we assume that introducing screening in the group 69-74 will shift the stage distribution amongst **all breast cancer** cases in the population to the stage distribution of the women 60-69 in the same period, using data for Flanders from the Belgian breast cancer registry.

This we call scenario 3

Stage distribution of cancers not found by screening

BCR data (Flemish population,70-74 years, 2004-2008)

Stage	%
I	31.6%
II	42.3%
III	16.6%
IV	9.5%

Stage distribution amongst all breast cancers if screening levels are similar to levels among 60-69 in Flanders

Stage	%
I	45.7%
II	35.9%
III	12.5%
IV	5.9%

3.3.7. Stage specific relative survival

Stage specific survival was taken from Belgian stage specific annual survival data (taken from KCE report 150A)[86]. We only have data up to 5 years. We used data from the Dutch cancer register taken from the website (http://www.cijfersoverkanker.nl) to supplement until 7 years (see appendix 4.1). We assumed that survival conditional on stage is similar in screened and unscreened breast cancer patients. As a sensitivity analysis we used also:

- Entirely the Dutch survival data for women above 70 years per stagegroup. The relative survival curve shows a lower relative survival for women above 70 compared with the overall survival. This may reflect the fact that older women support the invasive treatments less well but it is also possible that there is undertreatment of the elderly. Moreover, the data include patients that were treated more than 20 years ago, this may also explain the lower relative survival.
- British survival data coming from breast cancer research UK (http://info.cancerresearchuk.org/) They provide 10 years survival data but survival is considerably lower than the Dutch or Belgian data. One of the problems with 10 year survival data is the fact that it reflects survival of persons treated at least 11 years ago, given the fast evolution in breast cancer treatment this is a long time.

- Belgian survival data supplemented by French 10 year survival data coming from[87]. The problem of the evolution in breast cancer treatment apply here as for the British data.

We did not use US SEER data as they use an outdated staging method, so that survival curves per stage are not comparable to the other sources and difficult to incorporate in the model.

The survival curves can be found in the annex.

3.3.8. QALY

Number of life years was calculated for each stage and a stage and this was adjusted for the quality of life (QALYs), based on a literature search (see point 2.3).

We made some assumptions:

- Utility values at start of the model (before screening) were stratified by age but percentage changes relative to these values were assumed to not vary according to the age of the women (we did not have data on this). For the sensitivity analysis, we apply a 20% reduction or increase.
- Patients with negative results had utility values equal to the general population.
- In the assessment of utility values for true negative and false positive results, mobility and ability of self-care were assumed to be unaffected by screening.
- In the study of Lidgren et al.[80], non-metastatic patients were divided in only two groups, i.e. the first year of diagnosis and the following years. Therefore patients in stage I, II, III (grouped as non metastatic patients) were assumed to have the same utility. This assumption is supported by the fact that for years 2001-2006, the treatment was the same for patients in stage I, II, III according to the KCE report on quality indicators in breast cancer[86]. Note that more recent data may change this picture because many cancer found by screening are now treated by conservative surgery. Nevertheless, data to prove this assumption are not available and we found no study comparing the impact of partial versus total mastectomy on quality of life corresponding to our inclusion criteria.
- Utility values for non-metastatic patients (after the year of surgery) and metastatic patients were assumed to remain constant across years. For non-metastatic patients, this assumption is supported by an US study showing no significant differences at year 5, 10, and 15[81]. Nevertheless, as a sensitivity analysis we apply a 20% decrement in QALYs for taking into account a variation of utility values across years.
- At baseline, we did not discount QALYs. For the sensitivity analysis, discount rates of 1.5%, 3% and 5% were applied.

Parameters used in the model are shown in table 3.2.

Table 3.2: Parameters used in the model

	Parameters	No screening	Base case	Sensitivity analysis	
3.3.1	Age specific overall survival	Belgian life-table	Belgian life-table	Belgian life-table	
3.3.2	Breast cancer incidence	BCR data (Flanders population, 2004-2008)	BCR data (Flemish population, 2004-2008) increased by lead time. over-diagnosis invasive cancer of DCIS	BCR data (Flemish population, 2004-2008) increased by lead time. over-diagnosis invasive cancer of DCIS	
	Lead time		2 years	3 years	
	Over-diagnosis invasive cancer		10.0%	range from 3 to 30%	
	Over-diagnosis DCIS		40/100 000 women per year	40/100 000 women per year	
3.3.3	Participation rate		70.0%	range from 60% to 80%	
3.3.4	Proportion of screened detected cancers		70.0%	range from 60% to 80%	
3.3.5	Recall rate		3.5% (Flemish screening program)	range from 2% to 10%	
	Duration of period after positive result		45 days	range from 36 to 54 days	
	QALYs lost in this period		16.0%	estimated between 13% to 19%	
				Scenario 2	Scenario 3
3.3.6	Stage distribution	BCR data (Flemish population, 70-74years, 2004-2008)	Data of the DNETB screening report 2009	Stage distribution of Flemish screening programme (50-69)	Stage distribution of BCR (Flemish women 60-69 (screened and not screened)
	Stage I	31.6%	80.0%	62.5%	45.7%
	Stage II	42.3%	18.7%	31.6%	35.9%

	Stage III	16.6%	0.8%	5.6%	12.5%
	Stage IV	9.5%	0.5%	0.4%	5.9%
3.3.7	Stage specific relative survival	Belgian stage specific annual survival data supplemented until 7 years by Dutch data	Belgian stage specific annual survival data supplemented until 7 years by Dutch data	Dutch survival or British survival or Belgian/French survival	
3.3.8	QALY				
	Stage II III IV	-constant	- constant	-20.0%	
	Stage III IV	-constant	-constant	-20.0%	
	Stage IV	-constant	-constant	-20.0%	
	Age related QALY	-constant	-constant	range from + 20% to -20%	
	Discounted QALY			discounting rate + 1.5%. 3% and 5%	

3.4. Results

In the baseline scenario the model predicts that there would be 1307 years of life saved per 100 000 (13,1 per 1000) women invited for screening and 395 per 100 000 (3.9 per 1000) QALYs. The model also predicts that 128 deaths would be averted per 100 000 women screened (1.3 per 1000), being a reduction of 21% (number needed to be offered screening: 782).

Because of the considerable uncertainty surrounding the parameters and model structure we did an extensive sensitivity analysis. Most uncertainty is not due to random error but due to issues relating to the right choice of source of information on the parameter. We did not do a probabilistic sensitivity analysis, as it was not possible to choose appropriate probability-distributions in a meaningful way.

Table 3.4 shows the results of the sensitivity analysis, the plausible ranges used for this analysis was discussed in 3.3, description of the parameters and justification of chosen values.

The number of years of life gained remains fairly constant under different assumptions. The number of QALY gained or lost varies much more under different assumptions. This is partly due to the fact the a lot of the uncertain or variable parameters have an impact on the quality of life gained rather than on mortality, such as high recall rates, over diagnosis, apart from the values accorded to the QALYs.

Assumed degree of over-diagnosis has a strong impact on QALYs gained and under the higher assumed values of 20% or 30% even imply that QALY would be lost instead of gained. Years of life gained increases slightly, this due to the fact that an over diagnosed case cannot become a new case in the model, one could argue that this is somewhat of an artifact but the effect is very small.

Recall rates of 10% also can shift the balance towards a loss of QALYs. Ten per cent recall rates are actually found in some parts of Belgium.

Assumptions on the choice of the appropriate survival curve have both an impact on years of life gained and QALYs gained. The Dutch and the British survival data increase the number of life years gained but lead to a loss of QALYs in certain scenarios. Belgian survival data supplemented by French 10 year survival is somewhere in between.

Increasing the assumed lead time to 3 years has an impact on both years of life gained and QALYs.

The model's estimation of the number of QALYs gained or lost depends on the valuation of these QALYs. Diminishing the age related QALYs, this is the decrease in quality of life due to old age, decreases the number of QALYs gained, as could be expected.

The estimations coming from the Lidgren's paper are fairly uniform and do not vary much in function of the different stages. We introduced a larger decrement in quality of life due to increasing stage at diagnosis, with 3 scenarios: (i) decreasing stage II III and IV with 20%, (ii) decreasing stage III and IV with 20% or (iii) decreasing stage IV with 20%. This has the effect of increasing the number of QALYs gained, because there are also gains in QALYs due to the stage shift alone outside the effect on mortality, as persons in stage I have in this scenario a better assumed quality of life, in contrast to the Lidgren data.

As could be expected, introducing discount rates decreases the number of QALYs gained.

As a worst case scenario, we set the estimation of over diagnosis at 20%, recall rate at 10%, loss of QALYs per recall at 0.19 during a period of 54 days and using the stage distribution coming from the Flemish screening program (scenario 2). This gives a gain of 872 Years of Life but a loss of 307 QALYs per 100 000.

As a best case scenario, we set the estimation of over-diagnosis at 3%, recall rate at 2%, loss of QALYs per recall at 0.13 during a period of 36 days and using the stage distribution coming from the Dutch screening program (scenario 1). This gives a gain of 1704 Years of Life and a gain of 1626 QALYs per 100 000.

Applying the 22% reduction in mortality from the meta-analysis from Götzsche et al. to the Belgian life table, as described above, gives a very similar result, 139 cancers deaths due to breast cancer avoided and 1145 years of life saved.

Table 3.3 Modeling results: baseline, worst and best case scenario.

Scenario	Assumptions	Years of life Per 100 000 women	Quality adjusted years of life Per 100 000 women
Baseline	Over diagnosis: 10% recall rate at 3.5% loss of QALYs per recall 0.16 during a period of 45 days stage distribution coming Dutch screening program (scenario 1)	1307 gained	395 gained
Worst case	Over diagnosis: 20% recall rate at 10% loss of QALYs per recall 0.19 during a period of 54 days stage distribution coming Flemish screening program (scenario 2)	872 gained	307 lost
Best case	Over diagnosis: 3% recall rate at 2% loss of QALYs per recall 0.13 during a period of 36 days stage distribution coming Dutch screening program (scenario 1)	1704 gained	1626 gained

Table 3.4 Modeling results: sensitivity analysis.

	Stageshift scenario 1		Stageshift scenario 2		Stageshift scenario 3	
	Years of life	QALYs	Years of life	QALYs	Years of life	QALYs
Baseline	1307	395	1014	186	1246	420
Assumed overdiagnosis						
0.03	1304	526	1011	317	1245	551
0.05	1305	489	1012	280	1245	514
0.1	1307	395	1014	186	1246	420
0.2	1310	208	1018	0	1249	232
0.3	1314	22	1022	-187	1251	45
Recall rate						
0.02	1307	442	1014	234	1246	459
0.035	1307	395	1014	186	1246	420
0.05	1307	348	1014	139	1246	380
0.1	1307	190	1014	-19	1246	249
Period between false positive and confirmation test: duration						
36 days	1307	417	1014	208	1246	438
45 days	1307	395	1014	186	1246	420
54 days	1307	373	1014	164	1246	401
Period between false positive and confirmation test:% QALYs lost						
QALYs loss period 13%	1307	416	1014	207	1246	449
QALYs loss period 16%	1307	395	1014	186	1246	434
QALYs loss per period 19%	1307	374	1014	166	1246	420
Participation rate						
0.6	1120	281	869	102	na	na

0.7	1307	395	1014	186	na	na
0.8	1493	509	1159	270	na	na
Effectiveness screening amongst participants						
0.6	1120	281	869	102	na	na
0.7	1307	395	1014	186	na	na
0.8	1493	509	1159	270	na	na
Survival curve by stage from other sources						
Dutch survival	1607	710	1181	310	1481	505
Britisch survival	1714	399	1148	-3	1585	374
Belgian survival supplemented by French data	1460	473	1045	96	1365	477
Assumed lead time 3 years	1098	118	875	77	1169	187
All QALYs minus 20%	1307	-948	1014	-1089	1246	-787
All QALYs plus 20%	1307	1587	1014	1310	1246	1534
Stage II III IV -20	1307	903	1014	465	1246	na
Stage III IV -20	1307	648	1014	370	1246	na
Stage IV -20	1307	450	1014	241	1246	na
Discounted QALYs						
Discount rate 1.5% for QALYs	1307	297	1014	121	1246	274
Discount rate 3% for QALYs	1307	215	1014	67	1246	193
Discount rate 5% for QALYs	1307	138	1014	15	1246	114

3.5. Discussion

Under baseline assumptions, screening in the age group 70-74 has a limited impact on breast cancer deaths avoided and number of years of life saved, amounting to 1.4 death avoided per 1000 women offered screening in that period and 13 years of life saved per 1000 women, amounting to 4.7 days of life gained per women offered screening.

This results fall within the range that the modelers of the CISNET project found as reported by Mandelblatt et al in 2009[48], where years of life gained ranged from 9 to 22 per thousand women screened. This, despite the fact that completely different data and model structures were used.

Years of life gained remained fairly constant in the sensitivity analysis. We choose a worst and best case scenario, with years of life gained ranging from 872 to 1704. It corresponds also with the simplified estimation based on the meta-analysis of Götzsche et al.[4] and the Belgian life tables, despite that this estimation comes from a completely different source of data and estimation method. This indicates that the estimations of the years of life gained are fairly robust and consistent with other studies.

The gain in quality adjusted life years (QALYs) is considerably less, with only 3.9 QALYs per 1000 women (1.4 quality adjusted day of life per women) offered screening and uncertainty is larger. One can present these data in another way by stating that 250 women need to be offered screening for 5 years to gain one year of life. The sensitivity analysis shows that under certain assumptions introducing breast cancer screening in this age group would actually generate a loss of QALYs. The most important of these is an assumed recall rate of 10%, as is the case in certain parts of Belgium, so these high recall rates should certainly first be addressed before proceeding.

The worst case scenario would imply a loss of 3 QALYs per 1000 women screened, we made sure that the assumptions of this worst case scenario are still reasonable assumptions and not unduly extreme. A number of elements were not considered in the worst case scenario because the effect is sometimes mixed. Bringing down the baseline estimation for the quality of life per age-group increases or decreases the final number of QALYs gained depending on the chosen values of the other parameters. This is due to the fact that introducing screening induces losses due to false positives and over diagnosis but gains due to the stage shift.

Under the best case scenario one would gain 16 QALYs per 1000 women screened.

The higher variability seen in the estimation of the QALYs compared to years of life lost has a number of reasons. There is considerable uncertainty around key parameters that determine a loss of QALYs, in particular concerning over-diagnosis. Variability due to recall rates on the other hand rather reflects real underlying differences in practice between countries and in Belgium between regions. There is also considerable uncertainty around the valuation of the quality of life, and this is not only uncertainty concerning the quality of life surrounding different breast cancer states but also age specific quality of life in Belgium. If in the future cost-effectiveness analyses are considered, this problem should be addressed if we want to have meaningful results. Quality of life attributed to stage IV has only a limited impact on overall numbers of QALYs gained or lost as survival in this stage is short and proportion of stage IV patients is low.

Carles et al. 2011[38], in an adaptation of the CISNET model of Lee & Zeelen, found an incremental benefit for biannual screening 50-74 of 2.78 life years gained per 1000 women compared to a schedule 50-69. They did not report the QALYs gained with extending the screening to 50-74 from 50-69, as it was dominated by screening from 45- 69, but reported that 1.86 QALYs per 1 000 were gained by extending the screening to 45-74 from 45-69. Interestingly, they did not incorporate the results of Vilaprinyo et al, 2011[58] into their calculations, but used US survival data.

The model takes into account over-diagnosis and lead time bias. However, it does not take into account length bias, the fact that screen-detected cancers would have a slower clinical course and have a better survival because screening tends to pick-up slow growing tumors some of which are not life threatening. Follow up studies of screen detected cancers and non screen detected cancer in the literature show that survival of screen detected cases is better than cases among non participants, independent of stage, and that survival of interval cases is somewhat in between[88-91]. This indicates that there may indeed be a length time bias, through selection of less aggressive cancers by screening. However, the fact that interval cancers have a better survival than cancer among women not attending screening indicates that other factors also play a role, such as

selection bias (such as the social class or other health related factors amongst women non attending screening) and residual confounding after adjustment for stage. However length time has no direct impact on efficacy and on effectiveness, as detection of slow growing tumors does not necessarily negatively correlate with the ability to detect potentially life-threatening cancers at an earlier stage. Length time has indeed a negative impact on efficiency, as detection of indolent tumors means more harm and greater cost for no benefit to women.

We unfortunately did not have data on stage specific survival for tumors in unscreened women, tumors found by screening and interval cancers. Moreover, there is in general considerable uncertainty around the survival curves that will apply in the future, as treatment evolves and actual data on survival may be outdated.

Another major source of uncertainty is the right choice of stage distributions of the diagnosed cancers and stage-shift. The Flemish data on stage distribution show that the stage distributions of the interval cancers and the cancer amongst the people who do not participate in the screening are very similar.

We choose a modeling approach that is essentially based on the stage shift and its consequences, in contrast to most CISNET models that are essentially tumor growth models. This has the advantage that it allowed us to stay closer to the data and make less use of unobserved variables, incorporating parameters based on Belgian data, but has the disadvantage that the model is less flexible and has more simplifications. We model an overall effect of screening on the proportion of cancers that are screen detected based on Belgian data in the group 50-69. This implies that we can only evaluate the effect of the screening schedules actually in place in Belgium, we cannot vary the screening interval. We do not have the data however needed to parameterize the CISNET models and would be forced to use the same parameters that are already used in the published models, we would just merely replicate them.

In conclusion, there is considerable structural uncertainty around the right choice of the parameters, so a lot of caution is needed when interpreting the results. This uncertainty is reflected in the wide range of estimated Years of Life gained and QALYs gained in the end result. Nevertheless, there is evidence that continuing screening until the age of 74 years has modest effect on the number of Life Years Saved but there is considerable uncertainty on the effect on quality adjusted life years, and the data show that under reasonable assumptions the intervention may even lead to a loss of quality adjusted life years. It is important to bring the recall rates to acceptable level before extending screening.

4. ANSWER TO CLINICAL QUESTIONS

What are clinical benefits and specific harms of an extension of breast cancer organized screening in women between 70 and 74 years?

4.1. Breast cancer related mortality

What is the effect of an extension (70-74 years) of breast cancer organized screening on the breast cancer related mortality? The continued screening for breast cancer between the ages of 70 and 74 makes it possible to obtain an extra 13 years of life for 1,000 women screened. The model also predicts that 128 deaths would be averted per 100 000 women screened (1.3 per 1000), being a reduction of 21%.

4.2. Delay between the screening and the mortality reduction

How long is the delay between the screening and the associated breast cancer related mortality reduction? The mortality reduction appears between 4 and 7 years after screening

4.3. Overall mortality

What is the effect of an extension (70-74 years) of breast cancer organized screening on the overall mortality? The effect of an extension (70-74 years) of breast cancer organized screening on the overall mortality is unclear. Studies did not have statistical power to detect an all-cause mortality reduction.

4.4. Morbidity

What is the effect of an extension (70-74 years) of breast cancer organized screening on morbidity? We found no data related to the cancer morbidity in randomized control trials. In other words, on this basis we do not accept or reject the hypothesis that screening reduces the morbidity of the breast cancer disease. Aim of screening is to detect minor tumors. Consequently, morbidity may be diminish by less aggressive treatment. The Belgian data currently at our disposal do not enable us to ratify this assertion. Actually, the most recent data (KCE report 150)[86] show 58% of the interventions are conservative surgery versus 38% of total mastectomies in the least advanced stages (C Stage I and II). Nearly 90% of patients undergoing conservative surgery also receive radiotherapy treatment, 38% are given a treatment of neo-adjuvant chemotherapy, and 41% receive hormone treatment.

4.5. False positive or false negative results

What are the specific harms in terms of false positive or false negative results? The Belgian data currently at our disposal show a recall rate of 3,5% in Flanders and of 10% in Walloon and Brussels region per screening round. At this age group, performance of mammography is high and rates of false negative results are relatively low. For USA, rate of false negative results are 1.5 per 1000 women aged 70 to 79 years per screening round (BCSC-USA).

4.6. Additional diagnostic tests

What are the specific harms in terms of additional diagnostic tests? Twenty to forty additional punctures or biopsies may be expected per 1000 women offered screening (three rounds).

4.7. Over-diagnosis and over-treatment

What are the specific harms in terms of over-diagnosis and over-treatment? Based on selected studies, over-detection (excluding DCIS cases), ranged from (7% to 21%) to 35% (no data specific for women aged 70 to 79 years are available). Götzsche reported that the number of mastectomies and lumpectomies was significantly larger in the screened groups (no data specific for women aged 70 to 79 years are available). Three trials with adequate randomization showed a significant increase in mastectomies and lumpectomies (Relative Risk (RR) 1.31, 95% (CI) 1.22 to 1.42). Two trials with suboptimal randomization showed the same increase in interventions (RR of 1.42 (95% CI 1.26 to 1.61)). The RR for all five trials combined was 1.35 (95% CI 1.26 to 1.44).

4.8. What attitude should be recommended for women in case of self referral?

It is advisable that when a patient asks her doctor for a screening, the doctor should develop a strategy minimizing the drawbacks of screening[92]. In this way, an attitude structured around three phases can be recommended:

- Information specific to the age bracket[93]
- Decision making according to the patient personal assessment[94]
- Steering of the person who so wishes towards a screening involving methods that minimize the drawbacks.

The criteria defined in the framework of the European Programme notably make provision for the monitoring of the technical quality of the equipment used, the double reading of the mammographies, and an optimization of the recall rate[95]. In Belgium, the approved mammography units meet the criteria laid down in the context of the European Programme, and it is therefore logical to steer those women who explicitly request a screening towards these structures.

5. REFERENCES

1. Paulus D, Mambourg F, Bonneux L. [Breast cancer screening]. Good Clinical Practice (GCP). Brussels: Belgian Health Care Knowledge Centre (KCE); 2005 02/05/2005. KCE reports 11 Available from: http://kce.fgov.be/index_en.aspx?SGREF=5221&CREF=9348
2. Verleye L, Desomer A, Gailly J, Robays j. [Identifying women at risk for breast cancer/technical methods for breast cancer screening]. Good clinical Practice (GCP). Bruxelles: Centre fédéral d'expertise des soins de santé (KCE); 2012. KCE Reports 172 Available from: http://www.kce.fgov.be/publication/report/identifying-women-at-risk-for-breast-cancertechnical-methods-for-breast-cancer-sc
3. Mambourg F, Robays j, Camberlin C, Vlayen J, Gailly J. [Breast cancer screening with mammography for women in the agegroup of 40-49 years]. Good Clinical Practice (GCP). Brussels: Belgian Health Care Knowledge Centre (KCE); 2010 07/07/2010. KCE reports 129 Available from: http://kce.fgov.be/index_en.aspx?SGREF=14851&CREF=16581
4. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane Database of Systematic Reviews. 2011(1).
5. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. Ann Intern Med. 2009;151(10):727-37, W237-42.
6. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. Lancet Oncol. 2007;8(12):1129-38.
7. Jorgensen KJ, Götzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. Bmj. 2009;339.
8. Virnig BA, Tuttle TM, Shamliyan T, Kane RL. Ductal carcinoma in situ of the breast: a systematic review of incidence, treatment, and outcomes. J Natl Cancer Inst. 2010;102(3):170-8.
9. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med. 2002;137(5 Part 1):347-60.
10. Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O, et al. Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. Lancet. 1985;1(8433):829-32.
11. Dixon JM. Breast screening has increased the number of mastectomies. Breast Cancer Res. 2009;11 Suppl 3:S19.
12. Smith RA, Duffy SW, Gabe R, Tabar L, Yen AM, Chen TH. The randomized trials of breast cancer screening: what have we learned? Radiol Clin North Am. 2004;42(5):793-806, v.
13. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet. 2002;359(9310):909-19.
14. Tabar L, Vitak B, Chen TH, Yen AM, Cohen A, Tot T, et al. Swedish Two-County Trial: Impact of Mammographic Screening on Breast Cancer Mortality during 3 Decades. Radiology. 2011.
15. Maass N, Alkasi O, Bauer M, Jonat W, Souchon R, Meinhold-Heerlein I. Actual management of ductal carcinoma in situ of the breast. Arch Gynecol Obstet. 2009;280(5):699-705.
16. Duffy SW, Chen HH, Tabar L, Day NE. Estimation of mean sojourn time in breast cancer screening using a Markov chain model of both entry to and exit from the preclinical detectable phase. Stat Med. 1995;14(14):1531-43.
17. Fracheboud J, Groenewoud JH, Boer R, Draisma G, de Bruijn AE, Verbeek AL, et al. Seventy-five years is an appropriate upper age limit for population-based mammography screening. Int J Cancer. 2006;118(8):2020-5.

18. Fett MJ. Computer modelling of the Swedish two county trial of mammographic screening and trade offs between participation and screening interval. Journal of Medical Screening. 2001;8(1):39-45.
19. Shen Y, Zelen M. Robust modeling in screening studies: estimation of sensitivity and preclinical sojourn time distribution. Biostatistics. 2005;6(4):604-14.
20. Wu JC, Hakama M, Anttila A, Yen AM, Malila N, Sarkeala T, et al. Estimation of natural history parameters of breast cancer based on non-randomized organized screening data: subsidiary analysis of effects of inter-screening interval, sensitivity, and attendance rate on reduction of advanced cancer. Breast Cancer Research & Treatment. 2010;122(2):553-66.
21. Olsen AH, Agbaje OF, Myles JP, Lynge E, Duffy SW. Overdiagnosis, sojourn time, and sensitivity in the Copenhagen mammography screening program. Breast Journal. 2006;12(4):338-42.
22. Frisell J, Eklund G, Hellstrom L, Somell A. Analysis of interval breast carcinomas in a randomized screening trial in Stockholm. Breast Cancer Research & Treatment. 1987;9(3):219-25.
23. Paci E, Duffy SW. Modelling the analysis of breast cancer screening programmes: sensitivity, lead time and predictive value in the Florence District Programme (1975-1986). International Journal of Epidemiology. 1991;20(4):852-8.
24. Boer R, de Koning HJ, van der Maas PJ. A longer breast carcinoma screening interval for women age older than 65 years? Cancer. 1999;86(8):1506-10.
25. Duffy SW, Day NE, Tabar L, Chen HH, Smith TC. Markov models of breast tumor progression: some age-specific results. Journal of the National Cancer Institute. Monographs. 1997;22:93-7.
26. Duffy SW, Gabe R. What should the detection rates of cancers be in breast screening programmes? British Journal of Cancer. 2005;92(3):597-600.
27. Shen Y, Zelen M. Screening sensitivity and sojourn time from breast cancer early detection clinical trials: mammograms and physical examinations. Journal of Clinical Oncology. 2001;19(15):3490-9.
28. Spratt JS, Greenberg RA, Heuser LS. Geometry, growth rates, and duration of cancer and carcinoma in situ of the breast before detection by screening. Cancer Research. 1986;46(2):970-4.
29. Tabar L, Fagerberg G, Chen HH, Duffy SW, Smart CR, Gad A, et al. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. Cancer. 1995;75(10):2507-17.
30. Weedon-Fekjaer H, Lindqvist BH, Vatten LJ, Aalen OO, Tretli S. Estimating mean sojourn time and screening sensitivity using questionnaire data on time since previous screening. Journal of Medical Screening. 2008;15(2):83-90.
31. Weedon-Fekjaer H, Vatten LJ, Aalen OO, Lindqvist B, Tretli S. Estimating mean sojourn time and screening test sensitivity in breast cancer mammography screening: new results. Journal of Medical Screening. 2005;12(4):172-8.
32. Zappa M, Visioli CB, Ciatto S. Mammography screening in elderly women: efficacy and cost-effectiveness. Review 16 refs. Critical Reviews in Oncology Hematology. 2003;46(3):235-9.
33. Boer R, de Koning HJ, van Oortmarssen GJ, van der Maas PJ. In search of the best upper age limit for breast cancer screening. Eur J Cancer. 1995;31A(12):2040-3.
34. Mandelblatt JS, Silliman R. Hanging in the balance: making decisions about the benefits and harms of breast cancer screening among the oldest old without a safety net of scientific evidence. J Clin Oncol. 2009;27(4):487-90.
35. Karnon J, Goyder E, Tappenden P, McPhie S, Towers I, Brazier J, et al. A review and critique of modelling in prioritising and designing screening programmes. Health Technol Assess. 2007;11(52):iii-iv, ix-xi, 1-145.
36. Barratt A, Howard K, Irwig L, Salkeld G, Houssami N. Model of outcomes of screening mammography: information to support informed choices. BMJ. 2005;330(7497):936.
37. Berry DA, Inoue L, Shen Y, Venier J, Cohen D, Bondy M, et al. Modeling the impact of treatment and screening on U.S. breast

cancer mortality: a Bayesian approach. J Natl Cancer Inst Monogr. 2006;Monographs.(36):30-6.

38. Carles M, Vilaprinyo E, Cots F, Gregori A, Pla R, Roman R, et al. Cost-effectiveness of early detection of breast cancer in Catalonia (Spain). BMC Cancer. 2011;11.

39. Carter KJ, Castro F, Kessler E, Erickson BA. Simulation of begin and end ages for mammography screening. J Healthc Qual. 2005;27(1):40-7.

40. Feuer EJ. Modeling the impact of adjuvant therapy and screening mammography on U.S. breast cancer mortality between 1975 and 2000: introduction to the problem. J Natl Cancer Inst Monogr. 2006;Monographs.(36):2-6.

41. Fryback DG, Stout NK, Rosenberg MA, Trentham-Dietz A, Kuruchittham V, Remington PL. The Wisconsin Breast Cancer Epidemiology Simulation Model. J Natl Cancer Inst Monogr. 2006;Monographs.(36):37-47.

42. Gyrd-Hansen D. Cost-benefit analysis of mammography screening in Denmark based on discrete ranking data. Int J Technol Assess Health Care. 2000;16(3):811-21.

43. Hanin LG, Miller A, Zorin AV, Yakovlev AY. The University of Rochester model of breast cancer detection and survival. J Natl Cancer Inst Monogr. 2006;Monographs.(36):66-78.

44. Lee S, Zelen M. A stochastic model for predicting the mortality of breast cancer. J Natl Cancer Inst Monogr. 2006;Monographs.(36):79-86.

45. Lee SJ, Zelen M. Modelling the early detection of breast cancer. Ann Oncol. 2003;14(8):1199-202.

46. Mahnken JD, Chan W, Freeman DH, Jr., Freeman JL. Reducing the effects of lead-time bias, length bias and over-detection in evaluating screening mammography: a censored bivariate data approach. Stat Methods Med Res. 2008;17(6):643-63.

47. Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. J Natl Cancer Inst Monogr. 2006;Monographs.(36):47-55.

48. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms.[Erratum appears in Ann Intern Med. 2010 Jan 19;152(2):136]. Ann Intern Med. 2009;151(10):738-47.

49. Mandelblatt JS, Schechter CB, Yabroff KR, Lawrence W, Dignam J, Extermann M, et al. Toward optimal screening strategies for older women: Costs, benefits, and harms of breast cancer screening by age, biology, and health status. J. Gen. Intern. Med. 2005;20(6):487-96.

50. Neeser K, Szucs T, Bulliard JL, Bachmann G, Schramm W. Cost-effectiveness analysis of a quality-controlled mammography screening program from the Swiss statutory health-care perspective: quantitative assessment of the most influential factors (Structured abstract). 2007;10(1):42-53.

51. Plevritis SK, Sigal BM, Salzman P, Rosenberg J, Glynn P. A stochastic simulation model of U.S. breast cancer mortality trends from 1975 to 2000. J Natl Cancer Inst Monogr. 2006;Monographs.(36):86-95.

52. Rauner MS, Gutjahr WJ, Heidenberger K, Wagner J, Pasia J. Dynamic Policy Modeling for Chronic Diseases: Metaheuristic-Based Identification of Pareto-Optimal Screening Strategies. Operations Research. 2010;58(5):1269-86.

53. Rijnsburger AJ, van Oortmarssen GJ, Boer R, Draisma G, To T, Miller AB, et al. Mammography benefit in the Canadian National Breast Screening Study-2: a model evaluation. Int J Cancer. 2004;110(5):756-62.

54. Rojnik K, Naversnik K, Mateovic-Rojnik T, Primiczakelj M. Probabilistic cost-effectiveness modeling of different breast cancer screening policies in Slovenia. Value Health. 2008;11(2):139-48.

55. Rue M, Vilaprinyo E, Lee S, Martinez-Alonso M, Carles MD, Marcos-Gragera R, et al. Effectiveness of early detection on breast

cancer mortality reduction in Catalonia (Spain). BMC Cancer. 2009;9(326).

56. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. J Natl Cancer Inst. 2006;98(11):774-82.

57. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. J Natl Cancer Inst Monogr. 2006;Monographs.(36):56-65.

58. Vilaprinyo E, Rue M, Marcos-Gragera R, Martinez-Alonso M. Estimation of age- and stage-specific Catalan breast cancer survival functions using US and Catalan survival data. BMC Cancer. 2009;9(98).

59. Wang H, Karesen R, Hervik A, Thoresen SA. Mammography screening in Norway: Results from the first screening round in four counties and cost-effectiveness of a modeled nationwide screening. Cancer Causes Control. 2001;12(1):39-45.

60. Messecar DC. Mammography screening for older women with and without cognitive impairment. J Gerontol Nurs. 2000;26(4):14-24; quiz 52-3.

61. Wen YP, Sheu ML. A cost-benefit analysis of preventive care: The case of breast cancer screening. Taiwain J. Public Health. 2005;24(6):519-28.

62. Advisory Committee on Breast Cancer S. Screening for breast cancer in England: past and future. J Med Screen. 2006;13(2):59-61.

63. Anonymous. Landelijk bevolkingsonderzoek naar borstkanker volledig ingevoerd; resultaten van de implementatiefase 1990-1997. Landelijk Evaluatie Team voor bevolkingsonderzoek naar Borstkanker. Ned Tijdschr Geneeskd. 2000;144(23):1124-9.

64. Barratt A, Irwig L, Glasziou P, Salkeld G, Houssami N, Kerlikowske K, et al. Relative benefit of mammography reduces with age. Evid.-Based Healthc. 2002;6(4):156-7.

65. Barratt AL, Les Irwig M, Glasziou PP, Salkeld GP, Houssami N. Benefits, harms and costs of screening mammography in women 70 years and over: a systematic review. Med J Aust. 2002;176(6):266-71.

66. Bonneux L. De voor- en nadelen van borstkankerscreening: tijd voor evidence-based informatie. Nederlands Tijdschrift voor Geneeskunde. 2009;153.

67. Caplan LS. To screen or not to screen: the issue of breast cancer screening in older women. Public Health Rev. 2001;29(2-4):231-40.

68. Carney PA, Abraham LA, Miglioretti DL, Yabroff KR, Sickles EA, Buist DSM, et al. Factors associated with imaging and procedural events used to detect breast cancer after screening mammography. Am. J. Roentgenol. 2007;188(2):385-92.

69. De Koning HJ. Breast cancer screening; cost-effective in practice? Eur J Radiol. 2000;33(1):32-7.

70. Feuer EJ, Etzioni R, Cronin KA, Mariotto A. The use of modeling to understand the impact of screening on U.S. mortality: examples from mammography and PSA testing. Stat Methods Med Res. 2004;13(6):421-42.

71. Grivegnee AR, Autier P. Approche economique du depistage du cancer du sein en Belgique. Rev Med Brux. 2001;22(4):A277-81.

72. Habbema JD, Tan SY, Cronin KA. Impact of mammography on U.S. breast cancer mortality, 1975-2000: are intermediate outcome measures informative? J Natl Cancer Inst Monogr. 2006;Monographs.(36):105-11.

73. Mandelblatt J, Saha S, Teutsch S, Hoerger T, Siu AL, Atkins D, et al. The cost-effectiveness of screening mammography beyond age 65 years: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med. 2003;139(10):835-42.

74. Prevost TC, Abrams KR, Jones DR. Hierarchical models in generalized synthesis of evidence: an example based on studies of breast cancer screening. Stat Med. 2000;19(24):3359-76.

75. Rautenstrauch J. Is mammography screening only a pointless waste of money? MMW-Fortschr. Med. 2000;142(12):4-10.

76. Xu W, Vnenchak P, Smucny J. Screening mammography in women aged 70 to 79 years. J. 2000;49(3):266-7.
77. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med. 2005;353(17):1784-92.
78. Carter KJ, Castro F, Kessler E, Erickson B. A computer model for the study of breast cancer. Comput Biol Med. 2003;33(4):345-60.
79. Cleemput I, Van Wilder P, Vrijens F, Huybrechts M, Ramaekers D. Recommandations pour les évaluations pharmacoéconomiques en Belgique. Health technology Assessment (HTA). Bruxelles: Centre fédéral d'expertise des soins de santé (KCE); 2008. KCE Reports 78B (D/2008/10.273/24)
80. Lidgren M, Wilking N, Jonsson B, Rehnberg C. Health related quality of life in different states of breast cancer. Qual Life Res. 2007;16(6):1073-81.
81. Freedman GM, Li T, Anderson PR, Nicolaou N, Konski A. Health states of women after conservative surgery and radiation for breast cancer. Breast Cancer Res Treat. 2010;121(2):519-26.
82. Burstrom K, Johannesson M, Diderichsen F. Health-related quality of life by disease and socio-economic group in the general population in Sweden. Health Policy. 2001;55(1):51-69.
83. Gerard K, Johnston K, Brown J. The role of a pre-scored multi-attribute health classification measure in validating condition-specific health state descriptions. Health Econ. 1999;8(8):685-99.
84. Domeyer PJ, Sergentanis TN, Zagouri F, Zografos GC. Health-related quality of life in vacuum-assisted breast biopsy: short-term effects, long-term effects and predictors. Health & Quality of Life Outcomes. 2010;8(11):2010.
85. Borstkanker LETvbn. Landelijke evaluatie van bevolkingsonderzoek naar borstkanker in Nederland 1990-2007. 2010.
86. Stordeur S, Vrijens F, Beirens K, Vlayen J, Devriese S, Van Eycken E. Quality indicators in oncology: breast bancer. Good Clinical Practice (GCP). Brussels: Belgian Health Care Knowledge Centre (KCE); 2010. KCE reports 150C (D/2010/10.273/101) Available from: http://kce.fgov.be/index_en.aspx?SGREF=5211&CREF=18847
87. INC. Survie attendue des patients atteints de cancers en France : état des lieux. 2010.
88. Mook S, Van 't Veer LJ, Rutgers EJ, Ravdin PM, van de Velde AO, van Leeuwen FE, et al. Independent prognostic value of screen detection in invasive breast cancer. J Natl Cancer Inst. 2011;103(7):585-97.
89. Cortesi L, Chiuri VE, Ruscelli S, Bellelli V, Negri R, Rashid I, et al. Prognosis of screen-detected breast cancers: results of a population based study. BMC Cancer. 2006;6:17.
90. Joensuu H, Lehtimaki T, Holli K, Elomaa L, Turpeenniemi-Hujanen T, Kataja V, et al. Risk for distant recurrence of breast cancer detected by mammography screening or other methods. Jama. 2004;292(9):1064-73.
91. Olsson A, Borgquist S, Butt S, Zackrisson S, Landberg G, Manjer J. Tumour-related factors and prognosis in breast cancer detected by screening. Br J Surg. 2012;99(1):78-87.
92. Physicians AAoF. Summary of Recommendations for Clinical Preventive Services. In: AAFP Policy Action AAFP; 2010.
93. Woloshin S, Schwartz LM. The benefits and harms of mammography screening: understanding the trade-offs. JAMA. 2010;303(2):164-5.
94. Jorgensen KJ, Gotzsche PC. Content of invitations for publicly funded screening mammography. BMJ. 2006;332(7540):538-41.
95. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition--summary document. Ann Oncol. 2008;19(4):614-22.

cancer mortality reduction in Catalonia (Spain). BMC Cancer. 2009;9(326).

56. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. J Natl Cancer Inst. 2006;98(11):774-82.

57. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. J Natl Cancer Inst Monogr. 2006;Monographs.(36):56-65.

58. Vilaprinyo E, Rue M, Marcos-Gragera R, Martinez-Alonso M. Estimation of age- and stage-specific Catalan breast cancer survival functions using US and Catalan survival data. BMC Cancer. 2009;9(98).

59. Wang H, Karesen R, Hervik A, Thoresen SA. Mammography screening in Norway: Results from the first screening round in four counties and cost-effectiveness of a modeled nationwide screening. Cancer Causes Control. 2001;12(1):39-45.

60. Messecar DC. Mammography screening for older women with and without cognitive impairment. J Gerontol Nurs. 2000;26(4):14-24; quiz 52-3.

61. Wen YP, Sheu ML. A cost-benefit analysis of preventive care: The case of breast cancer screening. Taiwain J. Public Health. 2005;24(6):519-28.

62. Advisory Committee on Breast Cancer S. Screening for breast cancer in England: past and future. J Med Screen. 2006;13(2):59-61.

63. Anonymous. Landelijk bevolkingsonderzoek naar borstkanker volledig ingevoerd; resultaten van de implementatiefase 1990-1997. Landelijk Evaluatie Team voor bevolkingsonderzoek naar Borstkanker. Ned Tijdschr Geneeskd. 2000;144(23):1124-9.

64. Barratt A, Irwig L, Glasziou P, Salkeld G, Houssami N, Kerlikowske K, et al. Relative benefit of mammography reduces with age. Evid.-Based Healthc. 2002;6(4):156-7.

65. Barratt AL, Les Irwig M, Glasziou PP, Salkeld GP, Houssami N. Benefits, harms and costs of screening mammography in women 70 years and over: a systematic review. Med J Aust. 2002;176(6):266-71.

66. Bonneux L. De voor- en nadelen van borstkankerscreening: tijd voor evidence-based informatie. Nederlands Tijdschrift voor Geneeskunde. 2009;153.

67. Caplan LS. To screen or not to screen: the issue of breast cancer screening in older women. Public Health Rev. 2001;29(2-4):231-40.

68. Carney PA, Abraham LA, Miglioretti DL, Yabroff KR, Sickles EA, Buist DSM, et al. Factors associated with imaging and procedural events used to detect breast cancer after screening mammography. Am. J. Roentgenol. 2007;188(2):385-92.

69. De Koning HJ. Breast cancer screening; cost-effective in practice? Eur J Radiol. 2000;33(1):32-7.

70. Feuer EJ, Etzioni R, Cronin KA, Mariotto A. The use of modeling to understand the impact of screening on U.S. mortality: examples from mammography and PSA testing. Stat Methods Med Res. 2004;13(6):421-42.

71. Grivegnee AR, Autier P. Approche economique du depistage du cancer du sein en Belgique. Rev Med Brux. 2001;22(4):A277-81.

72. Habbema JD, Tan SY, Cronin KA. Impact of mammography on U.S. breast cancer mortality, 1975-2000: are intermediate outcome measures informative? J Natl Cancer Inst Monogr. 2006;Monographs.(36):105-11.

73. Mandelblatt J, Saha S, Teutsch S, Hoerger T, Siu AL, Atkins D, et al. The cost-effectiveness of screening mammography beyond age 65 years: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med. 2003;139(10):835-42.

74. Prevost TC, Abrams KR, Jones DR. Hierarchical models in generalized synthesis of evidence: an example based on studies of breast cancer screening. Stat Med. 2000;19(24):3359-76.

75. Rautenstrauch J. Is mammography screening only a pointless waste of money? MMW-Fortschr. Med. 2000;142(12):4-10.

76. Xu W, Vnenchak P, Smucny J. Screening mammography in women aged 70 to 79 years. J. 2000;49(3):266-7.

77. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med. 2005;353(17):1784-92.

78. Carter KJ, Castro F, Kessler E, Erickson B. A computer model for the study of breast cancer. Comput Biol Med. 2003;33(4):345-60.

79. Cleemput I, Van Wilder P, Vrijens F, Huybrechts M, Ramaekers D. Recommandations pour les évaluations pharmacoéconomiques en Belgique. Health technology Assessment (HTA). Bruxelles: Centre fédéral d'expertise des soins de santé (KCE); 2008. KCE Reports 78B (D/2008/10.273/24)

80. Lidgren M, Wilking N, Jonsson B, Rehnberg C. Health related quality of life in different states of breast cancer. Qual Life Res. 2007;16(6):1073-81.

81. Freedman GM, Li T, Anderson PR, Nicolaou N, Konski A. Health states of women after conservative surgery and radiation for breast cancer. Breast Cancer Res Treat. 2010;121(2):519-26.

82. Burstrom K, Johannesson M, Diderichsen F. Health-related quality of life by disease and socio-economic group in the general population in Sweden. Health Policy. 2001;55(1):51-69.

83. Gerard K, Johnston K, Brown J. The role of a pre-scored multi-attribute health classification measure in validating condition-specific health state descriptions. Health Econ. 1999;8(8):685-99.

84. Domeyer PJ, Sergentanis TN, Zagouri F, Zografos GC. Health-related quality of life in vacuum-assisted breast biopsy: short-term effects, long-term effects and predictors. Health & Quality of Life Outcomes. 2010;8(11):2010.

85. Borstkanker LETvbn. Landelijke evaluatie van bevolkingsonderzoek naar borstkanker in Nederland 1990-2007. 2010.

86. Stordeur S, Vrijens F, Beirens K, Vlayen J, Devriese S, Van Eycken E. Quality indicators in oncology: breast bancer. Good Clinical Practice (GCP). Brussels: Belgian Health Care Knowledge Centre (KCE); 2010. KCE reports 150C (D/2010/10.273/101) Available from: http://kce.fgov.be/index_en.aspx?SGREF=5211&CREF=18847

87. INC. Survie attendue des patients atteints de cancers en France : état des lieux. 2010.

88. Mook S, Van 't Veer LJ, Rutgers EJ, Ravdin PM, van de Velde AO, van Leeuwen FE, et al. Independent prognostic value of screen detection in invasive breast cancer. J Natl Cancer Inst. 2011;103(7):585-97.

89. Cortesi L, Chiuri VE, Ruscelli S, Bellelli V, Negri R, Rashid I, et al. Prognosis of screen-detected breast cancers: results of a population based study. BMC Cancer. 2006;6:17.

90. Joensuu H, Lehtimaki T, Holli K, Elomaa L, Turpeenniemi-Hujanen T, Kataja V, et al. Risk for distant recurrence of breast cancer detected by mammography screening or other methods. Jama. 2004;292(9):1064-73.

91. Olsson A, Borgquist S, Butt S, Zackrisson S, Landberg G, Manjer J. Tumour-related factors and prognosis in breast cancer detected by screening. Br J Surg. 2012;99(1):78-87.

92. Physicians AAoF. Summary of Recommendations for Clinical Preventive Services. In: AAFP Policy Action AAFP; 2010.

93. Woloshin S, Schwartz LM. The benefits and harms of mammography screening: understanding the trade-offs. JAMA. 2010;303(2):164-5.

94. Jorgensen KJ, Gotzsche PC. Content of invitations for publicly funded screening mammography. BMJ. 2006;332(7540):538-41.

95. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition--summary document. Ann Oncol. 2008;19(4):614-22.